文本

描述已自动生成**IN6299/IS6799/KM6399 CRITICAL INQUIRY**

**Research Proposal:**

**Use of Generative Artificial Intelligence (AI) by Professionals**

**Group ID LGP-05**

**Supervisor: Assoc Prwongof Pee Loo Geok**

Submitted by:

|  |  |  |
| --- | --- | --- |
| **Name** | **Matric No.** | **Email** |
| Cheng Hangfan | G2405675G | HANGFAN001@e.ntu.edu.sg |
| Liang Hou | G2401631D | HLIANG008@e.ntu.edu.sg |
| Liu Zihan | G2404273B | LIUZ0120@e.ntu.edu.sg |

**Date: 2025/02/04**

# Background to the research

Generative artificial intelligence (AI) has been one of the most trending technologies in recent years. It started with a significant breakthrough in machine learning, particularly the transformer architecture introduced in *Attention Is All You Need* by Vaswani et al. (2017). A few years later, in 2022, OpenAI’s ChatGPT series was released, attracting over one million users in just five days (Marr, 2024). Major technology companies have also developed AI models and integrated them into various professional domains, including marketing, software development, customer service, and education.

As AI adoption grows, so does its impact on workplace operations. According to McKinsey & Company’s 2023 report on AI adoption, one-third of surveyed organizations are already using generative AI tools regularly in at least one business function, accounting for 60% of organizations that have integrated AI into their workflows. However, few companies are prepared to address the associated risks, with only 21% of respondents having established policies governing employees’ use of generative AI tools at work (Chui et al., 2023).

A growing concern is the rise of shadow AI-the unapproved use of AI tools by employees without organizational oversight (Wiz Experts Team, 2025). A 2024 Cyberhaven Labs study analyzing AI usage patterns of 3 million employees found that the technology industry has the highest AI adoption rate. Many employees use AI without company authorization, exposing sensitive data such as legal documents, source code, employee records, and research materials (Coles, 2024). The rise of shadow AI poses significant risks to data security, corporate revenue, and reputation (Gupta, 2024). In response, organizations must implement policies to regulate AI usage, while individuals should take steps to protect sensitive information (PwC, 2023).

While previous research has explored shadow AI risks broadly, the specific factors contributing to knowledge leakage remain underexplored. Unapproved AI usage can lead to both intentional and unintentional exposure of confidential data, yet existing coping strategies remain incomplete. There is a need for further research into why employees resort to shadow AI, how it leads to knowledge leakage, and what measures organizations can take to mitigate this risk.

# Problem Statement

This study aims to discuss the factors affecting knowledge leakage in employees’ shadow use of generative AI. Unauthorized use of these tools, often due to a lack of corporate support or understanding, exposes organizations to risks such as data breaches, intellectual property exposure, and compliance issues. Despite its increasing importance, little empirical research has examined the elements that lead to knowledge leakage in shadow AI usage. Identifying these factors is crucial for developing effective strategies to mitigate risks and enhance organizational security.

# Objectives of the Research

1. To identify and analyse the factors influencing knowledge leakage due to employees’ shadow use of generative AI in professional settings.
2. To provide actionable insights for organizations to mitigate risks, promote ethical AI usage, and establish effective policies for using generative AI.

# Literature Review

**Definition**

The first section of this literature review looks at the ideas of *knowledge leaking* and the *shadow use of generative AI*. According to Durst (2014), there are two different aspects of knowledge leakage: exposure and lack. The most common cause of knowledge lack is staff turnover, which includes situations in which workers retire, move to another company, or depart for other reasons. These departures, regardless of the particular circumstances, frequently lead to the loss of relational capital and tacit knowledge, leaving the company without enough experienced or talented workers to cover the hole.

On the other hand, knowledge exposure—the main goal of this project—refers to the purposeful or inadvertent sharing of an organization's confidential information with outside parties. According to Jiang et al., this is "the degree to which partners purposefully appropriate or inadvertently transfer the focal firm's private knowledge." In this study, the term "knowledge leakage" specifically refers to knowledge exposure as defined above.

