PhD projects in the Department of Informatics, AY 25-26 — Human-centred computing (human-computer interaction)

The PhD projects listed below will be considered for 2025/26 studentships available in the Department of Informatics to start on 1 October 2025 or later during the 2025/26 academic year.

Please note that this list is not exhaustive and potential applicants can alternatively identify and contact appropriate supervisors outlining their background and research interests or proposing their own project ideas.

Each project is designated for a single student, meaning it can only be assigned to one successful applicant. Some projects come with allocated studentships, while others are eligible for "unallocated" studentships. Applicants who apply for projects with allocated studentships and are selected will be offered a full studentship. In the project list, these are marked as "studentship allocated." Applicants chosen for other projects will compete for the unallocated studentships.

We welcome applications from students who have secured, or are applying for, or plan to apply for other funding (within other studentships internal to the university or external schemes) and from self-funded students. See also this list of funding opportunities available at King's for post-graduate research in Computer Science.

PhD projects

- Debugging and runtime visualisation in a frame-based system (studentship allocated)
- AI and NLP for Multilingual Code-Switching in Education (studentship allocated)
- Leveraging Generative AI for Creativity Education (studentship allocated)
- Enabling Access to the Co-Design Process
- Embodied Approaches to Assistive Technology
- Enabling Accessible Remote Communication Environments
- Enabling Access to Digital Content
- Accessible Health Data Representation
- Visual live programming in scientific computing and similar domains
- Exploring the Diversity of Data Storytelling: A Comparative Analysis Across Cultures, Disciplines, and Media
- Understanding the Complexity of Negotiations
- Exploring Interactive Multi-Dimensional Approaches of Delivery of Communication in Patient Scenarios in Oral Health Education
- Predictive Profiling using Biometric Data in Educational Environment.
- Designing novel privacy IxD mechanisms in mobile health apps
- Inclusive and Accessible Cybersecurity Education
- Investigating LLM-based Generative AI Applications in Cybersecurity
- Designing and Developing a framework for responsible security and privacy practices for GenAI Tools

Debugging and runtime visualisation in a frame-based system

Supervisor: Prof Michael Kolling / Dr Neil Brown

Areas: Human-centred computing (human-computer interaction), Computing Education, Systems (software engineering, programming)

Project Description

Frame-based editing is a novel program manipulation paradigm that combines advantages from both text-based and block-based editors. It has been implemented in the Stride language [1], and in the online Strype system [2]. The visualisation possibilities embedded in frame-based editors provide opportunities for improved debugging and runtime visualisation functionality, beyond what is available in typical text-based or block-based systems. The goal of this project is to design and implement novel debugging/visualisation functionality in a frame-based system. Candidates for this project must have exceptional programming skills in multiple languages, a deep understanding of object-orientation, interest in HCI and in programming education.

References

[1] https://stride-lang.net [2] https://strype.org

AI and NLP for Multilingual Code-Switching in Education

Supervisor: Zheng Yuan

Areas: Artificial Intelligence (symbolic AI, logic, etc.), Machine learning / Deep learning, Natural Language Processing, Human-centred computing (human-computer interaction)

Project Description

The rapid growth of multilingualism has led to an increased prevalence of code-switching (CSW) -- the practice of alternating between two or more languages within a single conversation or utterance. Despite its common usage in multilingual communication, current Natural Language Processing (NLP) technologies struggle to handle CSW effectively, particularly in educational contexts. This project aims to address this gap by developing advanced NLP technologies and educational AI systems specifically designed to support multilingual CSW environments. The goal is to create a personalised, inclusive, and engaging AI-powered tutoring system that adapts to the unique linguistic needs of learners. This project will focus on one or more of the following key areas: 1) Development of NLP models that can accurately process and analyse CSW data, distinguishing code-switching from grammatical errors; 2) Creation of an Intelligent Tutoring System (ITS) that provides personalised feedback and assessment tailored to the needs of multilingual learners; 3) Leveraging multilingual Large Language Models (LLMs) to enhance the capabilities of AI in educational settings, particularly in low-resource languages; and 4) Evaluating the impact of educational AI systems in real-world settings, assessing improvements in learning outcomes, learner engagement, and satisfaction.

