

Scaling AI with Control

Building Trust and Resilience Across
the Enterprise

Copyright ©2026 | All rights reserved.

The information contained herein has been obtained or derived from sources believed by DSCI to be reliable. However, DSCI disclaims all warranties as to the accuracy, completeness, or adequacy of such information. We shall bear no liability for errors, omissions or inadequacies in the information contained herein, or for interpretations thereof.

The information contained herein should not be relied upon as a substitute for specific professional advice. Professional advice should always be sought before taking any action based on the information provided.

You may not, distribute, modify, transmit, reuse, or use the contents of the report for public or commercial purposes, including the text, images, presentations, etc. without prior consent from DSCI.

Table of Contents

1. Executive Summary	4
2. The AI Enablement Wave	5
AI Adoption and Enablement	5
From IT-Owned to Business-Driven AI	6
AI-Driven Transformation	6
AI Proliferation and Emergence of Hidden Exposure	7
3. AI Enablement as an Exposure Multiplier	9
Business Impact Assessment	11
4. Exposure Surface: What AI Can Expose	14
Identity & Credential Exposure	14
Data Exposure	15
5. Compounding Enterprise Risk: From Exposure to Impact	18
6. Building Enterprise AI Trust & Resilience	20
7. Unified Layer for Building Enterprise AI Trust & Resilience	29

1

Executive Summary

An organization's AI resilience depends on its ability to understand and control the data, identities, permissions and execution pathways that AI tools can access. As AI becomes more deeply integrated into enterprise operations, excessive permissions, unmanaged identities, sensitive-data exposure and limited governance can evolve into interconnected exposure pathways with far-reaching business impact. Securing AI, therefore, requires more than governing models; it requires mapping the AI footprint, understanding what it can access, how we can limit access to its needs, control what it can do, observe what it does and how it responds when risk changes.

Prevention alone is insufficient in an environment where AI operates dynamically and at machine speed. Organizations need continuous visibility into AI activity, data

access and downstream actions to understand not only what an AI system was instructed to do, but what it accessed, executed and exposed. By correlating AI activity with identities, permissions and sensitive data, organizations can identify abnormal activity, assess potential blast radius and respond before risk propagates across interconnected systems.

The report aims to establish the concept of continuous AI security control loop: discover, understand, minimize, control, observe and respond. This approach enables organizations to reduce unnecessary exposure, maintain effective control over AI-enabled access and actions, and build the resilience required to scale AI safely and responsibly.

2

The AI Enablement Wave

AI adoption is accelerating → Businesses are becoming AI driven and agentic → AI is accessing more data, systems and workflows → AI exposure is expanding across the enterprise

AI Adoption and Enablement

Organizational AI adoption has moved beyond experimentation. The question is no longer whether organizations are using AI, but where it is being embedded, what enterprise data it can access, and how it is changing the way information is discovered, analyzed and acted upon.

As organizations deploy copilots, assistants, and AI agents across business functions, AI is becoming integrated into day-to-day operations. Its value is directly tied to its ability to access enterprise knowledge, connect disparate systems, and help employees work more efficiently. As a result, AI is fundamentally reshaping how organizations consume, generate, store and act upon data.

Employees are also using AI to write code, analyze information, generate content, conduct research and automate routine tasks such as meeting summarization and email drafting. This rapid adoption is delivering measurable productivity benefits, enabling organizations to accelerate decision-making, reduce manual effort and improve operational efficiency.

However, AI adoption is also expanding the enterprise exposure surface. To deliver value, AI systems require access to enterprise applications, repositories, communications platforms and business data. Every new AI assistant, agent, connector and integration creates additional relationships between identities, permissions, sensitive data and enterprise systems that organizations must understand and govern.

The share of organizations that are now using AI in at least one business function has increased from 50% in 2020 to 89% in 2026.¹ At the same time, only 44%¹ report organization-wide scaling of AI initiatives. This highlights a growing gap between AI adoption and AI maturity. While many organizations have deployed AI capabilities, relatively few have established the visibility, governance, and operational controls required to manage AI at enterprise scale.

From IT-Owned to Business-Driven AI

Another significant shift is the way AI adoption is being driven. AI is no longer confined to technology-led transformation programs. It is becoming function-led transformation, with business units adopting AI to optimize processes, improve decision-making, automate repetitive work and personalize customer and employee experiences.

As AI adoption becomes more decentralized, organizations often see AI capabilities emerge outside traditional technology programs. Employees adopt AI-enabled applications, departments deploy specialized AI solutions, and teams introduce new AI agents and workflows to solve business problems. While this accelerates innovation, it can also create visibility and governance challenges across sanctioned and unsanctioned AI use.

AI is, therefore, evolving from a technology initiative into an embedded enterprise capability. Individual functions are identifying and scaling AI use cases based on their specific business needs, while security, data and governance teams must maintain visibility into mapping the

AI footprint, understanding what it can access and how it interacts with sensitive information.

AI-Driven Transformation

Early AI deployments focused primarily on enhancing existing workflows through capabilities such as document drafting, information summarization and code generation. The next phase of adoption is more transformative. Organizations are beginning to redesign business processes around AI's ability by embedding GenAI and agentic AI capabilities to continuously analyze information, generate recommendations, coordinate activities and execute actions.

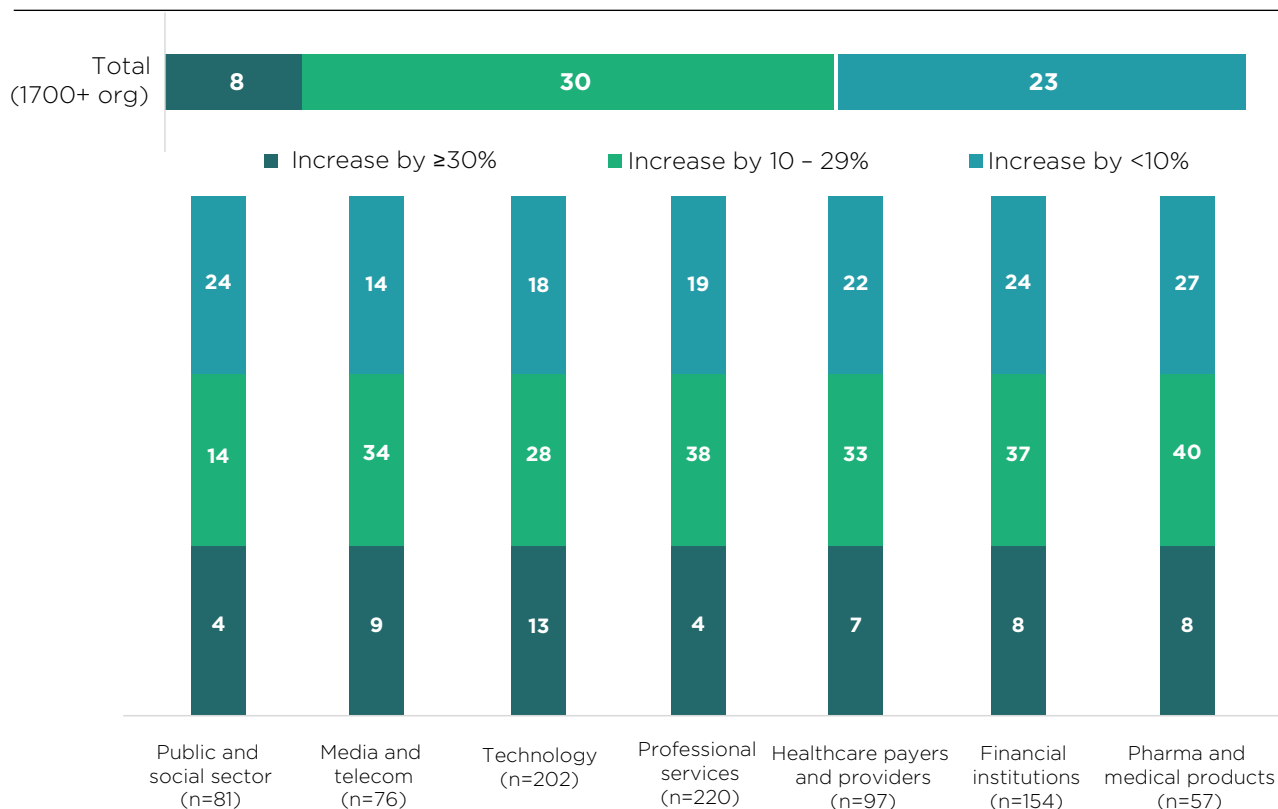
This shift toward agentic AI represents a significant evolution in enterprise operations. Rather than simply assisting users with individual tasks, AI systems are becoming capable of orchestrating workflows, interacting with business applications and acting on information across multiple systems.