Furthermore, the term "shadow AI " (IBM, n.d.) as used in this study, describes the unapproved or unofficial application of generative AI inside a company that does not fall under the jurisdiction of IT control. This phenomenon draws attention to the possible dangers of uncontrolled workplace technology adoption.

**Factors Identification**

The shadow use of generative AI tools in workplaces presents significant risks, particularly in terms of knowledge leakage. To explore the underlying factors that contribute to this phenomenon, we review existing research from three perspectives: the technical characteristics and vulnerabilities of generative AI tools, the organizational dynamics related to shadow IT, and the behavioral and organizational dimensions of knowledge leakage. Each perspective highlights unique drivers and challenges, providing a comprehensive understanding of the risks associated with shadow AI usage.

***Generative AI in Workplaces***

In terms of Generative AI in Workplaces, there are the following factors that may have an impact on the knowledge leakage of employees’ shadow use of generative AI.

* Unauthorized access to generative AI tools: Many generative AI tools are open platforms that users can access without authentication, potentially leading to unauthorized use or data leakage (SentinelOne, 2024). This can be assessed through questionnaires asking employees if and how often they use generative AI tools without explicit authorization from the organization.
* Frequency of Inputting Sensitive Information: Generative AI may not recognize sensitive information, and employees may enter private or sensitive company data when using it (Gupta, Akiri, Aryal, Parker, & Praharaj, 2023). This can be assessed through questionnaire surveys (e.g., ask employees how many times they input company data into the generative AI per week/month.)
* Data Storage Policies in AI: Some generative AI may store user input for model training or optimization, and employees who do not understand this may lead to the disclosure of sensitive information (Zhang et al., 2023). This can be assessed through questionnaire surveys (e.g., Whether employees understand the data storage policies of AI tools.)
* Data Protection Mechanisms in AI: Generative AI tools may lack explicit data protection mechanisms, especially in the areas of user authorization and data encryption (Goodman III & Garman, 2024). This factor can be measured as whether there are security mechanisms such as encryption, access control, etc. of the AI tools used in a company.

***Shadow IT and Knowledge Leakage***

To better understand the drivers of knowledge leakage in the shadow use of generative AI, this section identifies possible influencing factors after reviewing of prior research on shadow IT and knowledge exposure.

Kopper and Westner (2016) summarize several key factors that contribute to the cause of shadow IT in their report. These factors, which are highly relevant to the unapproved use of generative AI tools in professional settings, include:

* Shortcomings of IT Systems: Employees could turn to shadow AI, when they think their organization’s IT systems is not enough to finish their tasks.
* Rigidity of IT Systems: IT systems that are overly standardized and inflexible can drive users to adopt shadow AI as an alternative solution that is easier to adapt to their specific tasks.
* High IT User Competence: Employees with high levels of technical proficiency could be more likely to independently explore and utilize generative AI tools without the involvement or approval of the IT department, thus increasing the risk of knowledge leakage.
* Low Technical Complexity: Overall generative AI platforms, such as ChatGPT, have lower barriers to adoption, leading to their unapproved use for professionals.
* Lack of Restrictions: Lacking IT usage policies or restrictions encourages employees to use unapproved generative AI tools, heightening the chances of confidential information being exposed.
* Lack of Awareness of IT policies and security risks: Employees may have limited understanding of organizational IT policies and security risks associated with generative AI, including potential knowledge leakage.

***Behavioral and Organizational Dimensions of Knowledge Leakage***

Knowledge leakage, a critical organizational risk, has been explored in numerous studies, often emphasizing the behavioral and organizational factors that contribute to its occurrence.

Wong (2019) highlights that mitigating information leakage—a component of knowledge leakage—requires human governance mechanisms, such as fostering an ethical organizational climate and cultivating a robust information security culture. Factors contributing to knowledge leakage can be categorized into:

* Behavioral Dimensions: Intentional leakage may arise from jealousy, dissatisfaction with the company, vindictiveness, or personal obsession.
* Cognitive Dimensions: Unintentional leakage often stems from human errors, negligence, over-enthusiasm for innovation, and unclear role responsibilities for new employees.

