

VIKRAM JHA - DEFENSIBLE AI QUESTION BANK

REVIEWER QUESTION BANK

Version 1.0 - 24 July 2026

<https://vikramjha.work/artifacts/reviewer-question-bank>

The questions examiners and internal audit actually ask about agents, per regime, and the artifact that closes each one.

WHY IT EXISTS

Most preparation optimizes for the question teams expect - 'is the model accurate?' - and reviewers ask something adjacent that the program has no artifact for. The questions below are phrased the way they arrive: narrow, evidential, and usually about one specific past action rather than about the system in general. Each is paired with what the reviewer is really testing, because answering the surface question while missing the test is the most common way a well-run program still fails an examination.

WHAT IT CONTAINS

1. Questions grouped by regime, phrased as they are actually asked
2. What each question is really testing, which is rarely what it appears to ask
3. The artifact that closes it - and the answer that looks like evidence but is not
4. Cross-regime questions that arrive regardless of sector
5. A note on which citations move, and how to check them

ASKED REGARDLESS OF SECTOR

SHOW ME ONE ACTION THIS AGENT TOOK LAST QUARTER AND WHY IT WAS ALLOWED.

What it is testing: Whether evidence is reconstructible on the reviewer's timeline rather than reproducible on yours.

What closes it: A completed reconstruction: authority, policy version, inputs, model and prompt versions, outcome.

What does not: A walkthrough of how the system is designed to work.

WHO AUTHORIZED IT TO DO THAT?

What it is testing: Whether authority is granted and recorded, or merely inherited from a service account.

What closes it: A scoped, time-boxed grant traceable to a named human.

What does not: A named product owner with no recorded grant.

WHAT WOULD HAVE STOPPED IT?

What it is testing: Whether enforcement sits at the point of action or is monitoring wearing a control's name.

What closes it: A captured denial of the same action class.

What does not: An alert that fired after the record changed.

WHICH MODEL VERSION WAS IN EFFECT?

What it is testing: Whether you are pinned or riding a provider alias.

What closes it: An immutable version identifier on the action record.

What does not: A provider name, or an alias such as 'latest'.

WHAT DID THE HUMAN APPROVER SEE?

What it is testing: Whether human-in-the-loop is a control or a formality.

What closes it: The presented context, the decision, the identity and the dwell time.

What does not: A workflow status reading 'approved'.

HOW LONG DO YOU KEEP THAT EVIDENCE?

What it is testing: Whether retention is mapped per regime or inherited from a log default.

What closes it: A retention schedule per evidence class, with legal-hold handling.

What does not: The platform's default log retention.

BY REGIME

Regulatory instruments move. Dates are given so they can be checked, and every one of these should be verified against the primary source on the day it is used - particularly the Gulf and EU entries, where secondary summaries are unusually unreliable.

INTERAGENCY MODEL RISK GUIDANCE (OCC 2026-13, 2026)

The question: Is this agent in the model inventory, and as what?

What closes it: An inventory row representing action classes, tools and delegated authority - not a single row named after the assistant.

MODEL RISK - EFFECTIVE CHALLENGE

The question: Who challenged this, and could they reconstruct the decision?

What closes it: An independent review covering authority and containment, with its limitations stated.

MODEL RISK - ONGOING MONITORING

The question: What tells you the agent's behavior has changed?

What closes it: Thresholds tied to action risk, plus a trigger that fires on provider-side version changes.

OCC 2026-13 - AGENTIC SCOPE CARVE-OUT

The question: If the framework excludes agentic AI, what governs this agent?

What closes it: The obligations attached to the action itself, plus your own documented authority model. The specification is deferred; the duty is not.

OCC 2023-17 (THIRD-PARTY RISK)

The question: How is this provider overseen, and what is your exit?

What closes it: Executed audit-rights and version-notice clauses, plus a tested replacement or cease-use path. Note the guidance sets no dollar thresholds.

