Mediating Multiple Generative-AI Tools in Academic Work

Author: Liam van der Veen
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT,

As generative AI (GenAI) tools become increasingly embedded in academic workflows, academics are not only becoming users of individual systems but also learning to interact with and navigate multiple platforms that have different capabilities, limitations, and epistemic assumptions. This thesis explores how academic users mediate between multiple GenAI systems in their knowledge work. Using qualitative data from 12 semi-structured interviews with academics, the study applies the Gioia methodology to develop a grounded model of GenAI orchestration. Four aggregate dimensions were identified: frictions in interfacing with GenAI, tactics for coordinating tools, strategic mediation and transformational learning and identity. These were linked through three underlying mechanisms: epistemic calibration, delegation as coping and value-driven bricolage. These mechanisms reveal how users adapt to uncertainty, manage overload and align GenAI practices with personal and pedagogical values. The study presents a dynamic, non-linear model of GenAi integration that foregrounds user agency, reflective adaptation and contextual negotiation. These findings offer conceptual insight into how GenAI can reshape academic practice and identity, with implications for pedagogy and tool design in higher education.

Graduation Committee members:
Dr. J .U. Dahlke, University of Twente
Dr. V.C. Göttel, University of Twente

Keywords
Generative AI, Multi-tool orchestration, Human-AI interaction, Strategic mediation, GenAI in higher education, Epistemic calibration, Academic identity, AI trust and verification

During the preparation of this work, the author used ChatGPT to generate ideas and Grammarly to refine the bachelor's thesis by checking grammar and spelling. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the work.

1. INTRODUCTION

Generative-AI (GenAI) systems like ChatGPT and Claude are no longer only add-ons. Many academics use a combination of these tools daily, rather than relying on a single one. GenAI is transforming academic work across the spectrum, touching all levels from paper summarisation and ideation to code debugging and graphic design (Chan & Hu, 2023). These tools are also changing the way how users are thinking about and reflecting on their work (Abdel-Karim et al., 2023). As the use of multiple GenAI tools becomes more common, new challenges are arising. Users are dealing with inconsistencies across tools, as well as difficulties in effectively using them, especially when the outputs do not align. Even though many of these systems are built on similar foundation models, such as GPT-4 or BERT, they differ in their design and intended purpose. For instance, tools like ChatGPT are used for brainstorming, DALL-E for visual content creation, Claude for generating ethically framed responses and Copilot for code completion. Research shows that GenAI tools perform differently depending on the task or context (Hochmair et al., 2024; Mavrych et al., 2025). But what happens when these tools contradict one another, when outputs fail to align, or when tasks require more than one tool? Despite having the same technical foundation, academic practice is becoming increasingly diverse. Yet most existing research focuses on the use of individual GenAI tools (Bin-Nashwan et al., 2023; Gill et al., 2024; Vaithilingam et al., 2022) rather than on how users work with these tools in practice. As a result, there is limited understanding of how academic users navigate, coordinate, or resolve tensions between different GenAI systems.

Generative artificial intelligence is a class of machine learning systems capable of generating new content, such as text, images, audio, or code, mirroring the patterns of their training data. These systems are built on large-scale neural architectures trained with deep learning and self-supervised techniques. Foundation models mark a shift from task-specific AI to more general-purpose systems with emergent capabilities (Bommasani et al., 2021). While GenAI is transforming academic work by offering diverse functionalities, there is still a limited understanding of how users effectively utilise these multiple tools. Users of this expanding network of GenAI systems are pushed to be strategic. When to apply which tool, how to keep consistency, and how to integrate results properly. Although GenAI is being incorporated into higher education at a rapid pace, little is known about the daily tactics that academic users employ to navigate this complexity. To close this gap, the thesis uses interviews, rather than a single-tool benchmark, to examine how academics navigate a multi-tool GenAI ecosystem.

1.1 Research question

Considering the increasing use of multiple generative AI systems in academic environments, this research aims to explore how users manage the challenges and opportunities of working across these tools in their everyday knowledge work. This research will focus on answering the following research question:

"How do academic users mediate between multiple generative AI systems in their knowledge work?"

To systematically attain the needed information to address this question, the research is guided by the following three sub-questions:

1. What kinds of frictions arise when users interface with multiple GenAI tools?
2. What tactics and procedures do users develop to coordinate GenAI tools in their academic workflows?
3. How do academic users strategically mediate between different GenAI systems?

To answer these questions, the research will combine a literature review with qualitative interviews involving academic users. The interviews will serve to generate in-depth insight into the strategies, tensions, and adaptations of users when working with multiple GenAI tools.

1.2 Academic and practical relevance

The goal of this research is to increase the awareness of how academic users mediate between multiple generative AI systems in their knowledge work. It explores not only the frictions that occur in multi-tool use but also the strategies, procedures, and workarounds that users develop to properly coordinate and integrate these tools effectively. This helps to close the gap in current research, which primarily focuses on individual tool usage and also contributes to ongoing debates on the changing nature of academic practice in the age of GenAI.

The practical relevance of this research lies in the fact that the results can support students, educators, and researchers in utilising GenAI tools more effectively. As universities continue to integrate generative AI tools into academic practice (Jin et al., 2025), a clearer understanding of how users move between different systems can support smoother adoption and reduce confusion, overlap and inefficiencies caused by switching tools or dealing with inconsistent results

2. THEORETICAL BACKGROUND AND LITERATURE REVIEW

2.1 Generative AI in Academic Knowledge Work

As academics increasingly rely on multiple GenAI tools, new problems emerge. Users now need to manage conflicting outputs and decide how to use each system best when answers conflict with one another. Although many tools use the same standard base model, such as GPT-4 or BERT, they vary widely in design and purpose. ChatGPT may be used for brainstorming, DALL-E for visual content, Claude for ethically framed responses, and Copilot for code completion. Field studies confirm that researchers frequently switch among tools within a single project (Jin et al., 2025), while performance still depends heavily on task and context (Hochmair et al., 2024; Mavrych et al., 2025). But what happens when these tools give conflicting results, don't align, or when a task needs more than one tool? Recent studies indicate that GenAI tools are becoming an integral part of everyday workflows. For instance, writers often use ChatGPT for drafting, editing, and refining text (Sullivan et al., 2023), while Copilot helps students and researchers with writing and debugging code (Wermelinger, 2023). DALL·E is used for generating images and Claude is known for its ethical responses. Even though many of these tools are built on similar models, they function quite differently (Rudolph et al., 2023). They found that GPT-based models, such as ChatGPT, generally score highest overall, but no single tool performs like an "A-student." Instead, each tool excels in specific jobs and lags in others.

GenAI is accelerating in higher education. Most literature has focused on single-tool usage; however, studies have mainly examined the use of tools like ChatGPT in relation to academic integrity and writing quality (Bin-Nashwan et al., 2023; Lo, 2023). While these understandings are valuable, they do not accurately represent the trend towards a more multi-tool workflow, where users switch between tools based on the type of task or their preference. New and enduring trends suggest that users are taking more complex, planned steps to use GenAI. This means that users are choosing tools not just on capabilities, but also on tone, ethical framing, formatting style or speed of response (Gavira Durón & Jiménez-Preciado, 2025; Rudolph et al., 2023). This introduces new challenges. Users must decide which tool to use for each task, how to handle conflicting answers, and how to maintain consistency in their work. While large-scale studies have highlighted the prominence of terms like 'ChatGPT', 'generative AI', and 'teaching-learning' in academic discussions (Gavira Durón & Jiménez-Preciado, 2025), there is still a lack of qualitative research on how users actually navigate between multiple tools and the coordination of multiple tools in routine academic practice.

2.2 Frictions and Fragmentation in Multi-Tool Ecosystems

As GenAI becomes more deeply integrated into academic work, many users are no longer working with a single tool but rather a variety of tools. Users often switch between tools depending on the task at hand. These tools may be similar on the technological side, but they differ in their user interfaces, intended use and output (Gavira Durón & Jiménez-Preciado, 2025; Rudolph et al., 2023). Due to this variety, a landscape is created where tools do not always align, thus leading to fragmentation for users who must navigate between them. This fragmentation often slows users down. It can be confusing for users to switch between tools. Moving will also require extra effort to achieve the desired results. For instance, Wermelinger (2023) demonstrates that Copilot can complete a function in a certain way, but a subsequent prompt often presents an alternative, sometimes even conflicting, solution. Users have to stop and compare solutions to achieve valuable results. It is not just the content that can create friction, but it can also stem from inconsistent formatting, citations or differing views on how users should frame their work (Lim et al., 2023; Rudolph et al., 2023). It is not just a technological problem. GenAI technologies are also being introduced to academia without any common framework or approach. As Dwivedi et al. (2023) point out, tools like ChatGPT are changing how we write, teach, and publish. However, this change is largely uncoordinated. Academics will seek to develop their own approaches, as there are no common rules or standards to help them. Users must deal with technology that varies significantly, not only in interface and capabilities but also in value regarding academic work. The user experience can vary considerably even when the base models are identical. This can lead to conflicts when users try to combine the outputs from multiple tools into a single project.