Traditional information leakage studies inform our understanding of shadow AI-related knowledge leakage. Over-enthusiasm for innovation, amplified by generative AI capabilities, and unclear responsibilities among employees are particularly relevant. Additionally, shadow AI usage introduces unique drivers, such as efficiency gains offered by AI tools and the pressure of tight deadlines or urgent tasks. For instance, Stojan (2024) illustrates how a developer might use ChatGPT to debug code quickly, inadvertently exposing proprietary source code. This code could then be stored on external servers, potentially becoming part of future model training datasets or vulnerable to external security breaches.

Organizational factors such as cultural climate, emphasis on efficiency, production pace (especially in fast-moving sectors like IT), and adherence to security protocols play pivotal roles in understanding and addressing knowledge leakage risks associated with shadow AI. By examining these dimensions, organizations can develop actionable and effective strategies to manage this emerging challenge and minimize its potential impact.

***Conclusion***

The review of existing literature highlights multiple factors contributing to the knowledge leakage risks associated with the shadow use of generative AI. In order to facilitate the following questionnaire design and further research, the above variables were reclassified into three types of factors: Technical Factors focus on the limitations of the technology, such as access control and system shortcomings; Organizational Factors focus on the organizational problems in setting up adequate policies and educating employees about security and AI tool usage; Behavioral factors focus on how employees’ actions contribute to knowledge leakage(either intentional or accidental), such as jealousy or lack of awareness. The following are three types of factors:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Types of factors | Factor name | Constituent | | Reference | Reason for modification | Description |
| Original type | Original  factor |
| Technical | Unauthorized Access to Generative AI Tools | Generative AI in Workplaces | Unauthorized Access to Generative AI Tools | SentinelOne (2024) | This is a defect of the generative AI tool itself, which is a technical problem. | Open generative AI platforms allow users to access without authentication, leaving information exposed to dangerous areas. |
| Shortcomings and Rigidity of IT Systems | Shadow IT and Knowledge Leakage | Shortcomings of IT Systems. | Kopper and Westner (2016) | Both of them are the defects of IT system in practical application. | Organizational IT limitations and inflexibility can drive employees toward unapproved generative AI tools, highlighting the importance of system adaptability in mitigating shadow IT behavior. |
| Rigidity of IT Systems. |
| Low Technical Complexity | Shadow IT and Knowledge Leakage | Low Technical Complexity | This is a defect of the generative AI tool itself, which is a technical problem. | The low threshold for acquisition or operation of generative AI amplifies the risk of knowledge leakage |
| Organizational | Lack of Restrictions and Policies | Shadow IT and Knowledge Leakage | Lack of Restrictions | The lack of relevant policies or constraints is an organizational problem. | Organizations that do not have policies or have insufficient restrictions on employees using generative AI without approval will increase the likelihood that employees use AI privately, leading to potential knowledge leakage. |
| Behavioral | Behavioral Dimensions | Behavioral and Organizational Dimensions of Knowledge Leakage | Behavioral Dimensions | Wong (2019) | These are human elements that measure how employees' actions contribute to knowledge leakage | e.g., intentional leakage: jealousy, dissatisfaction, etc. |
| Cognitive Dimensions | Cognitive Dimensions | These are human elements that measure how employees' perceptions contribute to knowledge leakage | e.g., unintentional leakage: human errors, negligence, etc. |
| High IT User Competence | Shadow IT and Knowledge Leakage | High IT User Competence | Kopper and Westner (2016) | This is a problem caused by employee behavior | Employees with strong technical skills may bypass the IT department's involvement or approval and independently explore generative AI tools, which could heighten the risk of knowledge leakage. |
| Lack of Awareness of Security Risks and IT Policies | Shadow IT and Knowledge Leakage | Lack of Awareness of IT policies and security risks | Kopper and Westner (2016) | Employees' cognitive deficits in IT-related aspects are human elements | The limited understanding among employees regarding the risks associated with generative AI usage and neglecting organization’s IT/AI policies are the key driver of unintentional knowledge leakage. |
| Lack of Awareness of Data Storage Policies and Protection Mechanisms in AI | Generative AI in Workplaces | Data Storage Policies in AI | Zhang et al.(2023) | To facilitate the measurement of variables and further investigation and analysis. | Users' limited understanding of generative AI in terms of data storage and protection, such as: Not knowing that the generative AI being used will test the model with user-input data, can also inadvertently greatly increase the risk of leakage. |
| Data Protection Mechanisms in AI | Goodman III & Garman(2024) |

