The AI you didn't build: From black box to defensible risk

Thursday, 24 September
08:00-09:00 PDT
11:00-12:00 EDT
17:00-18:00 CEST

Welcomes and Introductions

Joanne Furtsch
VP, Knowledge & Global DPO
TrustArc

Hilary Wandall
Chief Ethics & Compliance Officer
Dun & Bradstreet

Darren Abernethy
Shareholder
Greenberg Traurig

TrustArc WEBINAR

LEGAL DISCLAIMER

The information provided during this webinar does not, and is not intended to, constitute legal advice.

Instead, all information, content, and materials presented during this webinar are for general informational purposes only.

Discussion Topics

1. The hidden AI problem
2. Finding and mapping the exposure
3. Scoring the risk: inherent → residual
4. From risk score to governance decision
5. Contracting and ongoing oversight
6. Q&A

Most Organizations Aren't Building Their Own AI

They're buying software with AI already embedded inside it — often as an increasingly invisible feature of an ordinary SaaS product, renewal, or business process.

Privacy, legal, security, and procurement teams may be responsible for evaluating AI they did not select, cannot inspect, and may not even know is processing organizational data.

Working Example

A Running Case Study

We'll carry one example through discovery, scoring, and contracting — so the framework feels connected, not like separate topics.

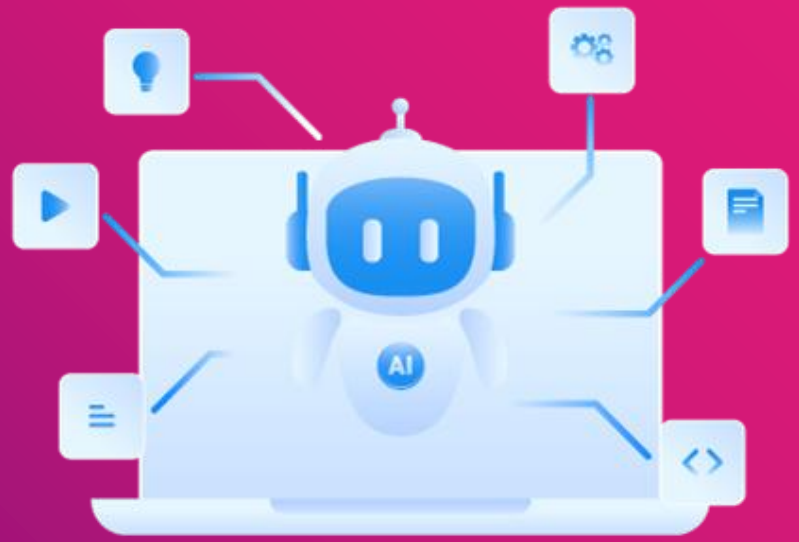
Fictional Case Study

Northwind Acme Co.

A mid-sized, multi-state consumer services company—a seller of widgets at ~40 locations, with a call center and web presence.

Northwind licenses a customer-support ticketing platform under a 3-year MSA. In Year 1, it activates the vendor's AI assistant module feature for web chat: order status, hours, FAQs, along with escalate-to-human for anything else.

Procurement and privacy ran a review. Confirmed: data in–name, email, order number. HITL for everything consequential. Legal signs off. Up and running.



The Challenge

When the Model is a Black Box, “High Risk” Isn't Enough

A Vague Label Isn't a Decision

- “High risk” with no explanation
- A checkbox answer on a vendor questionnaire
- No record of what data or process is actually involved

What Teams Actually Need to Know

- What data and processes are involved
- Which controls actually reduce the risk
- Whether to accept, mitigate, or reject the risk, and defend that call to leadership, an auditor, or a regulator.

You Don’t Need a Governance Committee to Start

Discovery, scoring, and governance can move in parallel — you don't need a standing committee or an approved policy before the first vendor gets reviewed.

The goal isn't a finished program on day one.

It's a repeatable way to find AI, size the risk, and make a defensible call — starting with what you already have.

The Framework

From Discovery to Defensible Action

1 Discover the AI

Where is AI embedded in your vendor and SaaS environment?

→

2 Understand the Exposure

What process, data, and people are affected?

→

3 Score the Inherent Risk

What would make it risky before mitigations?

4 Test the Controls

What evidence shows safeguards actually work?

→

5 Calculate Residual Risk

What risk remains, and is it accepted correctly?

→

6 Turn It into Action

Approve, restrict, escalate, or reject?

