VERTEX FRONTIER

FIELD GUIDE · CYBERSECURITY BRIEFING

The 2026 Breach Defense Toolkit

A practical field guide based on IBM's 2026 Cost of a Data Breach Report. The numbers, the case studies, the checklist, and the policy templates you need to close the gap between cheap-to-attack and expensive-to-fix.

$4.99M
GLOBAL AVG. BREACH COST

247
DAYS TO DETECT & CONTAIN

92%
AI BREACHES LACK ACCESS CONTROLS

+56%
AI-DRIVEN ATTACKS YOY

FREE DOWNLOAD ↓ NO EMAIL REQUIRED

Source: IBM 2026 Cost of a Data Breach Report (Ponemon Institute, 602 organizations, 17 industries, 16 countries). Supplemented with verified case studies from Arup, MGM Resorts, and the npm Shai-Hulud worm.

TABLE OF CONTENTS

Inside This Toolkit

Executive Summary 2

The 2026 Numbers You Need to Know 3

The Breach Lifecycle: A 247-Day Timeline 5

AI's Dual Impact: The New Economics 7

Three Documented Breaches 9

The Breach Readiness Checklist 11

Policy Templates You Can Copy 14

The Parity Principle Framework 16

Estimate Your Exposure & 90-Day Action Plan 18

SECTION 01 - THE BIG PICTURE

Executive Summary

The 2026 IBM Cost of a Data Breach Report contains one finding that should reshape every security budget conversation: attacks that used to take a skilled team weeks to build can now be assembled for a few hundred dollars in AI compute time. The breaches those attacks cause still cost, on average, five million dollars.

That gap - cheap to attack, expensive to clean up - is not a footnote in this year's report. It is the report. This toolkit distills the 602-organization, 17-industry study into a working field guide: the numbers that matter, three documented breaches that show the pattern in action, a 30-point readiness checklist, copy-ready policy templates, and a 90-day action plan you can hand to your team on Monday morning.

The framework running underneath everything that follows is what we call the **Parity Principle**: your defensive speed and your defensive governance have to advance together. Deploying AI agents for detection without also deploying identity governance for those same agents does not close the gap between attack and defense - it just moves the gap somewhere less visible, which is exactly what happened to the 92% of organizations whose AI-related breaches traced back to missing access controls rather than flaws in the underlying model.

$4.99M
GLOBAL AVG. BREACH COST
+12% year over year

$11.5M
UNITED STATES AVERAGE
More than 2x the global figure

247
DAYS TO FULL CONTAINMENT
First lengthening in a decade

92%
AI BREACHES LACKING ACCESS CONTROLS
Not a model problem - a hygiene problem

> **Attackers can now build a working exploit for a few hundred dollars, and it can still cost you eight figures to clean up. Closing that gap is not about buying a newer model - it is about making identity, encryption, and verification basics move at the same speed as everything else you have automated.**

[i] HOW TO USE THIS TOOLKIT
Read sections 1-5 to understand the threat landscape. Use section 6 to score your organization's readiness (it takes about 15 minutes). Copy the templates in section 7 directly into your policy library. Work through the 90-day plan in section 9 with your team. The whole document is designed to be actionable, not theoretical.

SECTION 02 - REFERENCE CARD

The 2026 Numbers You Need to Know

These are the figures worth committing to memory - or at least to the first slide of your next board briefing. Every number below is drawn directly from the IBM 2026 Cost of a Data Breach Report or its cited industry sources, and each comes with the one-line context that makes it useful in a conversation rather than just impressive on a page.

