AI Meets Cyber: Hype, Reality, and What Actually Matters

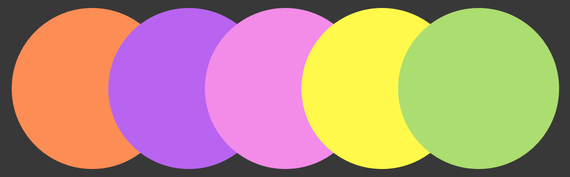
April 2 2026

Speaker:

Visesh Gosrani, FIA

Betty Zhu, FCAS, FCIA

Queen Mary University of London



Agenda

1. Introduction
2. Cyber Risk
3. Artificial Intelligence
    a. Overview
    b. Impact on Cyber Insurance
4. Q&A

Introduction

Betty Zhu, FCAS
Co-Founder · Solux Tech

Visesh Gosrani
Director · Cydelta

Cyber Risk

Man made risk

Types of Cyber Risk

Single entity

Data breach

- Historically single-entity
- Potential for systemic through
  - Managed Service providers
  - Vendors

Single entity

Ransomware

- Typically single entity
  - Often mass campaigns
  - AI as a force multiplier can enhance mass campaigns
- Viral elements can result in multi-entity

Systemic

Denial of Service

- DoS attacks have not increased to the degree others have
- Propensity to return
- Propensity for big bang
- Increased potential for collateral damage

Systemic

Cloud outage

- Exposure increasing significantly with increasing adoption
- Varies by provider, service location
- Increased potential for collateral damage
- Operational risks

Cyber evolves quickly

Exposure

Fast growth from a low base.

- Increasing numbers of policies sold
  - $4bn (2017)
  - $8bn (2020)
  - $15bn (2025)
  - $30bn (2030)
- Increasing penetration
  - Geographies
  - Industries
  - Smaller and Medium Co's

Likelihood

Never an if, always a when

- Trade off of attack strength and vectors versus defensive mechanisms
  - Offensive linked to firmographic profile
  - Defensive linked to technographic profile

Severity

When it occurs, how do we reduce the impact

- Greater consideration of the disruption from Cyber events
- Greater resilience measures including fire drills
- Governmental input to Cyber security

Differences to other lines

External models

Significant additional discussion

- Speeding up understanding by generating debate
- Industry forums comparing results