Back to Northwind Acme Co. Case Study

Requesting Voice

Eighteen months later, operations extends the same AI assistant module to the phone line.

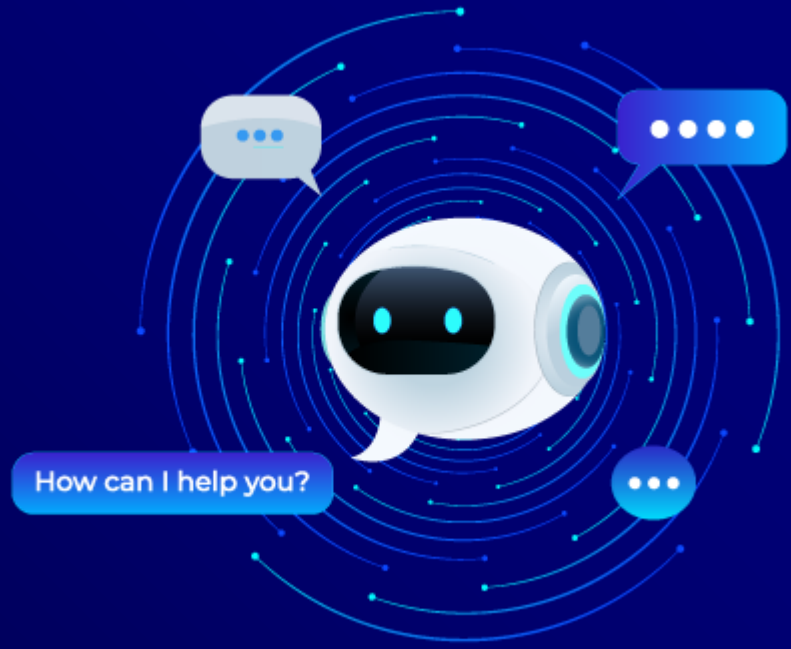
Fictional Case Study

Northwind Acme Co.

An AI voice agent handles inbound calls, with speaker verification to authenticate callers and sentiment scoring to route irritated customers to a supervisor.

The team treats this as a channel expansion of an already-approved tool. No new review.

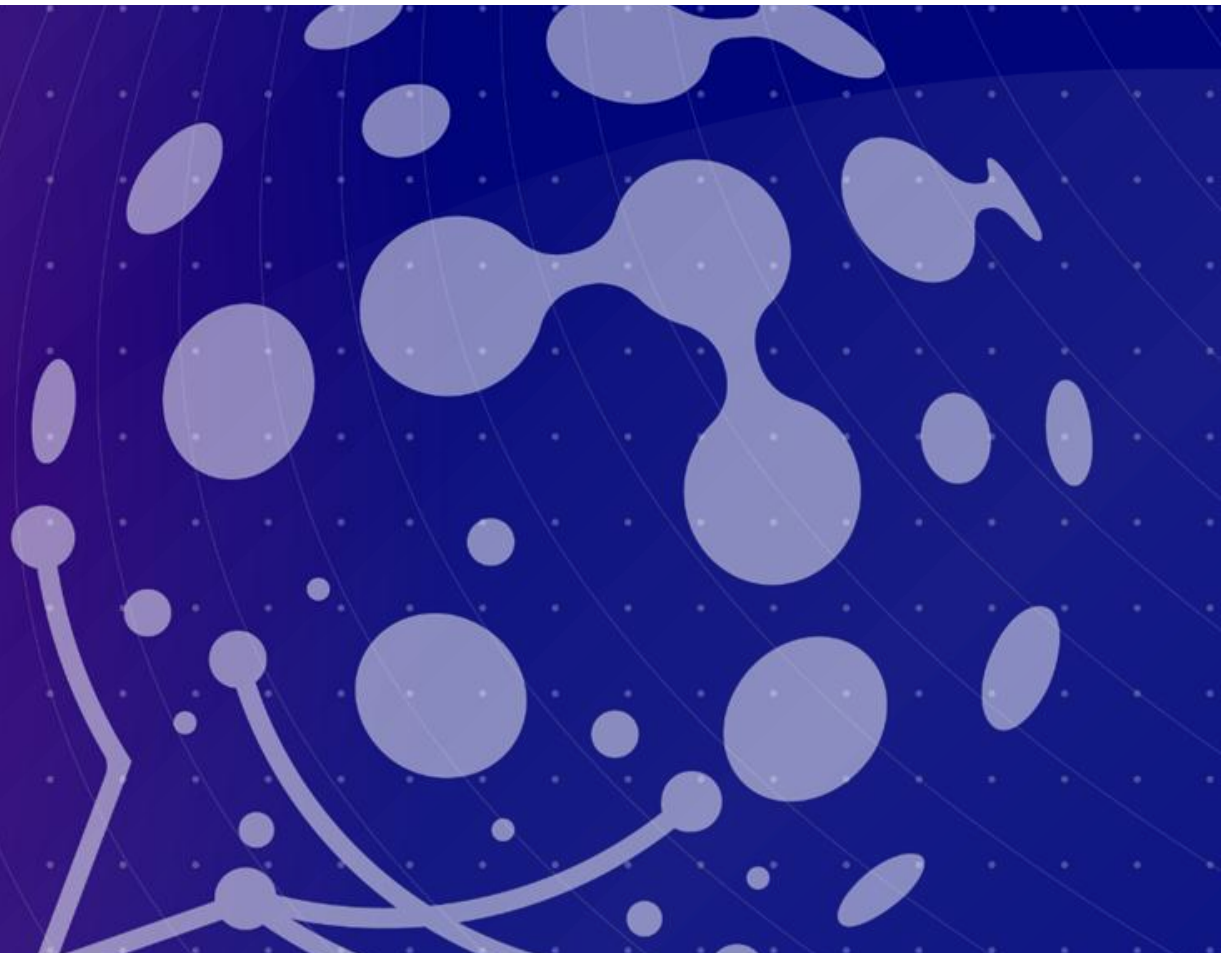
In fact the data categories changed completely…audio recordings, transcripts, a voiceprint, and an inferred emotional state, across states with different recording consent rules.