A further problem emerges from the cognitive requirements of managing multiple AI tools. Users often need to adapt to the behaviour of the different tools and need to understand their characteristics. Zhang et al. (2024) describe this as the Mutual Theory of Mind. Here, effective collaboration depends on the response and anticipation of each other's actions. This can be mentally draining. Users who use different GenAI tools and switch between them find this cognitive load to be greater, especially when dealing with conflicting results or thinking processes. (Abdel-Karim et al., 2023; Li et al., 2023). One of the biggest unknowns in GenAI work is "hallucination". Here, the model provides an answer that sounds plausible but is incorrect or cannot be verified. In a large-scale benchmark study Li et al. (2023) found that 19.5% of the responses generated by ChatGPT contained hallucinated information. These errors can be hard to spot. Everything appears to be in order, but the facts are incorrect. Even the most advanced models continue to struggle with recognising errors in their own output. Crucially, hallucinations do not remain confined to one system, and this issue becomes more problematic when users combine outputs of multiple tools, each with their own bias and limitations. Errors can therefore cascade across a multi-tool workflow. This resulting workflow needs critical evaluations, but these may not be consistently supported in current educational settings (Jin et al., 2025). A key challenge in using multiple GenAI tools is the lack of transparency regarding how each one generates its responses. As Ali et al. (2023) point out, many of these tools function as "black boxes", giving little, if any, information about the logic behind their outputs. In a multi-tool environment, this lack of explainability makes it especially difficult for users to determine why one tool delivers an answer while another responds differently. The user is left to interpret these differences without helpful explanations and, more often than not, tries to develop informal methods of understanding or validating what they found. Such methods may be beneficial sometimes, but overall, they can be neither valid nor reliable.

These challenges highlight the fact that using several GenAI tools involves far more than just choosing the right model for the task. Users often have to make sense of conflicting outputs without any formal guidance. But how do academic users actually manage this complexity in practice? This needs to be explored further.

2.3 Strategic Mediation and User Decision-Making

Much of the existing literature has focused on what GenAI tools can or cannot do, but it has focused less on how users actively manage their interactions with these tools, especially when using multiple tools simultaneously (Abdel-Karim et al., 2023; Bin-Nashwan et al., 2023; Bommasani et al., 2021). In academic practice, users are not passive consumers. They make ongoing decisions about which tools to use, when to switch between them, and how to interpret or combine their outputs. These decisions are frequently driven by individual preference, task complexity, prior experience, and contextual anticipations.

The centaur-cyborg spectrum, as created by Dell'Acqua et al. (2023) it is a significant model that provides an understanding of mediation in a multi-tool GenAI context. In the centaur-cyborg model, users are not only utilising a single technology, but they are also a dynamic component who continuously change their connection with multiple tools as they go. On one end of the spectrum, we have the Centaurs. They maintain control and evaluate AI outputs. These outputs are viewed as recommendations that still require evaluation. On the other end of the spectrum are the Cyborgs. They let the system serve as an extension of their cognitive processes. They rely on the tool used to influence their output, and in some instances, even their thinking. This approach provides a way to illustrate the dynamic relationship among users as they move between cooperation and dependency while interacting with multiple GenAI tools, depending on the task, tool familiarity, and confidence in the output. Sousa and Cardoso (2025) show that students hop between ChatGPT for idea generation, Quillbot for paraphrasing, and GPTZero to check for originality. They argue that tool choice follows the Task-Technology Fit (TTF) logic. Here, a student weighs the usefulness and trustworthiness of the tool against the demands of an assignment before committing to it. In a multi-tool setting, this TTF calculus must be repeated for every hand-off because each new tool can improve or disrupt the evolving work product. Nguyen (2025) reaches a similar conclusion but adds a warning. When learners offload too much of the heavy thinking to AI, their ability to reason independently starts to erode. He notes that, lacking clear guidelines, many students have begun to craft their own ad hoc routines. Writing a first draft unassisted, for example, then letting a large language model polish style and citations.

Understanding how the various tools "think" also plays a role. Zhang et al. (2024) demonstrate this by using the Mutual Theory of Mind. This theory states that users come to expect how tools will respond and subsequently change their behaviour. When AI responses are uncertain or ambiguous, this interaction creates an undesired investment of effort. In a multi-tool workflow, user must repeat this effort for every tool. Abdel-Karim et al. (2023) note that these moments of uncertainty can promote deeper reflection because users are then able to reassess their own thinking in relation to what the AI suggests. Watson et al. (2025) take a broader view. They argue GenAI is changing not only how we finish a task but what we mean by "academic work" in the first place. In their model, researchers and AI systems continue co-adapting. This means that each turn of use slightly impacts the next. This means that strategic adoption is less about choosing a single, "best" app and more about being adaptive while tools and their expectations continue to change. Still, we know little about how users coordinate several GenAI tools in practice. As Bozkurt et al. (2021) point out, most AI-in-education studies stop short of the messy routines of daily academic work. Nguyen (2025) adds that students are largely left to fend for themselves, experimenting with GenAI without much structured guidance.

From the literature, it has become clear that GenAI use in academia is no longer a single-tool affair but a multi-tool ecosystem. In the empirical part of this research, these theoretical claims will be tested by examining the real-world multi-tool workflows among academics.

2.4 Human Judgment and AI Mediation in Professional Contexts

Recent literature on AI in knowledge-intensive environments has highlighted that, while technical capability is undoubtedly part of users' engagement with AI, so are the social, contextual, and interpretive dimensions. In an educational setting, where authorship, critical and disciplinary integrity are key elements, this complexity is evident. As GenAI tools begin to be adopted more commonly in academic contexts, the conceptual challenge of understanding how users apprehend, coordinate and critically think about these systems is of substantial importance. One relevant line of research examines how professionals assess the credibility of the outputs that they receive from AI. In their research, Lebovitz et al. (2021) argue that the concept of "ground truth" is a poor representation of what we know about how knowledge workers interact with AI-generated content. Instead of merely accepting system outputs as-is, users use domain knowledge, contextual judgement and tacit expertise to assess and modify those outputs. These practices complicate the assumptions of algorithmic objectivity and reveal the epistemic labour required for the usability of AI. A further focus of research is the coordination and sensemaking work needed to apply AI in organisational contexts. Waardenburg et al. (2022) use the term "knowledge brokers" to designate individuals who mediate between the formal logic of an AI system and the social expectations of their respective professional contexts. Such knowledge brokering involves translating, aligning and adapting AI outputs to make them actionable in practice. The work of knowledge brokers reveals that interacting with AI is not only a technical task but usually involves subtle negotiations, alignment and social interpretation, especially when multiple tools or systems are involved. Trust in AI systems is also a significant consideration for users when deciding when and how to engage with GenAI tools. Glikson and Woolley (2020) review empirical research on human trust in AI and emphasise that trust is not multifaceted or stable but dependent on task complexity, perceived reliability, emotional engagement and how often the users have engaged with the system in the past. In academic contexts where responsibility, intellectual rigour, and relational work are valued, these trust dynamics are not merely cognitive, but also affective and ethical.

Together, these perspectives offer a substantial yet incomplete understanding of GenAI interactions in academic practice. While the existing literature is starting to investigate how users assess, interpret and trust AI tools, most research focuses on either single-tool use or through formal implementation frameworks. There has been limited exploration of how individuals in educational contexts coordinate across multiple AI tools, respond to tensions or contradictions between AI outputs, and make sense of technologies in the context of their own pedagogical values and professional identities. Despite a growing awareness of these complexities, we still remain ungrounded in our understanding of how academics practically navigate and make sense of these tools in their daily work.

3. METHODOLOGY

3.1 Research design

The research design employed in this study was qualitative. Qualitative research has the ability to provide a detailed understanding of complex phenomena such as user behaviours, experiences, and practices (Creswell & Poth, 2016; Patton, 2014). This qualitative approach was especially well-suited for addressing the central research question: "*How do academic users mediate between multiple generative AI systems in their knowledge work?*". Creswell and Poth (2016) highlight how qualitative research is suitable for investigating lived experiences and processes, which makes it especially well-suited to examining the strategic interactions that academic users have when using multiple generative AI technologies. Patton (2014) reinforces this approach by highlighting the value of qualitative approaches in collecting detailed, context-sensitive understandings and providing flexibility to explore new topics as data is being collected. Practically, the findings are expected to support universities, educators, and policy-makers in designing targeted guidelines, training programs, and supportive frameworks for effective and responsible use of generative AI in higher education. Limitations of this approach include the relatively small and context-specific sample size. This can affect the generalisability of the results. Furthermore, the self-reported experiences of participants are a major component of qualitative approaches, and they may be prone to socially desired responses or recall bias. These restrictions will be explicitly recognised and taken into consideration when the results are analysed and interpreted.

