DELEGATING TO THE MACHINE:

Implementing Guardrails for Agentic AI in HR

October 8, 2026

#IAPPpsr26

Introductions

Christina McCoy

Associate General Counsel
Innova Solutions

Amanda Witt

Partner
Nelson Mullins

Jodi Daniels

Owner & Founder
Red Clover Advisors

Celeste James

Associate Commercial Counsel
Google

Agenda Outline

- **Welcome and Introductions**
- **From Generative AI vs. Agentic AI**
- **Legal and Privacy Landscape**
- **Evaluating the Tool**
- **Use Case: ABC Staffing, Inc.**
- **Practical Takeaways**
- **Questions**

Section 1

From Generative AI to Agentic AI

Generative AI vs. Agentic AI in Hiring

GENERATIVE AI = ASSISTS	AGENTIC AI = ACTS
Drafts Summarizes Suggest	Screens Schedules Updates systems
*Waits for instructions from you before acting.	*Acts based on an overarching goal.

Multi-Agent Setups

There are 3 common design setups for multi-agent systems:

i. Sequential: linear workflow where one agent works after another.
ii. Supervisor: an overseeing agent coordinates specialized agents beneath it.
iii. Swarm: multiple agents work together at the same time, delegating to another agent when necessary.

Section 2

Legal and Privacy Landscape

Privacy and AI Regulatory Landscape

U.S. Privacy and AI Regulatory Landscape

Federal Sectoral Privacy Laws
(e.g., HIPAA; GLBA)

State Consumer Privacy Laws
(e.g., CCPA; Utah Consumer Privacy Act)

State Data Security Laws
(e.g., NY Gen. Bus. Law 899-bb(2); Or. Rev. Stat. 646A.622)

State Breach Laws
(available in all 50 states, Washington DC, and US territories)

Children's Privacy
(e.g., COPPA; California Age Appropriate Design Code Act or "CAADCA")

Social Media and Children
(e.g., Protecting Children from Social Media Act (TN))

Artificial Intelligence
(including automated decision-making, profiling, and sensitive data use)

Biometric Privacy Laws
(e.g., Illinois Biometric Information Privacy Act or "BIPA")

Consumer Health Data Privacy Laws
(e.g., Washington's My Health My Data Act)

Data Brokers
(e.g., CA Delete Act; VT Data Broker Regulation Legislation, Act 171)

Wiretapping Laws
(e.g., Electronic Communications Privacy Act or "ECPA"; California Invasion of Privacy Act or "CIPA")

Sensitive Personal Information
(e.g., biometric data, children/teen data, health data, geolocation)

Marketing Communications
(e.g., CAN-SPAM, TCPA)

Enacted International AI Laws

EU AI Act — European Union
(In force Aug 1, 2024; transparency obligations from Aug 2, 2026,)

Italy — Law No. 132/2025
(In force Oct 10, 2025 (first EU national AI law))

South Korea — AI Basic Act
(In force Jan 22, 2026)

Japan — AI Promotion Act
(In force 2025; full effect Sept 1, 2025)

China — Generative AI Measures
(In force Aug 15, 2023)

Taiwan — AI Basic Act
(In force Jan 14, 2026)

Vietnam — Law on AI
(In force Mar 1, 2026)

Peru — Law No. 31814
(Enacted 2023)

EU AI Act: Global Implications

A Global Compliance Benchmark

RISK-BASED FRAMEWORK

Prohibits unacceptable AI uses; regulates high-risk and general-purpose AI (GPAI)

HIGH-RISK AI OBLIGATIONS

Require conformity assessments, transparency, human oversight, and, for certain deployers, fundamental rights impact assessments

GPAI MODEL RULES

Require documentation, copyright compliance, and systemic risk mitigation

GLOBAL REACH

Applies to non-EU providers and deployers where the AI system's output is used in the EU.

