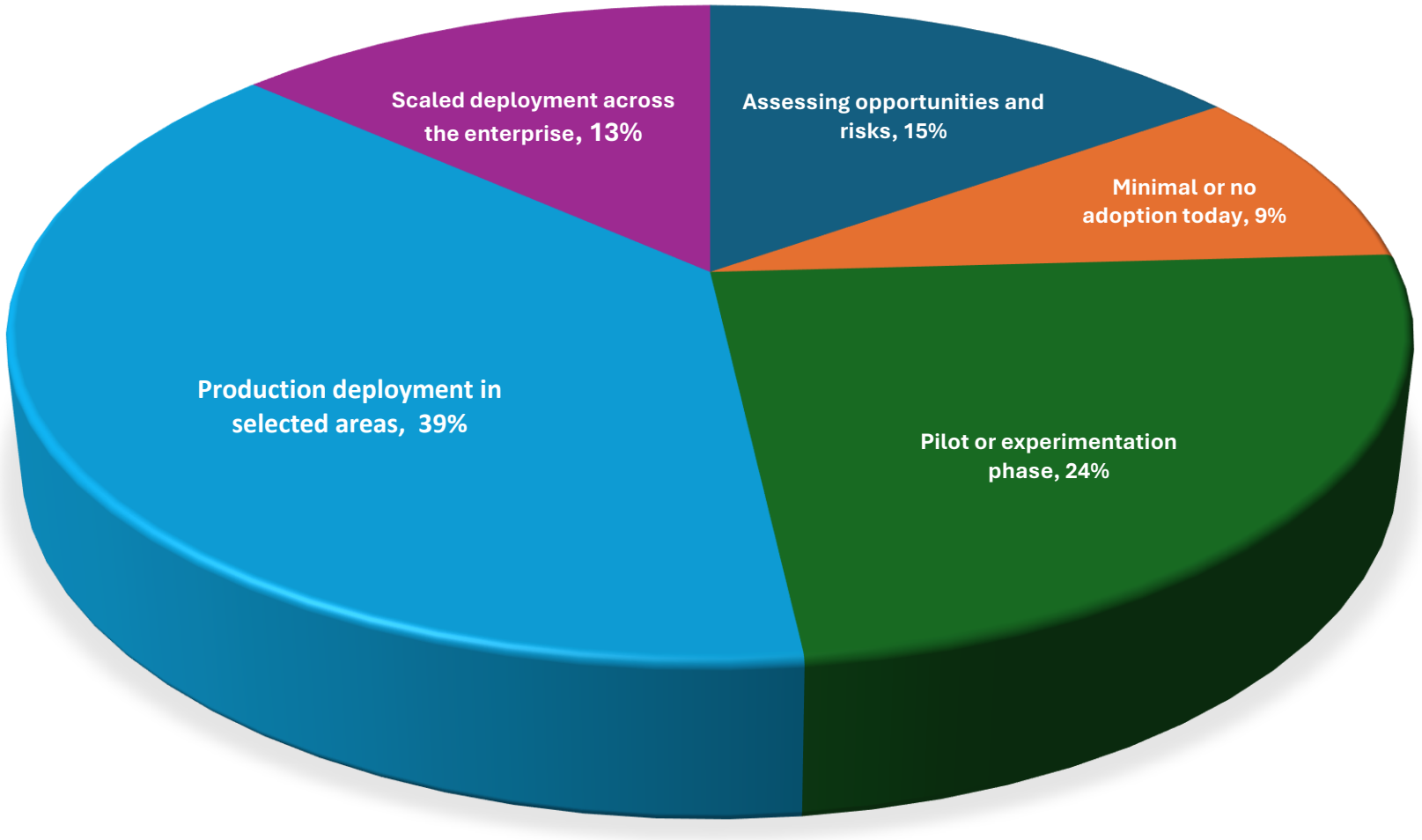
Exclusive Webinar Content

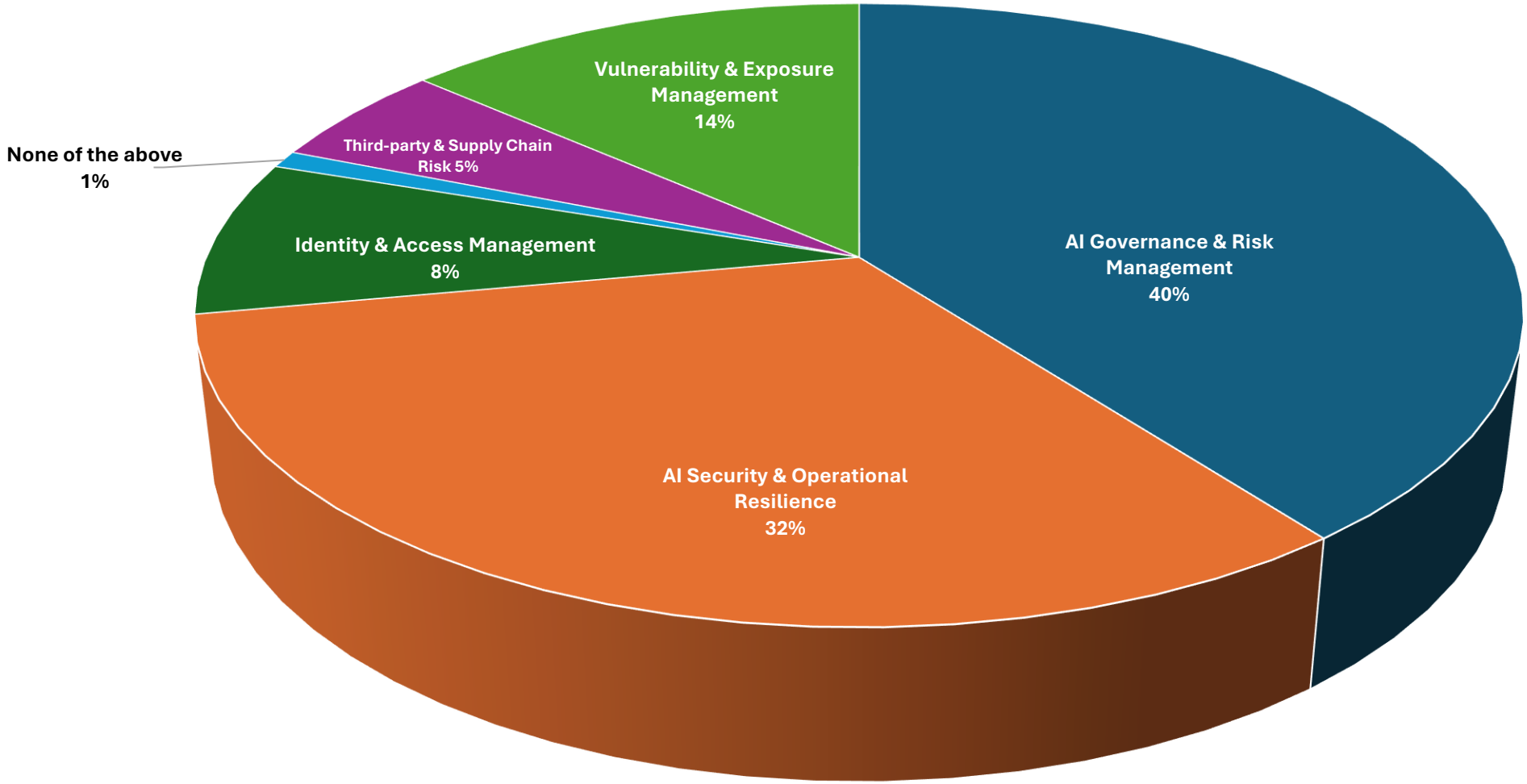
Frontier AI, Mythos & the Future of Trust in APAC Financial Services

