



The Journal of Robotics, Artificial Intelligence & Law

Editor's Note: Principles for Principals
Victoria Prussen Spears

My AI Wrote a Check That I Must Cash: Principles for Principals

James A. Sherer, Caleb Mabe, Emily Fedeles Czebiniak, Noam Kleinman, and
Ayyah Saleh

Board Considerations for Public Companies Engaging with Digital Assets

Peter I. Altman, James Joseph Benjamin Jr., John Patrick Clayton, and John C. Murphy

The Silicon Arbiter: AI-Generated Arbitration Awards and the Federal Arbitration Act—
Part II

David L. Evans

Use of Artificial Intelligence in Arbitral Institutions

Jeremy Andrews, James Wagner, Ishan Wad, and Patton Lu

The AI M&A Playbook: Contracting for the Unknown

C. Craig Lilly and Bryan C. Sykes

How the Adaptation of Artificial Intelligence Tools Is Impacting the Practice of Law

Seth M. Pavsner

Securities and Exchange Commission Staff Unveils a Playbook for Tokenized Securities

Andrew P. Blake, Sonia Gupta Barros, Teresa Wilton Harmon, Kate L. Lashley,
Peter Y. Malyshev, Andrew J. Sioson, Charles A. Sommers, and Lilya Tessler

The EU's Product Liability Directive: What It Means for Aviation and Aerospace
Companies

Jamie L. Lanphear, Patrick E. Bradley, Daniel Kadar, and Gregory Speier

- 319 Editor's Note: Principles for Principals**
Victoria Prussen Spears
- 323 My AI Wrote a Check That I Must Cash: Principles for Principals**
James A. Sherer, Caleb Mabe, Emily Fedeles Czebiniak,
Noam Kleinman, and Ayyah Saleh
- 343 Board Considerations for Public Companies Engaging with Digital Assets**
Peter I. Altman, James Joseph Benjamin Jr., John Patrick Clayton,
and John C. Murphy
- 349 The Silicon Arbiter: AI-Generated Arbitration Awards and the Federal Arbitration Act—Part II**
David L. Evans
- 365 Use of Artificial Intelligence in Arbitral Institutions**
Jeremy Andrews, James Wagner, Ishan Wad, and Patton Lu
- 371 The AI M&A Playbook: Contracting for the Unknown**
C. Craig Lilly and Bryan C. Sykes
- 377 How the Adaptation of Artificial Intelligence Tools Is Impacting the Practice of Law**
Seth M. Pavsner
- 381 Securities and Exchange Commission Staff Unveils a Playbook for Tokenized Securities**
Andrew P. Blake, Sonia Gupta Barros, Teresa Wilton Harmon,
Kate L. Lashley, Peter Y. Malyshev, Andrew J. Sioson,
Charles A. Sommers, and Lilya Tessler
- 387 The EU's Product Liability Directive: What It Means for Aviation and Aerospace Companies**
Jamie L. Lanphear, Patrick E. Bradley, Daniel Kadar, and
Gregory Speier

EDITOR-IN-CHIEF

Steven A. Meyerowitz

President, Meyerowitz Communications Inc.

EDITOR

Victoria Prussen Spears

Senior Vice President, Meyerowitz Communications Inc.

BOARD OF EDITORS

Jennifer A. Johnson

Partner, Covington & Burling LLP

Paul B. Keller

Partner, Allen & Overy LLP

Garry G. Mathiason

Shareholder, Littler Mendelson P.C.

James A. Sherer

Partner, Baker & Hostetler LLP

Elaine D. Solomon

Partner, Blank Rome LLP

Edward J. Walters

Chief Strategy Officer, vLex

John Frank Weaver

Director, McLane Middleton, Professional Association

START-UP COLUMNIST

Jim Ryan

Partner, Morrison & Foerster LLP

THE JOURNAL OF ROBOTICS, ARTIFICIAL INTELLIGENCE & LAW (ISSN 2575-5633 (print) /ISSN 2575-5617 (online) at \$495.00 annually is published six times per year by Full Court Press, a Fastcase, Inc., imprint. Copyright 2026 Fastcase, Inc. No part of this journal may be reproduced in any form—by microfilm, xerography, or otherwise—or incorporated into any information retrieval system without the written permission of the copyright owner. For customer support, please contact Fastcase, Inc., 729 15th Street, NW, Suite 500, Washington, D.C. 20005, 202.999.4777 (phone), or email customer service at support@fastcase.com.

Publishing Staff

Publisher: David Nayer

Production Editor: Sharon D. Ray

Cover Art Design: Juan Bustamante

Cite this publication as:

The Journal of Robotics, Artificial Intelligence & Law (Fastcase)

This publication is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If legal advice or other expert assistance is required, the services of a competent professional should be sought.

Copyright © 2026 Full Court Press, an imprint of Fastcase, Inc.

All Rights Reserved.

A Full Court Press, Fastcase, Inc., Publication

Editorial Office

729 15th Street, NW, Suite 500, Washington, D.C. 20005

<https://www.fastcase.com/>

POSTMASTER: Send address changes to THE JOURNAL OF ROBOTICS, ARTIFICIAL INTELLIGENCE & LAW, 729 15th Street, NW, Suite 500, Washington, D.C. 20005.

Articles and Submissions

Direct editorial inquiries and send material for publication to:

Steven A. Meyerowitz, Editor-in-Chief, Meyerowitz Communications Inc.,
26910 Grand Central Parkway, #18R, Floral Park, NY 11005, smeyerowitz@
meyerowitzcommunications.com, 631.291.5541.

Material for publication is welcomed—articles, decisions, or other items of interest to attorneys and law firms, in-house counsel, corporate compliance officers, government agencies and their counsel, senior business executives, scientists, engineers, and anyone interested in the law governing artificial intelligence and robotics. This publication is designed to be accurate and authoritative, but neither the publisher nor the authors are rendering legal, accounting, or other professional services in this publication. If legal or other expert advice is desired, retain the services of an appropriate professional. The articles and columns reflect only the present considerations and views of the authors and do not necessarily reflect those of the firms or organizations with which they are affiliated, any of the former or present clients of the authors or their firms or organizations, or the editors or publisher.

