INSIGHTS

By **Pedro Carreira,**
Principal at 33N

Gonçalo Borges,
Principal at 33N

And **Margarida Correia,**
Senior Associate at 33N

Generative AI: A (Secure) Opportunity Waiting to be Unlocked

At the tipping point

In the wake of a transformative year in 2023, where the promise of GenAI captured the world's imagination, 2024 dawned as the year to bring these lofty ideals into tangible reality. However, three months into this pivotal year, critical voices are now raising concerns about the balance between the benefits and the risks associated with adopting GenAI use cases.

What are the real impediments in the path of widespread adoption? Beyond the challenges of prioritising use cases, hiring/retaining the right technical talent, designing & building the tech stack, among others, there are still substantial technological hurdles yet to overcome (and enough space to build!).

Nascent Adoption of Large Language Models (LLMs)

Enterprises, while enthusiastic about the potential of LLMs, remain in the experimental phase, hesitating to transition from exploration to production. GSI Accenture announced bookings of a whopping $600 million in AI revenue, underscoring the significant investments being made in this arena.

According to Gartner, the initial response from enterprises was one of caution, actively blocking usage of public LLM's by its staff to avoid issues such as data leakage or IP protection, focusing on damage control measures(Figure 1). Notably, emerging security vendors like Prompt Security, among others, stepped up to address these concerns by offering 'Shadow AI detection' tools to monitor the usage of GenAI as well as LLM firewalls to protect the organisation from data leaks.

Medium to long-term, however, enterprises have started to embrace LLMs, initially for internal use cases which provide higher risk control, while building confidence for future external ones. At present, as a recent report from a16z shows, the primary use cases for LLMs are predominantly internal-facing, reflecting a cautious approach to deployment (Figure 2).

Persistent Challenges in LLM Security

Security remains a paramount concern in the adoption of LLMs, with issues such as bias, hallucinations, and vulnerabilities yet to be fully resolved. The Open Web Application Security Project (OWASP) has identified the top 10 threats posed by LLMs, with proactive measures being advocated to mitigate risks. Caleb Sima, Chair of the CSA AI Safety Initiative, provides detailed insights into the top-3 threats for LLMs today and most mature solution types the industry has so far found:

- Prompt Injection: Manipulating the-prompts provided to the LLM to induceunintended behaviour. Most promisingsolutions so far include prompt vulnerability scanners as well as LLM firewalls, dual LLM approaches or ChatML model.
- Data Poisoning: Attackers injecting malicious data into the training datasets, leading to skewed model outputs. Most considered solutions are still evolving but mainly rely on pre-GenAI tools for data verification, outlier detection and trusted domain enforcement.
- Data Leakage: Unauthorised access to model outputs or training data poses a significant risk. Access control mechanisms (built into the application's architectural design) as well as LLM firewalls are essential for mitigating this threat.

Several players have emerged to cover these issues (Figure 3) but the landscape is yet made up of very nascent players, while 'traditional ML' security players (eg. Robust Intelligence, Protect AI, HiddenLayer or CalypsoAI with different scopes across the ML dev pipeline) quickly evolved their offering or acquired smaller vendors to cover LLM security.

Navigating LLM Compliance Amid Regulatory Shifts

The regulatory landscape surrounding LLMs is evolving rapidly, with recent legislative developments demanding compliance from stakeholders. The European Union's AI Act, one of the most comprehensive regulatory frameworks, is set to have a profound impact on the adoption and deployment of AI technologies (ML and LLM broadly). Most importantly the act sets key references in terms of risks for AI systems and different timelines for application of the legislation for each risk level:

- Timeline Overview: The act outlines a phased timeline for implementation, but generally becomes fully applicable 24 months after coming into force
- Risk Overview: While use case dependent, certain industries such as insurance and financial services, are poised to be among the first to feel the impact of these regulatory changes (e.g., credit scoring used for loans were deemed as high risk) - industry leaders like Allianz recently highlighted the need for proactive measures to ensure compliance with the EU AI Act.

As stakeholders grapple with the complexities of compliance, collaboration and knowledge sharing will be essential in navigating the regulatory landscape effectively.

