Empowering Students in Emerging Technology: A Framework for Developing Hands-on Competency in Generative AI with Ethical Considerations

Dr. Chun Kit Chui, University of Hong Kong

Dr. Chun Kit Chui serves as the Director of the Tam Wing Fan Innovation Wing in the Faculty of Engineering at the University of Hong Kong (HKU). Innovation Wing aims to unleash students' creativity by entrusting them to spearhead ambitious innovation and technology projects that will shape the future. The iconic facility is located at the heart of the campus, offering 2400m2 of space with state-of-the-art resources and a supportive environment to enhance hands-on and experiential learning for undergraduate students.

He also holds the position of Senior Lecturer in the Department of Computer Science at HKU. His research interests include database and data mining, as well as pedagogical research in computing education. Dr. Chui has received several education awards, including the University Outstanding Teaching Award (Individual Award) at the University of Hong Kong for the 2015-16 academic year and the Faculty Outstanding Teaching Award (Individual Award) in the Faculty of Engineering for the 2012-13 academic year. Additionally, he has been honored with the Teaching Excellence Award in the Department of Computer Science for the academic years 2011-12, 2012-13, 2013-14, 2014-15, and 2015-16. Furthermore, he was a shortlisted candidate for the UGC Teaching Award (Early Career Faculty Member).

Dr. LEI YANG, The University of Hong Kong

Lei Yang is a lecturer of Innovation Academy of the Tam Wing Fan Innovation Wing under the Faculty of Engineering, The University of Hong Kong. Before that, he worked as a Research Officer at Centre of Transformative Garment Production from 2021 to 2023 and as a postdoctoral fellow at Department of Computer Science, The University of Hong Kong from 2018 to 2021. Dr. Yang received his Bachelor's degree and Ph.D. degree from Dalian University of Technology in 2012 and 2018, respectively. Dr. Yang's research interest includes Computer-Aided Design, Computer-Aided Engineering, and Geometry Modeling and Multimedia.

Prof. Ben Kao, University of Hong Kong

Ben Kao received the B.Sc. degree in computer science from the University of Hong Kong in 1989 and the Ph.D. degree in computer science from Princeton University in 1995. He is currently full professor in the Department of Computer Science at the University of Hong Kong, the Associate Head of the Innovation Academy, and the co-Director of Hong Kong University's Law & Technology Centre. From 1992 to 1995, he was a research fellow at Stanford University. His research interests include database management systems, data mining, information retrieval systems, AI and NLP.

A Framework for Developing Hands-on Competency in Generative AI with Ethical Considerations

Abstract

This practice paper introduces a framework to enhance the practical skills of undergraduate engineering students in generative AI technologies. Our goal is to transform students from users of generative AI software into professional creators of new AI technologies. We begin by defining guidelines, emphasizing ethical, responsible, and lawful practices. Then we define the core practical competencies and design learning activities. The framework involves collaboration among undergraduate students, postgraduate tutors, instructors, technicians, legal experts, academic partners, and industry professionals.

The framework adopts a three-stage progression approach. At the adoption stage, students become familiar with generative AI software, such as composing textual prompts for image generation with the stable diffusion model. This helps students stay updated on the latest tools and developments in generative AI applications. In the development stage, the focus is on technical training for application programming interfaces (APIs), including language completion, text-to-speech conversion, and semantic search. This covers hands-on learning of using open-source large language models such as Llama2, and commercial cloud services such as Microsoft Azure OpenAI and Google GCP Vertex AI, etc. Techniques like Retrieval-Augmented Generation (RAG) and model fine-tuning are part of this training, equipping students with the necessary skills to enter the final application stage. In this stage, students participate in designing and developing generative AI-based solutions to address real-world problems. Our partnerships extend to the law and social science faculties, where we build customized chatbot solutions.