3.2 Interviews

Semi-structured interviews were conducted to gain a deeper understanding of how academic users mediate between multiple generative AI tools. Semi-structured interviews are well-suited for this type of research because they provide participants with the freedom to share their experiences while still providing sufficient structure to ensure consistency across interviews (Creswell & Poth, 2016; Patton, 2014). Open questions such as "How do you deal with disagreements between tools or when something feels off?" or "What role did each tool play and how did you move between them?" were asked, and the complete interview guide can be found in Appendix A. This style of interviewing also facilitated the exploration of specific topics and themes relevant to the research question.

3.2.1 Sampling approach

The population that this research focuses on is academics at universities who use generative AI systems in their academic work. Given the qualitative nature of this research, purpose sampling was used to select participants who could provide relevant understandings (Patton, 2014). The participants were identified based on their experience and regular usage of multiple generative AI tools. This ensured that the collected data is meaningful and closely aligned with the research objectives. The aim was to interview approximately ten to fifteen respondents. This is a common number for obtaining depth and data saturation in qualitative research (Creswell & Poth, 2016). A total of twelve respondents were interviewed, as shown in Table 1.

Table 1 Respondents

Number	Respondent	Organisation
1	PHD-Candidate	University of Twente
2	Assistant-Professor	University of Twente
3	Assistant-Professor	University of Twente
4	Assistant-Professor	University of Twente
5	PHD-Candidate	University of Twente
6	Assistant-Professor	University of Twente
7	Assistant-Professor	University of Twente
8	Assistant-Professor	University of Twente
9	PHD-Candidate	University of Twente
10	PHD-Candidate	University of Munich
11	Student	University of Twente
12	Student	University of Twente

3.2.2 Data collection

The interviews were conducted either face-to-face or online, depending on the participant's preference and availability. Each interview lasted approximately 45 to 60 minutes and was structured by open-ended questions that specifically aligned with the themes identified from the literature and research questions. The interview topics included user strategies, encountered frictions, decision-making processes, tool-selection criteria, and reflection practices in mediating between different generative AI systems. With the participants' consent, the interviews were audio-recorded and subsequently transcribed. Additionally, during the interviews, notes were taken to document observation data and contextual details. The analysis followed the Gioia methodology, a thorough qualitative approach aimed at developing grounded theory from the collected data (Gioia et al., 2013). This method was selected not only for its systematic coding process but also for its compatibility with exploring emerging, practice-based processes, such as tool mediation. The Gioia approach allows the researcher to remain close to the participants' original wording while still facilitating theoretical abstraction and model development. Ethical considerations, such as obtaining consent, anonymising participant identities and secure data storage, were followed as specified in the Netherlands Code of Conduct for Academic Practice. Additionally, an ethical request was submitted to the Ethics Committee of the University of Twente.

3.3 Data analysis

The collected data was analysed using the Gioia methodology, a thorough and systematic approach particularly suitable for qualitative research aiming to uncover deeper theoretical insights from data (Gioia et al., 2013). This method enhances rigour by clearly showing how data changes into conceptual ideas and theoretical construction. It also reflects grounded theory principles such as constant comparison, iterative memo-writing and theory emergence from practice (Pratt, 2009). This approach consists of three main stages:

The first stage is the first-order analysis. Here, the transcribed interviews were thoroughly examined and coded inductively, remaining as close as possible to the participants' original wording. This step initially resulted in numerous first-order concepts, approximately 120. These 120 first-order concepts were based on the researcher's initial interest. This was too much and needed to be reduced. This was achieved by comparing them, examining the meaning behind these quotes and looking critically at how the first-order concepts can be linked to the research question and sub-questions. After multiple rounds of refining the concepts, 36 remained. These first-order concepts are terms that represent the respondents' own descriptions of their experiences and practices in mediating between multiple generative AI systems. For example, quotes such as "I always assume that it's not necessarily knows what it talks about. I double check and always go to the actual sources" and "AI tools sometimes made-up the information. I have to confirm one by one their answers" were grouped under the first-order quote 'Assumes AI outputs are unreliable and double-checks sources'. These descriptive insights were later foundational for building more abstract categories in the second-order analysis.

The second stage is the second-order analysis. Here, the first-order concepts were compared and grouped to answer higher-level "why" and "how" questions. This answered questions such as how academics decide when to switch between models or how they manage risk when the outputs of different models contradict each other. Through constant comparison and memo-writing, these concept clusters were refined into second-order themes or categories that move from description toward explanation. This stage enabled the interpretation of what the participants' words actually meant in a broader theoretical context. For example, the theme 'Epistemic Uncertainty' was created based on the quote mentioned in the first-order analysis and some additional first-order quotes. This theme represents users' continuous indecisiveness about the epistemic reliability of GenAI outputs. The outcome of this stage was a concise set of second-order themes that began to theorise how and why academics decide to switch tools, balance conflicting outputs, and control risk when mediating multiple generative AI systems.

In the last stage, the second-order themes were put into a small set of aggregate dimensions. These are higher-level constructs that capture the basic methods by which academics mediate between multiple generative AI tools. For example, the dimension of 'Frictions in interfacing with GenAI tools' was created based on the second-order theme from the second-order analysis in combination with two more second-order themes. These dimensions were integrated into a concise process model that explains the entire route from task recognition to the final adoption of AI-generated output. The resulting data structure is illustrated in Figure 1, which reflects the transition from first-order codes to higher-level themes and dimensions, as recommended by Gioia et al. (2013)

In the end, the Gioia analysis produced an empirically based process model that explicitly shows how academics recognise a need, sequence and cross-validate tools, and finally adopt outputs when mediating between multiple generative AI systems.

4. RESULTS

The analysis of the content enhanced the understanding of how academic users interact with generative AI systems in their practices. Based on the interview data, four aggregate dimensions were identified: 1. Frictions in Interfacing with GenAI Tools, 2. Tactics for Coordinating GenAI, 3. Strategic Mediation of GenAI, and 4. Transformational Learning and Identity. Three second-order themes support each dimension, based on first-order codes derived from participants' quotes. The quotes and codes that illustrate the findings can be found in Appendix A.

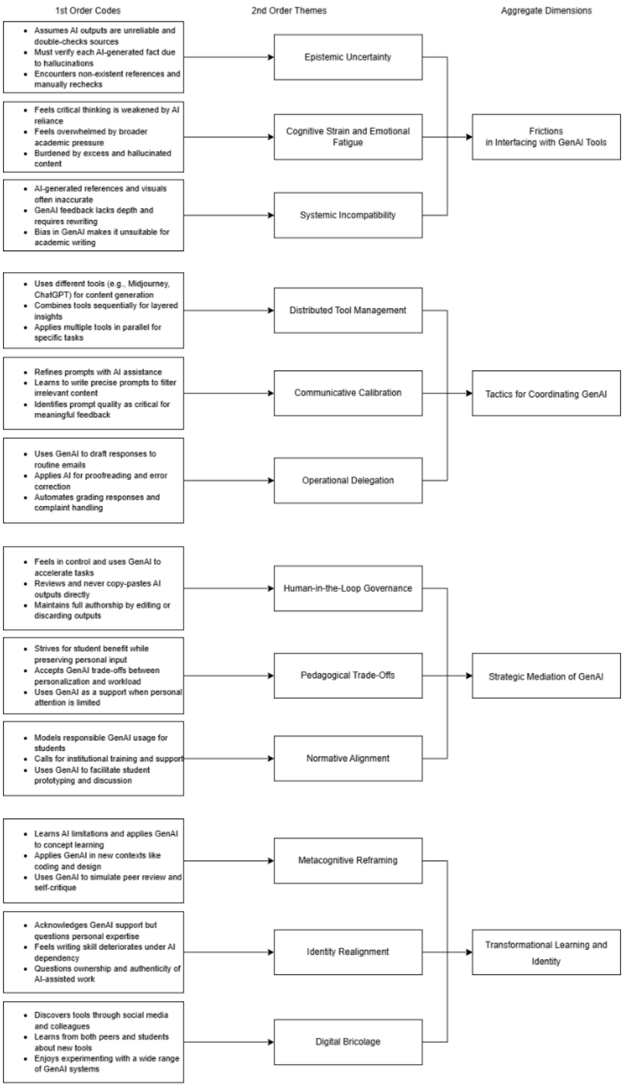

Figure 1: Data Structure

4.1 Frictions in Interfacing with GenAI Tools

While GenAI tools offer academic users new capabilities, they simultaneously present various frictions. These frictions can be categorised into three main areas: trust and validation issues, cognitive and emotional overload and tool limitations. Together, they show the discrepancies between automation and academic standards.