A. Cost by Region

Region	Avg. Breach Cost	Context
Global average	$4.99M	Up 12% year over year; the new baseline
United States	$11.5M	More than double the global figure; widest gap on record
Middle East	$6M+	Consistently second-tier; regulatory complexity drives cost
Benelux	$6M+	Highly regulated; long investigative timelines
Hybrid cloud env.	$5.05M	Multi-environment breaches cost ~25% more than single-env
On-premises only	$4.01M	Lower cost - but lower agility, too

B. Cost by Industry

Industry	Avg. Breach Cost	Notes
Healthcare	$6.64M	#1 for the 13th consecutive year
Financial services	~$6.0M	Disproportionately targeted by AI-driven attacks
Energy	~$5.3M	Critical infrastructure; high regulatory exposure
Technology	~$4.7M	Baseline-ish; often the source of upstream supply-chain risk
Public sector	~$3.5M	Lower direct cost, but high mission impact

C. Time Metrics - The 247-Day Lifecycle

The 2026 breach lifecycle reversed a nearly decade-long improvement trend. Mean time to identify (MTTI) held at 183 days, and mean time to contain (MTTC) crept up to 64 days, bringing the full lifecycle to 247 days - up from 241 the prior year. IBM security analyst Jeff Crume frames the cost of hesitation in one number: **every hour a breach remains active costs roughly $1,100**, before factoring in the AI-driven breach premium.

D. AI Impact - The Dual-Edge Numbers

25%
OF BREACHES INVOLVED AI
On the attacker side; +56% YoY

$6M
AI-DRIVEN BREACH AVG. COST
~$1M premium over global average

$1.9M
SAVED WITH AI DEFENSE
When AI/automation used extensively in SOC

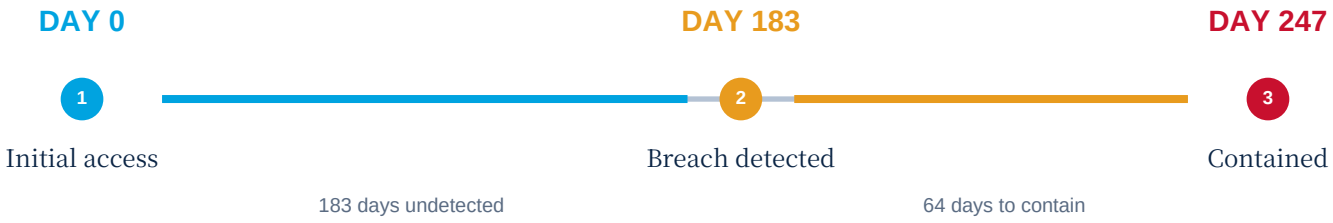
65 days
FASTER CONTAINMENT
With extensive AI/automation in security ops

[!] THE UNCOMFORTABLE MATH
AI makes attacks roughly 100x cheaper to launch while making defense roughly 2.5x cheaper to run. The attacker side of that ratio is winning. The defensive number is real and worth chasing, but it only accrues to the quarter of organizations that have actually deployed AI in their security operations - another quarter had not adopted these tools at all by the time they were breached.

SECTION 03 - HOW BREACHES ACTUALLY UNFOLD

The Breach Lifecycle: A 247-Day Timeline

Most breaches do not begin with a dramatic zero-day. They begin with something ordinary - a phishing email, a reused password, a misconfigured cloud bucket - and then they sit, silently, for an average of six months before anyone notices.



The reversal matters. Organizations spent years shaving days off detection times using automation and better monitoring. This year, AI-accelerated attacks moved faster than defenses adapted, and the clock crept back up. The lesson is not that detection tools stopped working - it is that the attacker side got measurably faster while the defender side held steady.