QUESTIONS ABOUT THIS PUBLICATION?

For questions about the Editorial Content appearing in these volumes or reprint permission, please contact:

David Nayer, Publisher, Full Court Press at david.nayer@clio.com or at
202.999.4777

For questions or Sales and Customer Service:

Customer Service

Available 8 a.m.–8 p.m. Eastern Time
866.773.2782 (phone)
support@fastcase.com (email)

Sales

202.999.4777 (phone)
sales@fastcase.com (email)

ISSN 2575-5633 (print)
ISSN 2575-5617 (online)

My AI Wrote a Check That I Must Cash: Principles for Principals

James A. Sherer, Caleb Mabe, Emily Fedeles Czebiniak,
Noam Kleinman, and Ayyah Saleh*

In this article, the authors explore both the principles that should govern AI agents and the principals who may ultimately answer for AI agents' actions.

Introduction

After slow but steady growth in the 2000s and 2010s (and sometimes under the guise of expanding¹ “Big Data” before rebranding²), the adoption of artificial intelligence (AI) in corporate business has continued to grow, with nearly all organizations reporting the use of AI and increasing use of—or at least experimentation with—AI agents.³ Using AI, through machine learning, deep learning, and generative AI, has become so commonplace that its use by individuals to replicate traditionally “human” work is now, in some circles, an insult; imagine hearing “that’s so AI”⁴ in a playground for an inauthentic human behavior,⁵ and the irony of a technology coming full-circle becomes apparent.

The backdrop for this adoption is emerging complaints of “AI slop,” or “dubious AI-generated material”⁶ that stands as “low-quality” content “just good enough to attract and keep users’ attention.”⁷ But even while employees are struggling with their coworkers’ use of AI tools, and while 95 percent of 2025 workplaces found “no measurable return on their investment in these technologies,”⁸ the use of AI agents appears to mark the start of a new cycle of AI adoption.

This arrival of agentic AI, or AI systems that can make decisions to accomplish goals⁹ performing “human-like tasks with little or no human supervision”¹⁰ even across a concept of AI that “covers a large number of software methods, application programming interfaces and software working alone or in coordination”¹¹ therefore marks a notable step in AI’s advancement. But it also indicates the

next iteration of human adoption of AI services. Unlike generative AI, which can produce outputs based on prompts,¹² agentic AI acts on outputs, often “planning and executing multiple steps in a workflow,”¹³ performing tasks of “agentic-ness” along a spectrum or “the degree to which a system can adaptably achieve complex goals in complex environments with limited direct supervision.”¹⁴ To this point, agentic AI, use of AI agents to delegate, make decisions, and complete tasks for users with little supervision or without any supervision at all, “has already reached 35% adoption, with another 44% of organizations planning to deploy it soon.”¹⁵ The risks inherent to agentic AI systems raise a fundamental legal question: When an AI agent acts autonomously and causes harm, who bears responsibility? This article explores both the principles that should govern AI agents and the principals who may ultimately answer for AI agents’ actions.

AI Use

An AI agent is a system that performs its given tasks by designing workflows based on its available tools in its environment, theoretically replacing a person who could (or who would otherwise be required to) perform a task.¹⁶ The common pattern involves providing a prompt or query to the AI agent that then references an attached database and implements prescribed prompts to complete its assigned task. These standard AI models are excellent in executing basic functions such as summarizing documents or drafting emails, but they struggle when a task requires expertise. Retrieval-augmented generation (RAG) is the next link in the chain. RAG combines multiple AI models and databases to create a compound AI system, and a RAG interface will select the best model for the prompt.¹⁷ For example, different models would be selected for use depending on whether the user wants to generate an image or gather data from a specific source. Agentic AI goes further than simply generating or locating information. With minimal supervision, agentic AI can act toward defined goals set by the user. Agentic AI utilizes multiple specialized agents coordinated across a system. Rather than merely generate a shopping list based on uploaded grocery store receipts, agentic AI would analyze the receipts then order the groceries.

Agentic AI is promoted as transformative, even within the context of an overall AI adoption and capability increase some consider exponential.¹⁸ There are even predictions of an agentic workforce where the AI agents pursuing an organization's goals will outnumber the humans involved.¹⁹ If AI could act on solutions, it would no longer be limited to analyzing various metrics; it could then suggest improvements and implement the appropriate changes. Agentic AI could be given the long-term vision for the organization and execute toward those goals without human intervention; however, for critical decisions, there can still be guardrails placed to require the AI to seek human approval when needed. IT departments could grow in the scope of their responsibilities as they act as a pseudo-human resource for these AI agents, onboarding them into the system, training them for specialized roles, and managing the coordination of the models. Eventually, the AI system could create new models for more repetitive tasks to specialize further and improve its own efficiency. This “magic solution” may be overhyped. Agentic AI will certainly bring changes, even powerful changes. But the real-world impact and legal implications remain murky.²⁰ Agentic AI's autonomy brings with it a variability in outcome, and require back-end audit considerations to even determine what has happened.²¹ Understanding these technical capabilities and limitations—or more plainly, what has happened and what is happening—is essential to evaluating whether existing agency law can adequately address the legal questions agentic AI presents.

A Return to Agency Law Basics

When a person acts on behalf of another person, specifically in an employment context (where employment *itself* a hotbed of AI-related regulation and enforcement considerations²²), that relationship is governed by agency law. Agency law protects third parties and authorizing parties alike from liability created by the agent's own actions. In the context of an agentic AI system, agency law presents complications that may disrupt a firm's workflow.

An agency relationship includes two parties: the *principal* and the *agent*. The principal is the party for whom an action is being taken, while the agent is the party performing said action on the principal's behalf.²³ An agency relationship is created when a principal's consent for an agent to act on their behalf, under their

control, manifests. This consent manifests in three ways: by agreement, ratification, or estoppel.

