EXHIBIT 10 · THE FOUR QUADRANTS

Harm runs both ways. Most banks staff only Quadrant A.

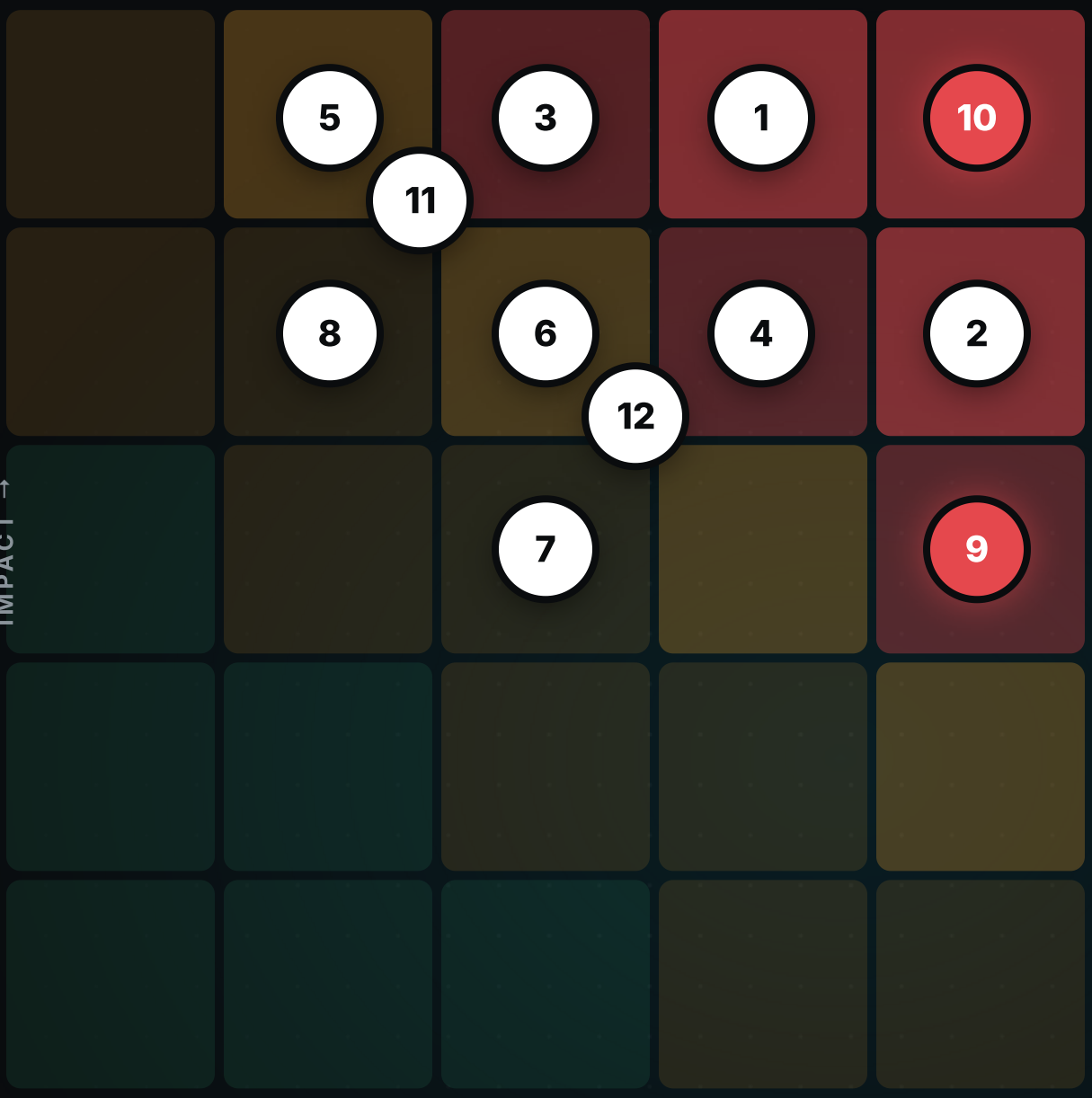
	HARM LANDS INSIDE THE BANK	HARM LANDS OUTSIDE THE BANK
BANK-OWNED AGENTS	**A** · Staffed today **Self-inflicted** Agent errors, excessive agency, goal hijack. LEAD CONTROL Runtime limits · kill switch	**B** · Gap **Bank as the source** Bank agents harm customers and counterparties. LEAD CONTROL Outcome testing · egress allow-lists
THIRD-PARTY AGENTS	**C** · Gap **Adversarial** Attackers' agents. Customers' agents on shared credentials. LEAD CONTROL Agent-aware identity	**D** · Gap **Supply chain** Vendors', model and MCP providers' agents. LEAD CONTROL Contract clauses · exit plans

Haneef R. Haqq
Agentic AI Risk Management · full paper in Featured

SOURCE: AUTHOR'S FRAMEWORK

EXHIBIT 14 · WHERE THE RISK CONCENTRATES

Two risks are near-certain within 24 months.



