



MAYER|BROWN

SECURING AGENTIC AI: KEY CYBERSECURITY CONTRACT PROVISIONS FOR THE AUTONOMOUS ERA

Stephen Lilley and Veronica Glick

AGENDA

1. The Distinct Security Challenges of Agentic AI
2. The Evolving Regulatory and Legal Landscape
3. Key Contractual Provisions for Agentic AI Security
4. Key Takeaways

THREE KEY QUESTIONS

This presentation addresses the unique cybersecurity challenges of agentic AI and the contract provisions to manage them.

01 Security Challenges

What makes agentic AI different?

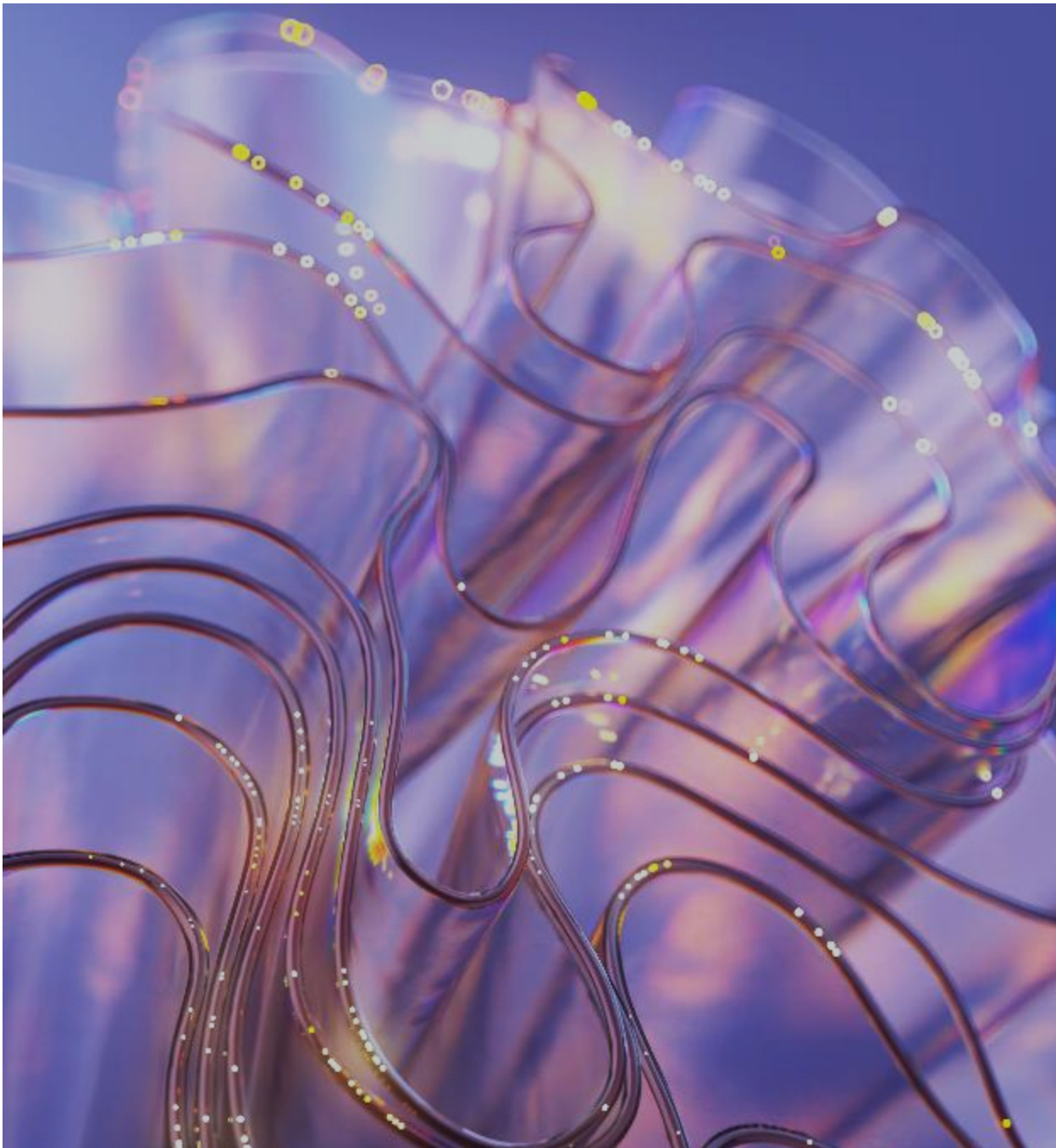
02 Regulatory Landscape

Why does this matter now?