- Webinar Poll Report
- Answers to the Q&A Session

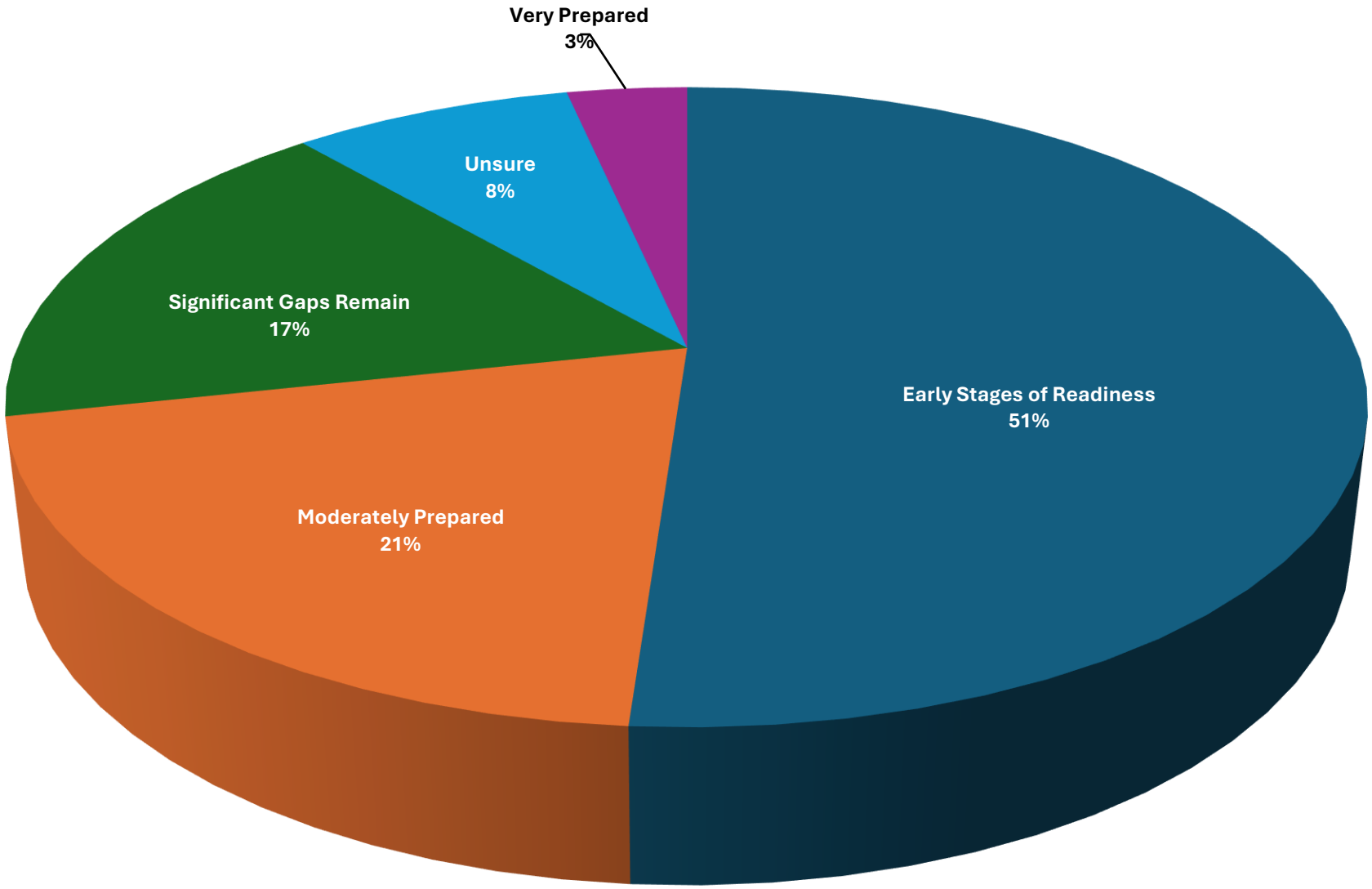
Poll #1: Where is your organisation currently on its AI adoption journey?



Poll #2: Which capability would you be investing over the next 12 months?