TrustArc WEBINAR

STEP 1 OF 6

Discover the AI



Step 1 - Discover

Vendor Responses Are One Input, Not the Answer

1 Feature-level AI

Auto-suggest, scoring, ranking, and generation features buried inside familiar tools

2 Sub-processor & 4th-party AI

Models your vendor's vendor runs, several layers removed from your contract

3 Silent updates

New AI capabilities rolled out with no customer notice or re-review

4 MCP & agent connections

A newer connector standard letting a vendor's AI reach into other tools and data on your behalf

TrustArc WEBINAR

STEP 2 OF 6

Understand the Exposure

Step 2 - Map

Map Vendor AI to Business Context

1. Business process
2. Data categories
3. Data subjects
4. Purpose
5. Geographic scope
6. Human involvement
7. Downstream recipients / subprocessors

Test Against Your Own Inventory

Vendor answers are a starting point, not the answer — verify against your own data and process inventory.

TrustArc WEBINAR

STEP 3 OF 6

Score the Inherent Risk

Back to Northwind Acme Co. Case Study

The Features Nobody Asked For!

Fictional Case Study: Northwind Acme Co.

Now, mid-contract, the vendor ships a quarterly release. Nobody re-ran a review, because nothing triggered one. There was no signature, no invoice change, no onboarding.

Three things arrive in the release notes, none of which were requested:

Auto-resolution: the assistant now closes tickets and issues account credits under a dollar threshold without human review.

Customer insights: every interaction generates a churn risk and value score, written back to the CRM over a new OAuth integration, and used to decide who gets retention offers and who gets the slower queue.

A quiet line stating that aggregated interaction data may be used to improve service quality.

Step 3 - Score

Six Factors, One Defensible Score

1 Data Sensitivity

What data feeds the model, and how sensitive is it?

2 Purpose

What decision or output is the AI producing?

3 Autonomy

How much human oversight sits between model and outcome?

4 Scale

How many records, systems, or workflows are involved?

5 Affected Individuals

Who is impacted by the AI's output?

6 Consequence

What's the impact if the output is wrong or misused?

TrustArc WEBINAR

STEP 4 OF 6

Test the Controls

Step 4 - Test

A Control That Exists Isn't a Control That Works

- Contractual
- Technical
- Organizational
- Operational

Model-training exclusions are one control — not the whole solution.

The broader question is how the vendor handles prompts, inputs, outputs, retention, human review, subprocessors, and changes to the service — and whether that handling can be verified, not just asserted.

TrustArc WEBINAR

STEP 5 OF 6

Calculate Residual Risk

Step 5 - Calculate

Two Different Numbers — Don't Confuse Them

Inherent Risk

The risk level before any controls, contract terms, or vendor safeguards are applied.

Residual Risk

What's left after controls are verified — not assumed — to actually reduce the risk.