References

- LLM-based Code-Switched Text Generation for Grammatical Error Correction. Tom Potter and Zheng Yuan. EMNLP 2024.
- Prompting open-source and commercial language models for grammatical error correction of English learner text. Christopher Davis, Andrew Caines, Øistein E. Andersen, Shiva Taslimipoor, Helen Yannakoudakis, Zheng Yuan, Christopher Bryant, Marek Rei and Paula Buttery. ACL 2024 Findings.
- Grammatical Error Correction for Code-Switched Sentences by Learners of English. Kelvin Chan, Christopher Bryant, Li Nguyen, Andrew Caines and Zheng Yuan. LREC-COLING 2024.
- Grammatical Error Correction. Christopher Bryant, Zheng Yuan, Muhammad Reza Qorib, Hannan Cao, Hwee Tou Ng and Ted Briscoe. Computational Linguistics; https://doi.org/10.1162/coli_a_00478
- Building Educational Technologies for Code-Switching: Current Practices, Difficulties and Future Directions. Li Nguyen, Zheng Yuan and Graham Seed. Languages; https://doi.org/10.3390/languages7030220

Leveraging Generative AI for Creativity Education

Supervisor: Zheng Yuan

Areas: Artificial Intelligence (symbolic AI, logic, etc.), Computer vision, Machine learning / Deep learning, Human-centred computing (human-computer interaction), Natural Language Processing

Project Description

Creativity is a crucial skill in today's world, driving innovation, problem-solving, and cultural expression. However, teaching and assessing creativity -- especially in fields like creative writing and visual arts -- pose significant challenges due to the subjective nature of creative outputs. The rise of Large Language Models (LLMs) and Generative AI provides new opportunities for enhancing creativity education by generating personalised, adaptive feedback and supporting learners in improving their creative skills across multiple modalities, such as writing and drawing. This project will explore the use of multimodal LLMs and Generative AI to enhance creativity education, offering new ways to assess creativity and helping learners across disciplines such as creative writing and digital arts. By leveraging the capabilities of multimodal models, this research will investigate how AI can support, nurture, and assess creativity in a personalised and scalable manner. Research Questions: 1) How can LLMs and Generative AI effectively assess creativity in different forms, such as written stories, poems, or drawings? 2) What are the most effective ways for AI systems to provide feedback that nurtures creativity, without stifling originality? 3) How can multimodal AI systems enhance cross-disciplinary creative education (e.g. combining writing and drawing) to create richer, more engaging learning experiences? 4) What metrics and frameworks can be developed to evaluate the success of AI-generated feedback and creativity assessment systems?

Enabling Access to the Co-Design Process

Supervisor: Timothy Neate

Areas: Human-centred computing (human-computer interaction)

Project Description

Co-design, the collaborative act of designing with end-users, often relies heavily on language-based methods like reading, writing, and discussion. This can pose challenges for individuals with language impairments, leading to the exclusion of these stakeholders' valuable insights and resulting in products and services that fail to meet diverse user needs. This PhD project will build on our group's work in co-created personas, generative AI, and tangible interfaces to develop novel tools and techniques that ensure equal participation in co-design for all users. The research will focus on creating accessible, innovative methods to broaden the reach and inclusivity of co-design across various contexts, improving design outcomes for marginalized groups.

Embodied Approaches to Assistive Technology

Supervisor: Timothy Neate

Areas: Human-centred computing (human-computer interaction), Machine learning / Deep learning, Computer vision

Project Description

Non-verbal expression plays a crucial role in everyday communication, whether nodding to indicate agreement or using vocal tone to imply a question. For individuals with language impairments, non-verbal cues are essential for both comprehension and expression. However, most assistive technologies overlook these vital communication methods (see our systematic review). This PhD project will extend our research on wearable devices such as smartwatches, smartbadges, and augmented reality (AR) tools, focusing on innovative solutions for non-verbal communication. You will collaborate directly with communities who experience language impairments to design technologies that support effective communication in real-world settings.

Enabling Accessible Remote Communication Environments

Supervisor: Timothy Neate

Areas: Computer vision, Human-centred computing (human-computer interaction)

Project Description

Remote communication, such as videoconferencing, has become an integral part of daily life. However, it often lacks essential elements like non-verbal cues that are critical for effective communication. This can pose significant challenges for people with language impairments. Building on our previous research into the accessibility of videoconferencing for individuals with aphasia, this PhD will explore new ways to enhance remote communication technologies. By collaborating directly with these communities, you will design and develop tools that make videoconferencing and other online environments more inclusive and supportive for all users.