**Table 1** *Three types of factors*

Most of the three types of variables summarized now are about Shadow IT, which is still necessary to discuss as the basis for exploring Shadow AI. Follow-up research will investigate which factors are more important and explore these dimensions in depth, aiming to identify effective organizational policies, technical safeguards, and employee training programs to address the challenges posed using shadow-generated AI.

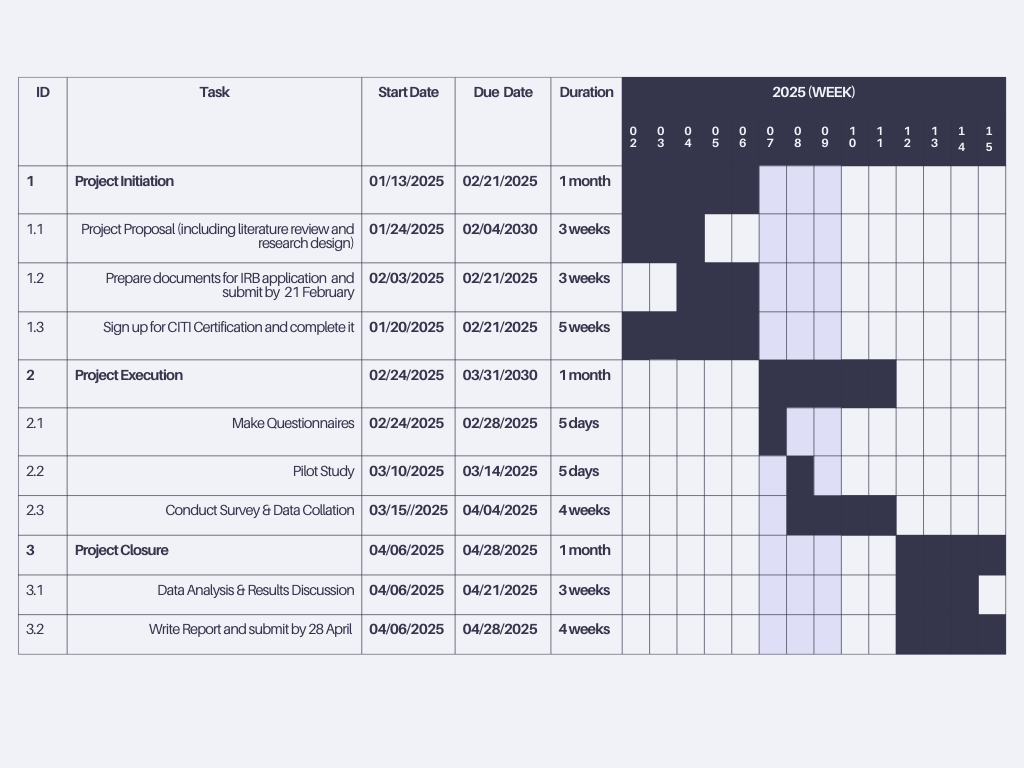
# Methodology

The methodology includes:

1. Recruit 150 individuals who have used unapproved generative AI for work tasks.
2. Data collection with structured survey.
3. Conduct in-depth interviews with a subset of 5 participants to gain deeper insights into their motivations, decision-making processes, and challenges associated with the unapproved use of generative AI tools. These interviews will complement the survey data by providing qualitative context and identifying nuances that may not emerge in survey responses.
4. Pilot testing with 5 respondents to refine the survey instrument and ensure clarity, reliability, and validity of the questions.
5. Analyse the survey results using Partial Least Squares (PLS) to assess relationships between factors, while qualitative data from interviews will be thematically analysed to identify recurring themes and support the interpretation of quantitative findings.