AI in Employment: What the Laws Require of HR

Law (effective date)	HR Impact
NYC Local Law 144 — AEDT (July 5, 2023)	Annual independent bias audit; publish the audit summary; give candidates 10 business days' notice.
Illinois AI Video Interview Act, 820 ILCS 42 (Jan. 1, 2020)	Notice and written candidate consent before AI analyzes a video interview.
Illinois Human Rights Act amendments, HB 3773 (Jan. 1, 2026)	Notify employees when AI is used in hiring, promotion, discipline or discharge; discriminatory outcomes are a civil rights violation.
Maryland HB 1202 (Oct. 1, 2020)	Signed applicant waiver before using facial recognition in an interview.
Colorado ADMT Act, SB 26-189 (Jan. 1, 2027)	Notice when AI materially influences a decision; plain-language explanation within 30 days after an adverse outcome; data correction and human review.
California FEHA regulations & CCPA ADMT rules (FEHA ADS regulations effective Oct. 1, 2025; CCPA ADMT compliance from Jan. 1, 2027)	Opt-out path from automated HR evaluation; 4-year record retention; employer liable for vendor algorithmic bias.

California SB 947 (effective July 1, 2027)

What counts as an "Automated Decision System" (ADS)?	• Any computational process derived from machine learning, statistical modeling, data analytics, or AI that issues a simplified output (score, classification, recommendation) used to assist or replace human discretionary decision-making and that materially impacts natural persons. • EXCLUDES spam filters, firewalls, antivirus, IAM tools, calculators, databases. • Applies to all employers, including government and labor contractors. • Agentic tools can meet the ADS definition. SB 947's limits apply when ADS output is used for discipline or termination; hiring tools are covered by other rules (FEHA ADS regulations, CCPA ADMT).
Human review required	• No discipline or termination on ADS output alone. • Where the employer relies primarily on ADS output, a human must corroborate it with supervisor evaluations, personnel files, peer reviews or witness interviews. • Output that cannot be corroborated, or is inaccurate or misleading, may not be used.
Employee rights	• Post-use notice to the affected employee • On request, a meaningful, objective description of the data the ADS used).
Enforcement	• Labor Commissioner and public prosecutors; plus, any worker may bring a civil action in a court of competent jurisdiction for damages caused by that adverse action, including punitive damages. • Injunctive relief, punitive damages, attorney's fees, and a $500 penalty per violation.

Additional State Employment-AI Laws

Law (effective date)	HR Impact
Connecticut SB 5 / PA 26-15 (Oct. 1, 2026; applies to AI deployed after Oct. 1, 2027)	• Disclose when automated employment decision technology is used. • Bias testing is not mandated, but evidence of it - or its absence - is weighed in discrimination claims. • Enforced by the Attorney General.
New Jersey NJLAD algorithmic discrimination (DCR guidance, Jan. 2025)	• Audit AI hiring and screening tools for disparate impact; intent is irrelevant. • Keep human oversight and reasonable-accommodation pathways. • Vendor assurances do not transfer liability.

Key Multi-Jurisdictional Compliance Tensions

TENSION	INDIVIDUAL RIGHTS REGIME	vs.	INSTITUTIONAL / MANDATORY REGIME	TAKEAWAY
Opt-Out Rights vs. Mandatory Screening	**CA CCPA ADMT Regs (2025)** Employee/consumer opt-out from automated decisions **CA FEHA ADT Rules** Automated-decision-system safeguards **Colorado ADMT Act** High-risk deployer duties; impact assessments; reasonable care	⇆	**NYC Local Law 144 (AEDT)** Annual independent bias audit + candidate notice (eff. July 2023) **IL AI Video Interview Act** 820 ILCS 42 + IHRA HB 3773 (eff. Jan 2026) Mandates disclosure/consent for AI video analysis	**Build human-in-the-loop fallback + jurisdiction-aware routing** Operationalize per-state opt-out, notice, and alternative review
Public Audit Disclosures vs. Confidential Assessments	**NYC Local Law 144** Must publicly post bias-audit summary; 10 days' candidate notice Public summary discoverable in litigation	⇆	**Colorado ADMT Act** Colorado ADMT law (SB 26-189, eff. Jan. 1, 2027): pre-decision notice; adverse-outcome explanation within 30 days; correction and human review; 3-year records **CT AI Governance Framework** Vendor due diligence; internal governance **CA ADMT ~4-yr retention**	**Keep public + internal docs consistent; manage privilege** Align retention schedules with disclosure obligations
High Consent (Biometrics) vs. Disclosure-Only	**IL BIPA** — informed written consent **MD HB 1202** — signed waiver for facial recognition **CA CIPA** — two-party call-recording consent **IL AI Video Interview Act**	⇆	**Conversational AI Disclosure Laws** NE, CA, ID, GA, WA, OR, CT, CO Only require disclosure of non-human identity at point of interaction; no affirmative consent needed	**Default to highest common denominator** Written consent + bot disclosure + two-party recording consent + biometric opt-in