Enabling Access to Digital Content

Supervisor: Timothy Neate

Areas: Human-centred computing (human-computer interaction)

Project Description

Digital content, such as streamed video and audio, is central to modern life, keeping us informed and entertained. However, people with communication challenges often struggle to access and understand this content. This PhD project will build upon our ongoing research and the EPSRC-funded CA11y project, which has explored the challenges faced by users with complex communication needs (such as aphasia) in consuming video content. We have developed accessibility interventions, such as independently adjusting audio channels or inserting pauses in dialogue for improved comprehension. Your work will extend these efforts, working closely with end users to design future content platforms that provide inclusive access for diverse communicators.

Accessible Health Data Representation

Supervisor: Timothy Neate

Areas: Human-centred computing (human-computer interaction)

Project Description

Health dashboards can empower individuals by giving them control over their health data, potentially improving health outcomes. However, the complexity of health data can be a barrier, especially for users with cognitive or language-based disabilities. This PhD project will build on our ongoing work developing a stroke dashboard for the South London Stroke Register, focusing on making health data more accessible. You will collaborate with stroke survivors to explore how we can represent complex data, as well as insights and explanations generated by machine learning systems, in ways that are accessible and easy to understand for all users.

Visual live programming in scientific computing and similar domains

Supervisor: Stephen Kell

Areas: Systems (software engineering, programming), Human-centred computing (human-computer interaction)

Project Description

Currently, working interactively with data means either using a pre-built application offering a fixed interface, which is visual and interactive but offers limited programmability, or using custom workflows built by programming/scripting or command-line wizardry, which are flexible but technically demanding and far less visual and interactive. Computational notebooks like Jupyter are in some senses a third way, being somewhat visual, and have proven approachable by those seeking to learn programming 'on the job'. However, they currently suffer many usability and reproducibility issues, and still present a 'walled garden' environment with poor integration into the surrounding system. This PhD is about ways to combine the interactivity of applications and the flexibility of command lines, possibly by designing a notebook system that works differently than Jupyter et al. We observe that crude operating system (OS) interfaces are the bottleneck to interoperable, visual programming, since they lack a rich data model on which to build visualisation as a system-wide service; this is what leads to smaller-scope walled-garden approaches. Recent work adding run-time type information to native code has shown that such limitations can be overcome without defining an entirely new platform. This project will pursue similar approaches encompassing file data and graphical user interface elements. The objective is to demonstrate a graphical workspace that is highly compatible and interoperable, dealing in files of existing formats, but can support working visually and programmatically via a palette of small, composable, user-tailorable graphical tools. Target audiences include computational scientists, data scientists, digital artists and the like. The project requires systems programming skills and an interest in human-- computer interaction topics.

Exploring the Diversity of Data Storytelling: A Comparative Analysis Across Cultures, Disciplines, and Media

Supervisor: Alfie Abdul-Rahman

Areas: Human-centred computing (human-computer interaction), Natural Language Processing

Project Description

Data storytelling is a rapidly evolving field that integrates data analysis, narrative techniques, and visualizations to convey complex insights. Despite its growing significance, the diversity of approaches to data storytelling across cultural contexts, disciplines, and media platforms remains largely unexplored. This PhD project aims to investigate how these variations influence data stories' effectiveness, ethics, and reception. By examining this diversity, this PhD aims to reveal the impact of cultural, disciplinary, and media-related factors on how data is interpreted and communicated. Traditionally, data storytelling has been viewed through the lens of information visualization and communication. However, recent studies have broadened the scope to include narrative techniques, audience reception, and the cultural context of both the storyteller and the audience. Most existing research focuses either on the technical aspects of data visualization or narrative elements, often neglecting cross-cultural and cross-disciplinary analysis. As media platforms diversify, understanding how these platforms mediate data stories and influence engagement becomes increasingly important. This research addresses these gaps by providing a comprehensive and comparative analysis of data storytelling. It explores how the cultural and disciplinary backgrounds of both creators and audiences and the media platforms used to shape the form and impact of data stories.