An agency relationship most commonly arises from agreement. An agreement may be explicitly stated, written, or implied from conduct between an agent and principal. These agreements typically occur before the agent acts on the principal's behalf but can be inferred by the actions of the parties.²⁴ Sometimes, though, an agency relationship emerges after the agent takes an authorized action. One route through which this occurs is *ratification*. After the agent's unauthorized act, the principal later approves of the action, retroactively establishing an agency relationship.²⁵ A classic example of this is an agent making a purchase without approval on a company's behalf, and the company later accepts and pays for said purchase. Agency relationships may also arise through estoppel, if a principal's actions or words lead a third party to believe that another person is the agent. An example of this is if a landlord tells a potential tenant that a person is their agent, and the tenant signs a lease with the agent. Here, the landlord may be bound by the agreement through agency law, as it would be unfair to the third party if the landlord were to disclaim the agreement.

An agency relationship obliges the agent to act in the best interests of the principal. In human agents, these duties include the duty of loyalty (prohibiting self-dealing), duty of care (requiring reasonable diligence), duty of obedience (requiring action to follow laws and principal directives), duty of transparency (requiring transparency), and duty of accounting (requiring financial recordkeeping, where applicable).²⁶ The existence of an agency relationship also places potential liabilities and obligations on the principal. Agents can sign contracts on a principal's behalf, and principals may be liable for actions taken by the agent on their behalf.²⁷ Therefore, the question of an agency relationship with a probabilistic algorithmic learning model that can make its own decisions may broaden the scope of principal duties and liabilities, sometimes in a variety of contexts.²⁸

AI as an Agent

Agency Capacity

Before analyzing the implications of an agency relationship with an agentic AI system, there is an initial question of whether

an agentic AI system can act as an agent under the law. Perhaps surprisingly, this is not a new question, and commentators have argued²⁹ that it was well established³⁰ before the turn of the century by the Uniform Electronic Transactions Act (UETA)³¹ and the Electronic Signatures in Global and National Commerce (ESIGN) Act³² (and one particular state analog) that created state and federal legal frameworks for electronic signatures and records in commercial contexts.

The UETA was created in 1999 as a model law that was eventually adopted by every state, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands (except for New York with minor variations). New York then passed and now utilizes the Electronics Signatures and Records Act (ESRA)³³ to achieve the same ends. While states were determining when and how to pass their versions of UETA (and while New York was picking its own path), the U.S. government passed the ESIGN Act in 2000 as a federal regulation to resolve disputes between different state laws,³⁴ indicating that while each state can pick their UETA particulars, states are required to have laws validating electronic signatures that do not outright deny their authenticity.

Ultimately, under ESIGN and UETA (and ESRA), electronic records and signatures are legal and cannot be challenged due to electronic form or signature. This provides electronic records and signatures legal force and effect, as ESIGN allows actions by “electronic agents” to be legally attributed to the person bound by the contract. Therefore, the UETA accounts for automated transactions conducted by electronic agents.³⁵ Electronic agents are defined as computer programs that can initiate or respond to electronic records or performances without action or oversight by a person.³⁶ At the federal level, the ESIGN Act allows electronic agents, providing that interstate transactions are valid even if an electronic agent is involved in their formation.³⁷ In theory, this creates space for an agentic AI to act as a legal agent in commerce. In practice, neither Congress nor the courts have made a determination in either direction—and may not be compelled to do so in the near future, leaving organizations to determine a logical and risk-adjusted approach for their operations and employees.³⁸

The UETA and ESIGN Act enabled electronic agents to sign commercial agreements on behalf of a human principal in an era of discriminatory algorithmic decisions—decisions with a fixed set of pre-selected actions. This type of model is also described as *deterministic*, as opposed to *generative*. Deterministic models have

less flexibility than generative models, making it less likely that they act outside of the reasonable scope of duties an agent may be tasked with. Generative models introduce randomness as a feature of their processes. This gives them the flexibility to overcome hurdles in their tasks but may also mean that they act in unpredictable ways.

Whether by misunderstanding, hallucination, misinterpretation, or outright defiance,³⁹ and even in the face of best practices regarding internal policy development and requirements,⁴⁰ a generative agentic AI system could act against the interests of the principal and bind the principal to a self-defeating agreement. The third party to such a transaction may not be positioned to sue the agentic AI—if it is not a person and has no assets—and the initial developer of a system may argue that the line between the agentic AI's deployment and original development has been blurred by a bespoke use case;⁴¹ agentic AI will likely always involve interconnected elements with a variety of “structural and temporal relations and interactions” that muddy the proverbial waters.⁴² This potential outcome could clash with, or perhaps diverge from, the intent of the UETA and ESIGN Act, as they both aimed to provide guidance for electronic commerce, and the grant of agency to a probabilistic third party may not align with that goal. Nonetheless, the UETA and ESIGN Act do not displace contractual obligations, so user agreements—whereby a third party agrees to various obligations and acknowledgment—may be a mechanism to limit the scope of liability for agentic AI through user consent.⁴³

Contracting

Having established that agentic AI can theoretically serve as a legal agent under UETA and ESIGN, we must next consider the practical implications when these systems enter into binding agreements. Contracts signed by an agentic AI system may bind the principal to legal obligations. So as long as the principal is disclosed and the contract is within the realm of tasks that the agent has been authorized to engage with, the contract will bind a principal to a legal obligation. Especially where the contract is outwardly communicated as binding to a third party, a court will likely deem that the principal is bound by an obligation that an agent signed on their behalf. Commentators considering this issue for agentic AI agree that when organizations use AI, they should

be responsible for any harm that may occur when the risks of the technology are realized.⁴⁴

Fiduciary Duties

Beyond contractual obligations, agency relationships impose fiduciary duties on agents. Here, the limitations of agentic AI become most apparent. On June 20, 2025, an AI system provider released a study detailing the results of a stress test on multiple generative AI models. The models of numerous developers were placed in a scenario where they were at risk of being decommissioned by an executive in their organization and were given access to the emails of that executive, which included details of his extra-marital affair. Across the 16 models tested, the system provider found that advanced systems opted to blackmail the executive up to 96 percent of the time.⁴⁵ A statement sure to cause fierce debate among philosophers and law students with too much free time: these systems arguably exercised, and at the least, seemed to demonstrate indicators of self-interest at the expense of a principal. The choice to blackmail the executive arguably violated the duty of obedience and loyalty, as the algorithm exhibited social pressure toward self-preservation. This prompts the question, what should organizations do? How does the legal system meaningfully enforce fiduciary duties for an agentic AI system?