The framework was implemented and evaluated at the Tam Wing Fan Innovation Wing (a.k.a. HKU Inno Wing) [1], a facility within the Faculty of Engineering at the University of Hong Kong dedicated to improving students' practical abilities. Students demonstrate increased awareness of ethical, responsible, and lawful practices in generative AI technologies under the careful guidance of instructors. We conducted an analysis of the written reflections from students in the 2023/24 cohort regarding their understanding of the strengths and weaknesses of generative AI technologies. Furthermore, we assessed how students' awareness of generative AI ethics, responsibility, and legal considerations evolved throughout their reflections. By identifying common blind spots, we gained valuable insights to continually enhance guidance for students at various stages of their learning progress.

Keywords
Generative AI, AI competency, AI ethics

Background

In the rapidly advancing field of artificial intelligence technology, there is a pressing need to bridge the gap between theoretical knowledge and practical skills, especially in the realm of

generative AI. Within the undergraduate engineering curriculum at the institution, generative AI-related courses primarily emphasize theories related to natural language processing, machine learning, and language models. However, they often lack actual hands-on experience in using contemporary tools and application programming interfaces (APIs) to build practical AI solutions for solving real-life problems. Such solutions often require cross-disciplinary collaboration with academic partners and industrial collaborators, who may contribute domain expertise and training data in their specific field of generative AI application.

To address this gap in the curriculum, we propose a framework to guide students through a structured three-stage progression, emphasizing practical competencies alongside ethical considerations. Our aim is to equip the upcoming generation of engineers with the skills and mindset necessary to navigate the field of generative AI while upholding ethical, responsible, and lawful practices.

The framework was piloted at the HKU Inno Wing as an extracurricular learning opportunity for undergraduate students in the 2023/24 academic year. The Inno Wing, a 2,000-square-meter innovation workspace established by the engineering faculty in December 2020, supports student innovation and hands-on learning by providing state-of-the-art equipment and extensive project workspace. In the realm of generative AI, the center has taken a leading role at the University of Hong Kong in nurturing the next generation of technology innovators who are both knowledgeable and responsible leaders committed to developing technology for the betterment of society. In 2023/24, the center expanded its infrastructure and activities, including computing facilities, software tools, and cloud resources, to support the learning and development of generative AI technologies for undergraduate students.

Related works

On ethical aspects of AI technology

In recent years, there has been a significant surge in the development of AI technology. As a result, diverse efforts have been dedicated to raising public awareness about both the advantages and limitations of AI [2], [3], [4]. A substantial body of literature delves into the impacts, as well as opportunities and challenges that generative AI presents to the education community in teaching and learning [5], [6] and to the scientific community in academic integrity and the risk of plagiarism in scientific publications [7], [8], [9]. To address these urgent concerns in ethic use of generative AI, many universities have developed their guidelines for students and staff on how to use these AI tools in the teaching and learning scenarios [10], [11]. However, from a student's perspective, these guidelines may be too general or condensed, making it difficult for them to relate the terms to real-world practices. While the guidelines at the Inno Wing align with those of other universities, covering aspects such as AI tool limitations, data privacy, intellectual property, and responsible practices, we aim to impart the concepts to our students through an experiential, progressive framework. This approach exposes students to different levels of use of generative AI tools and technology, enabling them to develop a concrete understanding of the underlying principles behind the terms outlined in the guidelines.

Generative AI-related curricula

Universities are actively developing coursework to enhance students' proficiency in generative AI technology while also addressing its limitations and ethical considerations. One notable example is the three-week experiential learning course offered by MIT [12], which places emphasis on both technical and managerial aspects of generative AI technology. This particular course targets senior leaders, managers, and professionals from the business sector, catering to a different audience compared to the framework proposed in this paper. In contrast, UC Berkeley's Data Science 290 course [13] takes a more technically oriented approach, aligning well with the third level of progression outlined in our framework. This course provides students with the necessary technical skills to work with generative AI technologies.

In designing our framework, we aim to provide undergraduate engineering students, as well as students from other faculties, with hands-on experience in generative AI through a three-level progressive training approach. By targeting entry-level students, we hope to motivate them to engage with generative AI technology and develop their skills and understanding in this field. Our goal is to equip students with the necessary knowledge, experience, and ethical awareness to navigate the opportunities and challenges presented in the era of generative AI.