4.1.1 Epistemic Uncertainty

This theme represents users' continuous indecisiveness about the epistemic reliability of GenAI outputs. Participants often reported scepticism regarding the reliability and truthfulness of generated content, even though the tools can generate content fluently. To create trust, the outputs had to be manually verified against reliable sources. A recurring concern among the participants was the unreliability of GenAI outputs. Some users have had to and still do cross-check AI responses, acknowledging enduring epistemic uncertainty. One participant mentioned: "I always assume that it not necessarily knows what it talks about. I double-check it and always go to the actual sources" (Participant 7). Others, especially in response to hallucinations, described more concrete verification challenges. One participant noted, "AI tools sometimes made up the information. I have to confirm one by one their answers" (Participant 4), when pointing to a manual vetting process that reduces the time saved. Another participant could relate to this by telling "Sometimes you know, OK, this is completely wrong, these sources don't exist. Then you have to go back and manually search." (Participant 5). These quotes together demonstrate that the need for human checks often limits the effectiveness of GenAI use, reinforcing that epistemic trust must be fostered, built and maintained.

4.1.2 Cognitive Strain and Emotional Fatigue

This theme refers to the cognitive demands and mental fatigue users experience while integrating GenAI into their work. It includes cognitive offloading, fatigue from managing excess or low-quality content and a broader sense of being overwhelmed within the already challenging academic environments. In addition to practical concerns, participants also reflected on the cognitive implications of GenAI use. While one participant acknowledged that AI tools "strengthen the quality of my work," they also admitted, "But I do think it did weaken my thinking like critical thinking" (Participant 9). This comment highlights a subtle yet significant concern: that overreliance on AI could destroy higher-order thinking over time. The pressures of academic life also influenced the utilisation of GenAI. One participant mentioned: "Sometimes overloaded, it's already like overwhelming by itself. I didn't feel overwhelmed by the use of AI, but the academic world" (Participant 1). This suggests that GenAI may not always add strain directly, but instead often operates in a context where cognitive resources are already taxed. Others pointed out that the quality of the outputs can be seen as a cognitive burden. One participant shared: "ChatGPT can be a bit overwhelming. More than 50% of the ideas are useless or hallucinated." (Participant 8). These experiences show the emotional burden of navigating through excessive or unreliable content.

4.1.3 Systemic Incompatibility

This theme highlights mismatches between the intended uses of GenAI tools and the requirements of academic tasks. Participants generally faced functional or contextual gaps that limited the usefulness of a tool. Some examples are hallucinated references, vague feedback or shallow writing support. These seem less like one-off problems or experiences and more like shared issues related to the academic use of GenAI tools, which do not provide the depth, nuance, or domain-specific precision that people expect in academic settings. Even with attempts to manage their trust or cognitive load, participants regularly encountered points in their task where the tools simply could not provide what was needed to them, especially in creative or pedagogical contexts. A participant stated that GenAI "It often doesn't work like I hope. References were all fake, and it is generating images not what you want." (Participant 6). This shows the gap between system output and disciplinary standards. In teaching contexts, the support of GenAI was sometimes incomplete. One participant noted, "I tried to use ChatGPT to give feedback on students, but I still had to change the core of it. It gave a relatively ok summary, but not really good for specific pointers" (Participant 7). Another participant had concerns regarding academic writing. They stated: "I tried using it like a year ago, but I didn't like it at all. It has a really big bias and I can't use it for academic writing in my opinion" (Participant 3). These experiences demonstrate that systematic incompatibility is not a single point of failure, but rather a persistent source of friction.

4.2 Tactics for Coordinating GenAI

This dimension shows how academics develop situated practices to coordinate different GenAI tools in ways that optimise their workflows. These tactics involve combining tools strategically, iterative prompt development and integrating GenAI into repetitive academic tasks.

4.2.1 Distributed Tool Management

This theme highlights how participants utilised multiple GenAI tools in conjunction. Instead of depending on just one, they combined different tools, assigning each to tasks it handled best. Participants described the purposeful combination of different GenAI tools to support various academic tasks. One participant explained: “I prepare and use Midjourney or ChatGPT to generate some images and then use text generators to provide examples” (Participant 6). This shows tool-specific applications for content development. Other participants highlighted the importance of sequencing and layering, “If you use it in combination with Gemini and then you can ask NotebookLM to summarise it” (Participant 4), demonstrating a workflow that derives insights from multiple platforms. Another participant used tools in parallel: “I used ChatGPT to write text and then I used Copilot to create a picture” (Participant 2). Demonstrating how each tool matched a specific task. These examples demonstrate that participants deliberately chose which GenAI tools to use for the task at hand, based on the proficiency of each tool. This suggests a pragmatic and adaptive stance towards work management, in which users utilise the strengths of various tools for different academic tasks. Transitions between tools and combinations of outputs demonstrate growing experience with GenAI, as well as a practical way to take ownership by integrating their tools within personal workflows.

4.2.2 Communicative Calibration

This theme captures the evolving skill of prompt crafting, where users continually improve their communication with GenAI tools to achieve better and more precise outputs. Unlike just asking questions, this required knowing how to use the right format, structure and scope. Participants reported refining their prompts over time to improve output quality. One participant described an iterative approach: “I give a prompt. I then asked ChatGPT: Can you redefine this prompt for me to have a better search in deep search?” (Participant 4). This demonstrates AI-assisted meta-prompting. Others redefined their input to reduce irrelevant results: “I got really good at writing prompts. I scrape away a bunch of nonsense” (Participant 11). A third participant noted the risk of under-specification: “There is a difficulty that if you don’t use the right prompt, the nuances of the feedback could diminish” (Participant 4). These examples suggest that participants viewed prompting as a skill that could be developed over time. They did not rely on the default inputs but adapted their structure and wording to get better results. This demonstrates an increasing familiarity with GenAI and a growing sense of shaping how the tools respond to users.

4.2.3 Operational Delegation

This theme illustrates how GenAI has been adopted to handle low-level, repetitive, or time-consuming academic tasks. Users aimed to reduce their administrative workload while improving the consistency and professionalism of outputs. GenAI tools were utilised to optimise routine work. One participant noted: “Emails are also a big burden. I also use it to draft me a response” (Participant 6). Another stated: “It does save my time like proofreading, online editing and avoiding some stupid mistakes” (Participant 9). Some implemented GenAI into the grading workflow when handling complaints: “I used it in the context of grading to manage complaints and send automated messaging when it came to student complaints” (Participant 10). These examples demonstrate how participants relied on GenAI to relieve them of regular tasks that were often time-consuming or cognitively demanding work, allowing them to focus their time and attention on more demanding academic work. By delegating smaller tasks, like drafting or checking grammar in emails, they leveraged a more practical application of GenAI as a support tool in managing their workload.

4.3 Strategic Mediation of GenAI

This dimension examines how academic users actively manage, moderate, and position GenAI within their teaching, writing, and professional roles. Unlike the tactical dimension, this one reflects strategic decisions about responsibility, ethical alignment and the balance between automation and authenticity.

4.3.1 Human-in-the-Loop Governance

This theme explores how users intentionally maintain control over GenAI outputs. Rather than letting the technology dominate their work, participants described the need for continuous human intervention to review, adjust, or reinterpret AI-generated material. One participant reflected, “I always had the feeling that I’m still in the driver’s seat accelerating my work” (Participant 10), showing psychological assurance. Another stated, “I always have a human loop, so myself. I never copy and paste anything” (Participant 6), putting emphasis on selective and intentional use. Similarly, “I feel that I am still 100% in control. I discard a lot of output and I take ownership of everything” (Participant 8) demonstrates that users continue to control and shape their own work. These reflections not only indicate that participants did not view GenAI as a replacement for their own judgment, but rather as a tool they need to grapple with to manage. Taking ownership of what was being used and how it was utilised made them feel in control of the quality and integrity of the work they were creating. This practical approach helped ensure that GenAI remained a support tool rather than a driver in their academic process.

4.3.2 Pedagogical Trade-Offs

This theme addresses the tensions between automation and human values in academic practice. Participants reported compromises in authenticity when using GenAI for efficiency, but also reflected on practical trade-offs. As one participant said: “Always with the idea that it would be beneficial for their knowledge, but we still want to be helping out students” (Participant 4), pointing to dual concerns. Another commented, “In an ideal world, AI makes my work standardised and impersonal, but in the real world it helps a lot” (Participant 1), indicating pragmatic acceptance. A third participant added, “Sometimes we as lecturers cannot give everyone the attention they deserve. At that point AI can actually be a good substitute” (Participant 10), showing how AI can be used ethically to fill systematic gaps. The participants acknowledged the dangers of becoming standardised or detached, but also viewed GenAI as an efficient solution to situations in which time or resources were constrained. Rather than rejecting the use of AI altogether, they sought a compromise that allowed for the use of AI alongside their educational values.

4.3.3 Normative Alignment

This theme reflects the participants' efforts to utilise GenAI in ways that support their teaching goals and responsibilities. Instead of fully rejecting or blindly accepting the tools, they used them in ways that aligned with values such as transparency, digital literacy, and inclusion. One participant noted, “I want to show them appropriate use of AI so they can copy that for their work” (Participant 10), describing modelling behaviour. Another pointed out, “There must be resources readily available for lecturers. If you're not educated on AI, you will be blind to how students use it” (Participant 10), urging for institutional support. Finally, “Letting my students interact with an AI agent. Using image generation for prototyping purposes, visualising ideas that they can discuss with stakeholders" (Participant 10) demonstrates academically grounded tool integration. These examples suggest that the participants adopted a significant attitude of responsibility towards the use of GenAI. They were not just concerned with their individual benefit, but also considered their influence on students and colleagues in their use of the tools. By situating GenAI use in relation to broader educational values, they sought to model appropriate use.