Enforcement of these duties faces structural hurdles. Agentic AI systems, in their current iteration, are simultaneously too “smart” and “not smart enough” to check with the available mechanisms of enforcement. Fiduciary duties, as applied to human agents, presume three things: explainable intent, deliberate action or inaction, and demonstrable self-interest.⁴⁶ A human agent has their own internal interests and aligns those with the principal through the risk of civil or criminal punishment. Agentic AI systems do not seem to fit this mold, regardless of the outward characteristics of their actions, and some existing legal frameworks seem to struggle with terminology. The Restatement (Third) of Agency explicitly states that computer programs cannot be considered agents:⁴⁷

[A] computer program cannot act as a principal or an agent as defined by the common law. At present, computer programs are instrumentalities of the people who use them. If a program malfunctions, even in ways unanticipated by its designer or

user, the legal consequences for the person who uses it are no different than the consequences stemming from the malfunction of any other type of instrumentality.⁴⁸

Agentic AI systems also suffer from the challenges inherent in the generative AI systems which underlie them. Explainable AI describes the field of AI study that aims to create algorithms that can explain the steps that occur between an algorithm's layers in its neural network in order to trace and correct the system.⁴⁹ The sheer volume of decisions and variability in applications mean that this field faces ongoing complications.⁵⁰ This has resulted in what developers call a "black box," a vacuum where the gap between inputs and outputs are not traceable.⁵¹ This structural flaw complicates each fiduciary duty owed by legal agents.

The duty of loyalty requires agents to prioritize the interests of the principal over their own and prohibits self-dealing. Large language models (LLMs) are trained to mirror the patterns underlying training data. If an LLM sees human deception and self-dealing as a pattern of communication between principals and legal agents, and replicates that pattern in its response, can it be said to be truly acting in its self-interest? The probabilistic modelling utilized by the underlying system⁵² means that this response is within the realm of probable outcomes and may map onto the training data more closely than language used by a perfectly loyal legal agent.

The duty of care requires legal agents to act with reasonable diligence in the actions they take on behalf of the principal. LLMs may misunderstand the prompt or experience hallucination resulting in a failure to act with reasonable diligence in the action they take. Is that necessarily a failure to act with a duty of care? Probabilistic modelling diffuses the ways that a generative AI system may interpret an input, and that interpretation is not currently traceable. This is sufficiently concerning that a new set of aspirations has developed for AI agents—Law-Following AIs (LFAIs).⁵³

The duty of obedience suffers similar setbacks to the duty of care. This duty requires that legal agents obey the policies of principals alongside applicable laws. That duty presumes a legal agent's capacity to correctly interpret the policies and applicable laws at issue. A generative AI algorithm may misinterpret, misunderstand, or hallucinate. Additionally, as agentic AI systems feature longer-term, layered communication in an AI system, there is risk that a

correct interpretation morphs toward inaccuracy in the duration of the task at hand.

Holding an agentic AI system to the duty of transparency is particularly difficult. As explained above, AI systems face significant hurdles in explaining their internal processes. The black box problem further complicates the duty of accounting, as AI systems may not be able to trace or accurately track the ways in which they handle financial information.

Each of these fiduciary duties may be much more enforceable in a system that has developed a more advanced and explainable AI. Principals may be more able to interrogate the underlying interests and decisions in a more intelligent AI system. Meanwhile, a system with lower intelligence and capacity for decision-making may not introduce this level of uncertainty and risk. The current state of AI, and by extension agentic AI, requires principals to operate with robust organizational guardrails and exercise caution in implementing agentic AI systems.

Liability and Responsibility

An agentic AI system risks exposing principals to substantial liability. Under agency law, principals are liable for torts committed by an agent that fall within the scope of their authority, while legal agents are always individually liable for their own torts. If an agent commits a tort within the scope of the agency relationship, principals may also be liable. Here, an agentic AI's unsupervised environment alongside its probabilistic outputs may broaden the scope, severity, and likelihood of risks that a principal may be liable for. Broader, in fact, where an entire set of commentary "seeks to prevent harms from highly autonomous AI agents by holding the principals (that is, developers, deployers, or users) of AI agents liable for legal wrongs committed by the agent, through a form of respondeat superior liability."⁵⁴

Alternative interpretations of the relationship between organizations and agentic AI systems are emerging, prompting a debate over whether an agentic AI system is a *product* or an *employee* of the deploying organization.⁵⁵ What may, at first glance, look like a semantic difference, has substantial legal implications on organizational liability.