NYDFS PART 500.11

The question: What access does the third party have?

What closes it: Access-scoped evidence - the obligation follows access, not contract value.

NAIC MODEL BULLETIN ON AI (ADOPTED IN 25 JURISDICTIONS, PLUS 4 WITH THEIR OWN AI REGULATION, AS OF 1 APRIL 2026)

The question: If an agent contributed to a declination, can you explain it in the terms the regulator expects?

What closes it: An adverse-action reconstruction from evidence, not a post-hoc rationalization.

STATE DOI - UNFAIR DISCRIMINATION

The question: How do you test for proxy discrimination through agent-selected features?

What closes it: Testing at the decision the customer experienced, repeated after model changes - not a one-off fairness dashboard on a model score.

HIPAA SECURITY RULE 164.312

The question: Minimum necessary - how does that hold when the agent retrieves broadly?

What closes it: Entitlement scoping at query time plus access evidence sized for agent volumes.

HIPAA - BUSINESS ASSOCIATES

The question: Does your agent platform process ePHI, and is its audit trail visible to you?

What closes it: An executed BAA and demonstrated visibility into the platform's own trail.

21 CFR PART 11

The question: Is the agent-written record attributable, and is the audit trail intact?

What closes it: Immutable action events with attribution to the agent and the authority behind it.

INDIA DPDP ACT (RULES NOTIFIED 13 NOVEMBER 2025; OBLIGATIONS COMMENCE 13 MAY 2027)

The question: What purpose was this processing bound to, and did the agent stay inside it?

What closes it: Purpose binding enforced architecturally, with sectoral overrides from RBI, SEBI or IRDAI applied where stricter.

CBUAE AI/ML GUIDANCE NOTE (ISSUED 11 FEBRUARY 2026, APPLIES TO LICENSED FINANCIAL INSTITUTIONS INCLUDING INSURERS)

The question: Where is the AI model inventory, and does it extend to third-party models?

What closes it: An inventory per 2f with name, purpose and risk rating, extended to third-party models per 9c. Note the Note is drafted in 'should' language - a supervisory expectation, not a penalty regime - but 2f ties to the CBUAE Model Management Standards (21 December 2022), which are mandatory for banks.

CBUAE 3C

The question: When did you last test for bias, and what triggered it?

What closes it: Testing at least annually and on every material change, connected to the customer-impacting decision.

CBUAE 9A / 6F

The question: What are your audit rights over the provider, and can you cease use immediately?

What closes it: Executed rights plus a tested cease-use path - a contractual promise until the exit has actually been run.

EU AI ACT - ARTICLE 50 TRANSPARENCY (APPLIES FROM 2 AUGUST 2026)

The question: Do people know they are dealing with an AI system?

What closes it: Disclosure at the point of interaction, evidenced in the outbound record.

EU AI ACT - ANNEX III HIGH-RISK (POSTPONED TO 2 DECEMBER 2027 BY THE DIGITAL OMNIBUS)

The question: Which of your uses fall in Annex III, and what is the plan?

What closes it: A scoping assessment with dates. The delay moves the deadline, not the design work.

DORA

The question: Is this an ICT service supporting a critical function?

What closes it: An early designation position - there is no advisory carve-out in the ESAs' January 2025 Q&A, so the classification should be asserted rather than assumed.

HOW TO USE IT

1. Run it as a mock examination, not a questionnaire. Have someone outside the delivery team ask, and refuse design explanations as answers.
2. Score each question as closed, partial or open on evidence alone.
3. Anything answered with a description rather than a link is open.
4. Give every open question an owner and a date before the next review cycle.
5. Re-check each cited instrument against its primary source before relying on it - several of these changed inside the last year.

Built from public standards and general practice. No client information appears anywhere in this document. Regulatory instruments move - check anything cited here against its primary source on the day you rely on it.

Free to use and adapt; attribution welcome, not required.