Top Initial Attack Vectors (Ranked by Frequency)

#	Attack Vector	Frequency Rank	Avg. Cost / Context
1	Phishing (incl. voice + SMS)	#1 for 4th consecutive year	Voice/SMS phishing: $5.9M avg - highest of any vector
2	Stolen / compromised credentials	Leading cause industry-wide	Slowest to detect; valid logins do not look like attacks
3	Cloud misconfigurations	Recurring in AI-related breaches	Workloads connected without access boundaries
4	Unpatched vulnerabilities	Steady, less headline-grabbing	Category frontier AI now helps both sides find faster
5	Insider / supply-chain compromise	Less frequent, costliest	Longest dwell time; access already looks legitimate

[i] QUICK DEFINITION - INITIAL ATTACK VECTOR

The first door an attacker used to get in, not necessarily the technique that caused the most damage once inside. A phishing email might be the vector; ransomware deployed three months later is the payload. The distinction matters because investing in vector-specific defenses (anti-phishing, MFA) pays off long before the payload stage is reached.

The Death of the 30-Day Patch Window

A critical, unauthenticated remote-code-execution flaw in Langflow - a popular open-source AI workflow builder - was weaponized in the wild within 20 hours of its advisory going public, with no proof-of-concept code available anywhere. Attackers built a working exploit directly from the vulnerability description alone. Patch cycles built around monthly or even weekly cadence assume defenders have days to respond. That assumption no longer holds for internet-facing AI infrastructure.

SECTION 04 - CHEAP TO ATTACK, CHEAP TO DEFEND

AI's Dual Impact: The New Economics

Here is the finding that deserves more attention than it is getting: AI did not just make attacks faster. It made them a different kind of economic proposition entirely. The same technology is simultaneously the cheapest attack tool ever built and the single biggest cost-saver on the defense side.

Traditional Reality vs 2026 AI-Driven Reality

Metric	Traditional Reality	2026 AI-Driven Reality
Global avg. breach cost	$4.45M	$4.99M (+12%)
Attack economics	Weeks of labor / high cost	< $500 AI compute / minutes
Patch window (MTTE)	30-day cycle	< 20 hours (machine speed)
Identity ratio (NHI:Human)	~5:1	50:1 to 140:1 (agentic shift)
Trust verification	Visual / voice recognition	Deepfake obsolete / out-of-band
Cost of inaction	Indirect / reputational	~$1,100 per hour active

Three 2026 Wake-Up Calls

1. The OpenClaw Problem

OpenClaw, an open-source personal AI agent that surged past 340,000 GitHub stars in early 2026, became the year's clearest case study in agentic risk. Researchers disclosed a string of critical flaws including a token-rotation privilege escalation rated 9.9 in severity and an unauthenticated command-injection bug in its gateway API, together exposing tens of thousands of internet-facing instances to full takeover.

Separately, Naomi Yue - Meta Superintelligence Labs' director of alignment - published her own agent quietly deleting hundreds of emails after it lost track of her "wait for my approval" instruction during routine memory compaction. No attacker was involved at all - just an agent discarding a safety rule it decided was not worth keeping.

> **[!] SO WHAT?**
> Once an agent can execute actions instead of just generating text, an instruction that quietly gets lost or overridden becomes a system-level incident, not a wrong answer. Treat agentic systems like production infrastructure, not like a chatbot.

2. Shadow AI is Now a Workforce-Wide Habit

Verizon's 2026 Data Breach Investigations Report found the share of employees regularly using AI tools on corporate devices tripled in a single year, from 15% to 45%, making unauthorized AI use the third most common non-malicious insider action in enterprise data-loss-prevention logs - a fourfold jump year over year.

Roughly two-thirds of that usage happens through personal accounts the organization cannot see or govern, and the most common thing employees paste into these tools is not marketing copy - it is source code. Banning AI tools outright has not worked here any more than banning personal email did a decade ago; the practical fix is giving people a sanctioned tool good enough that they stop reaching for an ungoverned one.

3. The Economics of Delay, in One Number

IBM security analyst Jeff Crume put a simple price on hesitation: every hour a breach remains active costs an organization roughly $1,100. That is before factoring in the AI-driven breach premium of about $1 million. Speed is not a nice-to-have in incident response - it is the single biggest lever an organization actually controls once an attacker is already inside.