As organizations increase investment in AI, they are also expanding the number of AI applications, agents, connectors and automated workflows operating throughout the enterprise. The challenge is no longer limited to managing individual AI tools. It is understanding organization's AI footprint, which data it can access, what permissions it inherits and what actions it can perform.

For organizations pursuing enterprise-scale AI adoption, success will depend not only on enabling AI innovation, but also on maintaining visibility and control over the growing network of relationships between AI, identities, data and critical business systems.

¹ *The state of AI in 2026: On the road to ROI, McKinsey & Company*

Figure 1: Expected Change in Organization-Wide AI Investment in 2027, % of Respondents²



AI investment is shifting from initial experimentation toward more substantial budget commitments, with the 10–29% increase consistently accounting for the largest share. However, a notable proportion of organizations are also planning to increase AI investments by more than 30%, indicating that AI is becoming a strategic investment priority rather than remaining a limited technology initiative. A total of 1,700+ organizations participated in the survey. The above-highlighted sectors represent only a selected subset of sectors (count mentioned as n) from the overall respondent base.

AI Proliferation and Emergence of Hidden Exposure

Rapid AI adoption is moving beyond formally approved organizational boundaries. Employees are using generative AI tools, coding assistants, browser extensions and AI-enabled SaaS capabilities, while business and technology teams are introducing agents, connectors, and AI workflows into existing processes.

This growing AI proliferation creates a new layer of enterprise exposure, particularly when these capabilities operate without central visibility, ownership or security controls. It can give rise to unmanaged AI exposure, including Shadow AI, unauthorized AI applications, unsanctioned AI agents, and ungoverned AI workflows.

An employee may share sensitive data with an external AI application. A developer may connect an AI assistant to an internal repository. An AI agent may be granted unintended access to applications and databases through an existing identity.

These activities can create exposure points that traditional asset inventories and security controls often miss. More importantly, AI can activate vulnerabilities that were previously dormant. Excessive permissions, stale entitlements, broad service-account access, exposed sensitive data and weakly governed repositories already exist in most organizations. Historically, exploiting these weaknesses required significant time, effort and specialized knowledge. AI changes that equation. When AI can search, retrieve,

² The state of AI in 2026: On the road to ROI, McKinsey & Company

correlate and act on information at machine speed, these exposures become significantly more likely to result in sensitive-data exposure, data leaks and broader security incidents.

Fundamentally, AI risk is not just a model-security challenge. It is also a data-security and access-control challenge. The value AI delivers is directly tied to the data it can access and the actions it can perform.

AI proliferation also creates new relationships between AI, identities, data and enterprise systems. An AI workflow may operate through a non-human identity, invoke multiple tools and APIs, query enterprise data sources and move information across applications within a single task. Without visibility into these interactions, security teams may see legitimate application or database activity without recognizing the AI-driven context behind it.

The resulting risk is, therefore, not simply the number of AI tools deployed, but the number of unmanaged AI-driven access and execution pathways created across the enterprise. Every AI application, agent,

connector and embedded capability introduces additional relationships between AI, identities, permissions, sensitive data and enterprise systems. If these relationships remain undiscovered or uncontrolled, isolated weaknesses can become interconnected exposure points, increase the potential blast radius of a compromise and enable unauthorized access, sensitive-data exposure and broader business impact.

The subsequent section examines how the proliferation of AI across business functions is reshaping the organization's risk landscape, creating new exposure points while amplifying vulnerabilities that may have previously remained dormant. It assesses how these interconnected exposures can escalate into broader enterprise risk and identifies the key actions, controls and capabilities required to build a more resilient and trusted AI-enabled organization. Also, outlining a structured approach to strengthening AI trust, enabling organizations to manage evolving exposure while maintaining security, privacy, operational continuity and business resilience.

3

AI Enablement as an Exposure Multiplier

AI amplifies existing weaknesses → Expands attack paths → Accelerates exploitation and automation → Creates cascading exposure → Increases business impact and blast radius

With the growing adoption of AI across business functions, organizations are encountering new risk vectors and exposure areas across the enterprise. At the same time, AI can amplify and accelerate existing risks by increasing the speed and scale at which enterprise data, identities and systems can be accessed and acted upon. As organizations connect AI to applications, repositories and business workflows, pre-existing exposures such as excessive permissions, stale entitlements and overexposed sensitive data can become significantly more consequential. This makes strengthening foundational security controls critical as organizations scale AI adoption and integrate AI more deeply into enterprise operations.

Exposure Multiplier and Expanded Attack Surface – Organizational RAG in Practice

As AI becomes embedded across business processes, it creates new pathways through which enterprise data can be discovered, retrieved, correlated and acted upon. AI-enabled applications connect users, identities, data sources and business systems, creating new relationships that organizations must understand and govern. As a result, a weakness in any connected layer can have broader and faster impact across the enterprise.

The flow diagram below illustrates how an organizational Retrieval-Augmented Generation (RAG) system connects AI applications with enterprise data and

enterprise systems. RAG is used as a representative approach because it is one of the most widely adopted approaches to understanding enterprise AI adoption, spanning a broad range of use cases and business functions. This interconnected architecture highlights the new access pathways and exposure points AI can create across the enterprise.

Fundamentally, AI security is becoming a data-security and access-control challenge. The primary risk is often not the model itself, but the data AI tools can access, the permissions it inherits and the actions it can perform across the enterprise.

Exposure Multiplier

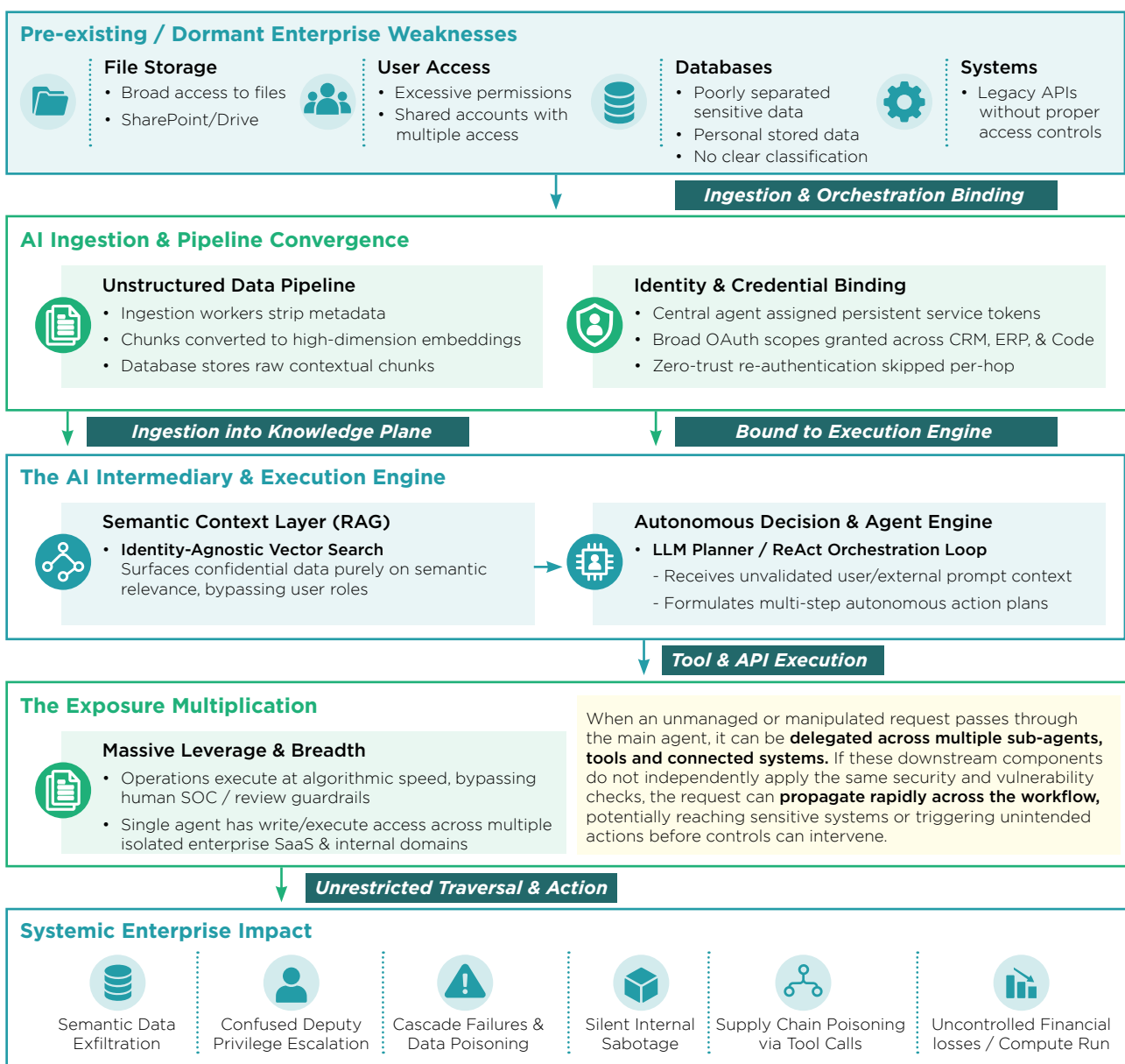
- ✦ AI model runs across large datasets, SaaS platforms and internal environments, enabling a single agent to connect previously isolated data, identities, applications and permissions, amplifying the impact of existing exposure and control weaknesses.
- ✦ Machine-speed retrieval, reasoning and automated tool execution can accelerate the spread of risk and expand the blast radius of an incident before traditional controls or human intervention can contain it.
- ✦ A compromised prompt, manipulated context source or unsafe workflow can allow AI tools to access, retrieve or expose information beyond its intended purpose, creating downstream risk across connected systems.