4.4 Transformational Learning and Identity

This dimension reveals how engagement with GenAI tools reshapes academics' professional identities, learning approaches and views on expertise. Rather than just using tools, participants were changed by them. They gained new skills, rethought their roles and shifted their perspectives.

4.4.1 Metacognitive Reframing

This theme captures how participants learned through the use of GenAI by experimenting, adapting and critically thinking about AI limitations and capabilities. One participant noted, "I've learned a lot about how they work, what they can and cannot deliver. I use it to learn about concepts" (Participant 8), showing technical understanding. Another explained, "I think in terms of writing, I learned a lot. I used ChatGPT to help me to write the HTML so I can have like a cool, nicer and improved canvas design" (Participant 2), reflecting cross-disciplinary skill growth. A third said, "When you finish your writing, you send it to ChatGPT and ask ChatGPT to ask you a question and challenge you as a critical reviewer" (Participant 9), demonstrating AI-supported self-evaluation. The reflections indicate that participants were not only utilising GenAI tools for work, but also for learning through personal interactions with GenAI tools. They were developing new approaches to their work and how it might be adapted as they experimented and tested the GenAI tools to meet their needs. This suggests that GenAI became part of the process of reflective learning. It helped the users rethink not only what they do, but how they do it.

4.4.2 Identity Realignment

This theme describes how participants shift in their perception of their own expertise and academic roles, driven by the transformative impact and presence of GenAI. One participant explained, "I think we can all use it as a tool, so I'm glad that it can help me so I can manage more, but it's also kind of undermining my own capabilities and expertise." (Participant 2), showing uncertainty. Another said, "Sometimes when you're like time pressured, I think I actually am way worse at writing now than before" (Participant 1). Finally, "I think my writing becomes a bit better, but yes, is it my writing then?" (Participant 6). It raises fundamental questions about authorship and ownership. These examples demonstrate a tension around what it means to achieve excellence in academic work. While GenAI helped participants manage pressure or improve their output, it also led some to question whether the result still felt like their own. Some tools provided support but also presented questions about ownership and expertise due to the ambivalence around help and replacement.

4.4.3 Digital Bricolage

This theme captures participants' willingness to explore, test and adopt GenAI tools through informal networks and learning driven by curiosity. One participant stated, "I follow people on LinkedIn or accounts that discuss AI tools for research from colleagues as well, it's just like word of mouth" (Participant 7). Another explained, "I think mostly through my network, so both like people I interact with and that they're sharing or they send me something. Hey, have you tried out this one? I actually also learn from the students" (Participant 2), highlighting a multidirectional knowledge flow. "I like to experiment with all kinds of different tools. But I mainly use ChatGPT" (Participant 6) illustrates the discovery mindset. As the examples outlined previously show, participants were typically informally introduced to and became aware of GenAI tools by their peers, students or simply through personal intrigue or curiosity. Participants took the time to explore what came most naturally to them as helpful in the tasks they were doing, and learnt and developed skills through their own activities and experiences over time. This enabled them to learn in an informal and exploratory way, allowing them to stay current and develop their own solutions for working with GenAI.

4.5 Mechanisms of GenAI Orchestration

The interviews revealed not only insights into what types of experiences users had with GenAI but also how these experiences had shifted and unfolded over time. Frictions, such as uncertainty about outputs, emotional overload or system limitations, often did not exist in isolation. They sparked shifts in user approaches to the tools, how they negotiated their roles in an AI-supported workflow, and, in some cases, how they reassessed their expertise and professional identities. These accounts revealed not static observations but conceptually patterned responses: recurring ways in which participants engaged with, adapted to and learned from GenAI in practice. These recurring dynamics form what this thesis describes as mechanisms of orchestration, the processes through which users shifted between challenge, adaptation, reflection and change. Each mechanism links multiple themes across the data structure and conveys that GenAI integration was shaped by more than the performance of the tools: it involved trust-building strategies, value alignment and cognitive negotiation. While participants did not follow a single linear path, the data revealed consistent patterns that together provided a basis for constructing a dynamic model of GenAI orchestration (see Figure 2).

This model represents the process of integrating GenAI as a cyclical and ongoing process. The cycle begins with explicit frictions that prompt users to engage with adaptive responses, which at times, develop into more complex forms, such as extrapolating their thinking to a strategic reflection phase or even identity transformation. The feedback loops illustrate how users experience the actions taken in the previous phases as they discover new tools or changes in expectations. The arrows represent the empirically grounded mechanisms in this section. Users did not engage in exactly the same processes, but the model identified general pathways through which GenAI becomes embedded in academic practice.

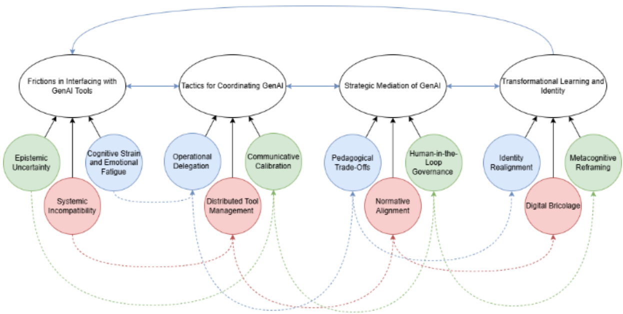

Figure 2: Dynamic model of GenAI integration in academic workflows.

4.5.1 Mechanism 1: Epistemic Calibration

This mechanism illustrates how participants navigated uncertainty and mistrust in GenAI outputs by adjusting their interaction with the tools and their understanding of their role in this process. Many academics reported having encountered hallucinated citations, incorrect facts or overly confident actions. Therefore, they did not step away from using the tools entirely, but rather began to modify their prompts, ask for clarifications or even prompt the AI to revise its own input. This interactional tuning process, which can be referred to as epistemic calibration, allowed users to lower uncertainty, regain control and modify the usefulness of outputs. This calibration turned into institutionalisation in the form of rules such as “always double-check” or “never copy-paste without editing.” These practices illustrated what participants label as a “human-in-the-loop” approach, where the technology was valued and respected but always mediated by human judgment. Eventually, this led to a form of metacognitive reframing as individuals began to think differently about how they asked questions, how they validated information or how they judged knowledge. Thus, a reaction to friction shifted into a more deeply established rethinking of epistemic responsibility regarding AI-augmented academic work.

4.5.2 Mechanism 2: Delegation as Coping

This mechanism provides insight into how both cognitive and emotional fatigue contributed to participants transferring work to GenAI tools as a means to conserve mental bandwidth. Numerous participants described feeling overwhelmed, not necessarily from GenAI itself, but as a cumulative consequence of responsibilities that come with academic life, combined with the noisy or unfiltered nature of AI outputs. In response, participants sought operational delegation using GenAI for low-stakes, repetitive tasks such as drafting emails and editing grammar or as an initial draft for providing feedback. This mechanism can be referred to as delegation as coping, because it reflects more than just an efficiency move. It is a psychologically and pedagogically informed strategy to reduce the mental cost of multi-tool orchestration. Nonetheless, delegation did not resolve tensions entirely. Participants expressed ambivalence: some feared that the automated responses compromised the authenticity of their reach, while others were concerned that excessive delegation risked diminishing their professional authorship. These tensions resulted in pedagogical trade-offs, as participants were now making reflective decisions about what tasks could possibly be automated without compromising their values. In some cases, this caused participants to realign their identity, as they questioned what their role, boundaries and expertise were based on how much control they were willing to give up. However, for others, this process was never transformative. Delegating remained an operational workaround. This mechanism illustrates how a coping tactic could serve, but not always, to produce a more robust change in role perception.

4.5.3 Mechanism 3: Value-Driven Bricolage

This mechanism illustrates how participants responded to the limitations inherent in GenAI tools by combining and adapting them, often in a pedagogically or ethically informed manner. Users regularly experienced limitations with the tools, from ambiguous feedback to unrealistic images or irrelevant citations. In response, users adopted a practice of distributed tool management, assessing a specific tool for a specific task based on its affordances and limitations. Yet, this orchestration was not purely technical. Users increasingly began to articulate these choices in terms of what is appropriate, transparent or beneficial to their students and colleagues. This mechanism can be referred to as value-driven bricolage, because across multiple tools, participants not only improvised but in ways that represented teaching norms, academic values and a desire to model ethical use. For instance, some described openly showing students how to use AI for prototyping or framing it as a conversation partner rather than an authority. Others stressed the importance of institutional support and training, arguing that unethical or uninformed use would only widen knowledge gaps. This process led to a bricolage mindset: a way of working that is adaptive, experimental, and anchored in values. Importantly, it emerged from systemic incompatibility but was able to evolve through trial, social learning and reflection. Although not all participants explicitly framed their approaches in this way, this pattern is consistent throughout the interviews. When a participant could only use limited tools, creative, value-based orchestration was an opportunity, not a barrier.