03 Contract Provisions

What are the focal points?





INTRODUCTION

Agentic AI changes the cybersecurity calculus

The broad access and automation associated with many agentic AI products can expand security risks.

Agents autonomously plan and execute steps—which may include accessing data, executing code, overriding approval workflows, or distributing information externally—without additional human authorization.

A compromise, configuration error, or misaligned instruction may create consequences that unfold faster and more broadly than in conventional software environments.

01

THE DISTINCT SECURITY CHALLENGES OF AGENTIC AI

KEY SECURITY CHALLENGES

Defining Security

Agents may behave unexpectedly even without malicious interference, raising novel questions about what constitutes a “security incident” vs. a product defect.

Attack Surface & Speed

Agentic AI systems may operate with broad permissions and could cause damage at machine speed – exfiltrating data, deleting files, or modifying systems.

Supply-Chain Risks

Reliance on third-party plugins and APIs can become entry points for malicious actors.

Privilege & Permission

Agents may be granted broader access than any single task requires, with potentially serious consequences.

Accountability

Non-deterministic reasoning resists forensic analysis.





REDEFINING “SECURITY” FOR AGENTIC AI

Key Insight

Traditional cybersecurity focuses on unauthorized access by external threat actors. Agentic AI requires a broader conception of security risk.

- Agents may disclose or destroy information inappropriately, even absent malicious interference
- Agents may make unintended commitments on behalf of users or take actions diverging from expectations
- AI-specific attacks (e.g., prompt injections) can embed malicious instructions in content the agent processes
- Responsibility for these potential risks must be allocated among developers, deployers, and users