Regulator focus

Forced exercises and benchmarking

- Similar benefits to external models
- Trying to not hamper market

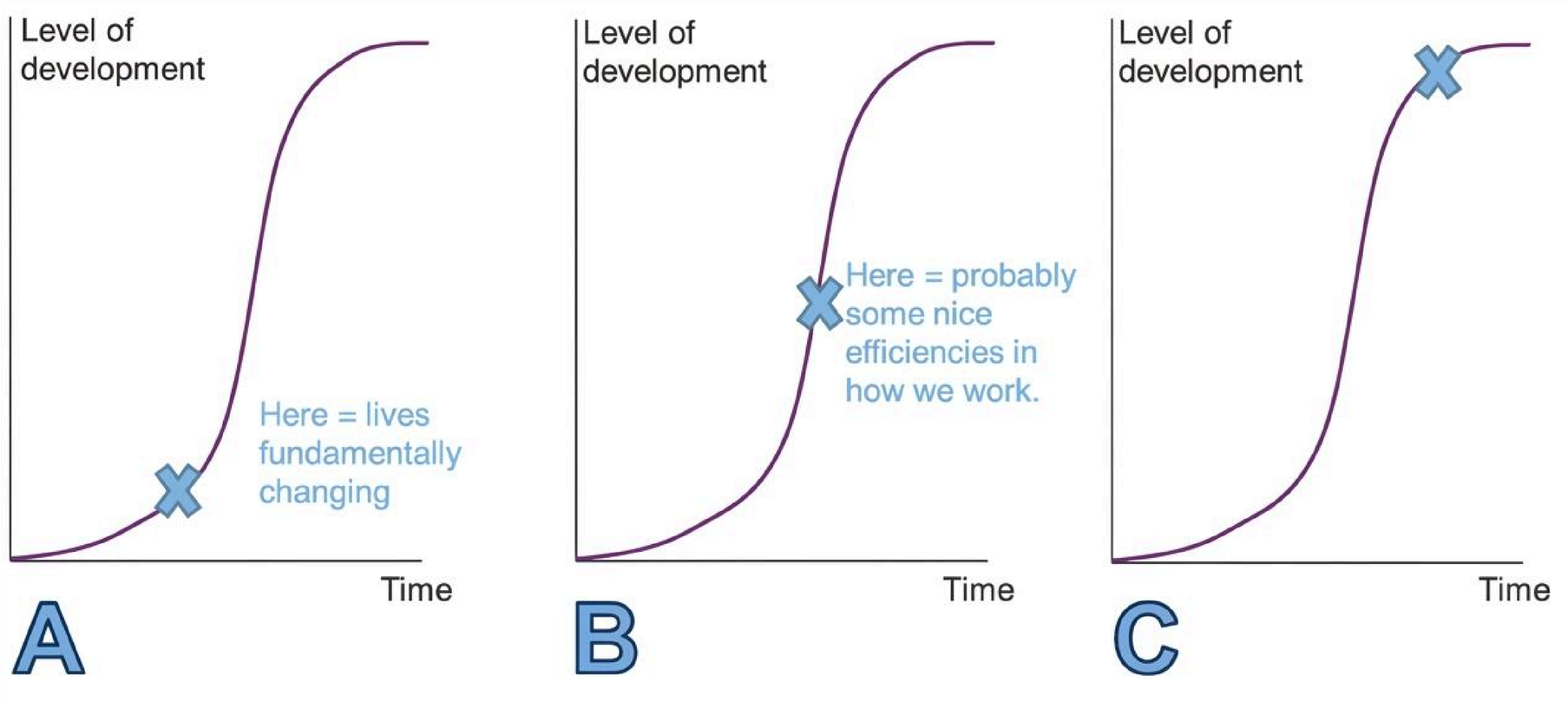
Analysis discrepancies

Significantly different pricing and reserving/ capital outcomes

- A real potential for markets to outperform
- However, tide remains high for an extended period of time

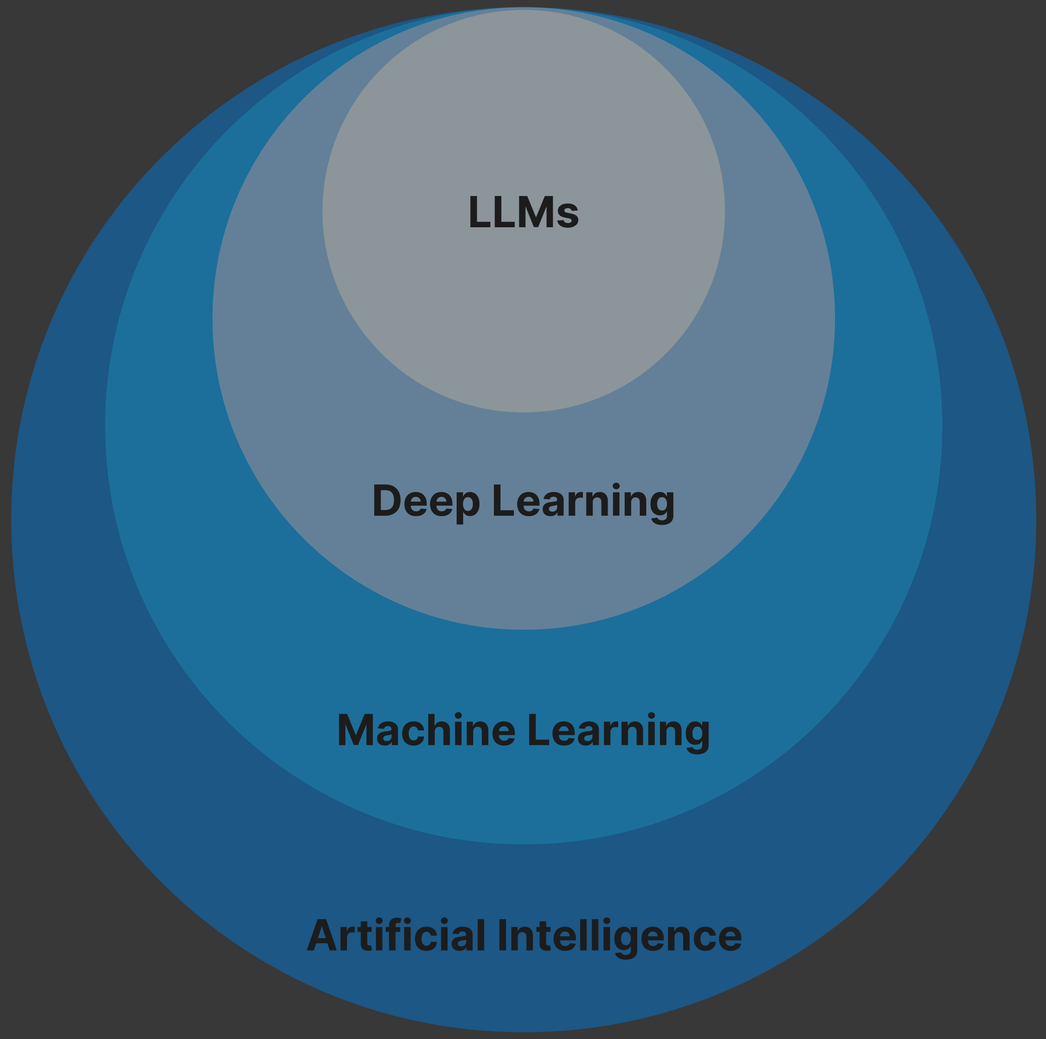
Artificial Intelligence

Where are we on the AI development journey?