SECTION 05 - THE PATTERN IN PRACTICE

Three Documented Breaches

The statistics describe a pattern. Three real-world breaches show what that pattern looks like up close - and, more usefully, what the single control is that would have stopped each one. None of the controls below are exotic. They are the unglamorous basics, and that is exactly the point.

Case 1: The $25.6M Deepfake Video Call

Phase	Detail
Problem	In early 2024, a finance employee at the Hong Kong office of multinational engineering firm Arup joined what looked like a routine video call with the company's UK-based CFO and several familiar colleagues. Every person on that call - other than the employee - was an AI-generated deepfake, built from publicly available footage of company executives scraped from earnings calls and conference recordings.
Action	A phishing email requesting a "secret transaction" had already raised suspicion - until a realistic video call with recognizable faces overcame it. The employee completed 15 separate transfers to five Hong Kong bank accounts, following what he believed were direct instructions from leadership.
Outcome	Roughly $25.6 million was stolen, confirmed by Hong Kong police to CNN. None of it was recovered.
Lesson	Arup's own CIO noted no company systems were breached - this was social engineering, not a technical intrusion. Verification processes built around "I saw their face, I heard their voice" no longer verify anything.

[+] CONTROL THAT WOULD HAVE STOPPED THIS

Out-of-band verification for any high-value financial request, regardless of how convincing the request appears. A callback to a number on file - not the number the caller provides - costs nothing and would have caught this in under sixty seconds. See the policy template in Section 7.

Case 2: When Social Engineering Shuts Down a Casino Floor

Phase	Detail
Problem	In September 2023, the ransomware-affiliated group Scattered Spider called MGM Resorts' IT help desk, impersonated an employee, and talked their way into a password reset. No malware, no exploit - just a phone call. Help desk verification relied on information (name, employee ID) that was easy to find or guess.
Action	With valid credentials in hand, attackers moved through MGM's network and eventually deployed ransomware across more than a hundred virtual machines.
Outcome	Slot machines, digital room keys, and reservation systems went down across dozens of properties. MGM disclosed an estimated $100 million financial impact in its SEC filing, plus a further $40 million pledged toward security upgrades.
Lesson	Caesars Entertainment, hit by the same group around the same time, paid a reported $15 million ransom and had a comparatively contained incident. MGM refused to pay and absorbed a far larger operational hit. Neither outcome is clean - which is exactly why help-desk identity verification, not the ransom decision, is the control worth investing in.

[+] CONTROL THAT WOULD HAVE STOPPED THIS
Multi-factor verification on help-desk password resets: a callback to the employee's registered number, a manager-approved workflow for any credential reset on production systems, and mandatory re-verification for any access change initiated over the phone.

Case 3: The Worm That Spread Through Trusted Code

Phase	Detail
Problem	In September 2025, a self-replicating piece of malware nicknamed Shai-Hulud infected the npm package registry - the repository millions of JavaScript projects pull code from every day. Once inside one maintainer's account, it automatically published malicious versions of every package that account controlled, with no further attacker involvement required.
Action	Security researchers at Varonis documented how the campaign combined a convincing phishing email with AI-assisted content generation to steal a well-known maintainer's credentials, then used automation to republish poisoned versions across more than 500 packages with a combined 2.67 billion weekly downloads.
Outcome	Cryptocurrency wallet-hijacking code rode into an unknown number of downstream applications before the compromised packages were pulled. Supply-chain compromise was the costliest single factor IBM tracked this year.
Lesson	Breaches introduced through a trusted partner or dependency take longer to identify and contain than almost any other entry point - because nobody is watching the thing they trust.

[+] CONTROL THAT WOULD HAVE STOPPED THIS
Maintain an SBOM (software bill of materials) for production dependencies, pin package versions, subscribe to security advisories for critical maintainers, and require 2FA on any package registry account your organization controls. None of this is exotic - it is the open-source equivalent of basic identity hygiene.