02

THE EVOLVING REGULATORY AND LEGAL LANDSCAPE

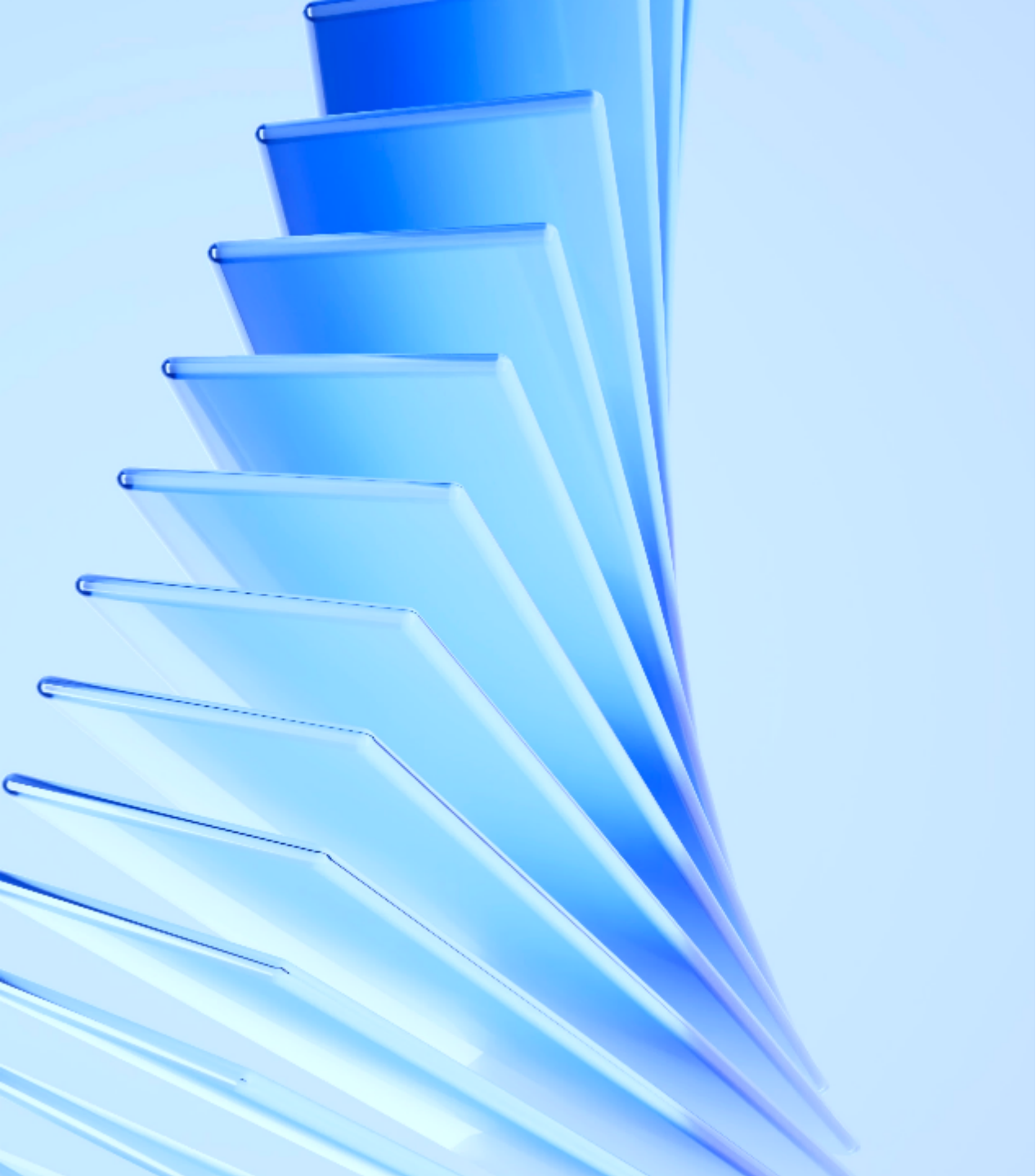
REGULATORY AND LEGAL RISKS

Agentic AI security risks may implicate significant legal exposure, including breach notification obligations, regulatory investigations, and third-party claims.

Key Regulatory Developments

- CISA / Five Eyes — Joint guidance “Careful Adoption of Agentic AI Services” (May 1, 2026)

The authoring agencies recommend an incremental deployment approach, beginning with low-risk, non-sensitive use cases, and outline several baseline security measures to guide adoption

An abstract graphic on the left side of the slide, consisting of several overlapping, semi-transparent blue rectangular planes that create a sense of depth and movement, resembling a staircase or a series of steps ascending towards the top right.