# Research Schedule

**Figure 1** *Gantt Chart of Our Group’s Schedule*



# Dissemination of Results

1. Feedback to participants at the end of the study
2. Concise report to highlight the main findings, recommendations, and actionable steps.

# References

Chui, M., Hall, B., Singla, A., Sukharevsky, A., & Yee, L. (2023, August 1). *The state of AI in 2023: Generative AI’s breakout year*. McKinsey & Company.

Coles, C. (2024, May 21). Shadow AI: How employees are leading the charge in AI adoption and putting company data at risk. Cyberhaven. <https://www.cyberhaven.com/blog/shadow-ai-how-employees-are-leading-the-charge-in-ai-adoption-and-putting-company-data-at-risk>

Durst, S., & Aisenberg Ferenhof, H. (2014). Knowledge leakages and ways to reduce them in small and medium-sized enterprises (SMEs). Information, 5(3), 440-450. January 25, 2025 at 10:45 AM

Goodman III, R., & Garman, J. (2024, November 15). Implement effective data authorization mechanisms to secure your data used in generative AI applications. AWS Security Blog. <https://aws.amazon.com/blogs/security/implement-effective-data-authorization-mechanisms-to-secure-your-data-used-in-generative-ai-applications/>

Gupta, A. (2024, February 6). The dangers of uncontrolled AI: Shadow AI and ethical risks. Securiti. https://securiti.ai/blog/dangers-of-uncontrolled-ai/

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of generative AI in cybersecurity and privacy. arXiv preprint, arXiv:2307.00691. <https://arxiv.org/abs/2307.00691>

Jiang, X., Li, M., Gao, S., Bao, Y., & Jiang, F. (2013). Managing knowledge leakage in strategic alliances: The effects of trust and formal contracts. Industrial Marketing Management, 42(6), 983-991. January 25, 2025 at 10:44 AM

Kopper, A., & Westner, M. (2016). *Deriving a framework for causes, consequences, and governance of shadow IT from literature*. Retrieved from <https://www.researchgate.net/publication/297698668_Deriving_a_Framework_for_Causes_Consequences_and_Governance_of_Shadow_IT_from_Literature>

Marr, B. (2024, February 20). A short history of ChatGPT: How we got to where we are today. *Forbes*. https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/

PwC. (2023, October). Decoding security risks: Challenges posed by generative AI. PwC China. <https://www.pwccn.com/zh/industries/telecommunications-media-and-technology/publications/decoding-security-risks-challenges-posed-generative-ai-oct2023.html>

SentinelOne. (2024, October 10). Generative AI security risks: Mitigation & best practices. SentinelOne. <https://www.sentinelone.com/cybersecurity-101/data-and-ai/generative-ai-security-risks/>

Stojan, J. (2024, August 16). Employees are leaking data over GenAI tools: Here's what enterprises need to do. VentureBeat. Retrieved January 25, 2025, from <https://venturebeat.com/business/employees-are-leaking-data-over-genai-tools-heres-what-enterprises-need-to-do>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is All you Need. *arXiv (Cornell University)*, *30*, 5998–6008. <https://arxiv.org/pdf/1706.03762v5>

Wiz Experts Team. (2025, January 22). *What is shadow AI? why it’s a threat and how to embrace and manage it*. wiz.io. [<https://www.wiz.io/academy/shadow-ai>](https://www.wiz.io/academy/shadow-ai)

Wong, W. P., Tan, H. C., Tan, K. H., & Tseng, M. L. (2019). Human factors in information leakage: mitigation strategies for information sharing integrity. Industrial Management & Data Systems, 119(6), 1242-1267.

Zhang, D., Xia, B., Liu, Y., Xu, X., Hoang, T., Xing, Z., Staples, M., Lu, Q., & Zhu, L. (2023). Navigating privacy and copyright challenges across the data lifecycle of generative AI. arXiv preprint, arXiv:2311.18252. <https://arxiv.org/abs/2311.18252>

IBM. (n.d.). What is Shadow AI? IBM - United States. Retrieved October 25, 2024, from https://www.ibm.com/think/topics/shadow-ai