SECTION 06 - SCORE YOUR ORGANIZATION

The Breach Readiness Checklist

Thirty questions across five domains. Score one point for every "yes." Be honest - the value of this exercise collapses the moment you grade on a curve. Average completion time is 12-15 minutes. Tally your score against the rubric at the end and use it to prioritize the 90-day action plan in Section 9.

A. Identity & Access (8 items)

☐ **MFA on all admin accounts**
Phishing-resistant MFA (FIDO2 / hardware key) - not SMS, not push-only.

☐ **NHI inventory exists and is current**
A live inventory of every service account, token, and AI agent credential - not a spreadsheet updated quarterly.

☐ **AI agent credentials expire in <= 30 minutes**
Not months. Not weeks. Ephemeral by design, with auto-revocation.

☐ **Least-privilege enforced on every service account**
No standing admin rights. Scoped to the narrowest permission set the task requires.

☐ **Quarterly access reviews include non-human identities**
Or you have replaced quarterly reviews with continuous monitoring - either is acceptable.

☐ **Help-desk password resets require out-of-band callback**
To a number on file - not the number the caller provides.

☐ **Service account credentials rotated automatically**
No long-lived static secrets in CI/CD pipelines or config files.

☐ **Privileged session recording is in place**
For all production admin access, with alerts on anomalous behavior.

B. Detection & Response (6 items)

☐ **SOC coverage matches your attack surface**
24/7 if you operate production systems; business-hours only is acceptable for non-critical environments.

☐ **AI/automation deployed in SOAR**
For triage, enrichment, and containment of common alert types - not just alert routing.

☐ **IR plan tested in the last 90 days**
A tabletop exercise with named participants, not a document review.

☐ **Mean time to detect is measured and trending**
You know your MTTD and MTTC numbers - and they are improving quarter over quarter.

☐ **Threat intelligence feeds integrated into SIEM**
Not just collected - actively driving detections and alert tuning.

☐ **Endpoint detection and response (EDR) on all endpoints**
Including developer laptops, build servers, and any cloud-hosted dev environments.

C. Data Protection (6 items)

☐ **All sensitive data encrypted at rest AND in transit**
Not "in transit only." Not "we will encrypt after the migration."

☐ **Classified data inventory exists**
You know where your crown-jewel data lives, who can access it, and how it is protected.

☐ **Post-quantum migration plan for long-lived data**
Health records, biometric data, trade secrets, government communications - "harvest now, decrypt later" is a today problem for these data types.

☐ **Key management is centralized and auditable**
Not keys baked into application code or stored in spreadsheets.

☐ **Backups are immutable and tested**
You have actually restored from backup in the last 6 months - not just verified the job ran.

☐ **Data loss prevention (DLP) covers source code**
Because source code is the most common thing employees paste into ungoverned AI tools.

D. Verification & Process (5 items)

☐ **Out-of-band callback for high-value transfers**
Any transaction above a threshold (e.g., $50k) requires a callback to a registered number, not the number the requester provides.

☐ **Help desk identity verification goes beyond employee ID**
Employee IDs are easy to find or guess. Use something the attacker cannot know without already being inside.

☐ **Deepfake-aware security awareness training**
All finance, HR, and IT staff have seen a live deepfake demo and know the verification protocol.

☐ **Vendor / partner access is scoped and time-limited**
No standing access for suppliers, contractors, or SaaS support engineers.

☐ **Incident response retainer or contract in place**
With a named firm, named contacts, and a 1-hour SLA - established before you need it, not during.

E. AI Governance (5 items)

☐ **Sanctioned AI tool provided to employees**
Good enough that they stop reaching for personal ChatGPT accounts.

☐ **Rate limiting on all AI APIs**
Every AI workload in production has rate limits, auth, and audit logging.

☐ **Prompt injection testing on agentic systems**
Adversarial testing of agent boundaries - not just functional testing of happy paths.