What do the words “Artificial Intelligence” mean?

The AI Landscape



Two fundamental approaches:

Traditional AI:

- Categorizes data by identifying patterns and learning the boundaries between different classes
- Handle typical classification and regression tasks
- Example popular model forms: KNNs, GBMs, Neural Networks (e.g., CNNs)

Generative AI (GenAI):

- Learn to generate new contents based on training data by capturing the underlying distribution of training data
- Often leveraging Deep Learning models (Neural Networks) as the foundation

Some Risks of AI… especially GenAI

01
Current LLMs will always hallucinate

- This is a known limitation of probabilistic language models
  - Recent paper from OpenAI researcher: https://openai.com/index/why-language-models-hallucinate/
  - Output is stochastic and probabilistic, not deterministic - They generate what’s likely, not what’s true

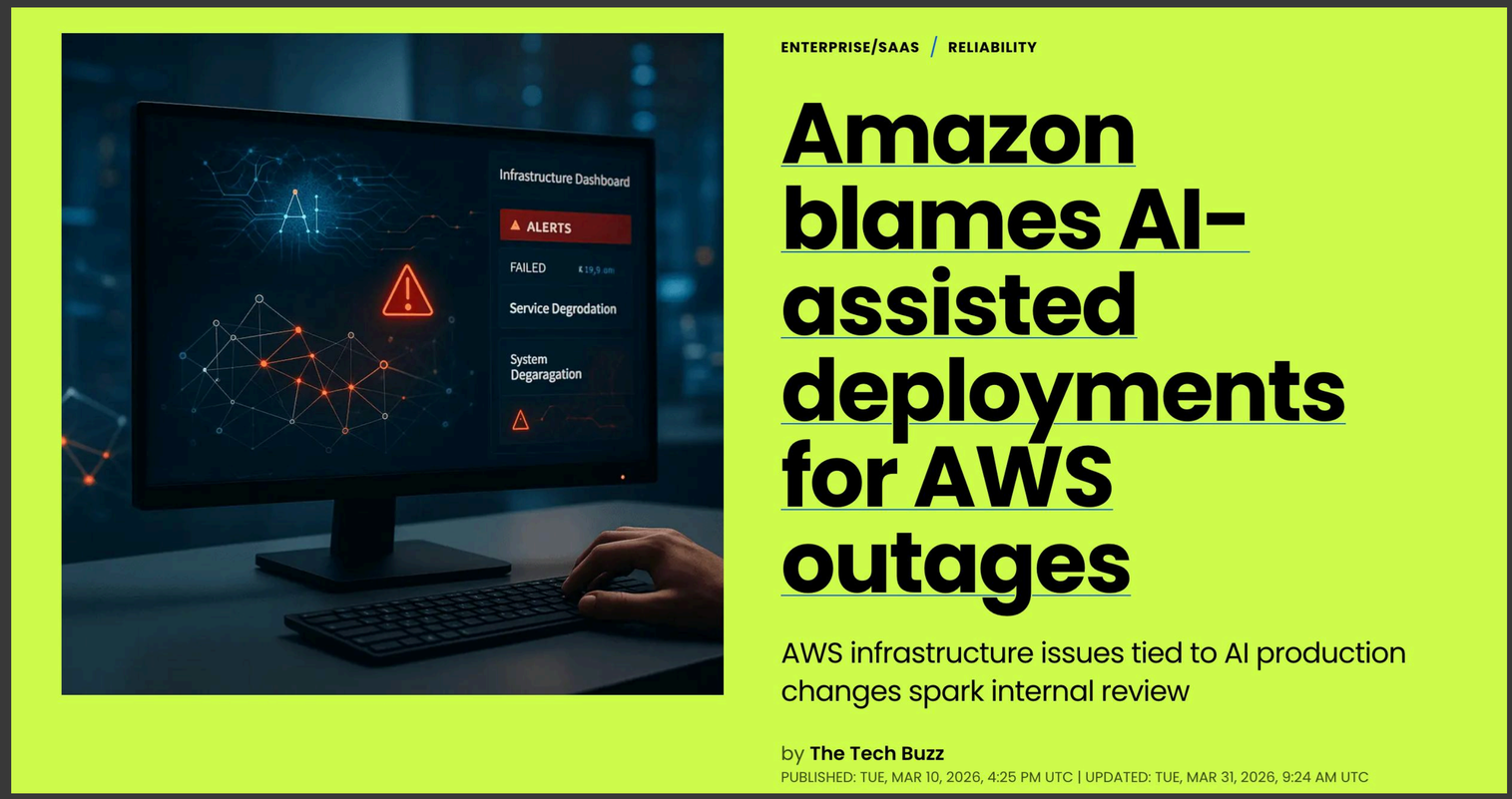
- This introduces real-world risk:
  - Wrong decisions
  - Misleading outputs
  - Overconfidence in AI