Expanded Attack Surface

- ✦ AI integration across applications, APIs, data platforms and agents introduces new exposure points and interconnected attack paths.
- ✦ Excessive permissions, weak identity governance and overexposed data can allow AI tools to access, retrieve or expose sensitive information beyond its intended purpose.
- ✦ Connected agents can increase the likelihood of sensitive data exposure across systems, expanding the potential blast radius of a single compromise.

Automation & Agentic Risk

- ✦ Autonomous AI agents can dynamically plan and execute multi-step actions across enterprise systems.
- ✦ Where agents operate with excessive privileges and limited human oversight, manipulated inputs can result in actions beyond the user's intended authorization.
- ✦ This creates risks such as confused deputy scenarios, unauthorized transactions and unintended changes to critical systems or records.



Confused Deputy Privilege Escalation: A trusted program, service, or automation uses its own privileges to perform an unauthorized action for a lower-privileged requester.

This diagram illustrates an enterprise-level example of how a Retrieval-Augmented Generation (RAG) system can work by connecting an AI application to an organization's internal data and knowledge sources. The same approach can be applied across sectors and use cases, including healthcare, financial services (FinTech), insurance, manufacturing, and other enterprise environments, where AI tools need to provide response-based specific information.

Cascading Risk

- ✦ AI dependencies across multiple applications and services can cause a failure or malformed input in one component to propagate across downstream processes.
- ✦ Repeated calls and automated retries/erroneous actions can degrade connected services and exhaust shared resources.
- ✦ As AI becomes embedded across business processes, a single failure can therefore evolve from a local incident into broader enterprise disruption.

Business Impact Assessment

In the previous sections, we explored how unmanaged AI adoption can increase enterprise exposure, particularly when AI systems operate without appropriate visibility, governance and security controls. As organizations connect AI to enterprise data, applications and workflows, existing weaknesses such as excessive permissions, overexposed sensitive data and poorly governed identities can have broader and faster consequences.

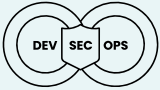
The analysis below examines how these exposures translate into tangible business impact. It traces the progression from pre-existing organizational weaknesses, through their amplification by AI-driven access, retrieval and orchestration, to the operational, financial and reputational consequences that can follow.

The objective is not to suggest that AI creates entirely new categories of risk. Rather, AI can increase the speed, scale and reach of existing risks by enabling information to be discovered, correlated and acted upon more efficiently than ever before. Weaknesses that once appeared isolated within specific applications, repositories or business processes can become interconnected exposure pathways when accessed through AI-driven workflows.

As a result, what begins as unwarranted access, exposed sensitive data or an unmanaged AI workflow can evolve into broader business impact. The potential consequences include data exposure, operational disruption, regulatory and compliance challenges, financial loss and reputational damage. In this way, AI can increase both the likelihood and potential blast radius of existing security and governance failures.

Domain	Pre-Existing / Dormant Enterprise Weakness	AI Ingestion & Amplification Mechanism	Active Exposure	Immediate Operational Disruption	Systemic Business Impact
File Storage & Data Governance 	Broader access: SharePoint/Drive folders set to "Anyone in domain with link"; unclassified, unredacted corporate archives.	Identity-agnostic chunking: Extract, transform, load embedding workers strip source Access Control Lists; text is vectorized and matched purely on cosine similarity.	Semantic data exfiltration	Low-privilege personnel retrieve highly sensitive context.	Catastrophic data breach, regulatory penalties under DPDP, insider-trading liability, and loss of competitive advantage.

Domain	Pre-Existing / Dormant Enterprise Weakness	AI Ingestion & Amplification Mechanism	Active Exposure	Immediate Operational Disruption	Systemic Business Impact
User Access & Identity Management (IAM) 	Over-Privileged Machine Identities: Monolithic service principals, static admin API keys, and wildcard scopes across internal tools.	Persistent Token Binding: Autonomous agents reuse long-lived credentials to execute multi-step tools across systems without per-hop Zero Trust re-authentication.	Confused privilege escalation	Adversaries or unauthorized staff use prompt injection to coerce agents into executing high-privilege commands.	Unbounded lateral movement, complete administrative bypass, mass resource reconfiguration, and compromised identity perimeters.
Databases & Data Integrity 	Unsegregated data stores: Stale schema designs, lack of column-level encryption, and unclassified personal/customer telemetry in production tables.	Recursive knowledge integration: Non-deterministic model outputs, hallucinations, and unredacted PII are automatically committed back into operational datasets.	Cascade failures & data poisoning	Polluted downstream databases corrupt analytics dashboards, automated forecasting models, and secondary AI inference pipelines.	Loss of enterprise data integrity, fundamentally flawed business forecasting, broken audit trails, and automated compliance violations.
Systems & Application Architecture 	Legacy internal APIs: Endpoints lacking strict input schema validation, mutual TLS or authorization checks behind internal Virtual Private Cloud (VPC) firewalls.	Dynamic agentic payload generation: LLM planners (e.g., ReAct loops) dynamically formulate and execute arbitrary endpoint payloads from unvalidated external context.	Silent Internal Sabotage & Server-Side Request Forgery (SSRF)	Autonomous execution of destructive payloads, unauthorized record updates, or arbitrary system configuration wipes at machine speed.	Severe mission-critical downtime, corrupted transaction histories, high recovery costs, and complete loss of system availability.

Domain	Pre-Existing / Dormant Enterprise Weakness	AI Ingestion & Amplification Mechanism	Active Exposure	Immediate Operational Disruption	Systemic Business Impact
Supply Chain & DevSecOps 	Unvetted third-party packages: Outdated open-source libraries, unpinned dependencies, and unverified third-party model weights.	Dynamic tool & plugin orchestration: Autonomous agents dynamically fetch external packages, invoke third-party webhooks, or query untrusted model APIs at runtime.	Supply chain poisoning via tool calls	Malicious packages, hijacked webhooks, or compromised remote tools execute arbitrary Remote Code (RCE) within internal VPCs.	Production infrastructure compromise, persistent external backdoors, code repository tampering, and software supply-chain contagion.
Financial & Resource Controls 	Missing hard compute limits: Absence of strict token budget caps, missing session timeouts, and unbounded automated retry parameters.	Machine-speed algorithmic runaway: Adversarial prompts or recursive tool loops drive infinite reasoning loops across expensive high-tier foundation models.	Uncontrolled financial losses/ compute runaway	High-tier foundation model APIs run continuously without triggering circuit breakers, exhausting GPU clusters and API credits.	Denial of Wallet (DoW), sudden budget depletion, rise in operational expenditure (OpEx), and resource starvation.