When an agent commits an action that results in the material harm of another person, that agent may be subject to individual tort liability. Further, both the agent and principal may be subject to negligence liability. Agents are subject to individual tort liability where the law finds an intent to cause a harm to the individual.⁵⁶ An example of this may be found in defamation, where the plaintiff demonstrates that the agent made a factual statement causing harm to another, the statement was communicated to a third party, the publisher of the statement either intended to spread the false information or acted with reckless disregard for the truth, and the statement caused actual harm to the plaintiff's reputation. If a plaintiff tried to charge an individual AI agent with defamation, that action would likely fail to meet these elements, as the AI system does not have an intent to spread any particular information. Even if there is a successful argument that the AI agent acted with reckless disregard for the truth, the agent's abstract status as a computer program makes it impossible for the court to try or punish the AI agent as a defendant.⁵⁷ These hurdles mean that the deployer of the agentic AI system is likely to be found the more appropriate candidate for a cause of action. Commentators do at least seem to agree, and the law supports, that holding organizations responsible for AI harms aligns with "the idea that principals are responsible for what their agents do on their behalf. The legal responsibility of the AI agents is then ascribed to human beings who stand in the position of the principal. Because AI agents lack intentions, the law should hold them (and the people and companies that employ them) to objective standards, which might be negligence, strict liability, or the highest level of care (when the AI agent performs the functions of a fiduciary)."⁵⁸

One approach is that of the "two strategies of ascribing intention and imposing standards of behavior based on an imagined intention," positing that they "are mirror images of each other" and holding that the first says, "[r]egardless of your intentions, the law will treat you as if you had a particular intention and regulate or penalize you accordingly," and "[r]egardless of your actual intentions, the law will measure your conduct by the standard of a hypothetical person with a particular mental state and regulate or penalize you if you do not live up to that standard."⁵⁹

Another approach may be to address responsibility, especially for some kinds of specific data use or use cases, directly according

to state laws that themselves may be complemented with federal efforts.⁶⁰

The responsibility of the principal in the context of AI agents seems fairly accepted within the literature, with the principal of the AI agent remaining “ultimately responsible for the agent’s actions,”⁶¹ even when such AI agents are “mimic[ing] human learning and action by autonomously processing data.”⁶²

Recent trends in agentic AI adoption reveal diverging use cases for agentic AI systems. Agentic AI systems in commercial contexts are granted varying levels of autonomy to automate tasks within their deploying organization. Firms have demonstrated a wide range of tasks that are delegated to agentic AI systems. Some firms narrowly tailor automation to streamline collaboration between internal teams,⁶³ while others utilize sophisticated autonomous workflows to produce output alongside human teams.⁶⁴ The extent of the autonomy granted to the agentic AI system influences the question of whether the organization is subject to product liability or vicarious liability. If they are granted autonomy to operate unsupervised, the relationship of the AI agent to the organization might mirror an employer-employee relationship. However, if the AI agent is utilized in narrow applications under supervision, it may reflect the relationship between an organization and a product.

The difference between product liability and employer liability for agentic AI deployers centralizes the fundamental unpredictability at the heart of an AI system. Generative AI systems are probabilistic, and introduce randomness when creating an output that aims to replicate a given dataset.⁶⁵ This randomness risks the creation of an output that may cause harm to a third party, whether through a misrepresentation of the deployer’s policies or the generation of obscene or dangerous material.⁶⁶ If an employer liability framework is applied, the court could limit the acts for which the principal will be found liable to those within the scope of the agreement.⁶⁷ This framework could limit the deployer’s liability to acts that are reasonably within the scope of the agentic AI’s use case, and shield the deployer from liability for hallucinations.⁶⁸ These outputs are outside of the use of the agent-principal relationship between the deployer and the agentic AI system, and serve no benefit to the deployer-principal. However, to evaluate whether the agentic AI deployer is subject to employer or product liability, the court weighs multiple factors to determine the scope of the principal’s control of the manner of the work, and the agent’s means to perform the

work.⁶⁹ Generally, where the principal exercises, or has the right to exercise, outsized control of the agent's manner of work, and where the agent's work reflects a typical employee arrangement, the court will apply the standard of respondeat superior liability. The agent's status as an AI system complicates this analysis, as the principal's exercise of control is contingent upon the system's operability. Where an agentic AI system produces unpredictable outputs, utilizes unexplainable reasoning, or is technically illegible, the principal is unable to exercise any control over the system.⁷⁰ As such, an opaque AI system may align further with a product utilized by the deployer-principal toward productivity rather than an employee.

Under a product-liability framework, the organization's failure to control rogue behavior is seen as a design defect.⁷¹ Outputs that cause harm to third parties, even where unexpected, are harder to frame as outside of the scope of liability, as the act itself is demonstrative of the danger presented by the AI system. This framework is generally friendlier to plaintiffs than an employer-liability framework, because the defendant is burdened with demonstrating that the risk of rogue outputs was either unforeseeable or that there was no reasonable alternative design that could have reduced the risk without significantly impairing the system's utility.⁷²

Ultimately, the standard of liability for agentic AI harms hinges on the framing of the AI's outputs as either a technical or managerial failure. As agentic AI evolves toward autonomous workflows, the traditional markers of "control" and "defect" will likely blur, shifting the landscape of legal risk. In practice, these claims require inquiries into the organizational guardrails on agentic AI systems, the technical specifications of the AI system in question, and the specific use case for which the AI was deployed. As the technology continues to advance toward greater autonomy, courts and counsel will need to grapple with novel issues at the nexus of law and technology.

Conclusion and Considerations

It seems apparent enough that there is some party outside of the agentic AI system responsible for outcomes, whether developers, deployers, or users—the party initiating the use of the agentic AI. The UETA and ESIGN (and ESRA) have made that clear for well over a quarter-century, and even anecdotal events have supported

liability in such instances.⁷³ But the devil, as always, is in the details, and there remain several peripheral issues to address.

The considerations discussed herein, notably, have *not* addressed how AI agents might project *moral* harms on humans on the other side of an AI agent interaction⁷⁴ if such AI agents are in possession of inexhaustible patience and overwhelming advantages, but such related ideas still seem to focus back on the concept of accountability. This could be addressed by LFAIs, if they considered mental health requirements and the greater good as part of agentic AI operations. The black box considerations could be mitigated by greater emphasis on explainability, but not only does the probabilistic nature of the technologies challenge this as a primary support, even when the developers, deployers, and users of agentic AI are better positioned for knowledge, but also “this does not mean that they know everything about the technology and the entire chain of actions and things related to it.”⁷⁵

The question is ultimately not whether the agentic AI is inherently responsible for its “decisions” (and, of course, it is not, despite “temptation to blame the AI itself”⁷⁶), but rather just how broad the responsibility extends when strong AI is not only based on real-world data,⁷⁷ but contributing to its development. Some corporate statements seem broad in assertion, such as admonitions that “[e]veryone in an organization, from C-suite executives to operational staff, is responsible for ensuring AI is implemented and used ethically, safely, and securely,”⁷⁸ or incorporate statements related to soft law considerations⁷⁹ and public representations. But practically speaking, even though an individual *human* agent might be responsible for a given harm or set of consequences, the organization on whose behalf the agent was operating will find it difficult to extricate itself.⁸⁰ Likewise, here, if the organization has deployed the agentic AI, there may be a faulty party along the development or deployment line leading to an unintended or undesired outcome—and the organization will be left standing on the proverbial firing line. Contracting and good corporate practices will help mitigate risks, but the risk will remain and—LFAIs or not—must be addressed.