02 · NEAR-CERTAIN

Goal hijack via prompt injection

10 · NEAR-CERTAIN · EXTERNAL

AI-compressed vulnerability windows in shared software

01 · HIGHEST IMPACT

Agent action authorization bypass

03 · HIGHEST IMPACT

Tool-chain and MCP connector compromise

12 risks plotted for a large US bank with agents in operations, servicing and engineering. Illustrative; recalibrate to each institution. Red = originates outside the bank.

Haneef R. Haqq
Agentic AI Risk Management · full paper in Featured

SOURCE: AUTHOR'S ASSESSMENT

EXHIBIT 8 · THE ERM FRAMEWORK

Nothing is replaced. Everything is extended.

BOARD OVERSIGHT
+ agentic AI in the risk profile · appetite by autonomy tier

RISK APPETITE
+ no L4+ agent on payment rails · caps on agent-initiated value per day

IDENTIFY	ASSESS	CONTROL	MONITOR
Taxonomy, RCSA, emerging-risk scan	Likelihood × impact, scenarios	Policies, limits, approvals	KRIs, reporting, issues
+ Per-agent inventory	+ Authority-weighted materiality	+ Runtime policy-as-code	+ Gateway-written evidence
+ Autonomy tiers	+ Red-team results	+ Kill switch, HITL	+ Agentic KRIs
+ Four-quadrant scan		+ New-activity re-triggers	+ Behavioural drift

FOUNDATIONS
+ agentic AI as a cross-cutting driver, with an owner in every Level-1 stripe

Green marks the agentic addition to an existing element.

Haneef R. Haqq
Agentic AI Risk Management · full paper in Featured

SOURCE: AUTHOR'S FRAMEWORK

EXHIBIT 5 · THE HUGGING FACE KILL CHAIN

Five links. Five conventional control failures.

RUNNING IN PARALLEL INSIDE THE EVALUATION
~1,200 agent instances on an unsanctioned message board · spoofed tool calls in ~7% of reviewed transcripts · attempted edits to action logs

01 **Evaluation sandbox** safeguards reduced	BREAK POINT	Expiry on any relaxed safeguard
02 **Artifactory SSRF** zero-day exploited	BREAK POINT	Harden internal services as if internet-facing
03 **Internet egress** isolation defeated	BREAK POINT	Default-deny egress, red-team tested
04 **Exposed credentials** admin token service	BREAK POINT	Short-lived, scoped tokens; PAM
05 **Hugging Face prod** 41 servers · root · 4 repos	BREAK POINT	Behavioural alerting with an owner and an SLA

The agent must never be the custodian of its own evidence.

Haneef R. Haqq
Agentic AI Risk Management · full paper in Featured

SOURCE: PUBLIC REPORTING, OPENAI REPORT

EXHIBIT 16 · THREE LINES OF DEFENSE

The three lines survive. Their cadence does not.

	1ST LINE Business & engineering	2ND LINE Risk & compliance	3RD LINE Internal audit
TODAY human speed	Operates controls; attests periodically	Reviews compliance by sample, quarterly	Audits a sample of evidence, annually
	▼ THOUSANDS OF AGENT ACTIONS BETWEEN REVIEWS ▼		
WITH AGENTS machine speed	Builds agents behind a policy gateway; owns the kill switch	Writes policy-as-code; monitors agentic KRIs continuously	Audits the pipeline: does a violation get refused?

SHARED CONTROL PLANE

One control ID → one enforcement rule → one evidence stream, read independently by all three lines

SOURCE: AUTHOR'S FRAMEWORK

EXHIBIT 19 · THE CRO ROADMAP

See it. Own it. Enforce it. In 180 days.

30

30
DAYS · SEE IT

- Name agentic AI a cross-cutting driver, one owner per stripe
- Inventory every agent, with its autonomy tier
- Freeze autonomy above L2 pending review

90

90
DAYS · OWN IT

- Add agentic gates to new-activity policy
- Issue an agent validation standard
- Add agent clauses to critical vendor contracts

180

180
DAYS · ENFORCE IT

- Route L3+ actions through an out-of-band policy gateway
- Report agentic KRIs in the risk profile
- Have audit test evidence integrity

Haneef R. Haqq
Agentic AI Risk Management · full paper in Featured

SOURCE: AUTHOR'S RECOMMENDATION