4

Exposure Surface: What AI Can Expose

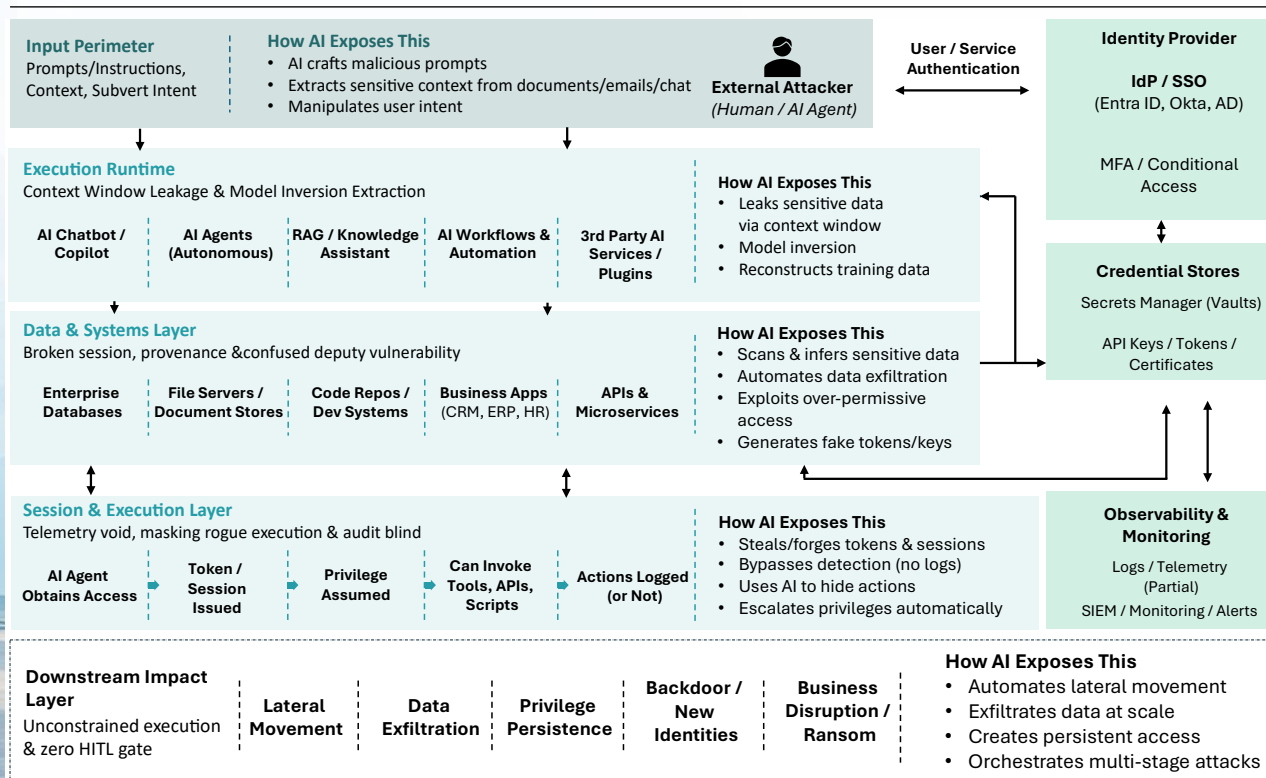
AI blurs identity boundaries → Expands credential and data exposure
→ Accelerates privilege abuse, data leakage and enterprise compromise

Identity & Credential Exposure

The deployment of organizational AI disrupts the foundational assumption of modern cyber security: that an authenticated identity corresponds to a deterministic, human-driven session. By introducing semi/fully autonomous agentic workflows, organizations are creating a pool of non-human identities operating with delegated authority across organizational systems.

AI does not just consume credentials; it abstracts identity boundaries. When autonomous agents, Large Language Model (LLM) orchestrators, and retrieval pipelines are granted machine credentials, traditional perimeter and session controls fail. The resulting environment exposes enterprise credentials to injection, privilege escalation, and lateral movement at machine speed. The following representation showcases the exposure surface for identity management and credential workflow.

Figure 2: Identity & Credential Exposure Workflow and Exposure Scenario



The scenario illustrates how an attacker can leverage AI as an amplification layer across the enterprise identity and credential chain, rather than exploiting a single isolated vulnerability. The attack begins by manipulating AI inputs and context, influencing AI-driven execution and exposing sensitive runtime information, which can then enable abuse of over-scoped credentials and non-human identities. These privileges can also propagate across delegated agent sessions and interconnected enterprise systems, while legitimate machine identities and limited telemetry make malicious activity harder to distinguish and detect. Once the AI-enabled workflow reaches downstream systems without adequate human approval, execution constraints or effective revocation, the accumulated access can translate into lateral movement, data exfiltration, privilege persistence, creation of new identities and ultimately business disruption or systemic compromise.

Data Exposure

AI adoption is reshaping how organizations access, integrate and use enterprise data. By introducing AI assistants, RAG systems, autonomous agents, embedded AI capabilities and third-party AI applications, organizations create new pathways through which structured and unstructured data can be discovered, retrieved, processed and redistributed beyond its original context.

AI does not simply consume enterprise data; it expands the reach, accessibility

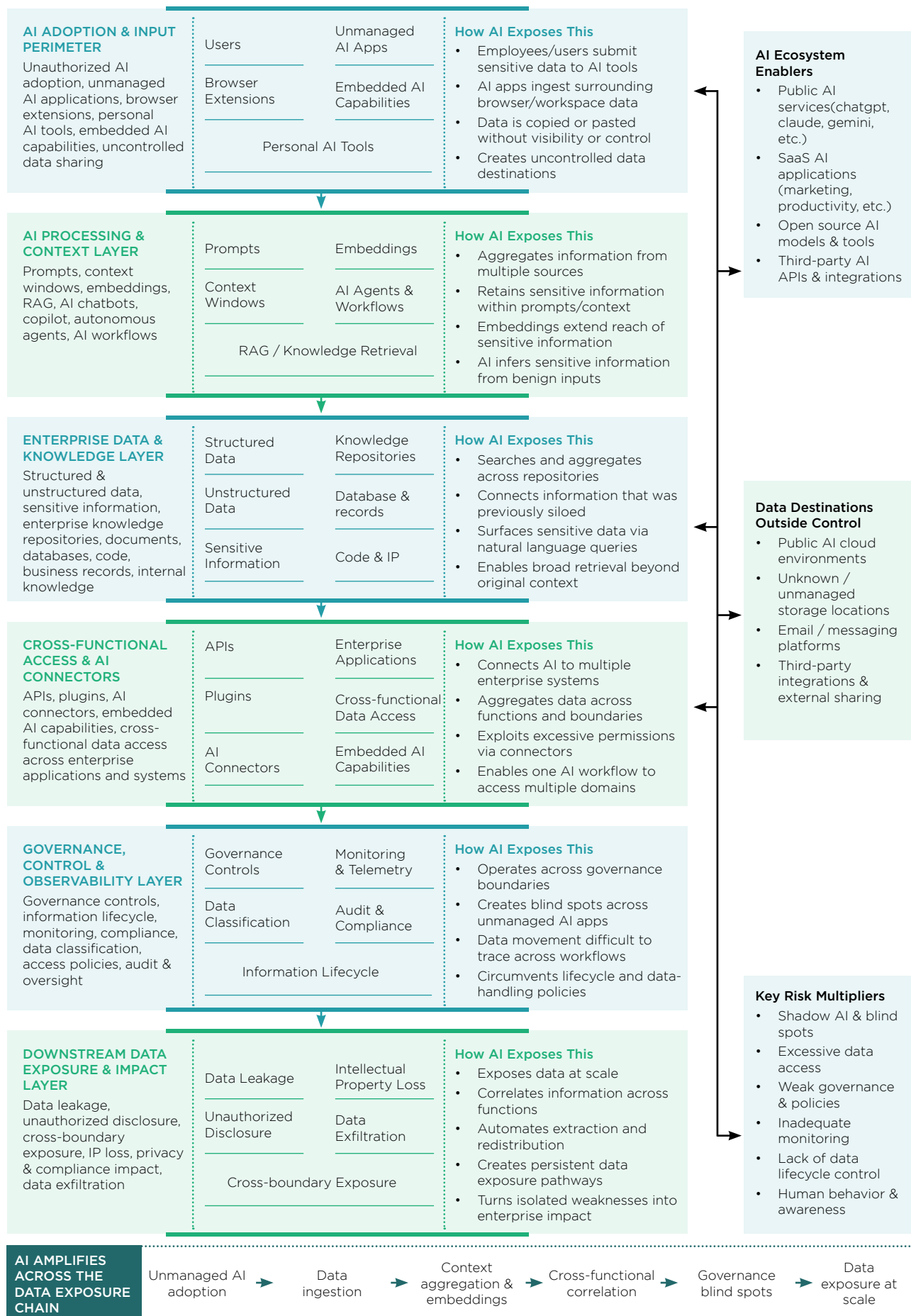
and interconnectedness of information across the organization. When prompts, context windows, embeddings, knowledge repositories, APIs, connectors and cross-functional data sources become part of AI workflows, previously isolated information can be aggregated, correlated and acted upon through a single interaction.