A vendor's certification, contract clause, or "responsible AI" statement is not automatic mitigation. Avoid treating vendor assurances as risk reduction by default — verify them like any other control before they move the score.

TrustArc WEBINAR

STEP 6 OF 6

Turn the Result Into Action

Step 6 - Decide

From Risk Score to Governance Decision

- Approve with standard monitoring
- Approve with restrictions
- Require additional controls or commitments
- Escalate for executive / committee review
- Reject or suspend deployment

ILLUSTRATIVE RED FLAGS

- Vendor won't disclose sub processors performing AI processing
- Vendor can't confirm whether customer data trains the model
- AI functionality changed mid-contract with no notice
- No ability to obtain audit evidence or documentation

Governance

Contracting and Ongoing Oversight

Contract terms and evidence that support defensible vendor governance

1. Limits on use of customer or employee data
2. Restrictions on model training and product improvement
3. Retention and deletion requirements
4. Subprocessor and service-provider transparency
5. Security and incident notification obligations
6. Rights to obtain documentation or audit evidence
7. Notice of material changes to AI functionality
8. Suspension, termination, or deletion rights where risk changes materially

Closing Principle

A Risk Score Is Not the Output

The output is a transparent, defensible decision trail.

1. What the AI does
2. What data it touches
3. What controls are in place
4. What uncertainty remains
5. Why the org accepted, mitigated, restricted, or rejected the risk

Recap

What You'll Walk Away With

1. Identify AI embedded in third-party products, even where vendor documentation is incomplete or ambiguous
2. Map AI-enabled vendor functionality to business processes, data elements, data subjects, and jurisdictions
3. Distinguish inherent risk from residual risk, without treating vendor assurances as mitigation by default
4. Use a repeatable structure to score third-party AI: data sensitivity, purpose, autonomy, scale, affected individuals, consequence
5. Assess whether controls actually reduce risk, rather than simply documenting that a control exists
6. Translate a risk assessment into procurement decisions, contract terms, monitoring, escalation, and termination conditions
7. Identify the evidence needed to defend a risk rating when asked, “How did you arrive at that number?”
8. Recognize red flags that justify delaying, restricting, or rejecting a vendor AI deployment

Questions and Answers

Joanne Furtsch
VP, Knowledge & Global DPO
TrustArc

Hilary Wandall
Chief Ethics & Compliance Officer
Dun & Bradstreet

Darren Abernethy
Shareholder
Greenberg Traurig

TrustArc @ PSR 2026

01

MEET US AT BOOTH 208

Expo Hall — Seattle Convention Center — Arch Building
October 8 | 8:00 AM–6:00 PM
October 9 | 8:00 AM–2:00 PM

Stop by for live demos, exciting giveaways, and conversations with our privacy experts about what's next for AI, privacy, and compliance.

02

WOMEN LEADING PRIVACY

Thursday, October 810:15–11:15 AM Connection Center • 4C-4

Join fellow women and allies advocating for advancement in privacy — and create a custombag charm with us!

03

HAPPY HOUR

SPIN

Thursday, October 86:00–9:00 PM SPiN Seattle • 1511 6th Ave

Join TrustArc + Filerskeepers + DAA for craft cocktails, great food, and a little friendly competition on the ping pong tables.

Seattle • October 8-9

RSVP NOW >
TrustArc.com/PSR2026

Web Conference
Participant Feedback Survey

Please take this quick (2 minute) survey to let us know how satisfied you were with this program and to provide us with suggestions for future improvement.

Click here: https://iappwf.questionpro.com/t/AbBPvZ9iYY

Thank you in advance!

For more information: www.iapp.org

Attention IAPP Certified Privacy Professionals:

This IAPP web conference may be applied toward the continuing privacy education (CPE) requirements of your AIGP, CIPP/US, CIPP/E, CIPP/A, CIPP/C, CIPT or CIPM credential worth 1.0 credit hour. IAPP-certified professionals who are the named participant of the registration prior to the live webinar will automatically receive credit. After the broadcast date, individuals may submit for credit by completing the continuing education application form here: submit for CPE credits.

Continuing Legal Education Credits:

The IAPP provides certificates of attendance to web conference attendees. Certificates must be self-submitted to the appropriate jurisdiction for continuing education credits. Please consult your specific governing body's rules and regulations to confirm if a web conference is an eligible format for attaining credits. Each IAPP web conference offers either 60 or 90 minutes of programming.

For questions on this or other IAPP Web Conferences or recordings please contact:

livewebconteam@iapp.org