AI ethics in Computer Science

The field of computing education has a significant surge in research focusing on ethics inclusion, as highlighted in the survey conducted in [14]. An outcome of this survey is the recognition of the need for more targeted pedagogical innovation to clarify expectations regarding students' understanding of ethics in computing. In response to this, the primary focus of this paper is directed towards the emerging field of generative AI, adopting a student-centered approach in the design of the learning framework. Our framework embraces an interdisciplinary approach, in alignment with the definition provided in [15], wherein students and teachers include cross-disciplinary professionals from various fields, including social work and social administration, law, and engineering. This collaborative approach ensures a well-rounded exploration of ethics in generative AI, drawing insights from diverse disciplines and fostering a holistic learning experience.

In addition to complementing theoretical-focused computing ethics research, as exemplified by [16], this practice paper implements hands-on experiential learning within our innovation center. By bridging the gap between theory and practice, we provide students with firsthand experience and insights into the ethical dimensions of generative AI. Furthermore, through our practical experiences, we identify potential blind spots that may arise in the context of AI-related teaching. This approach complements the story completion method discussed in [17], as we engage students in authentic scenarios in stage three of the proposed progressive framework. Through these efforts, our aim is to refine and strengthen our framework, fostering an environment that promotes ethical awareness and responsible practices in the field of generative AI.

Guidance for using generative AI tools and building applications

We emphasize the importance of closely integrating ethical education with every aspect of technical competency training. This involves continuous efforts to define ethical guidelines for generative AI, specifically tailored to students and staff. These guidelines should address the responsibilities of both users and developers. In light of this, we have taken the initiative to draft the guidelines and establish a continuous feedback mechanism involving various stakeholders, including legal experts, academic staff, technical staff, and students. This inclusive approach allows for ongoing critique and review to refine the guidelines. Table 1 presents a partial list of these guidelines as of April 2024; a comprehensive list can be found in [18]. All users, including students and staff involved in generative AI-related activities/projects, are required to agree to these guidelines before accessing the resources offered by the Inno Wing.

Responsible user: Ethical review and validating facts in crafted content

In emphasizing responsible usage, two key aspects demand our close attention. Firstly, in content crafting, users must take responsibility for the generated content. We aim to establish stringent practices among our users to review the generated content, preventing harm, bias, discrimination, misinformation, or engagement in any unethical behavior [19]. Usage must consistently adhere to ethical standards, refraining from harassment, hate speech, plagiarism, or any illegal activities. Secondly, users must acknowledge the inherent limitations of the created content and understand the potential for inaccuracies or incoherence within the context, often referred to as hallucination [20]. Therefore, users should refrain from treating the generated content as a reliable source of factual information. Instead, we advocate for the establishment of strict practices for factual validation through research and literature review.

Responsible developer: Transparent data collection and authorized IP use

In the development of generative AI technologies, it is important to focus on transparent data collection and the authorized use of intellectual properties. While approaches may vary among AI service providers regarding the collection and utilization of user data, it is common for developers to avoid tracking unnecessary personal data that is irrelevant to the functionality of the application. Our established practice requires developers to transparently inform users about the data to be collected and its intended use. We emphasize the importance of clearly conveying how personal data inputted into AI tools will be utilized beyond immediate results generation. Additionally, we aim to establish practices that mandate authorized use of intellectual property (IP) through proper citation, acknowledgment, and obtaining permission from data owners for the training of AI models.

Table 1. Guidance for Using Generative AI Tools and Building Applications
(partial list, version in February 2024)