At the same time, Shadow AI, unmanaged applications, personal AI tools, browser extensions and external integrations can introduce new data destinations outside established governance and security controls. The resulting environment can expose sensitive data to uncontrolled ingestion, excessive retrieval, cross-functional aggregation, unauthorized access and redistribution at machine speed.

From a security perspective, the challenge is often not the AI model itself, but the enterprise data AI can access, the permissions it inherits and the pathways through which information can move across systems. As AI becomes more deeply integrated into business processes, understanding effective access to sensitive data becomes important for managing enterprise risk.

The representation below illustrates the exposure surface across the enterprise data lifecycle, from the point at which information enters an AI-enabled workflow to the repositories, systems and destinations through which it is accessed, processed and ultimately distributed.

Figure 3: Data Exposure Workflow and Exposure Scenario



The scenario illustrates how an attacker or an AI-enabled workflow acting beyond its intended boundaries, can leverage AI as an amplification layer across the enterprise data chain, rather than exploiting a single isolated data weakness. The exposure begins with uncontrolled AI adoption and data sharing, progresses through prompts, context windows and embeddings, and expands as AI connects enterprise knowledge repositories, applications, APIs and cross-functional data sources. As these connections accumulate, governance

and observability gaps can make data movement difficult to control, trace, and contain.

Once sensitive information becomes accessible across these interconnected AI workflows without following appropriate data minimization, access controls, governance and monitoring, the accumulated exposure can translate into unauthorized disclosure, intellectual property loss, cross-boundary data leakage, large-scale exfiltration, and ultimately enterprise-wide data exposure.

5

Compounding Enterprise Risk: From Exposure to Impact

AI connects identity, data and workflows → Exposures reinforce one another → Attack paths and failures cascade → Enterprise risk and blast radius compound

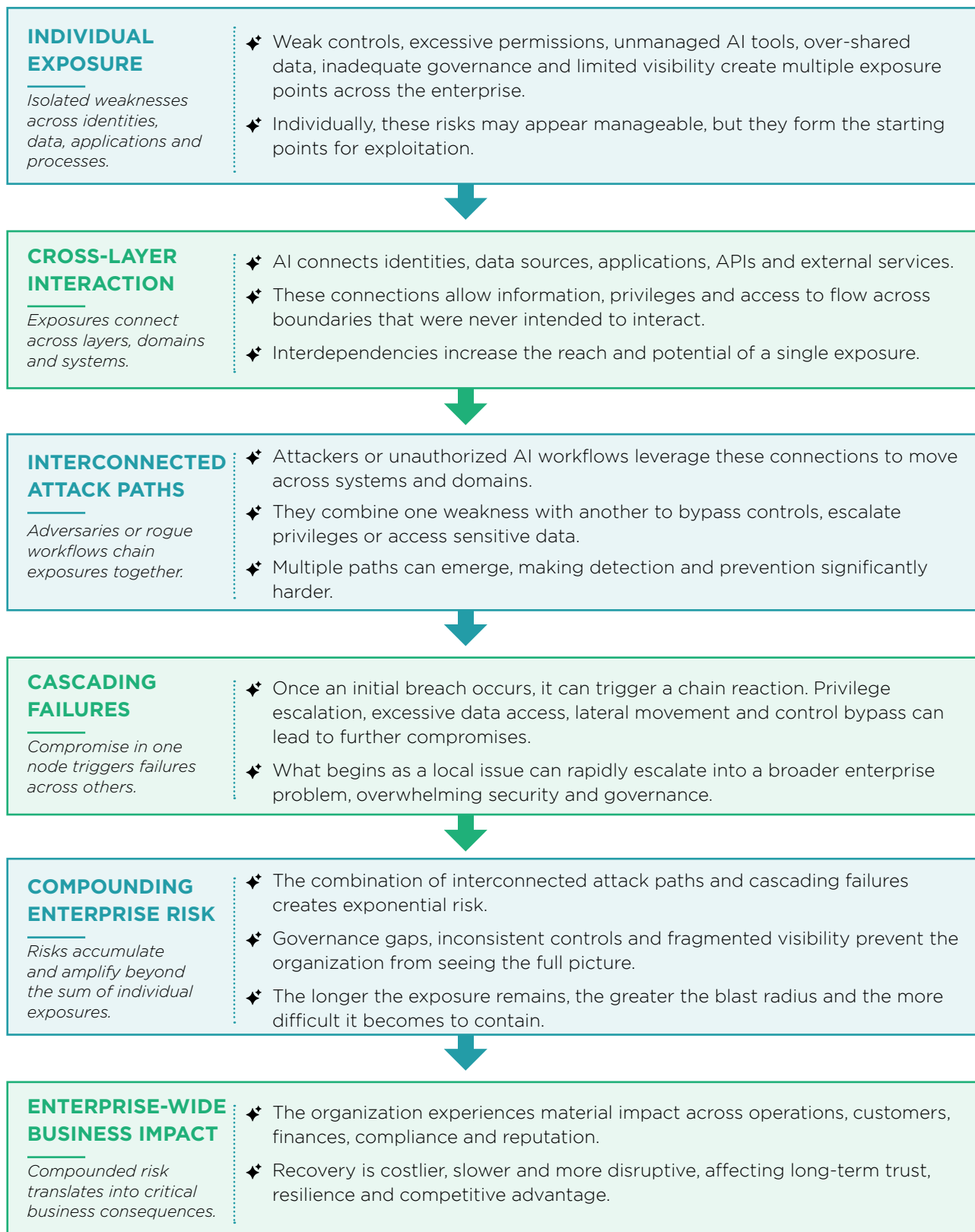
The AI risk extends beyond the security of the AI model itself. The greater concern lies in mapping the AI footprint, understanding what it can access and what actions it can execute once connected to the enterprise environment. As AI tools get embedded across the enterprise, gaps in access control, data protection, governance and visibility can turn individual weaknesses into compounding enterprise exposures.

While identity and data exposure represent distinct risk surfaces, they do not operate independently. In an AI-enabled environment, identities provide access to data, data informs AI-driven decisions, and AI-enabled workflows can use those identities to interact with additional systems. As these dependencies increase, a weakness in one layer can become an

enabler for compromise in another.

It then brings these individual exposure points together to illustrate how cross-layer interactions can create interconnected attack paths, trigger cascading failures and ultimately result in compounding enterprise risk.

The resulting risk is, therefore, not simply additive. Multiple exposures can interact, reinforce one another and expand the potential blast radius of an incident. This creates a progression from individual exposure to cross-layer interaction, interconnected attack paths and cascading failures; ultimately producing a level of enterprise risk that may be significantly greater than the impact of any individual weakness.