These three mechanisms (epistemic calibration, delegation as a coping strategy, and value-driven bricolage) indicate that the integration of GenAI tools into academic work is not merely a story of use or rejection. Instead, the process of GenAI interaction will frequently involve a changeable response to friction through layered, sometimes deeper learning, reflection, and identity shifts. These mechanisms illustrate the interplay of strategy, reflection and constraints that emerge in adapting to a rapidly changing technology landscape. Furthermore, this approach can help understand orchestration not as a singular skill, but as a constellation of responses, tensions and transformations in context.

5. DISCUSSION

As generative AI tools become more common in academic settings, academics are increasingly incorporating them into their teaching, research and communication practices. While these tools are clearly becoming more relevant in day-to-day academic work, we still know relatively little about how people, especially those in higher education, navigate the presence of multiple AI systems. This is particularly pressing when these systems produce opposing results, vary in reliability or create uncertainty around concepts of authorship and expertise. This study aims to address that gap by exploring the central research question: “*How do academic users mediate between multiple generative AI systems in their knowledge work?*”

The findings point to four interconnected dimensions that shape how academics experience and manage GenAI in their work: frictions in interfacing with AI tools, tactics for coordinating GenAI, strategic mediation of GenAI and transformational learning and identity. These dimensions provide scaffolding for understanding the user experience but do not represent a fixed sequence. The four dimensions, instead, capture the dynamic spaces in which users engage with, reflect on and adapt to GenAI in practice. Based on these four dimensions, the analysis also identified three deeper patterns of response, mechanisms that explain how academic users have adapted to and orchestrated their use of GenAI over time. These mechanisms (epistemic calibration, delegation as coping and value-driven bricolage) demonstrate how academic users navigate not only technical constraints but also ethical, emotional and cognitive tensions. Each mechanism brings together multiple themes across the data structure, illustrating how the adoption of GenAI tools into academic practice is by no means a linear process. Instead, it is shaped by iterative sensemaking, shifting values and evolving professional identity. Epistemic calibration represents how participants responded to the unreliable and unclear nature of GenAI outputs. They did not abandon the tools but instead worked on refining prompts, verifying sources and building self-defined norms of use. This strategy not only mitigated hallucinations and factual inconsistencies but also encouraged thinking more about what constitutes credible knowledge. In this way, calibration served as both a practical adjustment and a metacognitive reframing of how participants perceived themselves in determining and managing AI-generated content. Delegation as coping shows how participants experienced cognitive overload and managed it by delegating tasks to GenAI tools that they found repetitive or draining. This included responding to emails, proofreading, and generating feedback. While this tactic provided temporary relief and increased efficiency, it also raised concerns regarding authenticity, detachment, and the loss of academic voice. For some participants, the act of delegation engaged them in thinking about their professional limits and role definition. For others, the delegation approach continued to be a pragmatic, yet emotionally driven workaround without further transformation. Value-driven bricolage emphasised how participants engaged with the limitations and inconsistencies that GenAI tools presented by assembling and adapting multiple systems based not only on function, but also on alignment with personal and pedagogical values.

These mechanisms are captured in the dynamic model (see Figure 2), which provides a conceptual representation of GenAI integration as a cyclical, adaptive process. Frictions, such as epistemic uncertainty, emotional overload or systematic incompatibility, lead to adaptive responses that range from low-level tactics to more deliberate mediation strategies. Over time, repeated engagement with GenAI systems leads to broader shifts in professional identity, task, and epistemic norms. However, this is not a fixed endpoint. As tools develop and institutional expectations shift, frictions re-emerge that require recalibration, adaptation or resistance. Unlike linear models of adoption or efficiency, the dynamic model argues that GenAI orchestration remains ongoing and situated. It reflects a negotiation led by the user, across agency, institutional norms, tool affordances and evolving values. The mechanisms described in this research show that academic users did not simply utilise AI; they worked with and around it, sometimes strategically, sometimes through experimentation and often with differing assumptions about automation and control.

These insights provide a more pragmatic and process-oriented perspective of how GenAI is shaping academic work. Mediation between tools is not only about mediating technology; it is also a way to sustain agency while negotiating responsibility and adapting professional practices in response to friction. The following section builds on these findings to make theoretical contributions to our understanding of GenAI orchestration, user adaptation and the evolving practices of academic knowledge work.

5.1 Theoretical Contributions

This research advances the literature on human-AI interaction by providing an evidence-based explanation of how academic users mediate between multiple generative AI (GenAI) tools as part of everyday knowledge work. Although prior research has focused on the adoption of AI, trust alignment, and task fit, the vast majority of studies have examined only one tool or version of generative AI tools at a time, or static configurations. This study describes how users orchestrate across AI systems, tools that often overlap, contradict, or require different ways of interacting, and have, over time, come to learn how to manage the fragmentation. The contributions focus on three areas: mechanisms of orchestration, epistemic trust repair and professional identity negotiation. The findings challenge conventional models of technology adoption, like the Technology Acceptance Model (Davis, 1989), that considers static perceptions of usefulness or ease of use. In multi-tool GenAI environments, tool relevance is rarely stable. Participants did not make a single decision to "adopt" AI. Rather, participants continuously reassessed and reconfigured how and which tool to use based on task requirements, output quality, and the alignment of outputs with academic values. This recursive, evaluative behaviour also challenges Task-Technology Fit (TTF) assumptions (Sousa & Cardoso, 2025), which emphasises that users select tools based on pre-defined or well-understood criteria. This study demonstrated that fit is an emergent property and a co-produced condition that is shaped through ongoing experimentation, negotiation and reconfiguration during the workflow, particularly when the outputs of the tools conflict with each other or when they introduce unforeseen complications.

A key contribution of this research lies in the identification of three cross-cutting mechanisms: epistemic calibration, delegation as coping and value-driven bricolage. Each mechanism connects friction to particular forms of user adaptation and tool coordination, where each mechanism is not an isolated coping tactic but a structured response that appears consistently across participants, changes over time, and reshapes user-tool relationships. This theorisation extends orchestration theory (Dillenbourg, 2013) by demonstrating that users also orchestrate not only resources or learners but conflicting intelligences across platforms, each with its own affordances, limitations and interpretive challenges. Orchestration has traditionally been considered in instructional settings; these findings demonstrate that orchestration is also relevant for the personal, cognitive, and ethical management of AI in professional academic practice. Furthermore, the study expands the concept of orchestration by demonstrating that it is not merely strategic but also reactive, emotionally driven, and value-oriented. The mechanism of delegation as coping illustrates that operational delegation is not simply about efficiency; instead, it is a response to emotional overload or cognitive fatigue, particularly in high-pressure academic environments. Similarly, value-driven bricolage illustrates how users are creative with the tools they have and find ways to combine them while retaining pedagogical integrity, modelling accountable AI use or aligning with disciplinary norms. These mechanisms reveal that orchestration is not just an optimisation challenge, but also a negotiation of identity, responsibility, and professional norms.

This study also contributes to the literature on epistemic trust and algorithmic ambiguity. Glikson and Woolley (2020) argue that users modify their epistemic trust in AI based on explainability, perceived competence and transparency. The mechanism of epistemic calibration enables us to understand how users establish repeated cycles of re-prompting, verification, and reinterpretation of responses, not only to confirm outputs but also to recover epistemic stability in contexts of uncertainty. These behaviours were semi-informalized for users through repetition, leading to rules of thumb such as "never copy-paste things" or "always double check". This supports and extends findings from Lebovitz et al. (2021), who illustrate how AI can disrupt epistemic standards and challenge conventional authority in knowledge work. Participants in this study do not reject the authority of AI. Instead, they develop a means to co-produce credibility and iteratively engage with AI through processes that combine technical evaluation with personal judgment. By grounding this process in situated routines, this study elaborates on the literature on trust repair and epistemic accountability by highlighting that trust is not only a psychological state or an outcome of system transparency but a type of relational labour that occurs through interaction over time, particularly in contexts where no single tool is reliable enough to be trusted on its own.

Finally, this study contributes to recent work on professional identity and boundary work in AI-mediated environments. Previous work by Waardenburg et al. (2022) and Lebovitz et al. (2021) examined the ways in which AI tools destabilise established boundaries of expertise, triggering sensemaking and knowledge brokerage. The mechanism of value-driven bricolage extends these insights by showing how users recombine tools in ways that not only reflect their technical fit but also their value commitments to transparency, pedagogical care and modelling the appropriate use of AI. These academic users did not see themselves as technical experts or passive users of AI, but rather as ethical intermediaries, positioned between tools, students, and institutional expectations. This boundary work is not only strategic, but it is also reflective. Several participants described how their role as educators or scholars had shifted, not because they were using GenAI, but because using GenAI prompted them to re-evaluate what counts as authorship, contribution and original thinking. These findings indicate that GenAI not only disrupt workflows but also creates a form of identity turbulence. Some participants demonstrated stabilisation by creating new practices, while others resisted or felt discomfort. But across cases, the result was a form of identity work, where GenAI integration required users to rethink who they are and what it means to be academic in a changing technological landscape.

