FOR PROCUREMENT, IT, SECURITY AND RISK TEAMS

AI agent vendor due diligence

32 questions to ask before you buy an AI agent product, with the evidence to request and what a strong answer includes.

01 Is it really an agent?

02 Records and auditability

03 Control and limits

04 Data and security

05 Testing and quality

06 Commercials and exit

07 Compliance

Gautam Mehra · Aman Khanna

Whiteboxx AI

HOW TO USE THIS QUESTIONNAIRE

Ask every vendor the same questions.

Send the questions to each shortlisted vendor, ask for the evidence listed, and score each answer. The same questions make vendors comparable and keep the answers on record for the contract.

Scoring

SCORE	MEANING
2	A direct answer, backed by the evidence requested
1	A partial answer, or evidence still to come
0	A vague answer, or none

Reading the total

Compare vendors section by section, not only on the total. A low score on records (section B) or control (section C) matters more than a low score elsewhere if the agent will act in your systems.

The seven sections

A. Is it really an agent?	5
B. Records and auditability	6
C. Control and limits	5
D. Data and security	5
E. Testing and quality	4
F. Commercials and exit	4
G. Compliance	3
Total questions	**32**

An editable spreadsheet version, with scoring formulas, is available from the same page as this PDF.

SECTION A

Is it really an agent?

Gartner estimated in 2025 that only about 130 of the thousands of vendors claiming agentic AI offered real agents. Start by establishing what the product does on its own.

#	QUESTION	EVIDENCE TO REQUEST	A STRONG ANSWER INCLUDES	SCORE
1	**Show a live run of the product on a sample of our data.**	A recorded or live session on data you supplied	The run completes a whole task, not one answer	
2	**Which steps are fixed in advance, and which does the AI decide while it runs?**	A diagram of one workflow	A clear split between fixed steps and agent decisions	
3	**What happens when a step fails or a tool returns an error?**	A run where a step failed	Retries, another approach or a hand-off to a person	
4	**Which tools and systems can the agent use to complete a task?**	The list of available tools and connectors	Named tools, with what each can read and change	
5	**How does the agent know when a task is finished?**	Product documentation	A defined end state or a person confirming completion	

NOTES

SECTION B

Records and auditability

A reviewer should be able to reconstruct what the agent did without asking the person who ran it. The EU AI Act requires high-risk systems to keep logs (Article 12).

#	QUESTION	EVIDENCE TO REQUEST	A STRONG ANSWER INCLUDES	SCORE
6	**Is every run recorded with its request, steps, tool calls and output?**	A sample run record	One record per run, linking all four	
7	**Does each result show the sources it was based on, with dates or versions?**	A sample output with sources	Source names plus retrieval dates or versions	
8	**Can we see the calculations behind a numeric result?**	A sample calculation trace	Inputs, method and result a reviewer can redo	
9	**Who started and who approved each run, and is that recorded?**	A sample approval record	Named people and timestamps	
10	**Can we export run records and admin audit logs to our own systems?**	Export or API documentation	A documented API or export to your security tools	
11	**How long are records kept, and can we set the retention period?**	Retention settings	Configurable retention that meets your policy	

NOTES

SECTION C

Control and limits

Excessive agency, an agent with more access than its task needs, is one of the top risks in the 2025 OWASP Top 10 for LLM Applications.

#	QUESTION	EVIDENCE TO REQUEST	A STRONG ANSWER INCLUDES	SCORE
12	**Can we limit each agent to specific data, tools and actions?**	Permission settings for one agent	Per-agent limits, not one setting for the whole product	
13	**Can we require a person to approve an action before it happens?**	Approval configuration	Approval rules per action type	
14	**Can we set a budget or usage limit per agent, task or month?**	Budget settings	Hard limits that stop the run	
15	**How does the product defend against instructions hidden in documents or web pages?**	Security documentation and test results	Specific controls, plus results of testing	
16	**Can we stop a running agent immediately?**	Documentation of the stop control	A stop control available to named roles	

NOTES

SECTION D

Data and security

An agent can only be as safe as the data access behind it. Treat each connection like any other integration.

#	QUESTION	EVIDENCE TO REQUEST	A STRONG ANSWER INCLUDES	SCORE
17	**Where is our data processed and stored?**	Data processing addendum and regions	Named regions you can choose	
18	**Is our data used to train or improve any model?**	Contract terms	A clear no, written into the contract	
19	**Whose permissions does the agent use when it accesses our systems?**	Authentication documentation	The user's own permissions or a scoped service account	
20	**Which third parties, including model providers, process our data?**	Sub-processor list	A current list with notice of changes	
21	**Which security certifications and audits apply to the product?**	Reports such as SOC 2 or ISO 27001	Current reports covering the product you are buying	

NOTES

SECTION E

Testing and quality

The NIST AI Risk Management Framework makes measurement one of its four core functions. A vendor should be able to show how it measures quality.