Unlocking the Potential

Despite the formidable challenges on the horizon, the journey towards widespread LLM adoption presents a unique opportunity to unlock unprecedented productivity gains for enterprises, economies and our society as a whole.

At 33N, we remain optimistic in the cyber ecosystem's ability to address the security and compliance hurdles on the horizon. As specialized investor, 33N continues to be deeply grounded grounded on the pieces of the technology stack needed to deliver the GenAI promise, supporting entrepreneurs in the space via:

- Domain expertise, namely on identifying key pain-points in cybersecurity, evolving threat landscape, innovative tech approaches and bottlenecks to scaling
- Direct access to market, via a network of CISOs and tech executives with deep expertise in key domains across cyber, while at the same time regional network in key regions globally
- Indirect access to market, via a network of resellers and services partners across key regions globally

In innovative topics such as security and compliance for GenAI, corporates have also seen benefits in collaborating with innovative vendors and specialized investors. One of the largest Iberian banks, for example, took a strategic position in 33N's fund, contributing not only with a sizeable investment but also appointing a member to 33N's Strategic Committee. Such collaborations bring valuable insights into the needs of a major player in the financial industry to 33N, while reciprocally magnifying the bank's insight into the cybersecurity ecosystem and innovative vendors through 33N's network.

Pan-European service providers also play a significant role in the adoption of these innovative cybersecurity solutions. Christophe Bianco, Venture Partner at 33N and co-founder of Luxembourg-based Excellium services, reinforces his view: "Specialized service providers and investors are crucial partners in this endeavor, partnering with emerging vendors, in an increasingly challenging cybersecurity landscape."

As the industry navigates through these challenges, the potential for transformative advancements in GenAI adoption remains very promising. ●

Figure 1: Five Tactics to Quickly Adapt to Uncontrolled LLM App Consumption (source: Gartner)

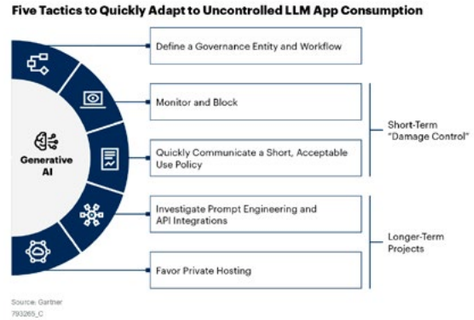

Figure 2: Enterprise LLM use cases (source: a16z).

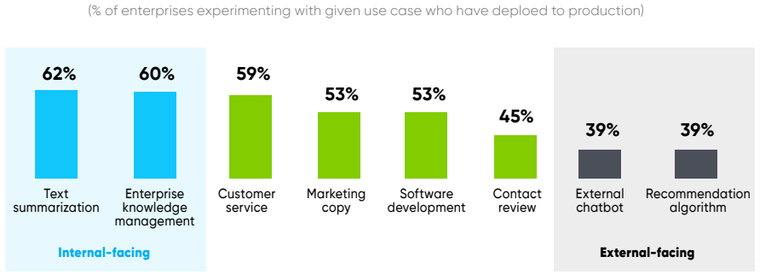

Figure 3: Security for AI market map in Feb-2024 (source: Menlo ventures)

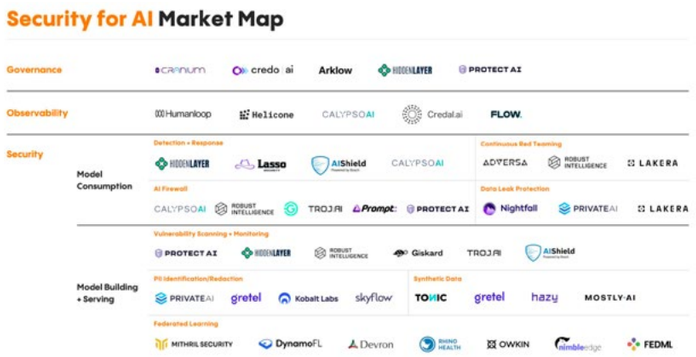

Figure 4: Regulatory Framework's 4 levels of risk for AI systems (source: EU AI Act)