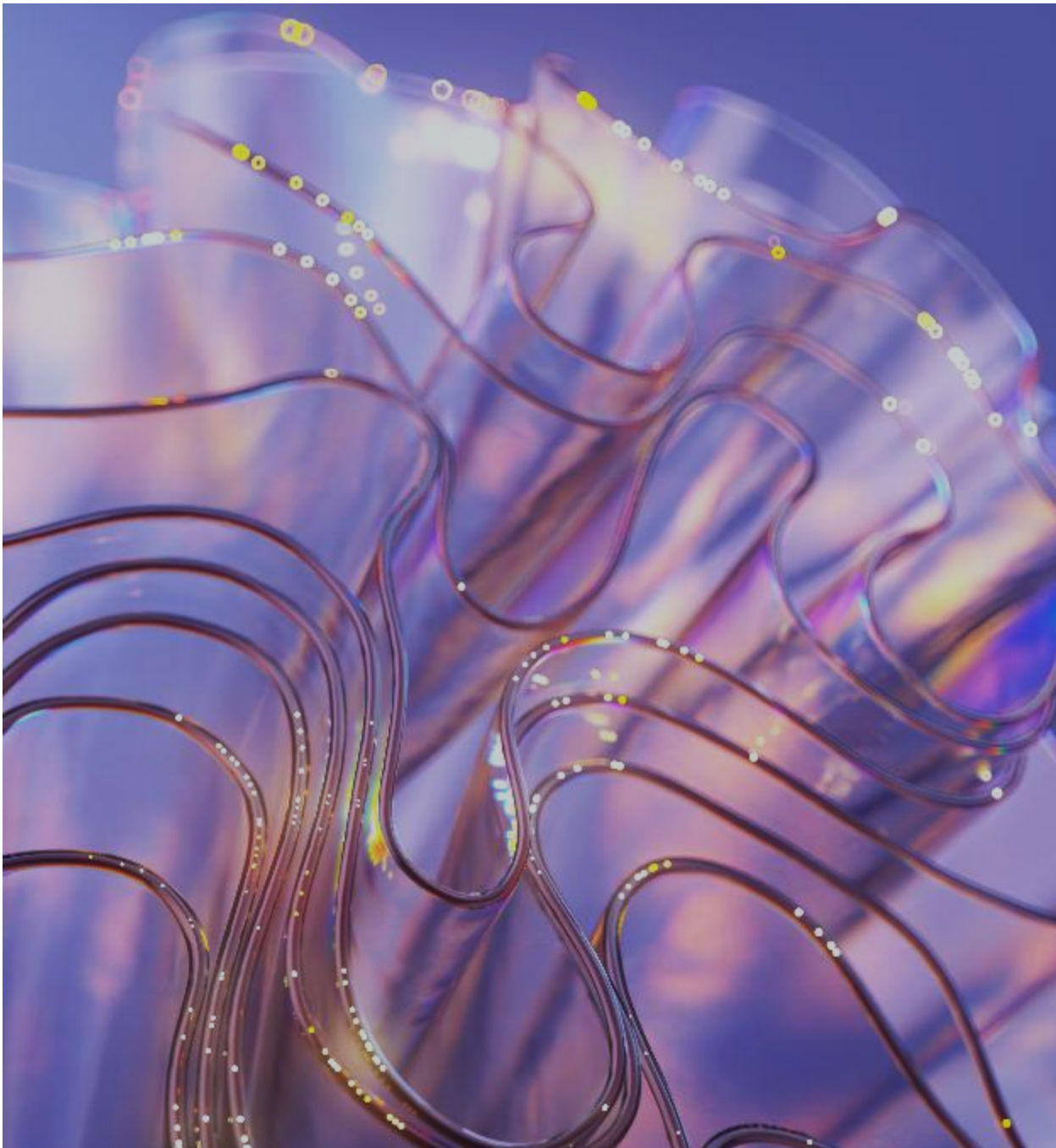
REGULATORY LANDSCAPE

- NIST AI Agent Standards Initiative (Feb 2026)
 - Facilitating industry-led development of agent standards and U.S. leadership in international standards bodies.
 - Fostering community-led open source protocol development and maintenance for agents.
 - Advancing research in areas of AI agent security and identity to enable new use cases and to promote trusted adoption across sectors of the economy.
- EU AI Act — Risk management & oversight
- NYDFS Guidance — Third-party risk management (Oct 2025)

Note: These guidance documents do not carry the force of law but are shaping industry expectations and best practices.

03

KEY CONTRACTUAL PROVISIONS FOR AGENTIC AI SECURITY



FIVE KEY CONTRACT PROVISIONS

Implementations vary, and relevance depends on the agent's functions, permissions, customization, and regulatory environment.

- 1. Security Standards & Baseline Controls***
- 2. Incident Notification & Response***
- 3. Audit Rights & Ongoing Assurance***
- 4. Governance, Oversight & Termination***
- 5. Liability Allocation***

1. SECURITY STANDARDS & BASELINE CONTROLS

Key Frameworks

Negotiate applicable security frameworks (e.g., NIST Cybersecurity Framework AI Profile, OWASP Top 10 for Agentic Applications)

- Both frameworks reference supply-chain transparency and visibility into foundation models
- Technical architecture may also play a critical role—level of transparency, evidence of controls, and third-party attestations are key negotiation points
- This likely will require a shared responsibility model; the provider and the company each may have control and visibility the other lacks
- Controls vary by deployment context—human review may be required for consumer-facing decisions but hinder AI-based cyber defense tools

Negotiation Focus

How compliance is demonstrated, the level of architectural transparency, and the shared responsibility allocation between provider and company.