Understanding the Complexity of Negotiations

Supervisor: Alfie Abdul-Rahman

Areas: Natural Language Processing, Human-centred computing (human-computer interaction)

Project Description

A negotiated text is the product of a formal decision-making process where a text has been negotiated and drafted over a period of time. Many of the foundational texts of the modern world have not been written by individuals, by negotiated by groups of people in formal settings. For example, treaties between states such as the Universal Declaration of Human Rights or the Treaty of Versailles; or constitutions, such as the one negotiated by the American states in the Constitutional Convention of 1787. During such negotiations, it is important for us to keep track of the delegations and their involvements in order to grasp their influence on the negotiation process either using techniques such as close reading, distance reading, or machine learning. Even relatively short historical documents written collectively in this way have been the product of thousands of specific proposals and decisions. This project will apply a visual analytics approach towards the understanding of the complexity of a negotiation and the influence of the delegations during a negotiation process. Possible research questions: - Developing new static and interactive visualization to assist with data discovery and insight generation in large datasets of events within interacting timelines. - Developing new approaches to show the evolution of complicated, technical documents over the period of months or years. - Developing new approaches for indexing the datasets related to the negotiation of documents, and more intuitive displays of the results. - Developing natural-language-based approaches to relate information captured in 'informal' archives (such as private diaries, letters, social media feeds etc.) to the formal records of a negotiation.

Exploring Interactive Multi-Dimensional Approaches of Delivery of Communication in Patient Scenarios in Oral Health Education

Supervisor: Informatics: Dr Alfie Abdul-Rahman & Dr Lin Gui FoDOCS: Dr Melanie Nasseripour & Dr Ana Angelova

Areas: Artificial Intelligence (symbolic AI, logic, etc.), Natural Language Processing, Human-centred computing (human-computer interaction), Machine learning / Deep learning, Education

Project Description

This is a joint project between the Department of Informatics and the Faculty of Dentistry, Oral & Craniofacial Sciences (FoDOCS). Communication in patient scenarios in oral health education can be cost-intensive in terms of time and resources. In this project, we propose exploring interactive multidimensional approaches such as immersive technology, text-to-text, and voice-to-voice communication delivery in patient scenarios in oral health education. These approaches enhance the learning experience and offer a cost-effective solution, making the delivery of communications in patient scenarios in oral health education more feasible and sustainable. This project aims to design and create adaptable, contextually relevant patient scenarios, offering engaging and realistic interactions for students. The beauty of these approaches is their adaptability. Whether it is a VR interactive tool, text-based, or voice-based conversation, they can all respond to students' inquiries and actions in real-time, mimicking the interaction they would normally have in the clinic. This adaptability ensures the relevance and effectiveness of the project in various educational settings. We aim to examine the Generative Language Models (GLMs) to generate customized case studies and simulation scenarios so that each learner can practice specific skills repeatedly in a controlled environment. This encourages the acquisition and refinement of skills, such as explaining the importance of oral hygiene and discussing dietary habits. The critical focus is patient-clinician communication, behaviour change, professionalism, etc.

Predictive Profiling using Biometric Data in Educational Environment.

Supervisor: Tasmina Islam

Areas: Machine learning / Deep learning, Education, Human-centred computing (human-computer interaction)

Project Description

Mental health and well-being of students is very important in achieving their full potential during academic studies in university [1]. Predicting their mental and emotional status can be very useful in monitoring student's well-being and providing the appropriate support at the time when needed. Although the principle focus of biometrics is identification/verification of individuals, biometric data can be used to predict some lower level (age, gender, ethnicity, etc.) and higher level (mental state, emotion etc.) individual characteristics [2]. Different biometric modalities (e.g., face, voice, EEG signals, keystroke, handwriting etc.) can be explored utilising this predictive capability to predict students' mental and emotional status that may have impact on their academic performance. As well as monitoring well-being, both physiological and behavioural biometrics can play a big role in facilitating education, for example, tracking attendance, monitoring engagement, and learning behaviour (especially when learning remotely). These could be beneficial for both students and educators. Due to the wider use of biometrics, the analysis of biometric data poses some challenges if the biometric data is captured under unconstrained environment, for example, voice recognition in a crowd or with noise/echo, full or partly covered mouth (e.g., wearing a mask), face recognition in limited/unevenly distributed light, pose variations of individuals, noise like other people in the background, where some parts of the face is occluded (e.g., wearing a mask or a sunglass) and many more. This project aims to explore different factors that affects the biometric recognition performance and investigate how to manage and improve the performance in facilitating education. The project will also explore the predictive capabilities of biometric data under both constrained and unconstrained environment. Prospective students can discuss about different modalities and options with the supervisor.

References

1. Smith, A.P., 2019, November. Student workload, well-being and academic attainment. In International Symposium on Human Mental Workload: Models and Applications (pp. 35-47). Springer, Cham.
2. Fairhurst, M., Li, C. and Da Costa-Abreu, M., 2017. Predictive biometrics: a review and analysis of predicting personal characteristics from biometric data. IET Biometrics, 6(6), pp.369-378.

