



VERSION 1.0 · SEPTEMBER 2026

BEACON: A Practitioner's Framework for Assessing AI Risk in Third-Party Vendors

Six categories. One anchor question each. Run live with the vendor in under an hour.

B Bias	E Explainability	A Access & Data	C Contractual Accountability	O Oversight	N Non-determinism
------------------	----------------------------	---------------------------	---	-----------------------	-----------------------------

01 THE PROBLEM

Every procurement pipeline now has an AI vendor in it somewhere — a contract-review tool, a customer support bot, a resume screener, a fraud-detection API. Most of these vendors get run through the same third-party risk questionnaire your organization has used for a decade: SOC 2 status, data encryption, access controls, incident response history.

None of that tells you whether the vendor's model hallucinates under load, whether its outputs can be explained to a regulator, or who is on the hook when it gets something wrong. Traditional TPRM tooling was built for a world where third parties stored and moved your data. It wasn't built for a world where third parties make decisions with your data — and increasingly, decisions *about* your customers.

The result: risk teams either skip AI-specific due diligence entirely, or bolt on an ad hoc set of questions that varies by whoever happens to be running the assessment that week. Neither is defensible when a regulator, auditor, or board asks how AI vendor risk is actually being managed.

02 WHY EXISTING TOOLS FALL SHORT — NOT WRONG, JUST THE WRONG SHAPE

This isn't a knock on the broader AI risk literature. Resources like NIST's AI Risk Management Framework and various industry taxonomies are thorough, well-researched, and genuinely useful for teams building AI governance programs from the ground up.

But "thorough" and "fast" are usually in tension. A risk analyst sitting across from a vendor in a 45-minute call doesn't need — and doesn't have time for — an exhaustive academic mapping exercise. They need a small number of pointed questions that surface the risks specific to *this* vendor's AI, right now, in the room.

BEACON is built for that moment. It doesn't replace deeper governance frameworks; it's the fast, memorable layer that sits in front of them — the questions you ask before deciding whether deeper diligence is even warranted.

03 THE BEACON FRAMEWORK

BEACON is six categories, each anchored to a single question a vendor should be able to answer clearly and specifically. **A vague, evasive, or "we haven't thought about that" answer to any of these is itself a data point.**



B Bias

ASK *How do you test outputs for discriminatory or skewed results affecting our customers?*

A vendor with a real answer will describe specific test sets, protected-class analysis, or fairness metrics they monitor. A vendor without one will talk about “quality” in general terms and never mention a protected characteristic.

A Access & Data Handling

ASK *Does our data train your model? Is it segregated from other customers' data, and what's the retention and deletion policy?*

This is the question most likely to reveal a gap between marketing language (“enterprise-grade privacy”) and actual architecture. Ask for specifics, not adjectives.

O Oversight

ASK *Is there a human in the loop, and can we override or escalate a bad output?*

Full automation without an escalation path is a risk multiplier, not a feature. Ask specifically what happens when something goes wrong — not just whether a human is technically “available.”

E Explainability

ASK *Can you document or justify how the model reached a given decision, for audit purposes?*

This matters most anywhere the AI's output affects an individual — credit, employment, healthcare, eligibility. If the vendor can't produce a decision trail after the fact, you're the one left explaining it to a regulator.

C Contractual Accountability

ASK *Who is liable when the AI produces a harmful or incorrect output, and what does the SLA actually say about AI-specific incidents?*

Most standard SaaS contracts were never drafted with model failure in mind. If the vendor's contract team hasn't touched this clause recently, that's worth knowing before, not after, an incident.

N Non-determinism

ASK *What's your measured hallucination or failure rate, and how do you monitor drift over time?*

Unlike traditional software, AI systems can degrade or shift in behavior after deployment without any code changing. A vendor that measures this proactively is fundamentally different from one that finds out from customer complaints.

04 HOW TO RUN A BEACON SESSION

BEACON is designed to run in under an hour, ideally live with the vendor rather than as a written questionnaire — verbal answers are harder to script and easier to probe.

GREEN

Specific, complete answer the vendor can evidence.

YELLOW

Partial answer, or a gap with limited relevance to this use.

RED

Vague, evasive, or no answer — or a gap that matters here.

Who's in the room

The risk/TPRM lead, and ideally someone from the business unit sponsoring the vendor relationship. The vendor should bring someone who can speak to the model itself, not just sales.

How to score it

BEACON isn't pass/fail. Score each category Green / Yellow / Red based on the specificity and completeness of the answer. **Two or more Reds should trigger deeper review before the vendor moves forward**, regardless of how the rest of the standard security questionnaire scored.

How it runs

Work through the six letters in order. For each, ask the anchor question, then one follow-up based on the answer. Document the response in the vendor's own words where possible — vague language is itself the finding.

How it fits your existing process

BEACON runs alongside your standard vendor risk tiering, not instead of it — an additional lens applied where a vendor's offering includes an AI/ML component. It slots in after your existing intake, before final sign-off.



05 CASE WALKTHROUGH

ILLUSTRATIVE SCENARIO — A mid-size legal team is evaluating an AI-powered contract review tool that flags risky clauses automatically. Running BEACON surfaces the following:

Bias	Tests for accuracy but has never run a fairness analysis — not obviously relevant to contract review.	YELLOW
Explainability	Confidence scores, but no rationale for <i>why</i> a clause was flagged. A legal function that has to justify decisions internally has a real gap.	RED
Access & Data	Uploaded contracts train the vendor's model unless the customer opts out in a separate settings panel most buyers never find.	RED
Contractual Accountability	The standard SLA has no language addressing incorrectly flagged clauses.	RED
Oversight	A human always reviews flagged clauses before action.	GREEN
Non-determinism	Tracks a documented false-negative rate and shares it on request.	GREEN

Result: a standard security questionnaire would have passed this vendor — SOC 2 compliant, encrypted at rest, no red flags on paper. BEACON surfaced three Reds specific to how the AI itself behaves, prompting a renegotiated data-training clause and an internal review process before rollout — the exact kind of risk a generic questionnaire is structurally unable to catch.

06 WHERE BEACON FITS IN YOUR TPRM PROGRAM

BEACON is not a replacement for your organization's vendor risk methodology, your evidence-sharing infrastructure, or your existing certifications and control frameworks. It's a fast, targeted addition — the layer you apply specifically to the AI-enabled portion of a vendor's offering, run early enough to inform the rest of the assessment rather than as an afterthought once the contract is already signed.

RUN BEACON WITH US

Third Party Threat Hunting runs BEACON assessments on AI-enabled vendors — on its own or alongside your team for the first few, so the method sticks.

thirdpartythreathunting.com/beacon · info@thirdpartythreathunting.com · 919-592-7757

07 ABOUT THIRD PARTY THREAT HUNTING LLC

Third Party Threat Hunting LLC helps organizations identify and manage the risks introduced by their vendor ecosystem, with a growing focus on AI-specific third-party risk. Gregory C. Rasner, CEO & Founder, hosts the *Third Party Threat Hunters* podcast and speaks regularly on TPRM, AI governance, and vendor risk strategy.

Contact: info@thirdpartythreathunting.com · **Podcast:** thirdpartythreathunting.com/podcast

BEACON™ is a trademark of Third Party Threat Hunting LLC. © 2026 Third Party Threat Hunting LLC. This paper may be shared in full with attribution.