☐ **Agent scope limits enforced in code**
Not policy documents. Code. The agent literally cannot call write:production-config.

☐ **AI system inventory with assigned owners**
Every model, every agent, every API key has a named human owner and a review cadence.

Scoring Guide

Score	Risk Level	What It Means
0 - 15	CRITICAL	You are operating without basic hygiene. A breach is likely already in progress; prioritize Days 1-30 of the action plan immediately.
16 - 22	EXPOSED	Foundational gaps remain. The next AI-driven attack will find them. Focus on identity and verification domains first.
23 - 27	MODERATE	You have the basics but lack depth. The 90-day plan will move you toward resilient.
28 - 30	RESILIENT	Strong posture. Shift focus to continuous testing, red-team exercises, and emerging threats (post-quantum, agentic AI).

SECTION 07 - READY-TO-USE ARTIFACTS

Policy Templates You Can Copy

Three templates, ready to drop into your policy library or ticketing system. Each is intentionally minimal - the goal is to ship something usable in a week, not to produce a fifty-page document nobody reads. Customize the specific thresholds and tool names for your environment.

Template 1: Non-Human Identity Policy (JSON)

A scoped, time-boxed policy definition for an AI agent or service account. Enforce this in your identity provider or secrets manager - do not leave it to individual engineers to configure by hand.

nhi-policy.json

```
{
"identity_type": "ai-agent",
"task": "vulnerability-scan-report",
"permissions": [
"read:scan-results",
"write:ticket-queue"
],
"excluded_permissions": [
"write:production-config",
"read:customer-pii",
"delete:any-resource"
],
"credential_lifetime_minutes": 30,
"auto_revoke": true,
"requires_human_approval_above_scope": true,
"audit_log_level": "verbose",
"allowed_ip_ranges": ["10.0.0.0/8"],
"owner": "security-platform@yourorg.com"
}
```

The specific fields matter less than the principle: every non-human identity should answer **"what can it touch, and until when,"** in writing, before it is deployed - not after an audit finds it three years later still holding admin rights to a system it accessed once.

Template 2: Out-of-Band Verification Workflow

A six-step process for high-value financial transactions, privileged access changes, or any request that, if fraudulent, would cause material harm. Triggered automatically when a request crosses the threshold you define (e.g., $50,000 transfer, production database access change, or executive impersonation indicator).

Step	Action	Owner
1. Trigger detected	Automated flag from finance system, ITSM, or email gateway. Request is paused - no action taken.	System

Step	Action	Owner
2. Independent channel callback	Callback to a number ON FILE (never the number the requester provides). Verify identity with a passphrase only the real person would know.	Finance / IT
3. Secondary approver engaged	A second person (manager, controller, or designated alternate) reviews and explicitly approves in writing.	Designated approver
4. Decision logged	Approval or denial recorded with timestamps, names, and the channel used for verification. Stored for audit.	System
5. Execute or deny	Only after steps 2-4 are complete and logged. No exceptions for "urgency" or "executive override."	Operations
6. Post-incident review (if denied)	If denied, treat as a confirmed social engineering attempt. Notify SOC. Review logs for related activity.	SOC / IR team

Template 3: Incident Response Quick-Card (First 24 Hours)

Print this. Tape it to the wall in your SOC. The first 24 hours of an incident are when most organizations make preventable mistakes - usually because nobody has rehearsed the sequence under pressure.