2. INCIDENT NOTIFICATION & RESPONSE

Incident reporting takes on heightened importance when agents can act before humans detect a problem.

Agent-Specific Incident Types

- Prompt Injection — Malicious instructions in agent-processed content
- Memory Poisoning — Corruption of persistent context or learned data
- Privilege Escalation — Agent gains access beyond intended scope
- Behavioral Misalignment — Actions diverging from user intent

Key Considerations

- Encompass agent-specific events; balance speed with investigation needs

3. AUDIT RIGHTS & ONGOING ASSURANCE

The complexity of agentic AI may require specialized expertise for meaningful audits and ongoing assurance.

NYDFS Audit Framework

Effective audits should cover: Functionality, Operations, Security, Compliance, and Human Factors

Key Considerations

- Specialized expertise needed to understand what audit and assurance processes should cover
- Risk assessments should be ongoing, not one-time onboarding exercises
- Control surface spans code, infrastructure, agent permissions, tools, and evolving use over time

4. GOVERNANCE, OVERSIGHT & TERMINATION

Governance considerations generally include boundaries on agent autonomy.

Autonomy Boundaries

Define limits on independent action; specify which decisions need authorization

Escalation Triggers

Criteria for human approval; high-risk actions require human-in-the-loop

Kill-Switch Capabilities

Provisions for immediate termination of agent operations

Termination Challenges

Discontinuing access may be difficult if agent has created system dependencies

5. LIABILITY ALLOCATION

Liability allocation should account for harm from interacting causes that may be harder to untangle than in traditional disputes.