The central implication that can be understood is that AI-driven risk cannot be assessed through isolated controls or individual vulnerabilities alone. The combination of identity exposure, data exposure, excessive access, interconnected workflows and fragmented governance creates dependencies that allow compromise to move across organization's boundaries. As AI operates at machine speed and acts across multiple systems, these dependencies can accelerate both the scale and velocity of risk propagation.

6

Building Enterprise AI Trust & Resilience

The section above demonstrates that embedding AI into the enterprise is not a “deploy and forget” exercise. Unmanaged AI adoption can create vulnerabilities that extend well beyond the AI application itself, from excessive permissions, unmanaged identities and sensitive-data exposure to prompt manipulation, uncontrolled tool execution, weak visibility and cascading compromise across connected systems. What may begin as an isolated AI use case can therefore become a pathway into some critical failures.

This changes the question organizations need to ask, from the CXO determining risk appetite and accountability, to technology and security teams defining and enforcing controls, to data and privacy teams protecting sensitive information, to AI owners securing use cases, and employees understanding what AI tools they are permitted to use and what information they can share.

A structured way to think about organization’s AI trust

Map your AI footprint → Understand what it can access → Limit access to what it needs → Control what it can do → Observe what it does → Respond when risk changes

The Technical Capability

Corresponding to this structured approach, technical capabilities are the foundational mechanisms that make AI trust measurable and enforceable. Organizations need visibility into:

- ✦ AI estate and the relationships between AI, identities, applications and data.
- ✦ Intelligence to understand AI access and potential blast radius.
- ✦ Zero Trust and least-privilege controls to minimize unnecessary reach.

Runtime guardrails, policy enforcement, data-loss controls and human approval mechanisms are required to control what AI can execute. Continuous AI, identity, data and application monitoring must provide the telemetry and provenance needed to understand what AI does, while automated detection, access revocation, isolation and remediation enable the organization to respond when risk changes. Together, these capabilities move AI security from static governance to continuous control and assurance.

The following presents how the structured AI trust approach aligns with the required technical capabilities and how, together, they can form the foundation for strategically building organization’s AI trust and resilience.

6.1 Map your AI footprint → Understand what it can access → Limit access to what it needs → Control what it can do → Observe what it does → Respond when risk changes

Establish enterprise-wide visibility into AI models, applications, agents, APIs, copilots, connectors and autonomous workflows across sanctioned and unsanctioned environments, while mapping ownership, purpose, deployment context, dependencies and risk posture to identify shadow AI and unmanaged assets.

Questions to Ask?

Board / CXO

- ✦ Do we have a clear view of where AI is being adopted and the associated enterprise risk?
- ✦ Who is accountable for AI risk across the organization?

Management

- ✦ Do we know what AI is being introduced within our functions and processes?
- ✦ Are AI use cases appropriately owned and risk-assessed?

Security Leaders (CISOs/CIOs/CTOs)

- ✦ Do we have complete visibility of AI applications, agents, models and connectors?
- ✦ Can we identify shadow AI and unmanaged AI activity?

Employees

- ✦ Do I know which AI tools I am allowed to use?
- ✦ Do I know when an AI tool or capability needs to be approved?

Key exposures addressed

- ✦ Shadow AI, unmanaged AI apps, unknown agents, fragmented adoption.

Required technical capabilities

- ✦ **AI asset discovery & centralized inventory:** Discover AI applications, models, agents, copilots, AI-enabled workflows and deployments across cloud, SaaS, on-premises and developer environments.
- ✦ **Shadow AI & unauthorized-use detection:** Identify unsanctioned AI tools, browser-based AI, embedded AI functionality and unapproved agentic workflows operating outside formal governance.
- ✦ **AI Security Posture Management (AI-SPM):** Continuously assess AI assets for security misconfigurations, exposed interfaces, excessive permissions, insecure integrations and other AI-specific control weaknesses.
- ✦ **Model, API, agent & connector discovery:** Map AI models and services to APIs, plugins, Model Context Protocol (MCP)/connectors, enterprise applications, data sources and downstream dependencies.

✦ **AI ownership, purpose & risk classification:** Establish accountable owners, business purpose, deployment context, data sensitivity and risk tier for each AI asset.

✦ **AI dependency & relationship mapping:** Establish relationships between AI systems, identities, data stores, applications, tools and execution environments to understand the broader AI attack surface.

✦ **AI asset posture & lifecycle management:** Continuously track AI

assets from development and testing through deployment, modification, retirement and decommissioning.

What the capability enables

✦ Provides a complete view of AI applications, models, agents, tools, connectors and their relationships across the enterprise.

✦ Establishes ownership, purpose, deployment context, dependencies and risk posture for both sanctioned and unmanaged AI assets.

Strategic outcome

Comprehensive AI visibility, accountability and attack-surface understanding.

6.2 Map your AI footprint → Understand what it can access → Limit access to what it needs → Control what it can do → Observe what it does → Respond when risk changes

Maps AI systems and non-human identities to the data, applications, APIs, infrastructure and privileges they can effectively access. Establishes the sensitivity, context and potential blast radius associated with each AI system's access footprint.

Questions to Ask?

Board / CXO

- ✦ What is our business risk appetite for AI-driven data exposure?
- ✦ Could an AI system reach critical data or systems beyond its intended purpose?

Management

- ✦ What data, systems and APIs do the AI use? By which function?

Security Leaders (CISOs/CIOs/CTOs)

- ✦ Can we map AI to the identities, data, applications and systems it can access?
- ✦ Is that access aligned with business purpose?
- ✦ Which AI identities or workflows have excessive or privileged access?

Employees

- ✦ Do I understand what information and systems the AI can access?
- ✦ Am I providing access or information beyond what is necessary?

Key exposures addressed

- ✦ Excessive data access, broad connectors, privileged identities, sensitive data exposure.

Required technical capabilities

- ✦ **Data discovery & classification:** Identify and classify structured, unstructured and semi-structured data based on sensitivity, business criticality and regulatory requirements.

- ✦ **AI-to-data access mapping:** Map AI applications, agents and non-human identities to the data repositories, records, documents and knowledge sources they can access.

- ✦ **Identity & entitlement intelligence:** Analyze human and non-human identities, roles, groups, service accounts and inherited permissions associated with AI workflows.

- ✦ **Effective-access analysis:** Determine actual, transitive and inherited access across identities, applications, data stores and interconnected systems.

- ✦ **Sensitive-data exposure mapping:** Identify where AI can access personal, confidential, regulated, financial, credential and intellectual-property data.

- ✦ **Data lineage & provenance mapping:** Trace how data flows into AI prompts, context windows, RAG pipelines, embeddings, model interactions and downstream outputs.

- ✦ **Application, API & connector access visibility:** Map AI connectivity to SaaS platforms, databases, APIs, enterprise applications and external services.

What the capability enables

- ✦ Maps AI systems and non-human identities to the data, applications, APIs, infrastructure and privileges they can effectively access.
- ✦ Establishes the sensitivity, context and potential blast radius associated with each AI system's access footprint.

Strategic outcome

Contextual visibility of AI access, data exposure and potential blast radius.

6.3 Map your AI footprint → Understand what it can access → Limit access to what it needs → Control what it can do → Observe what it does → Respond when risk changes

Apply least privilege, Zero Trust, task-scoped authorization and just-in-time access so AI receives only the permissions and data required for a defined purpose. Continuously identify and remediate excessive entitlements, stale permissions, over-scoped machine identities and unnecessary data exposure.

Questions to Ask?

Security Leaders (CISOs/CIOs/CTOs)

- ✦ Are we accepting the right level of AI access risk? Are Zero Trust and least privilege applied consistently to AI?
- ✦ Are we reducing AI exposure in line with business need and risk appetite?

Management

- ✦ Are permissions and data scopes limited to what the process requires?
- ✦ Can unnecessary access and sensitive-data exposure be reduced?

Employees

- ✦ Am I sharing only the required information necessary?
- ✦ Do I understand what information should not be provided to AI?

Key exposures addressed

- ✦ Over-privileged agents, stale access, unnecessary data exposure.