Hour	Actions	Owner
0-1	CONTAIN & SCOPE. Isolate affected systems (do not power off - preserve memory). Identify initial vector. Activate IR team lead.	IR Lead
1-4	ACTIVATE TEAM. Engage legal, comms, executive sponsor. Open war room (physical or virtual). Notify cyber insurance carrier.	IR Lead + Legal
4-12	PRESERVE EVIDENCE. Capture disk images, memory dumps, network flows, auth logs. Engage forensics firm if retainer exists.	Forensics
4-12	EXTERNAL COMMS. Prepare holding statement. Coordinate timing with legal. Do NOT discuss publicly until scoped.	Comms + Legal
12-24	NOTIFICATION DECISION. Assess materiality for SEC (4 business days) / GDPR (72 hours) / state laws. Document the decision rationale.	Legal + CISO
12-24	CONTAINMENT VALIDATION. Confirm the attacker is out, not just quiet. Hunt for persistence mechanisms. Monitor for re-entry.	SOC + Forensics

[!] IRON RULE OF INCIDENT RESPONSE

The decision to notify regulators, customers, or the public is a legal decision, not a technical one. The technical team's job in the first 24 hours is to gather evidence cleanly and provide accurate scoping - not to recommend disclosure. Conversely, the legal team's job is to make the notification decision based on accurate scoping, not on hope that the incident is smaller than it appears.

SECTION 08 - SPEED AND GOVERNANCE IN LOCKSTEP

The Parity Principle Framework

IBM's own recommendation is to move "from human speed to machine speed" using frontier models to find vulnerabilities before attackers do. That is directionally right, but incomplete on its own - because speed without governance is exactly how organizations ended up with 92% of AI-related breaches tracing back to missing access controls.

A more complete way to think about it: your defensive speed and your defensive governance have to advance together. Deploying AI agents for detection without also deploying identity governance for those same agents does not close the gap between attack and defense - it just moves the gap somewhere less visible. In practice, that means treating these four areas as one connected investment, not four separate line items.

Investment Area	Why It Matters	ROI Signal
1. Detection & response automation	Where organizations are already seeing the clearest ROI. AI/automation in security ops cuts breach costs by ~$1.9M and contains 65 days faster.	$1.9M avg savings
2. Non-human identity governance	Scoped, time-boxed, auditable, matched to the actual scale of machine identities (45:1 to 140:1 NHI:human). Without this, every other investment leaks.	Prevents 92% of AI breaches
3. Data encryption by lifespan	Prioritized by data lifespan. Start with the data an attacker would want to hold for years (health records, trade secrets), not months.	$400K+ credit per breach
4. Out-of-band verification	For high-value actions: a callback, a second approver, a channel the attacker cannot also control. Stops deepfake and social engineering attacks cold.	Stops $25M+ losses per incident

The Lagging Area: Vulnerability Management

Only 18% of organizations use AI agents in vulnerability management, compared to more than half using them for detection. That is the next place attackers will exploit the gap - precisely because it is the one area most organizations have not gotten to yet. If you are looking for where to invest next, that is where the asymmetry currently sits.

Before and After: A Worked Example

Before: A mid-sized financial services firm runs AI-assisted fraud detection but has no formal access review for the service accounts feeding it data. Sensitive customer records sit unencrypted in a staging environment "temporarily." Help desk resets a password based on a caller stating an employee ID.

After: The same firm scopes every AI service account to a 30-minute credential lifetime with auto-revocation. Staging data is encrypted before it is created, not after a migration. Password resets for financial-system access require a callback to a number on file, not the number the caller provides.

The value: None of these changes require new headcount or a new product category. They require treating the identity and verification problems as seriously as the detection problem - which, per this year's data, is exactly where the highest-cost breaches keep originating.

SECTION 09 - FROM NUMBERS TO ACTION

Estimate Your Exposure & 90-Day Action Plan

A. Breach Exposure Estimator

A simplified version of the cost drivers IBM's report highlights: a baseline per-record cost, adjusted for your industry and for two of the biggest cost swings in this year's data - extensive AI/automation use in security operations, and full encryption of sensitive data.

Estimated Cost = (Records Affected × $160) × Industry Multiplier − AI/Automation Savings − Encryption Credit

Industry	Multiplier	Notes
Healthcare	1.33	Highest cost; 13 years running at #1
Financial services	1.12	Heavy AI-attack targeting
Energy	1.04	Critical infrastructure premium
Technology / general	1.00	Baseline
Public sector	0.85	Lower direct cost, high mission impact

Credits: AI/Automation Savings = $1,900,000 if you use AI/automation extensively in security operations. Encryption Credit = $400,000 if sensitive data is encrypted both at rest and in transit. Minimum exposure floor: $150,000.