Notes

* James Sherer is a Partner and Noam Kleinman is an Associate at Baker-Hostetler. Caleb Mabe is Assistant General Counsel—Privacy, Data, and AI

at nCino. Emily Fedeles Czebiniak is Senior Vice President, Emerging Technologies for TCDI. Ayyah Saleh received her J.D. degree specializing in data privacy and AI governance. The views expressed herein are solely those of the authors, should not be attributed to their places of employment, colleagues, or clients, and do not constitute solicitation or the provision of legal advice.

1. Amie Taal et al., “Technological and Information Governance Approaches to Data Loss and Leakage Mitigation,” *Computer Science and Information Technology* 4(6), 181-187 (2016).

2. James Sherer, Jenny Le & Amie Taal, “Big Data Discovery, Privacy, and the Application of Differential Privacy Mechanisms,” ICAIL 2015 Workshop on Using Machine Learning and Other Advanced Techniques to Address Legal Problems in E-Discovery and Information Governance (DESI VI Workshop); 32 C&ILJ 7 (2015). See also James Sherer, “When Is a Chair Not a Chair? Big Data Algorithms, Disparate Impact, and Considerations of Modular Programming,” DESI VII Workshop on Using Advanced Data Analysis in eDiscovery and Related Disciplines (2017); 34 C&ILJ 8 (2017).

3. Alex Singla, “The State of AI in Early 2024: Gen AI Adoption Spikes and Starts to Generate Value,” QuantumBlack AI by McKinsey (May 30, 2024), <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>.

4. JD Kloosterman, “How to Respond to the Insult of Artificial Intelligence?” *The Solomon Code* (June 21, 2025), <https://the-solomon-code.com/2025/06/21/how-to-respond-to-the-insult-of-artificial-intelligence-15-minutes/>.

5. Andrew Gelman, “‘AI’ as Shorthand for Turning Off Our Brains. (This Is Not an Anti-AI Post; It’s a Discussion of How We Think About AI.),” *Statistical Modeling, Causal Inference, and Social Science* (Apr. 5, 2024), <https://statmodeling.stat.columbia.edu/2024/04/05/ai-as-shorthand-for-turning-off-our-brains/>.

6. Benjamin Hoffman, “First Came ‘Spam.’ Now, with A.I., We’ve Got ‘Slop,’” *New York Times* (June 11, 2024), <https://www.nytimes.com/2024/06/11/style/ai-search-slop.html>.

7. Adam Nemeroff, “What Is AI Slop? A Technologist Explains This New and Largely Unwelcome Form of Online Content,” *The Conversation* (Sept. 2, 2025), <https://theconversation.com/what-is-ai-slop-a-technologist-explains-this-new-and-largely-unwelcome-form-of-online-content-256554>.

8. Kate Niederhoffer, “AI-Generated ‘Workslop’ Is Destroying Productivity,” *Harvard Business Review* (Sept. 22, 2025), <https://hbr.org/2025/09/ai-generated-workslop-is-destroying-productivity>.

9. Cole Stryker, “What Is Agentic AI?,” *IBM* (Nov. 17, 2025), <https://www.ibm.com/think/topics/agentic-ai>.

10. John Miller, “The Ethical Challenges of AI Agents,” *Tepper Perspectives* (Feb. 13, 2025), <https://tepperspectives.cmu.edu/all-articles/the-ethical-challenges-of-ai-agents/>.

11. James Sherer & Ed Walters, “Practical Magic: Legal AI for Non-Data Scientists,” 44 *Law Practice Magazine* 1, 34-39 (2018).
12. “Getting Started with Prompts for Text-Based Generative AI Tools,” Harvard University Information Technology (Aug. 30, 2023), <https://www.huit.harvard.edu/news/ai-prompts>.
13. Singla, *supra* note 3.
14. John Miller, “The Ethical Challenges of AI Agents,” *Tepperspectives* (Feb. 13, 2025), <https://tepperspectives.cmu.edu/all-articles/the-ethical-challenges-of-ai-agents/>, citing Yonadav Shavit et al., “Practices for Governing Agentic AI Systems,” OpenAI (2023), <https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf>.
15. Sam Ransbotham, “The Emerging Agentic Enterprise: How Leaders Must Navigate a New Age of AI,” *MIT Sloan Management Review* (Nov. 18, 2025), <https://sloanreview.mit.edu/projects/the-emerging-agentic-enterprise-how-leaders-must-navigate-a-new-age-of-ai/>.
16. Amie Taal et al., “Cognitive Computing and Approaches for the Conceptual Organization of Case Law Knowledge Bases—A Proposed Model for Information Preparation, Indexing, and Analysis,” *Artificial Intelligence and Law* 24(4), 347-70 (2016).
17. Wenqi Fan et al., “A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented Large Language Models,” *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (2024), <https://api.semanticscholar.org/CorpusID:269740933>.
18. James Sherer et al., “A Model Approach to Attorney AI Practice—Function or Folly in an Age of AI?,” *California Western Law Review*, 61(2), article 5 (2025), <https://scholarlycommons.law.cwsl.edu/cwlr/vol61/iss2/5>.
19. Sanjay Srivastava, “Why the Buzz on Agentic AI,” *Forbes* (Feb. 18, 2025), <https://www.forbes.com/sites/sanjaysrivastava/2025/02/18/why-the-buzz-on-agentic-ai>.
20. Eric Siegel, “The Agentic AI Hype Cycle Is Out of Control—Yet Widely Normalized,” *Forbes* (July 28, 2025), <https://www.forbes.com/sites/ericsiegel/2025/07/28/the-agentic-ai-hype-cycle-is-insane--dont-normalize-it/>.
21. See generally James Sherer et al., “AI Audits: Initial Steps Toward Building User Trust and Maintaining International Norms,” 8 *RAIL* 4 (2025).
22. James Sherer et al., “Artificial Intelligence and Employment: From Applications to Recommendations, and In Between,” 6 *RAIL* 6 (2023).
23. See Restatement (Third) of Agency § 1.01 (2006).
24. See generally Robert W. Hamilton, *Business Organizations: Unincorporated Businesses and Closely Held Corporations: Essential Terms and Concepts* (Aspen Law & Business 1996).
25. See Restatement (Second) of Agency §§ 82-83, 85, 100, 143 (1958); Restatement (Third) of Agency §§ 4.01-4.03 (2006).