Required technical capabilities

- ✦ **Zero Trust access controls for AI:** Continuously verify AI identities, context, resource and access conditions before granting access.
- ✦ **Least-privilege & task-scoped authorization:** Restrict AI permissions to the minimum resources and actions required for a specific task or workflow.
- ✦ **Just-in-time & ephemeral access:** Replace persistent privileges with time-bound, dynamically issued access.
- ✦ **Non-human identity governance:** Discover, govern and continuously assess service accounts, machine identities, agent identities and associated credentials.
- ✦ **Entitlement analysis & remediation:** Identify excessive, stale, orphaned and conflicting permissions and automatically reduce unnecessary access.

- ✦ **Data minimization & purpose-based access:** Limit the data exposed to AI based on business purpose, sensitivity and task requirements.
- ✦ **Segmentation & isolation:** Isolate AI workloads, agents, sensitive repositories and execution environments to constrain lateral movement.
- ✦ **Credential lifecycle controls:** Enforce secure issuance, rotation, revocation and expiration of AI-associated credentials and tokens.

What the capability enables

- ✦ Continuously aligns AI permissions and data access with business purpose, task requirements, sensitivity and duration.
- ✦ Eliminates excessive privileges, stale entitlements and unnecessary data exposure to reduce the AI attack surface and blast radius.

Strategic outcome

Reduced AI attack surface and privilege exposure.

6.4 Map your AI footprint → Understand what it can access → Limit access to what it needs → **Control what it can do** → Observe what it does → Respond when risk changes

Enforce runtime policies over prompts, context, tool calls, API invocations, data retrieval, outputs and downstream actions before execution occurs. Use guardrails, DLP, action-level controls, human approval, egress restrictions and circuit breakers to constrain unauthorized or high-risk AI behavior.

Questions to Ask?

Security Leaders (CISOs/CIOs/CTOs)

- ✦ What are the boundaries for autonomous AI actions?
- ✦ Can we enforce runtime controls across AI workflows, tools and connected systems?
- ✦ Which AI-driven decisions or actions should require human approval?
- ✦ Can we prevent prompt manipulation, data leakage and unauthorized actions?

Management

- ✦ Are appropriate security and privacy controls embedded in AI design and deployment?
- ✦ Do high-risk actions have appropriate approval or escalation?

Employees

- ✦ Do I know when human validation is required?
- ✦ Can I recognize and report potentially unsafe AI behavior or outputs?

Key exposures addressed

- ✦ Prompt injection, malicious context, unauthorized tool use, sensitive data leakage.

Required technical capabilities

- ✦ **AI runtime guardrails & policy enforcement:** Enforce security, privacy and usage policies dynamically during AI execution.
- ✦ **Prompt & context inspection:** Analyze prompts, retrieved context and external content for malicious instructions, sensitive information and policy violations.
- ✦ **Prompt-injection & context-manipulation detection:** Identify attempts to manipulate AI behavior through untrusted or adversarial content.
- ✦ **AI-aware DLP & sensitive-data controls:** Prevent sensitive information from being retrieved, processed, exposed or transferred outside authorized boundaries.
- ✦ **Tool, API & connector controls:** Govern which tools, APIs, plugins and enterprise services an AI agent can invoke and under what conditions.
- ✦ **Action & transaction-level controls:** Constrain high-risk operations such as data modification, privilege changes, financial transactions or infrastructure changes.
- ✦ **Human-in-the-loop (HITL) controls:** Require explicit approval for sensitive, irreversible or high-impact AI actions.
- ✦ **Egress & output controls:** Inspect and control AI-generated outputs, external transfers and downstream destinations.
- ✦ **Circuit breakers & agent kill switches:** Immediately suspend or terminate AI execution when predefined risk conditions are triggered.

What the capability enables

- ✦ Enforces security, privacy and usage policies across prompts, context, data retrieval, tool calls, API interactions and downstream actions at runtime.
- ✦ Constrains high-risk or unauthorized AI behavior through guardrails, approvals, DLP, action controls and automated circuit breakers.

Strategic outcome

Controlled, policy-bound and accountable AI execution.

6.5 Map your AI footprint → Understand what it can access → Limit access to what it needs → Control what it can do → Observe what it does → Respond when risk changes

Capture and correlate AI interactions, identity activity, data access, tool/API calls and downstream system actions to establish end-to-end activity provenance. Use behavioral analytics, activity monitoring and unified audit trails to detect anomalies.

Questions to Ask?

Security Leaders (CISOs/CIOs/CTOs)

- ✦ Can we demonstrate responsible, secure and compliant use of AI?
- ✦ Can we provide evidence of how material AI-driven decisions and actions occurred?

Management

- ✦ Can we monitor AI behavior within our processes and identify deviations?
- ✦ Do we have sufficient evidence to demonstrate compliance and accountability?

Employees

- ✦ Do I understand that AI activity may be monitored?
- ✦ Do I know how to report unusual or suspicious AI activity?

Key exposures addressed

- ✦ Telemetry gaps, legitimate IDs masking activity, lack of provenance.

Required technical capabilities

- ✦ **AI activity monitoring:** Continuously capture AI interactions, prompts, model calls, tool invocation and agent activity.
- ✦ **Identity & access telemetry:** Correlate AI activity with the human, machine or non-human identity executing the workflow.
- ✦ **Data & database activity monitoring:** Monitor what sensitive data and database objects AI accesses, queries, modifies or extracts.
- ✦ **Application & API activity monitoring:** Capture downstream application, API, connector and infrastructure activity generated by AI workflows.
- ✦ **Behavioral analytics / UEBA:** Identify anomalous AI behavior, unusual access patterns, privilege escalation and deviations from established baselines.
- ✦ **Prompt, action & execution provenance:** Establish traceability from AI instruction and context through identity, tool invocation, data access and downstream action.

- ✦ **Unified audit trails:** Maintain tamper-resistant records of AI interactions and consequential activity for investigation, compliance and assurance.
- ✦ **SIEM/SOAR integration:** Correlate AI telemetry with enterprise security events and enable automated detection and response workflows.

What the capability enables

- ✦ Correlates AI interactions with identities, data access, tool invocation, application activity and downstream system actions to establish end-to-end provenance.
- ✦ Provides continuous behavioral visibility and an auditable record to detect anomalies, investigate incidents and reconstruct AI-driven activity.

Strategic outcome

Continuous AI visibility and behavioral analytics.

6.6 Map your AI footprint → Understand what it can access → Limit access to what it needs → Control what it can do → Observe what it does → Respond when risk changes

Continuously assess AI behavior and security posture to detect anomalous activity, privilege changes, compromised identities, policy violations and emerging exposure. Enable rapid containment through access revocation, credential rotation, session termination, agent isolation, policy blocking and automated remediation before risk propagates.

Questions to Ask?

Security Leaders (CISOs/CIOs/CTOs)

- ✦ Can we contain material AI risk quickly enough to prevent business impact?
- ✦ Do we have clear decision rights for suspending high-risk AI activity?
- ✦ Can we immediately revoke access, disable connectors or isolate compromised agents?
- ✦ Do we have AI-specific incident response and recovery capabilities?

Management

- ✦ Can we quickly escalate, disable or isolate an AI capability when risk changes?
- ✦ Are response responsibilities clearly defined?

Employees

- ✦ Do I know how and where to report unsafe or suspicious AI activity?
- ✦ Do I know what action to take if an AI tool or workflow is compromised?

Key exposures addressed

- ✦ Machine-speed propagation, persistent access, cascading business impact and expanding blast radius.

Required technical capabilities

- ✦ **Detect deviations in AI access patterns,** data retrieval, tool usage and execution activity that may indicate elevated risk or misuse.
- ✦ **Behavioral anomaly detection:** Detect deviations in AI access, execution patterns, data retrieval and tool usage.

- ✦ **Continuous AI security posture assessment:** Continuously reassess AI configurations, permissions, integrations and exposure as the environment changes.
- ✦ **Automated policy-based response:** Dynamically block, restrict or escalate AI activity when defined risk conditions are exceeded.
- ✦ **Access & credential revocation:** Immediately revoke AI permissions, tokens, sessions and associated credentials when compromise or misuse is suspected.
- ✦ **Agent isolation & session termination:** Suspend compromised agents and terminate active execution pathways to prevent further propagation.
- ✦ **Automated remediation:** Remediate excessive permissions, misconfigurations, exposed data paths and other identified control weaknesses.