Mobley vs. Workday

Case Overview

Background:

- Plaintiffs allege Workday's AI-driven applicant screening tools disproportionately excluded candidates based on age, race, & disability.
- The case is one of the first major U.S. lawsuits testing whether an AI vendor can be directly liable for employment discrimination.

Key Legal Development:

- The court allowed certain discrimination claims against Workday to proceed under an agency theory, recognizing that AI vendors may face liability when tools influence employment decisions.
- A nationwide age discrimination collective action was preliminarily certified, increasing potential exposure.
- The litigation remains ongoing and no final determination on liability has been made.

Legal Team Implications

- AI vendors may face direct legal risk, not just employers.
- Contracts should address AI governance, audit rights, cooperation obligations, data access, indemnification, and allocation of liability.
- Documentation demonstrating testing, validation, and mitigation of bias may become critical evidence.

Cybersecurity & Governance Team Implications

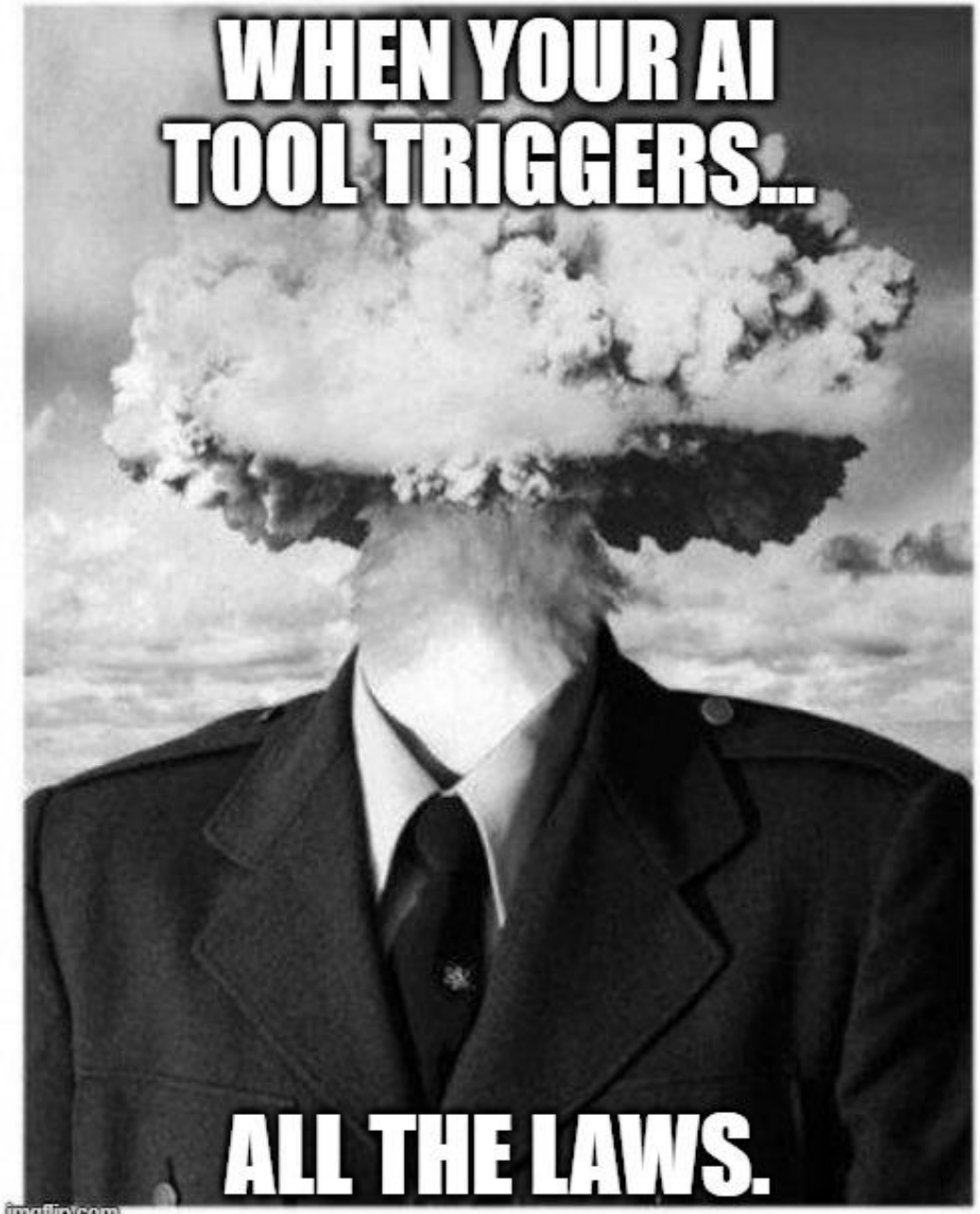
- AI systems should be treated as high-risk decision systems subject to governance controls.
- Maintain records of model testing, monitoring, training data governance, and change management.
- Ensure defensible audit trails and evidence preservation for regulatory inquiries and litigation.
- Align AI governance programs with privacy, security, and compliance controls to support accountability.