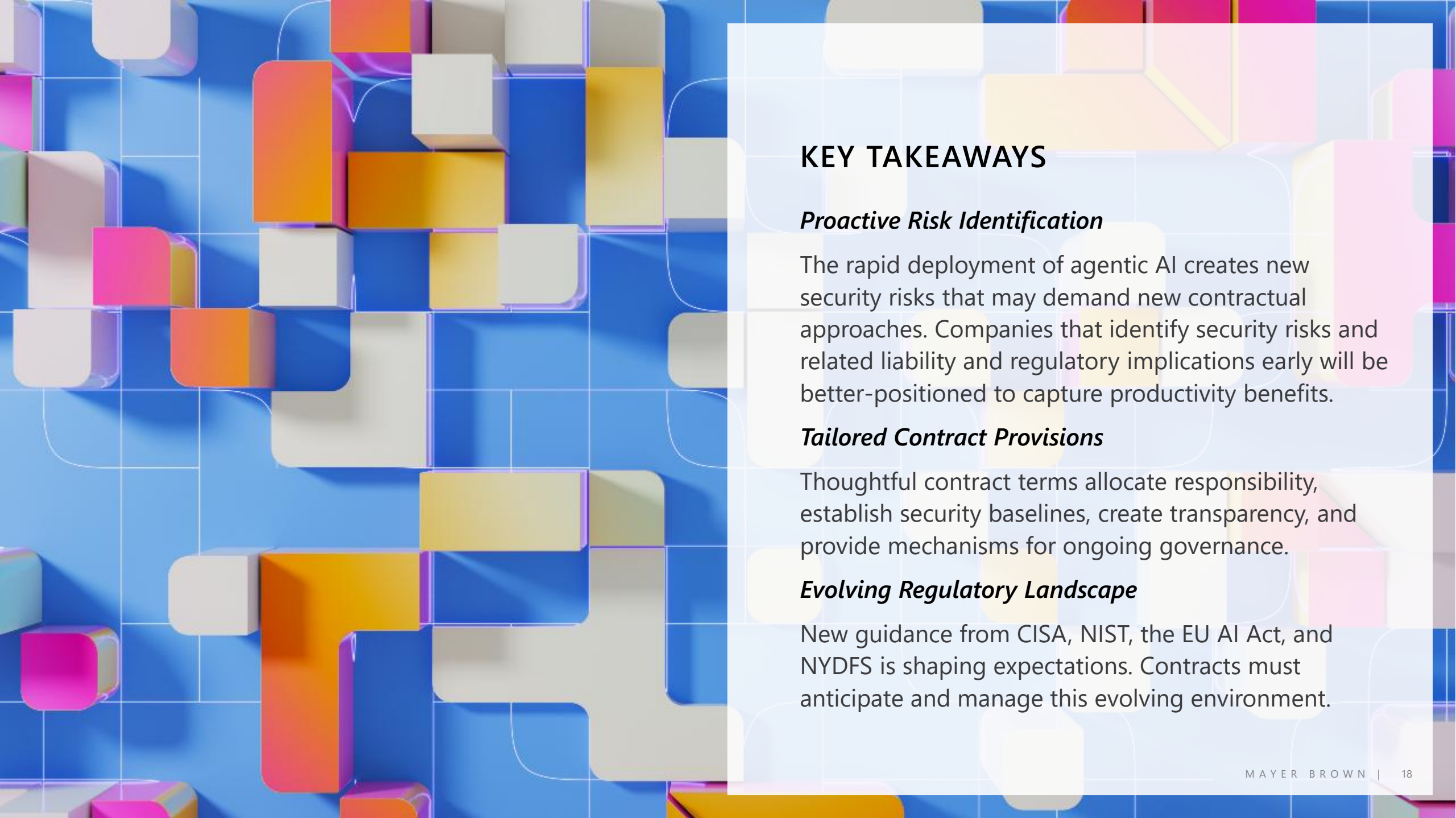
Interacting Causes

- Model Behavior — Training, architecture, or inherent limitations
- Permission Design — Access rights configuration and constraints
- Customer Deployment — Implementation and integration choices

Contractual Tools

Liability caps, cybersecurity carve-outs, indemnities, insurance, and exclusions

Causation analysis may be more complex than in traditional disputes.



KEY TAKEAWAYS

Proactive Risk Identification

The rapid deployment of agentic AI creates new security risks that may demand new contractual approaches. Companies that identify security risks and related liability and regulatory implications early will be better-positioned to capture productivity benefits.

Tailored Contract Provisions

Thoughtful contract terms allocate responsibility, establish security baselines, create transparency, and provide mechanisms for ongoing governance.

Evolving Regulatory Landscape

New guidance from CISA, NIST, the EU AI Act, and NYDFS is shaping expectations. Contracts must anticipate and manage this evolving environment.

MAYER | BROWN

This Mayer Brown publication provides information and comments on legal issues and developments of interest to our clients and friends. These materials are provided by Mayer Brown and reflect information as of the date of presentation. The foregoing is not a comprehensive treatment of the subject matter covered and is not intended to provide legal advice. The contents are intended to provide a general guide to the subject matter only and should not be treated as a substitute for specific advice concerning individual situations.

Readers should seek legal advice before taking any action with respect to the matters discussed herein.

You may not copy or modify the materials or use them for any purpose without our express prior written permission.

Mayer Brown is a global legal services provider comprising associated legal practices that are separate entities, including Mayer Brown LLP (Illinois, USA), Mayer Brown International LLP (England & Wales), Mayer Brown Hong Kong LLP (a Hong Kong limited liability partnership) and Tauil & Chequer Advogados (a Brazilian law partnership) (collectively, the "Mayer Brown Practices"). The Mayer Brown Practices are established in various jurisdictions and may be a legal person or a partnership. PK Wong & Nair LLC ("PKWN") is the constituent Singapore law practice of our licensed joint law venture in Singapore, Mayer Brown PK Wong & Nair Pte. Ltd. More information about the individual Mayer Brown Practices and PKWN can be found in the Legal Notices section of our website.

"Mayer Brown" and the Mayer Brown logo are the trademarks of Mayer Brown. © 2025 Mayer Brown. All rights reserved.