#	QUESTION	EVIDENCE TO REQUEST	A STRONG ANSWER INCLUDES	SCORE
22	**How are results tested before release, and what accuracy have you measured?**	Test method and results	Measured results on tasks like yours	
23	**Can we run our own test cases against the product?**	Evaluation tools	A way to run your cases and keep the results	
24	**How do you detect a drop in quality after a model or data change?**	Monitoring documentation	Regular re-testing and alerts	
25	**Which model or models does the product use, and how are changes announced?**	Model and change policy	Named models and advance notice of changes	

NOTES

SECTION F

Commercials and exit

Pricing tied to usage can grow faster than value, and an exit that loses your records or logic is expensive.

#	QUESTION	EVIDENCE TO REQUEST	A STRONG ANSWER INCLUDES	SCORE
26	**How does the price change with usage, and what would a busy month cost?**	A pricing model for your volumes	Worked figures for your expected volumes	
27	**Who maintains the product and our configuration in year two, and at what cost?**	Support and maintenance terms	Named responsibilities and costs	
28	**What do we keep if we leave: configuration, logic, data and run records?**	Exit terms	Export of all four in usable formats	
29	**What would it cost to build the same capability ourselves?**	The vendor's own comparison	A frank comparison, including what the vendor does not cover	

NOTES

SECTION G

Compliance

EU AI Act high-risk rules apply to uses such as hiring and credit from 2 December 2027, after amendments adopted in 2026.

#	QUESTION	EVIDENCE TO REQUEST	A STRONG ANSWER INCLUDES	SCORE
30	**Could our intended use count as high-risk under the EU AI Act, and how will you support us?**	The vendor's regulatory assessment	A clear view and a support plan	
31	**How does the product support human oversight, such as understanding, overriding or stopping the system?**	Oversight features	Features that map to Article 14	
32	**Which other regulations or standards does the product help us meet?**	Compliance documentation	Specific mappings, not general claims	

NOTES

SUMMARY

Score summary

Add up each section for each vendor. Maximum scores are shown in the first column.

SECTION	MAX	VENDOR 1	VENDOR 2	VENDOR 3
A. Is it really an agent?	10			
B. Records and auditability	12			
C. Control and limits	10			
D. Data and security	10			
E. Testing and quality	8			
F. Commercials and exit	8			
G. Compliance	6			
Total	64			

Before you sign

Attach the vendor's written answers to the contract, and agree how they will tell you about changes to models, sub-processors or data handling.

ABOUT WHITEBOXX

Evidence OS, extended for your workflows.

Whiteboxx builds Evidence OS, an enterprise AI platform, and extends it for each customer with forward deployed engineers. We are a vendor too: send us this questionnaire and we will answer every question in writing.

START WITH THE ARCHITECTURE SPRINT

30 days · USD 15,000, fixed

We review your AI tools and pilots, choose one workflow with you and deliver a build plan with costs. You can use the plan with us, with your own team or with another partner.

Book a scope call: **whiteboxx.ai/contact** · info@whiteboxx.ai

Sources and method

1. Gartner on agentic AI projects and agent washing, 25 June 2025
   https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027
2. OWASP, Top 10 for LLM Applications, 2025 edition
   https://genai.owasp.org/llm-top-10/
3. NIST, AI Risk Management Framework
   https://www.nist.gov/itl/ai-risk-management-framework
4. EU AI Act, Article 12: record-keeping
   https://artificialintelligenceact.eu/article/12/
5. EU AI Act, Article 14: human oversight
   https://artificialintelligenceact.eu/article/14/
6. Loyens & Loeff, “EU AI Act: key amendments adopted under Digital Omnibus package”, 2026
   https://www.loyensloeff.com/insights/news--events/news/eu-ai-act-key-amendments-adopted-under-digital-omnibus-package/

Method

The questions draw on the OWASP Top 10 for LLM Applications, the NIST AI Risk Management Framework, the record-keeping and human oversight articles of the EU AI Act, and Whiteboxx’s experience building agent workflows. They are a starting point: add questions for your sector and use case.

This document is general guidance, not legal, financial or investment advice. Edition 1 · November 2026, published 2026-11-20. Authors: Gautam Mehra and Aman Khanna, Whiteboxx AI.

You may share this document in full, unchanged, with attribution to Whiteboxx AI. The latest edition is at whiteboxx.ai/resources/ai-agent-vendor-questionnaire.