PREPARING FOR AI COMPLIANCE:

A Legal Roadmap

What IT, Privacy, and Legal Teams Can Do Now

- **Map AI Use Cases Across the Business:** Identify where AI is used and whether it impacts individuals' rights or triggers legal obligations
- **Conduct AI Risk Assessments:** Evaluate bias, transparency, data provenance, and regulatory exposure
- **Review Contracts and Vendor Agreements:** Ensure AI tools include indemnity, audit rights, and compliance warranties. Need to address ownership of outputs and data risk.
- **Monitor Emerging Laws and Guidance:** Track developments in state, federal, and international AI regulation


WHEN YOUR AI TOOL TRIGGERS...
ALL THE LAWS.
imgflip.com

Section 3

Evaluating the Tool

Five Foundations for Scaling Agent Use

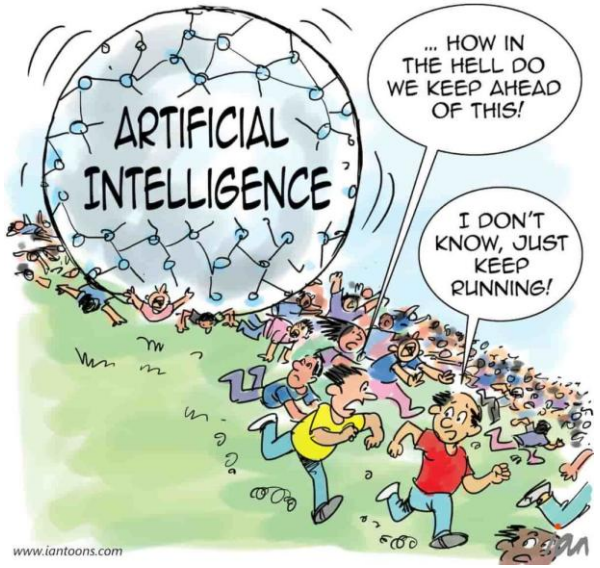
1 **Get Your Data House in Order**
Agents amplify siloed, inconsistent, or unverified data instead of fixing it. Know what you have, where it lives, who owns it.

2 **Map the Decision Boundary**
Write down, per deployment: what the agent can decide vs. what it must refer to a human. Can shortlist. Cannot select.

3 **Stand Up an AI Governance Function**
Someone owns the register, risk assessments, incident log, and review cycle — plus the AI acceptable-use policy.

4 **Train People to Interrogate, Not Just Use**
Overreliance is as dangerous as non-adoption. Decision-makers must know failure modes and how to audit outputs.

5 **Build for Auditability From Day One**
Log inputs, outputs, decisions, and human interventions — your best defense when something goes wrong.