02
Models can be poisoned

- “A small number of samples can poison LLMs of any size” - research from Anthropic
  - https://www.anthropic.com/research/small-samples-poison
  - Models can learn malicious patterns without you knowing

- Model Poisoning = Silent Risk
  - Small changes → disproportionate impact
  - Outputs look normal — but are biased/manipulated
  - Hard to detect, harder to fix

03
AI enables bad outputs to scale instantly



- AI makes risk scalable very fast
  - Humans trust fluent outputs
  - Content is reused and amplified
  - Errors propagate across systems
  - Small mistakes → system-wide impact

04
AI systems create new security vulnerabilities and data leakage risks

McKinsey realises the risk of rapid adoption of AI after hackers gain access to 46.5 million employee chat messages, 728000 ‘sensitive files’ and ...

TOI Tech Desk / TIMESOFINDIA.COM / Mar 13, 2026, 18:08 IST

Comments Share AA

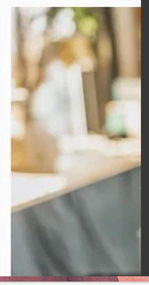
McKinsey & Company rushed to patch a serious security flaw in its internal AI platform after a cybersecurity researcher gained access to tens of millions of employee chat messages and hundreds of thousands of sensitive files – all within two hours. According to a report by The Financial Times (via CodeWall), the target was Lilli, the management consultancy’s in-house AI platform used daily by its 40,000 employees to plan strategy, analyse data, and build project plans and client presentations.

- Input risk → sensitive data and malicious instructions enter the model
- Output risk → unintended disclosure
- System risk → insecure integrations (RAG, APIs, tools)
- Control failure → access and governance gaps

More Horror News…

AI • CODING

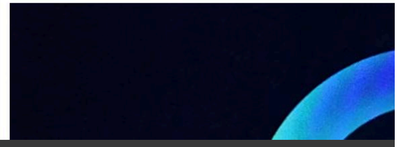
An AI agent destroyed this coder's entire database. He's not the only one with a horror story



By **Beatrice Nolan**
Tech Reporter
March 18, 2026, 11:36 AM ET

Meta AI agent's instruction causes large sensitive data leak to employees

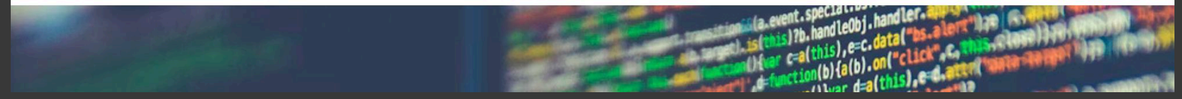
Artificial intelligence agent instructed engineer to take actions that exposed user and company data internally



HOMEPAGE › INNOVATION › ARTIFICIAL INTELLIGENCE

ARTIFICIAL INTELLIGENCE / CYBERCRIME / INNOVATION / SECURITY

Hacker Steals Huge Data Trove From Mexico Using Anthropic's Claude



How AI impacts Cyber Insurance

Impact of AI on Cyber Insurance models

01

Offence

- Force multiplier
- Smoother operator
- Investigation tool
- Code generation tool

02

Defence

- Network monitoring
- Connecting dots at scale

03

Policies

- Strengthening of wording
- Cybersecurity requirements including provision and monitoring of tools
- Underwriting standards
- Partnership model replacing insurer and insured

New coverages for AI-related events and AI Agents

Regulatory intervention

Takeaway - Equipping you for the path ahead

- Cyber is growing and developing quickly.
- Only hindsight will tell us where we currently are on this AI journey.
- Current LLMs will always hallucinate.
  - Know the limits before you trust the outputs.
- Stay grounded. Get educated. Keep learning.
  - This space is evolving fast - judgment matters more than tools
- Don’t panic.
  - A huge growth area when considering careers
  - You don’t need to become an AI expert - just a smart user
  - It will not take over the world!