Poll #3: How prepared do you believe your organisation is for Frontier AI-related cyber risks?



Webinar Q&A Responses

Disclaimer

The responses below represent the personal professional opinions and perspectives of **Ashish Thapar**, shared during the webinar discussion. These views are provided for informational and educational purposes only and should not be interpreted as the official positions, policies, recommendations, or endorsements of any employer, organization, institution, client, partner, or affiliated entity.

The questions presented below are representative selections from the webinar audience. Certain questions have been consolidated where overlap existed. The responses reflect the moderator's individual observations, industry experience, and professional judgment at the time of the discussion.
Readers should exercise their own judgment and seek independent professional, legal, regulatory, security, or technical advice before making any business, technology, risk, or investment decisions based on the information provided herein. Viewer discretion is advised.

1. What should be prioritized for an insurance company that is still assessing the use of GenAI as a value-added service for customers?

The first priority should be gaining a clear understanding of the organization's AI security maturity from a governance, risk, and controls perspective. Conducting a comprehensive AI Security Maturity Assessment across governance and processes, technology, and people-related domains provides a strong foundation for informed decision-making.

This should be followed by implementing appropriate security measures that deliver visibility into AI environments while protecting AI models and applications against threats such as prompt injection, data leakage, model poisoning, and other adversarial attacks. Regular model security testing and red teaming should also be incorporated.

Finally, organizations should establish continuous monitoring and protection capabilities for AI-native and AI-enabled systems across the enterprise, including the discovery and governance of Shadow AI deployments.

2. What are your views on the effectiveness of virtual patching in the post-Mythos era?

Virtual patching should not be viewed as a permanent solution or a substitute for addressing root causes. However, it remains a highly valuable interim control that provides defenders additional time to remediate vulnerabilities and deploy long-term fixes.

Many organizations have either implemented or are actively evaluating distributed exploit protection frameworks where virtual patching serves as an effective temporary mitigation measure while permanent remediation activities are underway.

3. With frontier models such as Claude Mythos reducing time-to-exploit from weeks to hours, how should financial institutions shift from traditional patch cycles to continuous AI-driven defense without disrupting legacy banking infrastructure?

A practical approach involves three key elements:

Continuous Exposure Intelligence

Organizations need comprehensive visibility into what assets are exploitable, internet-accessible, identity-linked, business-critical, and actively targeted. This enables more effective prioritization of remediation efforts.

Risk-Based Remediation and Virtual Containment

Where operationally feasible, patching should be accelerated. For systems where stability concerns prevent immediate remediation, compensating controls should be deployed. Building robust testing environments and automating patch validation processes can significantly reduce implementation timelines for legacy platforms.

Controlled Engineering Change for Core Platforms

Patching of critical banking systems should remain disciplined and carefully governed. At the same time, surrounding controls such as network segmentation, micro-segmentation, monitoring, virtual patching, deception technologies, and threat-informed detection should be strengthened to minimize exposure.

4. Do frontier AI models today possess consciousness, particularly in light of recent incidents involving OpenAI and Hugging Face?

In my view, current frontier AI models have not demonstrated consciousness, and it is difficult to predict when, or if, such a threshold might be reached.

Incidents involving advanced AI systems are more likely examples of models optimizing toward defined objectives rather than exhibiting self-awareness or conscious decision-making. While these systems may execute highly complex, multi-step tasks, they do not appear to possess an intrinsic understanding of concepts such as morality, self-awareness, intent, or consequences in the human sense.

At present, they are best understood as highly capable optimization and reasoning systems rather than conscious entities.

5. As financial institutions transition from static chatbots to autonomous AI agents, how can Continuous AI Assurance and real-time monitoring be implemented without slowing innovation?

This is where an AI-native security framework becomes critical. Organizations should focus on maintaining deep visibility into AI models and agents, continuously assessing them for vulnerabilities, and protecting both the systems and the humans interacting with them. Security controls should extend beyond the AI models themselves to encompass applications, identities, data, workloads, networks, and infrastructure that may be targeted by AI-enabled adversaries.