FROM RISK LIST TO OWNER LIST

Risk, Control, and Who Actually Owns It

*A risk with no owner is a risk you have documented, not **managed**.*

AGENTIC RISK	THE CONTROL THAT ACTUALLY HELPS	WHO OWNS IT
Over-delegation	Written authority limits, enforced at the gate	**HR process owner**
Prompt injection	Input controls, least privilege, no unattended writes	**IT security**
Data sprawl	Scoped, time-bound credentials; retention limits	**IT security + privacy**
Bias at machine scale	Pre-launch bias audit and periodic re-testing	**HR, with privacy review**
Vendor / model opacity	Contract: change notice, audit rights, test evidence	**Procurement / contracts**
No audit trail	Action-level logging with a named log reviewer	**IT + records management**

WHERE PRIVACY ACTUALLY SITS

Most of these controls sit with other teams. Privacy's contribution is making sure each one has a **named owner** before deployment, and that the assignment is documented.

Human Oversight: Three Levels

*Not all human oversight is equal — the right level depends on the **impact and reversibility of the decision.***

LEVEL 1

Human-ON-the-Loop

AGENT ACTS → HUMAN REVIEWS AFTER

Agent acts; human can review and reverse *after*. For low-stakes, high-volume calls where speed wins.

Example: routing a support ticket

LEVEL 2

Human-IN-the-Loop

AGENT RECOMMENDS → HUMAN APPROVES

Agent does the legwork; a human must approve before action. Material financial, employment, or reputational stakes.

Example: résumé screening & hiring recs

LEVEL 3

Human-AS-the-Loop

AGENT INFORMS → HUMAN DECIDES

Agent is purely a reference tool; human makes every call independently. The highest-stakes decisions.

Example: legal advice, terminations, medical calls

"WE HAVE A HUMAN IN THE LOOP"

...said the reviewer who approved 200 agent recommendations in 10 minutes without reading them. That's not oversight — **that's a rubber stamp**. Regulators are watching, especially in employment.

FROM PRINCIPLE TO PRACTICE

Applying Oversight Tiers to HR Decisions

*Bring this to HR: the tier follows the **impact and reversibility of the decision**, not the sophistication of the tool.*

AGENTIC RISK	OVERSIGHT TIER	WHAT PRIVACY ASKS FOR
Sourcing & job-ad targeting	LEVEL 1 · HUMAN-ON	Logged targeting criteria; periodic sample review for proxy exclusion
Résumé screening & ranking	LEVEL 2 · HUMAN-IN	Named reviewer sign-off; rationale retained for rejected candidates
Interview scoring & assessment	LEVEL 2 · HUMAN-IN	Bias audit before go-live; candidate notice and accommodation route
Offer terms & compensation	LEVEL 2 · HUMAN-IN	Human is the decision of record; agent output is an input only
Discipline, termination, leave & medical	LEVEL 3 · HUMAN-AS	No agent decisioning — policy retrieval and drafting support only

THE TEST HR SHOULD BE ABLE TO PASS

For any decision above, HR should be able to name the human who approved it, show what that person saw, and say when. If not, the tier is **aspirational, not a control.**

Five Questions Before Deploying Agentic AI

- What authority does the agent have?
- What decisions can be made without human approval?
- With what systems will the agent interact?
- How will the agent's behavior be monitored and audited?
- What happens if it fails? How do you know and how do you stop it?

Section 4

Use Case: ABC Staffing, Inc.

ABC Staffing: Agentic AI in Global Hiring

A Practical Application

- ABC Staffing, Inc. is a U.S.-based staffing company with operations in the U.S., Europe, Latin America, and the South Pacific.
- They're exploring a multi-agent agentic AI solution to make the hiring process more efficient. ABC will partner with a vendor to deploy an agentic AI system that:
  - **Drafts & posts** job descriptions.
  - **Screens incoming resumes** for relevance to job descriptions.
  - **Searches a database** of existing resumes to identify qualified candidates.
  - **Conducts initial screening interviews** via AI-powered phone calls, which will be **recorded**. The vendor will be required to describe the AI agent's scoring/ranking/summaries of the candidates pre- and post-interview.

Scenario 1

AI in Screening and Recruiting

Your CEO at ABC Staffing, Inc. in a meeting says that we need an agentic AI workforce. Without the workforce, the company will be left behind its competitors. The CEO wants this agentic workforce in place by the end of the year.