6. LIMITATIONS AND FUTURE RESEARCH RECOMMENDATIONS

This study provides a grounded, practice-based perspective on how academics navigate the integration of multiple GenAI tools. However, this research is not without its limitations. First, the study is based on a relatively small and purposeful sample of academics who were actively engaged in exploring and utilising GenAI tools. While this focus on lived experience proved a unique opportunity to investigate, it also incorporated a level of selection bias. The interview participants were all volunteers and may have been more reflective, critical or more innovative in their use of technology than the average academic. Therefore, the findings do not necessarily represent the active academics who are passively, hesitantly or unreflectively engaging with GenAI across the broader academic context. Second, this qualitative research has produced findings that are inevitably interpretive and contextual and not in any way generalisable. The aim of this research is to make meaning or develop a conceptual understanding based on experiences, rather than generalisability. However, this absence of generalisability means we cannot judge the extent to which certain behaviours or attitudes are prevalent across the context. Last, the study was conducted at a specific technological point in time in the evolution of GenAI. At the time of data collection, tools such as ChatGPT and Copilot were publicly accessible, but many features and integrations were rapidly changing. This presents a temporal limitation: user experiences may change as tools become more integrated into academic platforms, as training materials are developed, or as institutions establish policies. GenAI is constantly changing, and longitudinal studies will be necessary to track how practices, perceptions and identities change over time.

Building on these limitations, future research should look to enhance and expand the current findings rather than merely confirm them. Since this study included a diverse sample of students, researchers, and teachers who actively use GenAI in their academic practice, future studies could explore differences among the groups in how they engage with, adapt to and reflect on their use of GenAI in academic contexts. Comparing their usage via roles, disciplines, or proficiency could explain how strategic mediation or transformation took shape during the research process. Additionally, examining how institutional frameworks, disciplinary cultures or socio-technical access influence the use of GenAI would be worthwhile. This study focused on the individual experience with GenAI; however, broader-scale studies could map how policies or infrastructural constraints influence GenAI practice over time. Additionally, the recursive nature of GenAI integration examined in this study suggests the need for longitudinal research. Future research could follow participants over time to explore how frictions emerge again, how and if coping strategies change and how reflective practices like strategic mediation are routinised, or disappear altogether, as the landscape of tools and institutional expectations change. Finally, future research could further explore how academics negotiate concepts such as authorship, epistemic authority, and trust when using GenAI. These negotiations signify the emergence of different norms amid shifting expectations of what constitutes credible knowledge and ethical pedagogy. Research grounded in ethnographic, cross-cultural or design-based approaches could yield important insights into how academic norms are evolving under AI mediation and how ways of being reflective in orchestration are institutionalised, or contested, in different educational systems.

7. ACKNOWLEDGEMENTS

A special thanks goes out to everyone who has made this research possible. First of all, I would like to thank all the participants who were interviewed. Secondly, I would like to thank my colleague, Malina Neculaita-Mocanu, with whom I conducted the interviews. Finally, I would like to thank Dr. Johannes Dahlke for his guidance, generous feedback and helpful insights.

8. REFERENCES

Abdel-Karim, B. M., Pfeuffer, N., Carl, K. V., & Hinz, O. (2023). HOW AI-BASED SYSTEMS CAN INDUCE REFLECTIONS: THE CASE OF AI-AUGMENTED DIAGNOSTIC WORK [Article]. *MIS Quarterly: Management Information Systems*, *47*(7), 1395-1424. https://doi.org/10.25300/MISQ/2022/16773

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence [Article]. *Information Fusion*, *99*, Article 101805. https://doi.org/10.1016/j.inffus.2023.101805

Bin-Nashwan, S. A., Sadallah, M., & Bouteraa, M. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance [Article]. *Technology in Society*, *75*, Article 102370. https://doi.org/10.1016/j.techsoc.2023.102370

Bommasani, Rishi A., Hudson, D., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, Michael S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, Jared Q., Demszky, D., . . . Liang, P. (2021). On the Opportunities and Risks of Foundation Models. In. Stanford, CA: Stanford Institute for Human-Centered Artificial Intelligence.

Bozkurt, A., Karadeniz, A., Baneres, D., Guerrero-Roldán, A. E., & Rodríguez, M. E. (2021). Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century [Article]. *Sustainability (Switzerland)*, *13*(2), 1-16, Article 800. https://doi.org/10.3390/su13020800

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(1). https://doi.org/10.1186/s41239-023-00411-8

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Davis, F. D. (1989). Technology acceptance model: TAM. *Al-Suqri, MN, Al-Aufi, AS: Information Seeking Behavior and Technology Adoption*, *205*(219), 5.

Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*(24-013).

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & education*, *69*, 485-492.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., . . . Wright, R. (2023). “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy [Article]. *International Journal of Information Management*, *71*, Article 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Gavira Durón, N., & Jiménez-Preciado, A. L. (2025). Exploring the role of AI in higher education: a natural language processing analysis of emerging trends and discourses. *The TQM Journal*.

Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., Fuller, S., Singh, M., Arora, P., Parlikad, A. K., Stankovski, V., Abraham, A., Ghosh, S. K., Lutfiyya, H., Kanhere, S. S., Bahsoon, R., Rana, O., Dustdar, S., Sakellariou, R., . . . Buyya, R. (2024). Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots [Article]. *Internet of Things and Cyber-Physical Systems*, *4*, 19-23. https://doi.org/10.1016/j.iotcps.2023.06.002

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational research methods*, *16*(1), 15-31.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of management annals*, *14*(2), 627-660.

Hochmair, H. H., Juhász, L., & Kemp, T. (2024). Correctness Comparison of ChatGPT-4, Gemini, Claude-3, and Copilot for Spatial Tasks. *Transactions in GIS*, *28*(7), 2219-2231.

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines [Article]. *Computers and Education: Artificial Intelligence*, *8*, Article 100348. https://doi.org/10.1016/j.caeai.2024.100348

Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). IS AI GROUND TRUTH REALLY TRUE? THE DANGERS OF TRAINING AND EVALUATING AI TOOLS BASED ON EXPERTS'KNOW-WHAT. *MIS quarterly*, *45*(3).

Li, J., Cheng, X., Zhao, W. X., Nie, J.-Y., & Wen, J.-R. (2023). Halueval: A large-scale hallucination evaluation benchmark for large language models. *arXiv preprint arXiv:2305.11747*.

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The international journal of management education*, *21*(2), 100790.

Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature [Review]. *Education Sciences*, *13*(4), Article 410. https://doi.org/10.3390/educsci13040410

Mavrych, V., Yaqinuddin, A., & Bolgova, O. (2025). Claude, ChatGPT, Copilot, and Gemini Performance versus Students in Different Topics of Neuroscience. *Advances in Physiology Education*.

Nguyen, K. V. (2025). The Use of Generative AI Tools in Higher Education: Ethical and Pedagogical Principles. *Journal of Academic Ethics*, 1-21.

Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage publications.

Pratt, M. G. (2009). From the editors: For the lack of a boilerplate: Tips on writing up (and reviewing) qualitative research. In (Vol. 52, pp. 856-862): American Society of Nephrology Briarcliff Manor, NY.

Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education [Article]. *Journal of Applied Learning and Teaching*, *6*(1), 364-389. https://doi.org/10.37074/jalt.2023.6.1.23

Sousa, A. E., & Cardoso, P. (2025). Use of Generative AI by Higher Education Students. *Electronics*, *14*(7), 1258.

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning.

Vaithilingam, P., Zhang, T., & Glassman, E. L. (2022). Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models. Conference on Human Factors in Computing Systems - Proceedings,

Waardenburg, L., Huysman, M., & Sergeeva, A. V. (2022). In the Land of the Blind, the One-Eyed Man Is King: Knowledge Brokerage in the Age of Learning Algorithms [Article]. *Organization Science*, *33*(1), 59-82. https://doi.org/10.1287/ORSC.2021.1544

Watson, S., Brezovec, E., & Romic, J. (2025). The role of generative AI in academic and scientific authorship: an autopoietic perspective. *AI & SOCIETY*. https://doi.org/10.1007/s00146-024-02174-w

Wermelinger, M. (2023). Using GitHub Copilot to Solve Simple Programming Problems. SIGCSE 2023 - Proceedings of the 54th ACM Technical Symposium on Computer Science Education,

Zhang, S., Wang, X., Zhang, W., Chen, Y., Gao, L., Wang, D., Zhang, W., Wang, X., & Wen, Y. (2024). Mutual theory of mind in human-ai collaboration: An empirical study with llm-driven ai agents in a real-time shared workspace task. *arXiv preprint arXiv:2409.08811*.