Designing novel privacy IxD mechanisms in mobile health apps

Supervisor: Ruba Abu-Salma

Areas: Cybersecurity, Human-centred computing (human-computer interaction)

Project Description

Mobile health (mHealth) apps provide a wide range of benefits. However, they collect a significant amount of sensitive user medical data, posing privacy risks to users. Whilst legislation exists to protect users' medical and health data, levels of protection vary across countries. Users may also overestimate these legal protections and, as a result, trust mHealth apps unduly with their health data---or may lose trust due to lack of transparency, and avoid mHealth despite its benefits. Efforts have been made to restrict mobile apps' data practices and improve transparency, but are not always efficacious at addressing problematic app behavior or improving users' understanding of app data practices. This project aims to design and evaluate interaction design (IxD) mechanisms that enable users to mitigate the privacy risks associated with the use of mHealth apps while allowing them to continue benefiting from such apps.

References

https://kclpure.kcl.ac.uk/ws/portalfiles/portal/251441290/chi24-626-21.pdf

Inclusive and Accessible Cybersecurity Education

Supervisor: Tasmina Islam

Areas: Cybersecurity, Human-centred computing (human-computer interaction), Education

Project Description

People are spending more time online due to the increasing digitisation of society. This also means the security measures need to be stronger and awareness of the cybersecurity risks, and implications is crucial, particularly for vulnerable and underrepresented groups, such as children, ethnic minority communities, and people with disabilities, who may face increased risks due to limited access to tailored resources and education. This project aims to address these challenges by developing adaptive learning environments that cater to individual needs, offering age-appropriate content for children and culturally relevant, accessible material for diverse communities. Learners will engage with realistic cybersecurity scenarios, equipping them with essential skills like identifying phishing attacks, managing privacy, and securing personal information. Using a combination of qualitative methods, such as interviews and focus groups, and quantitative surveys, the project will study the specific cybersecurity challenges faced by these populations. The knowledge gained from this research will inform the development of scalable educational tools, integrating AI for personalised learning and immersive technology to create engaging, hands-on experiences. By addressing the cybersecurity education gap, this initiative seeks to empower all learners with the knowledge and skills to safely navigate the digital world, enhancing digital safety, privacy, and security for all while promoting a more inclusive digital landscape. Prospective students can discuss options with the supervisor.

References

1. Hedges, M., & Islam, T. (2024). VirSec — Immersive Security Training within Virtual Reality. In 17th International Conference on Advanced Visual Interfaces: 2nd International Workshop on CyberSecurity Education for Industry and Academia (CSE4IA 2024)
2. Islam, T & Zou, Y 2023, ChildSecurity: A Web-based Game to Raise Awareness of Cybersecurity and Privacy in Children. in Cybersecurity Challenges in the Age of AI,Space Communications and Cyborgs.

Investigating LLM-based Generative AI Applications in Cybersecurity

Supervisor: Ievgeniia Kuzminykh/Hannan Xiao

Areas: Artificial Intelligence (symbolic AI, logic, etc.), Cybersecurity, Natural Language Processing, Human-centred computing (human-computer interaction)