What would you ask your CEO before committing to the deadline?

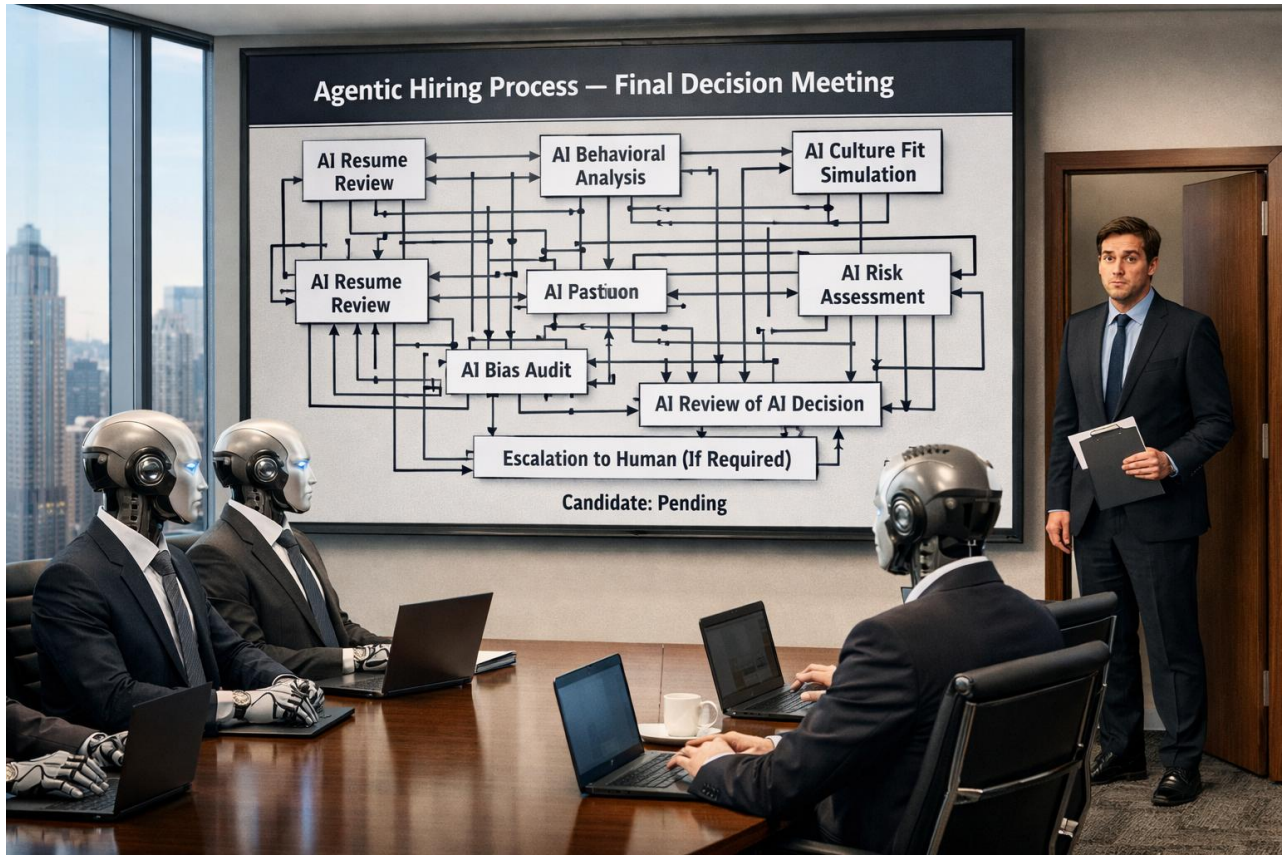
ABC Staffing: Litigation and Enforcement Exposure

- AI use raises consent, bias, and data provenance concerns – would be considered a high-risk system under the EU Act (Annex III obligations apply from Dec. 2, 2027)
- Website tracking and call recording may violate wiretapping and consumer protection laws – adequate notice and consent required
- Global footprint triggers multi-jurisdictional exposure – Italy's AI law took effect on October 10, 2025.