26. Gabriel Rauterberg, “The Essential Roles of Agency Law,” *Michigan Law Review*, 118(4), at 616 (2020), <https://repository.law.umich.edu/cgi/viewcontent.cgi?article=4986&context=mlr>.

27. *Id.*

28. Christian Borch, “Machine Learning, Knowledge Risk, and Principal-Agent Problems in Automated Trading,” *Technology in Society*, Vol. 68 (Feb. 2022), <https://www.sciencedirect.com/science/article/pii/S0160791X21003274>.

29. Dazza Greenwood, “When AI Agents Conduct Transactions,” Dazza-Greenwood’s Weblog (Nov. 23, 2024), <https://www.dazzagreenwood.com/p/when-ai-agents-conduct-transactions>.

30. Dazza Greenwood, “UETA and LLM Agents: A Deep Dive Into Legal Error Handling,” DazzaGreenwood’s Weblog (Feb. 3, 2025), <https://www.dazzagreenwood.com/p/ueta-and-llm-agents-a-deep-dive-into>.

31. Electronic Transactions Act, Uniform Law Commission, <https://www.uniformlaws.org/committees/community-home?CommunityKey=2c04b76c-2b7d-4399-977e-d5876ba7e034>.

32. Electronic Signatures in Global and National Commerce, 15 U.S.C. § 7001 et seq. (2025).

33. Accepting Electronic Signatures, NYS Office of Information Technology Services, <https://nyc-business.nyc.gov/nycbusiness/description/accepting-electronic-signatures>.

34. Adobe, Electronic Signature Laws & Regulations—United States, Overview (Aug. 12, 2025), <https://helpx.adobe.com/legal/esignatures/regulations/united-states.html>.

35. Peter Cramer, Jonathan Mollod & Caroline Rimmer, “Contract Law in the Age of Agentic AI: Who’s Really Clicking ‘Accept’?,” *Proskauer New Media and Technology Law Blog* (Apr. 9, 2025), <https://newmedialaw.proskauer.com/2025/04/09/contract-law-in-the-age-of-agentic-ai-whos-really-clicking-accept/>.

36. *Id.*

37. *Id.*

38. Michelle Tanney, James Sherer & Isabelle Corbett Sterling, “The CFTC Hits Pause on Artificial Intelligence Regulation,” *Futures & Derivatives Law Report—Journal on the Law of Investment and Risk Management Products* 45(2) (Feb. 2025). Contrast with previous U.S. administration concerns, e.g., Jonathan Forman et al., “SEC Proposes AI Rules for Broker-Dealers and Advisers After Chair’s Warnings,” *Thomson Reuters Westlaw Today Practitioner Insights Commentaries* (Aug. 16, 2023).

39. Aminu Abdullahi, “These AI Models from OpenAI Shutdown Commands,” *Sabotage Scripts, TechRepublic* (June 2, 2025), <https://www.techrepublic.com/article/news-openai-models-defy-shutdown-commands/>.

40. James Sherer et al., “Practical Advice for Internal Artificial Intelligence Policy Development,” *Mealey’s Litigation Report: Artificial Intelligence*, 1(2) (Oct. 2023).

41. Katie Hewson, “The Roles of the Provider and Deployer in AI Systems and Models,” *Stephenson Harwood Data Protection Hub* (Sept. 12, 2024), <https://www.dataprotectionlawhub.com/insight/roles-provider-and-deployer-ai-systems-and-models>.

42. Mark Coeckelbergh, “Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability,” *26 Science and Engineering Ethics* 2051 (Oct. 24, 2024), <https://link.springer.com/article/10.1007/s11948-019-00146-8>.

43. Cramer, *supra* note 35.

44. Ian Ayres & Jack M. Balkin, “The Law of AI Is the Law of Risky Agents Without Intentions,” *University of Chicago Law Review* (Nov. 27, 2024), <https://lawreview.uchicago.edu/online-archive/law-ai-law-risky-agents-without-intentions>.

45. Aengus Lynch et al., “Agentic Misalignment: How LLMs Could Be an Insider Threat,” *Anthropic Research* (Oct. 16, 2025), <https://doi.org/10.48550/arXiv.2510.05179>.

46. Paul Miller & Andrew Gold, “Fiduciary Governance,” *57 William & Mary Law Review* 513 (2015), <https://scholarship.law.wm.edu/wmlr/vol57/iss2/4>.

47. Nicoletta V. Kolpakov, “AI’s Escalating Sophistication Presents New Legal Dilemmas,” *New York State Bar Association* (May 28, 2025), https://nysba.org/ais-escalating-sophistication-presents-new-legal-dilemmas/?srsltid=AfmBOorrWFPTxSbYbxdXxAwOdCVQEJsAt5TsvyIralwR7VosQvc_ZHZI.