Project Description

The rapid development and deployment of GPT-based agents in cybersecurity mark a significant leap forward in approaching digital security challenges from a practitioner standpoint. Under this topic you can explore the ways generative AI is impacting the cybersecurity industry, from both sides, such as gen AI for security and security of gen AI. From one side, malicious attackers are seizing the potential of generative AI to launch cyber attacks that are harder to detect and defend against. OWASP top 10 for Generative AI [1] lists out the top 10 vulnerabilities impacting the applications usilising LLM. Prompt Injection, Insecure Output Handling and Data Poisoning take the top 3 spots and are also the root causes for the other type of vulnerabilities (Overreliance, Insecure Plugins etc) as shown in research paper [2]. From another side, Gen AI is also helping make security teams more accurate, efficient, and productive in defending their organisations. Examples of utilising of generative AI for security operations could be [3] : - Supplementing understaffed security teams - Detecting threats in real time - Improving incident response. The potential topics in this project area could include but not limited to: 1. Optimisation of prompts for security related topics. Through clever prompt engineering (called jailbreaking [4]), LLMs can be made to reveal internal mechanisms, share private data, produce offensive speech, or perform unintended workloads. LLMs thus pose a security risks [1, 5, 6]. 2. Prompt injection detection mechanisms. 3. Ensuring online safety using LLM. AI seems like the perfect response to the growing challenges of content moderation on social media platforms: the immense scale of the data, the relentlessness of the violations, and the need for human judgments without wanting humans to have to make them. The paper [7] elaborates on the topic of prompt/response classifiers. The prompts and answers could be classified into groups such as safe and harmful. Typical examples of a harm would be Child Safety, Exfiltrating PII/SPII, Sexually Explicit Content, Malicious/Dangerous content. 4. Content moderation using LLM. Similar to previous but can be extended to the detection of harassment and throlling [8]. 5. Understanding Generative AI for Cloud Security. Generative AI can make new data from existing patterns. For cloud security, this means it has the potential to: 6a. Simulate Threat Scenarios: Generative AI can create realistic threat scenarios, allowing security teams to test and validate their Cloud infrastructure's resilience. By simulating potential attack vectors, organizations can proactively identify vulnerabilities and take steps to ensure they are protected against them before they are exploited. 6b. Optimize Security Configurations: AWS offers a number of services, each with its own set of security configurations. With Generative AI, we can analyze existing configurations, simulate various combinations, and ask Generative AI to provide recommendations based on our specific needs. 6c. Enhance Monitoring and Alerts: By training on historical security logs and events, Generative AI can predict potential security breaches or anomalies. The key word here is "potential." Knowing what "could" happen allows security teams time to prepare and allows for more rapid action to be taken. 6. Understanding Generative AI for firewall optimisation Generative AI could simulate web traffic patterns based on your historical log data and compare that to your existing WAF or firewall rules, ensuring that malicious requests are blocked while legitimate traffic flows seamlessly. 7. 10. Understanding of Gen AI for Qualitative audit of security policies Each organisation is governed by a security policy, which technically or conceptually specifies a number of guidelines for ensuring IT security. You will investigate whether GenAI can be employed to translate a security policy for wider staff [9, 10]..

References

[1] https://owasp.org/www-project-top-10-for-large-language-model-applications/
[2] Knowledge Bases and Language Models: Complementing Forces. https://link.springer.com/chapter/10.1007/978-3-031-45072-3_1
[3] https://secureframe.com/blog/generative-ai-cybersecurity
[4] Liu, Y., et al.: Jailbreaking ChatGPT via prompt engineering: an empirical study. arXiv preprint arXiv:2305.13860 (2023)
[5] https://simonwillison.net/2023/May/2/prompt-injection-explained/
[6] https://www.geeksforgeeks.org/what-is-jailbreak-chat/
[7] Llama Guard: LLM-based Input-Output Safeguard for Human-AI Conversations https://arxiv.org/html/2312.06674v1
[8] Content moderation, AI, and the question of scale https://journals.sagepub.com/doi/full/10.1177/2053951720943234
[9] Introducing Rules Genie: Generative AI for Automating Policy Creationhttps://www.opsmx.com/blog/introducing-rules-genie-generative-ai-for-automating-policy-creation/
[10] The LLM Police: Between Firewalls and Policies for Generative AI https://www.fairly.ai/blog/the-llm-police-between-firewalls-and-policies-for-generative-ai

Designing and Developing a framework for responsible security and privacy practices for GenAI Tools

Supervisor: Maher Salem

Areas: Cybersecurity, Human-centred computing (human-computer interaction), Artificial Intelligence (symbolic AI, logic, etc.)

Project Description

Generative Artificial Intelligence (GenAI) technologies have transformed human life, impacting areas such as healthcare, education, and social interactions. While GenAI tools offer creative content generation, data synthesis, and automation benefits, they also introduce significant challenges that may negatively impact individuals and communities, particularly vulnerable groups. These challenges include privacy and ethical issues, legal risks, bias and discrimination, misinformation, and inaccurate outputs. Despite the importance of these issues, there is limited empirical research on users' experiences and views regarding GenAI security and privacy. This project aims to apply both qualitative and quantitative methods to investigate how users interact with GenAI tools, the reasons behind their use, and how these experiences shape perceptions of GenAI. A key component of the study will involve understanding users' mental models of the benefits and risks of GenAI, the educational resources they use to assess potential risks, and the protective measures they adopt. By examining users' learning processes and resources, this research will provide insights into gaps in GenAI literacy. The study's educational goals include developing targeted resources and practical guidelines to improve GenAI literacy among diverse groups. Empirical insights from this research will guide the design of safeguards for GenAI technologies and inform curriculum and policy recommendations, enabling institutions to equip students, educators, and users with the skills to navigate GenAI responsibly and securely.