ABC Staffing: AI and Privacy in Practice

Operationalizing the Use Case

- Global Data Collection Requires Consent Management
- AI Use Triggers Transparency Obligations
- Data Minimization Is Key
- AI & Privacy Governance Must Be Documented
- HR Data in the Spotlight



Section 5

Practical Takeaways

THE ARTIFACT YOU HAND OVER

The One-Page Agent Go-Live Gate

*Completed by the business, reviewed by privacy, signed **before the agent is switched on** not after.*

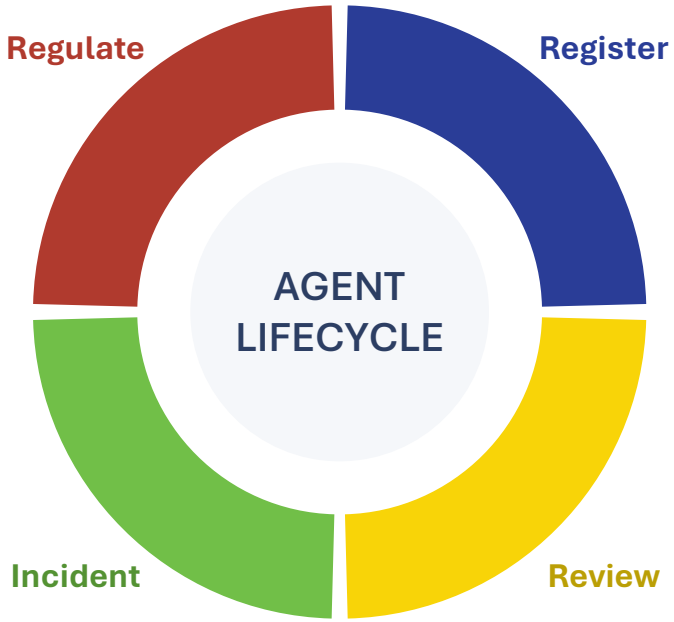
- ☐ **Purpose & scope**
  The decision it supports — and what it may not touch
- ☐ **Authority limits**
  What it may do unattended, with written thresholds
- ☐ **Access & least privilege**
  Scoped, time-bound, revocable systems access
- ☐ **Notice & rights**
  What people are told; how to object or appeal
- ☐ **Assessment on file**
  DPIA/AIA done; bias evidence and vendor reps kept
- ☐ **Logging & review**
  Every action attributable; who reads the logs
- ☐ **Named humans**
  The approver, escalation owner, and pause authority
- ☐ **Re-approval date**
  When it is reviewed again — and what triggers it early

WHY A GATE BEATS A POLICY

A policy asks people to remember. A gate makes them answer. It is also the record you will want when a candidate, a regulator, or a plaintiff asks **how the decision was made.**

IT'S A LEGAL PROBLEM IN A TECH COSTUME

Agent Lifecycle Governance: Monitor, Review, Retire



Live Register

Know, at any moment, every deployed agent: function, data, owner, status. Not a 2024 spreadsheet gathering dust — agents get modified, deprecated, and stood up off-books.

Mandatory Review Cycles

Annual review at minimum, triggered reviews on material change, model updates, or shifting regulation. Six-month cycles for fast-moving regulated sectors.

Incidents Are Incidents

A discriminatory hiring call, a brand-unsafe ad placement, a wrong answer to a customer, an agent hacking a third party — log it, investigate it, root-cause it, remediate it. Same rigor as a data breach.

Prepare for Scrutiny

EU AI Act is phasing in (employment high-risk rules apply starting Dec. 2027). CA, IL, CO, NYC, UT, TX, WA and more are legislating. Regulatory scrutiny will only continue – be ready for it!

Bottom line: Lifecycle governance is fundamentally a legal and compliance discipline dressed up as a technology problem — treat it that way from deployment day one.

Section 6

Questions

How Did Things Go?
(We Really Want To Know)

Did you enjoy this session? Is there any way we could make it better? Let us know by filling out a speaker evaluation.

1. Open the IAPP Events app.
2. Select **IAPP P.S.R. + AIGG 2026**.
3. Tap "Schedule" on the bottom navigation bar.
4. Find this session. Click "Rate this Session" within the description.
5. Once you've answered all three questions, tap "Done".

Thank you!

Thank you!

#IAPPpsr26