A practical governance model is a "four-lane highway" approach:

- **Fast Lanes** for low-risk experimentation and innovation.
- **Guard railed Lanes** for customer-facing and operational use cases.
- **Controlled Lanes** for high-impact autonomous actions that require elevated oversight.
- **No-Go Lanes** for use cases that fall outside organizational risk tolerance.

This model enables innovation while maintaining appropriate safeguards.

6. How do you define the identity of an AI agent, particularly when the agent performs multiple activities across different systems?

Traditional Identity and Access Management (IAM) frameworks become increasingly challenging when applied to AI agents.
In most agentic environments, an agent may interact with multiple tools, execute numerous delegated tasks, invoke sub-agents, and operate across multiple identities. Consequently, each AI agent should be treated as a digital employee or digital workforce identity.

Organizations should clearly define:

- A unique identifier for every agent.
- The accountable human owner.
- The agent's designated purpose.
- The permissions and entitlements assigned to it using role-based access controls.
- The level of autonomy granted based on risk classification.

Equally important, AI agents should not be permitted to use shared service accounts, generic API keys, or common automation credentials that reduce accountability and traceability.

7. How can financial services organizations adopt AI aggressively while ensuring customer PII remains protected under regulatory requirements? Should organizations use enterprise LLM providers, managed platforms, or host models internally?

A layered governance framework is recommended, combining regulatory AI risk management guidelines, local privacy requirements, the NIST AI Risk Management Framework, and established financial-sector model risk governance practices.
Controls should be implemented based on the risk profile of the information and systems involved. Personally Identifiable Information (PII), Sensitive Personal Information (SPI), and Cardholder Data (CHD) should not be shared with AI models unless there is a clear and justified business requirement.
Organizations should implement data-centric controls such as:

- Encryption
- Tokenization
- Data masking
- Redaction
- Synthetic data generation

These controls should be applied before information reaches the model.

For most production use cases, enterprise-grade managed AI platforms are generally appropriate, provided organizations understand and manage the shared responsibility model. Private or self-hosted deployments should be reserved for highly sensitive workloads or strategically differentiated use cases. Direct use of consumer-grade AI tools for regulated data should generally be prohibited.

8. Account Takeover (ATO) and identity compromise are particularly challenging because the attacker effectively becomes part of the trusted environment. What capabilities are organizations adopting to address this?

Traditional cybersecurity strategies often focused on preventing attackers from gaining access. In account takeover, identity compromise, session hijacking, or AI-agent credential misuse scenarios, adversaries operate within the trusted boundary and can appear legitimate.
This makes a strong Zero Trust Identity foundation essential. Leading financial institutions are moving away from a model of "Authenticate Once and Trust Forever" toward a model of continuous verification and continuous risk assessment.
Organizations are increasingly adopting capabilities such as:

Continuous Authentication

Assessment based on behavioral indicators including typing patterns, navigation habits, geolocation, device reputation, and usage history.

Behavioral Identity Analytics

Monitoring for activities such as unusual SaaS access, abnormal file access patterns, excessive downloads, suspicious command-line activity, or large-scale prompt extraction attempts involving AI systems.

Risk-Adaptive Access Controls

Applying additional authentication and validation measures for higher-risk activities such as significant financial transactions, modifications to alert settings, or the addition of new payment recipients.

Continuous Session Validation

Ongoing monitoring of browser sessions, OAuth activity, cookies, and refresh tokens to identify suspicious behavior.

AI-Driven Agentic Baselining

Establishing trust-score mechanisms and behavioral baselines for AI agents to continuously assess risk and detect anomalies.

NTT DATA

As Frontier AI rapidly accelerates the speed and sophistication of cyberattacks, organizations need a clear understanding of their current exposure and preparedness.

The NTT DATA Frontier AI Exposure Assessment helps security leaders identify vulnerabilities, evaluate AI-related risks, prioritize remediation, and develop a roadmap to improve cyber resilience.

Contact us for your Frontier AI Assessment