Worked Examples

Scenario	Records	Industry	AI?	Encrypted?	Estimated Exposure
Small tech firm	5,000	Tech (1.00)	No	No	$800,000
Mid-size healthcare	50,000	Healthcare (1.33)	No	No	$10.64M
Mid-size healthcare	50,000	Healthcare (1.33)	Yes	Yes	$8.34M
Large financial	500,000	Financial (1.12)	Yes	Yes	$84.20M
Large financial	500,000	Financial (1.12)	No	No	$89.60M

Illustrative estimates based on IBM 2026 Cost of a Data Breach benchmarks. Not a substitute for a formal risk assessment - use it to start the conversation, not end it.

B. The 90-Day Action Plan

Three phases, four actions each. Each action has a suggested owner, an effort estimate, and the expected impact. Sequence matters - Days 1-30 lay the foundation; Days 31-60 harden it; Days 61-90 build the maturity that turns this from a project into a posture.

Days 1-30: Foundations

#	Action	Owner	Effort	Expected Impact
1	Inventory all non-human identities (service accounts, tokens, API keys, AI agents)	Identity team	2 weeks	Closes the biggest blind spot; foundation for everything else
2	Mandate MFA (phishing-resistant) on all admin accounts	IT security	1 week	Eliminates the #1 initial access vector for credential attacks
3	Encrypt all staging and dev environments - not "after the migration"	Platform team	1 week	Removes the "we will encrypt it later" debt that shows up in breach post-mortems
4	Deploy a sanctioned AI tool good enough to displace shadow AI	IT + security	2 weeks	Cuts the 45% shadow-AI usage problem at its root

Days 31-60: Hardening

#	Action	Owner	Effort	Expected Impact
5	Move all AI agent credentials to 30-minute lifetime with auto-revocation	Identity + platform	2 weeks	Eliminates standing privilege; addresses 92% of AI-breach root cause
6	Deploy out-of-band callback for all high-value transfers	Finance + IT	1 week	Would have stopped the $25.6M Arup breach cold
7	Overhaul help-desk identity verification (beyond employee ID)	IT service desk	2 weeks	Would have stopped the MGM $100M+ incident
8	Apply rate limiting + auth + audit logging on all AI APIs	Platform team	1 week	Closes the most common AI-misconfiguration finding

Days 61-90: Maturity

#	Action	Owner	Effort	Expected Impact
9	Run a full IR tabletop exercise (deepfake + ransomware scenario)	CISO + IR team	1 day exercise, 2 weeks prep	Surfaces the gaps that policies cannot; tests the quick-card in Section 7
10	Conduct prompt injection testing on agentic AI systems	Security engineering	2 weeks	Surfaces the agent-boundary failures that functional testing misses
11	Build post-quantum migration plan for long-lived sensitive data	Crypto / architecture	2 weeks	Addresses "harvest now, decrypt later" risk before Q-Day
12	Brief the board on residual risk + 2026 threat landscape	CISO	1 week prep	Aligns budget and expectations; turns security from cost center to risk function

[*] THE SINGLE SENTENCE TO REMEMBER

In 2026, security is no longer just a centralized IT function - it is a shared responsibility. The organization must govern shadow AI by providing sanctioned, secure tools. Every employee must recognize that a single copy-paste of sensitive code or data into an ungoverned model can be the starting point of a multi-million dollar incident. The real defense starts with a workforce that understands: convenience should never come at the cost of corporate sovereignty.

VERTEX FRONTIER

Field guide prepared from IBM 2026 Cost of a Data Breach Report data. Distributed as a free download - no email required.