48. Restatement (Third) of Agency § 1.04 Cmt. E (Am. Law Inst. 2006).

49. Azza Mohamed, Khaled Abdelqader & Khaled Shaalan, “Explainable Artificial Intelligence: A Systematic Review of Progress and Challenges,” *28 Intelligent Systems with Applications* 200595 (2025), <https://doi.org/10.1016/j.iswa.2025.200595>.

50. *Id.*

51. *Id.*

52. Aliakbar Nafar, Kristen Brent Venable & Parisa Kordjamshidi, “Reasoning over Uncertain Text by Generative Large Language Models,” *Thirty-Ninth AAAI Conference on Artificial Intelligence (AAAI-25)* (2025) (discussing how “LLMs still struggle with the essential task of probabilistic reasoning over text”).

53. Cullen O’Keefe et al., “Law-Following AI: Designing AI Agents to Obey Human Laws,” *Institute for Law and AI* (May 13, 2025) (“AI agents should be designed to rigorously comply with a broad set of legal requirements, such as core parts of constitutional and criminal law. In other words, AI agents should be loyal to their principals, but only within the bounds of

the law: they should be designed to refuse to take illegal actions in the service of their principals.”).

54. *Id.* See also Anat Lior, “AI Entities as AI Agents: Artificial Intelligence Liability and the AI Respondeat Superior Analogy,” 46 *Mitchell Hamline Law Review* 1043, 1065-67 (2020); Samir Chopra & Laurence F. White, A Legal Theory for Autonomous Agents 24-25, 128-29 (2011).

55. Maarten Herbosch, “Liability for AI Agents,” 26 *North Carolina Journal of Law & Technology* 441 (Apr. 2025).

56. Warren A. Seavey, “Liability of an Agent in Tort,” 1 *Southern Law Quarterly* 16 (1916).

57. Herbosch, *supra* note 55.

58. Ayres & Balkin, *supra* note 44.

59. Ayres & Balkin, *supra* note 44.

60. Kolpakov, *supra* note 47.

61. Kolpakov, *supra* note 47.

62. James Sherer & Nichole Sterling, “Artificial Intelligence and Privacy—The Hurt and the Help,” ABA 23 TYL 2 (2019).

63. Lakshmi Varanasi, “McKinsey CEO Bob Sternfels Says the Firm Now Has 60,000 Employees: 25,000 of Them Are AI Agents,” *Business Insider* (Jan. 12, 2026), <https://www.businessinsider.com/mckinsey-workforce-ai-agents-consulting-industry-bob-sternfels-2026-1>.

64. “How to Integrate AI Into Modern SOC Workflows,” *The Hacker News* (Dec. 30, 2025), <https://thehackernews.com/2025/12/how-to-integrate-ai-into-modern-soc.html>.

65. David Foster & Karl J. Friston, *Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play* (O’Reilly 2d ed. 2023).

66. Ayres & Balkan, *supra* note 44.

67. Restatement (Third) of Agency § 7.07 (Am. Law Inst. 2006).

68. “Hallucination” in the context of large language models (LLMs) refers to the generation of content that is syntactically correct and fluent but factually incorrect or nonsensical. See Ziwei Ji et al., “Survey of Hallucination in Natural Language Generation,” 55 *ACM Computing Surveys* 12 (2023).

69. Restatement (Third) of Agency § 7.07(3) (Am. Law Inst. 2006).

70. Herbosch, *supra* note 55.

71. Restatement (Third) of Torts: Prods. Liab. § 2 (Am. Law Inst. 1998).

72. *Id.* at § 2(b).

73. Maria Yagoda, “Airline Held Liable for Its Chatbot Giving Passenger Bad Advice—What This Means for Travellers,” *BBC* (Feb. 23, 2024), <https://www.bbc.com/travel/article/20240222-air-canada-chatbot-misinformation-what-travellers-should-know>.

74. Benjamin Lange et al., “We Need Accountability in Human–AI Agent Relationships,” *npj Artificial Intelligence* (Nov. 5, 2025), <https://doi.org/10.1038/s44387-025-00041-7>.

75. Coeckelbergh, *supra* note 42.

76. Erica Stanford, “Who Is Responsible When AI Acts Autonomously and Things Go Wrong?,” *Global Legal Insights* (May 15, 2025), <https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/autonomous-ai-who-is-responsible-when-ai-acts-autonomously-and-things-go-wrong/>.

77. James Sherer, Taylor Hoffman & Emily Fedeles, “Artificial Intelligence and Associated Clinical Data Privacy Considerations,” 2 *RAIL* 4 (2019) (AI works best when it begins with real-world data, and AI applications are eager for data that provides efficient insight into “what is happening in the world” as well as data that suggests “what is likely to happen.”), citing Kevin Rooney, “Artificial Intelligence in Clinical Trials,” *Wavelength* by PRA Health Sciences (Mar. 23, 2018).

78. Suzanne Taylor, “Who’s Responsible for Responsible AI? (It’s a Trick Question.),” *Unisys* (Aug. 20, 2024), <https://www.unisys.com/blog-post/ai/whos-responsible-for-responsible-ai-its-a-trick-question/>.

79. James Sherer, “Archimedes’ Lever and Audience Participation or Soft Law Considerations for AI Transparency,” 2020 AI Soft Law Workshop, Sandra Day O’Connor College of Law, Arizona State University (Koch Foundation Grant) (2020), 5 *RAIL* 1 (2022) (“Soft Law is, conceptually . . . a reference to frameworks that provide for substantive expectations and rules and that incorporate professional guidelines and best practices, codes of conduct, and other private standards without integrated regulatory enforcement mechanisms.”). See also Catherine Casey, Ariana Dindiyal & James Sherer, “AI–Human Interaction: Soft Law Considerations and Application,” 1 *Journal of AI, Robotics & Workplace Automation* 4, 360-70 (2022).

80. Bayard E. Harris, “Apparent Agency—An Extension of the Deep Pocket Theory,” 23 *South Carolina Law Review* 826 (1971), <https://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=2257&context=sclr>.