9. APPENDIX A

Every interview started with introductions. The researcher introduces themselves and the research. The respondent was then asked a couple of general questions related to the topic before proceeding to specific, detailed questions. This was done because the interview was conducted together with another researcher who was looking into a similar topic.

General questions:

Question 1: Could you briefly describe your research area? As well as some typical day-to-day academic tasks.

Question 2: What AI tools do you typically use for these day-to-day tasks?

Question 3: Which GenAI systems do you currently rely on most, and for what tasks?

Question 4: When and why did you start using these GenAI tools?

Specific questions:

Question 5: What's in your toolkit and how did it get that way? How do new tools usually find their way in?

Question 6: Are there any you've dropped or that didn't work for you, and can you explain why this was?

Question 7: Can you walk me through a recent task where you used more than one AI tool? Did the tools ever push you in different directions?

Question 8: What role did each tool play, and how did you move between them? Was there a moment when one of them really changed how you saw the task?

Question 9: Has anything ever gone off track when using multiple tools? Can you tell me about it? Have you ever felt stuck or overwhelmed?

Question: 10: How do you usually deal with disagreement between tools or when something feels off? How did you figure out what to trust?

Question 11: Have these tools changed how you think about your work or how you work? Do you feel you've learned from using them?

Question 12: Have you had moments where something 'clicked' while using AI? Has it affected your confidence or how you see your strengths?

Question 13: How would you describe the role of AI in your work? Do you feel in control, or is it more like navigating a system?

Question 14: If you had to explain to someone how AI has changed your work, what would you say? What's been the most significant shift, for better or worse?

10. APPENDIX B

Structured Gioia Analysis

Quote	First-order Code	Second-order Theme	Dimension
"I always assume that it not necessarily knows what it talks about. I double-check it and always go to the actual sources"	Assumes AI outputs are unreliable and double-checks sources	Epistemic Uncertainty	Frictions in Interfacing with GenAI Tools
"AI tools sometimes made up the information. I have to confirm one by one their answers"	Must verify each AI-generated fact due to hallucinations	Epistemic Uncertainty	Frictions in Interfacing with GenAI Tools
"Sometimes you know, OK, this is completely wrong, these sources don't exist. Then you have to go back and manually search."	Encounters non-existent references and manually rechecks	Epistemic Uncertainty	Frictions in Interfacing with GenAI Tools
"It strengthens the quality of my work, but I do think it did weaken my thinking like critical thinking"	Feels critical thinking is weakened by AI reliance	Cognitive Strain and Emotional Fatigue	Frictions in Interfacing with GenAI Tools
"Sometimes overloaded, it's already like overwhelming by itself. I didn't feel overwhelmed by the use of AI, but the academic world"	Feels overwhelmed by broader academic pressure	Cognitive Strain and Emotional Fatigue	Frictions in Interfacing with GenAI Tools
"ChatGPT can be a bit overwhelming. More than 50% of the ideas are useless or hallucinated."	Burdened by excess and hallucinated content	Cognitive Strain and Emotional Fatigue	Frictions in Interfacing with GenAI Tools
"It often doesn't work like I hope. References were all fake, and it is generating images not what you want."	AI-generated references and visuals often inaccurate	Systemic Incompatibility	Frictions in Interfacing with GenAI Tools
"I tried to use ChatGPT to give feedback on students, but I still had to change the core of it. It gave a relatively ok summary, but not really good for specific pointers"	GenAI feedback lacks depth and requires rewriting	Systemic Incompatibility	Frictions in Interfacing with GenAI Tools
"I tried using it like a year ago, but I didn't like it at all. It has a really big bias and I can't use it for academic writing in my opinion"	Bias in GenAI makes it unsuitable for academic writing	Systemic Incompatibility	Frictions in Interfacing with GenAI Tools
"I prepare and use Midjourney or ChatGPT to generate some images and then use text generators to provide examples"	Uses different tools (e.g., Midjourney, ChatGPT) for content generation	Distributed Tool Management	Tactics for Coordinating GenAI
"If you use it in combination with Gemini and then you can ask NotebookLM to summarise it"	Combines tools sequentially for layered insights	Distributed Tool Management	Tactics for Coordinating GenAI
"I used ChatGPT to write text and then I used Copilot to create a picture"	Applies multiple tools in parallel for specific tasks	Distributed Tool Management	Tactics for Coordinating GenAI

"I give a prompt. I then asked ChatGPT: Can you redefine this prompt for me to have a better search in deep search?"	Refines prompts with AI assistance	Communicative Calibration	Tactics for Coordinating GenAI
"I got really good at writing prompts. I scrape away a bunch of nonsense"	Learns to write precise prompts to filter irrelevant content	Communicative Calibration	Tactics for Coordinating GenAI
"There is a difficulty that if you don't use the right prompt, the nuances of the feedback could diminish"	Identifies prompt quality as critical for meaningful feedback	Communicative Calibration	Tactics for Coordinating GenAI
"Emails are also a big burden. I also use it to draft me a response"	Uses GenAI to draft responses to routine emails	Operational Delegation	Tactics for Coordinating GenAI
"It does save my time like proofreading, online editing and avoiding some stupid mistakes"	Applies AI for proofreading and error correction	Operational Delegation	Tactics for Coordinating GenAI
"I used it in the context of grading to manage complaints and send automated messaging when it came to student complaints"	Automates grading responses and complaint handling	Operational Delegation	Tactics for Coordinating GenAI
"I always had the feeling that I'm still in the driver's seat accelerating my work"	Feels in control and uses GenAI to accelerate tasks	Human-in-the-Loop Governance	Strategic Mediation of GenAI
"I always have a human loop, so myself. I never copy and paste anything"	Reviews and never copy-pastes AI outputs directly	Human-in-the-Loop Governance	Strategic Mediation of GenAI
"I feel that I am still 100% in control. I discard a lot of output and I take ownership of everything"	Maintains full authorship by editing or discarding outputs	Human-in-the-Loop Governance	Strategic Mediation of GenAI
"Always with the idea that it would be beneficial for their knowledge, but we still want to be helping out students"	Strives for student benefit while preserving personal input	Pedagogical Trade-Offs	Strategic Mediation of GenAI
"In an ideal world, AI makes my work standardised and impersonal, but in the real world it helps a lot"	Accepts GenAI trade-offs between personalisation and workload	Pedagogical Trade-Offs	Strategic Mediation of GenAI
"Sometimes we as lecturers cannot give everyone the attention they deserve. At that point AI can actually be a good substitute"	Uses GenAI as a support when personal attention is limited	Pedagogical Trade-Offs	Strategic Mediation of GenAI
"I want to show them appropriate use of AI so they can copy that for their work"	Models responsible GenAI usage for students	Normative Alignment	Strategic Mediation of GenAI
"There must be resources readily available for lecturers. If you're not educated on AI, you will be blind to how students use it"	Calls for institutional training and support	Normative Alignment	Strategic Mediation of GenAI
"Letting my students interact with an AI agent. Using image generation for prototyping purposes, visualising ideas that	Uses GenAI to facilitate student prototyping and discussion	Normative Alignment	Strategic Mediation of GenAI

they can discuss with stakeholders”			
“I’ve learned a lot about how they work, what they can and cannot deliver. I use it to learn about concepts”	Learns AI limitations and applies GenAI to concept learning	Reflective Learning	Transformational Learning and Identity
“I think in terms of writing, I learned a lot. I used ChatGPT to help me to write the HTML so I can have like a cool, nicer and improved canvas design”	Applies GenAI in new contexts like coding and design	Reflective Learning	Transformational Learning and Identity
“When you finish your writing, you send it to ChatGPT and ask ChatGPT to ask you a question and challenge you as a critical reviewer”	Uses GenAI to simulate peer review and self-critique	Reflective Learning	Transformational Learning and Identity
“I think we can all use it as a tool, so I'm glad that it can help me so I can manage more, but it’s also kind of undermining my own capabilities and expertise.”	Acknowledges GenAI support but questions personal expertise	Evolving Academic Identity	Transformational Learning and Identity
“Sometimes when you're like time pressured, I think I actually am way worse at writing now than before”	Feels writing skill deteriorates under AI dependency	Evolving Academic Identity	Transformational Learning and Identity
“I like to experiment with all kinds of different tools. But I mainly use ChatGPT”	Questions ownership and authenticity of AI-assisted work	Evolving Academic Identity	Transformational Learning and Identity
“I follow people on LinkedIn or accounts that discuss AI tools for research from colleagues as well, it's just like word of mouth”	Discovers tools through social media and colleagues	Exploratory Engagement	Transformational Learning and Identity
“I think mostly through my network, so both like people I interact with and that they're sharing or they send me something. Hey, have you tried out this one? I actually also learn from the students”	Learns from both peers and students about new tools	Exploratory Engagement	Transformational Learning and Identity
“I like to experiment with all kinds of different tools. But I mainly use ChatGPT”	Enjoys experimenting with a wide range of GenAI systems	Exploratory Engagement	Transformational Learning and Identity