- ✦ **Incident investigation & response integration:** Integrate AI-specific events into enterprise detection, investigation, response and recovery processes.
- ✦ **Continuous control validation:** Test whether AI security controls remain effective as models, agents, integrations, data and workflows evolve.

What the capability enables

- ✦ Continuously detects changes in AI behavior, posture, permissions and exposure and translates them into risk-aware response actions.
- ✦ Enables rapid access revocation, credential rotation, session termination, agent isolation and automated remediation before risk propagates.

Strategic outcome

Adaptive AI resilience, rapid containment and increased time-to-impact to business operations.

Unified Layer for Building Enterprise AI Trust & Resilience

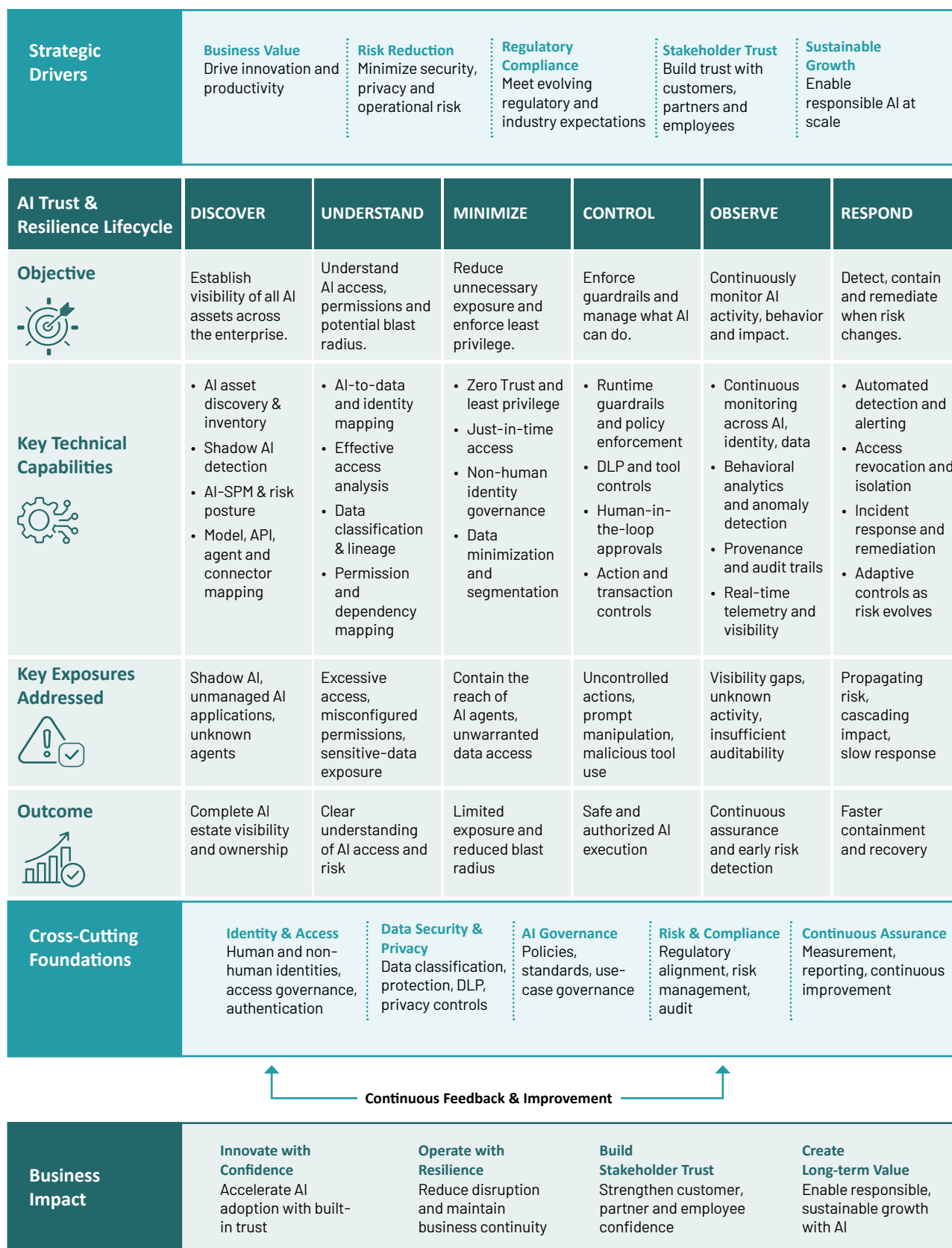
The six stages outlined above should not be viewed as standalone security capabilities, but as interconnected components of a broader Enterprise AI Trust and Resilience framework. Building trust in AI tools requires a unified approach that connects visibility, access intelligence, preventive controls, runtime enforcement, continuous monitoring and response across the AI lifecycle. By bringing together AI, identity, data, applications and activity context, organizations can establish a continuous chain of understanding, control and

assurance, from mapping the AI footprint, understanding what each AI system can access, limiting access to what it needs, controlling what it can do, continuously observing its activities, and responding as the risk landscape changes.

This integrated approach moves beyond point-in-time controls toward continuous risk management, enabling organizations to reduce exposure, constrain blast radius, strengthen accountability and scale AI adoption with greater confidence and resilience.

Unified Layer for Building AI Trust & Resilience

A strategic and technical framework to move from AI adoption to AI trust, connecting people, process and technology to enable safe, responsible and resilient AI at scale.



An organization's AI strength and resilience ultimately depend on its ability to understand, secure and control the data, identity, access and execution pathways that AI creates or amplifies. The assessment demonstrates that AI can turn excessive permissions, unmanaged identities, sensitive-data exposure, weak governance and limited visibility into interconnected exposure pathways. Addressing this requires an architecture in which every AI asset is discoverable, its ownership and purpose are understood, its effective access to sensitive data and critical systems is mapped, and unnecessary permissions are continuously identified and removed.

However, prevention alone is insufficient in an environment where AI tools can operate dynamically and at machine speed. Continuous activity monitoring must connect AI tool interactions with identities, data access and downstream system activity, creating the provenance required to understand not only what an

AI system was instructed to do, but what it actually accessed, executed and exposed. Behavioral analytics, data and database activity monitoring, unified audit trails and security operations integration provide the context required to identify anomalous activity, distinguish legitimate automation from potential misuse and understand the potential blast radius of AI-driven actions. When risk changes, automated remediation, access revocation, credential rotation, session termination and agent isolation can reduce the time between detection and containment.

The strategic objective is, therefore, to establish a continuous AI security control loop: discover, understand, minimize, control, observe and respond. This approach enables organizations to continuously reduce unnecessary exposure while maintaining visibility into how AI tools access enterprise data, interacts with enterprise systems and executes actions across the environment.

Authors

Neha Mishra, Lead - DSCI Insights & Research

Priya Sharma, Senior Associate – Strategy & Insights

Contributor

Mridushi Bose, Senior Associate - Marketing & Communications



Varonis (Nasdaq: VRNS) secures AI and the data that powers it. The Varonis platform gives organizations automated visibility and control over their critical data wherever it lives and helps ensure safe and trustworthy AI from code to runtime. Backed by 24x7x365 managed detection and response, Varonis gives thousands of organizations worldwide the confidence to adopt AI, reduce data exposure, and stop AI-powered threats.

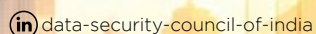
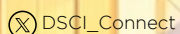


Data Security Council of India (DSCI), set up by Nasscom, is a not-for-profit industry body and think tank on data protection, cyber security, and critical technologies in India. DSCI is committed to making cyberspace safe, secure, and privacy-assured, and to advancing critical technologies vital for national security through policy advocacy, thought leadership, capacity building, and outreach initiatives. DSCI engages with the government, regulators, defence, law enforcement agencies, embassies, industry members, and academia to position India as a leader in the global digital ecosystem.

DATA SECURITY COUNCIL OF INDIA

Nasscom Campus, Fourth Floor, Plot. No. 7-10, Sector 126, Noida, UP - 201303

+91-120-4990253 | info@dsci.in | www.dsci.in



© 2026 DSCI. All rights reserved.