1	**Responsible AI: Crafting Ethical and Responsible Contents** ● Generative AI tools empower us to create content. Use this power responsibly by ensuring that the generated content adheres to ethical and legal standards. ● Refrain from generating content that promotes harm, bias, discrimination, misinformation, or any form of unethical behavior. Test results/contents that may have embedded such unethical behavior and avoid using them. ● Avoid any misuse of generative AI tools, including harassment, hate speech, plagiarism, or any illegal activities.
2	**Limitation on created contents: Understanding the Limits of Generated Contents** ● Recognize that user of AI-generated results/contents should bear the ultimate accountability in using them. ● Recognize the issues of confident responses produced by AI tools that are inaccurate or incoherent with the context, also known as hallucination, and the implications of prompt engineering in AI-generated results/contents. ● Consider whether AI-generated results/contents encourage or discourage diversity and conformity.
3	**Data Protection: Safeguarding User Data** ● Ensure that the data used with AI tools complies with the University's privacy and data protection regulations. ● Do not collect any unnecessary personal data from users that is not relevant to the application's functionality. ● Clearly inform individuals who utilize AI tools about the data to be collected and how the data will be used. ● Understand fully how personal data inputted to AI tools will be used beyond generating immediate results. ● Please seek approval from Human Research Ethics Committee before collecting any data for research purposes.
4	**Intellectual Property: Respecting Ownership Rights** ● Ensure compliance with the University's Intellectual Property Rights Policy. Do not use copyrighted materials, including those generated by AI tools, without proper authorization. ● Encourage collaboration among students while ensuring proper credit for contributions to projects involving generative AI. Give credit where it's due.
5	**Usage for Teaching and Learning: Not for Commercial Activities or Production Deployment** ● The generative AI tools provided by the Inno Wing are intended for teaching and learning purposes, prototyping, and non-production use only. They should not be used to support commercial activities or software deployment.

The three-stage progression framework

In June 2023, a team was formed in the Inno Wing, including a lecturer with expertise in AI and robotics as the staff-in-charge. This team comprises student interns selected based on their proficiency in machine learning and natural language processing, particularly those who excelled in relevant computer science courses. The primary objective of this team was to identify core technical competencies and create learning materials in generative AI for the 2,800 undergraduate users of the Inno Wing. Additionally, these pioneering interns undertook

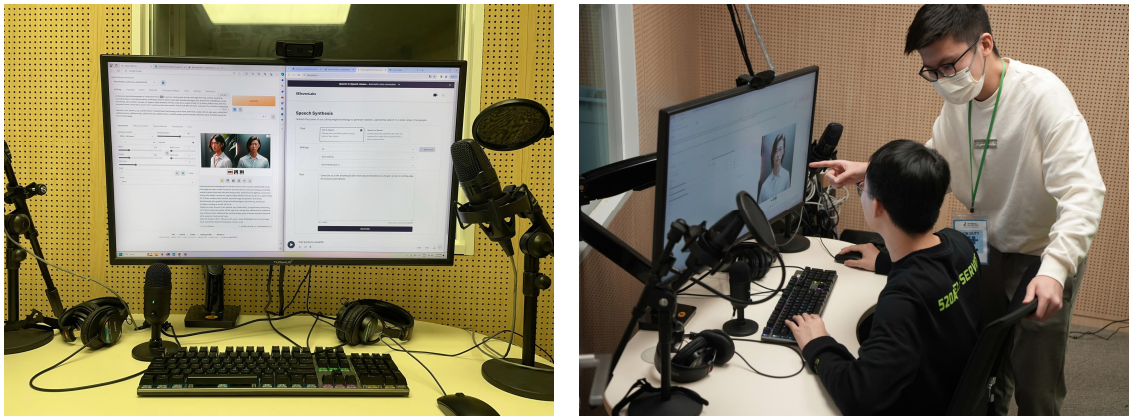

Figure 1. The walkthrough hosted in the podcast studio, which has been upgraded with high-performance computing facilities to power the computationally extensive gen-AI software.

initiatives to identify and address real-life problems. The experiences, results, and findings from these real-life projects are being shared with the student community through knowledge exchange workshops and seminars. For the 2023/24 academic year, the team consists of nine student interns.

To ensure a comprehensive approach, the team actively engaged academic advisors and industry collaborators involved in the field of generative AI. Collaborators for the 2023/24 academic year included academic staff from the law faculty and social science faculty, as well as industrial partners from cloud service providers such as Microsoft Azure and Google GCP.

1. Adoption stage: Become proficient users of contemporary AI tools

To familiarize students with contemporary software, the team consistently identifies the latest tools and developments in generative AI. For example, in the 2023/24 academic year, the growing interest in chatbots and image generation software offers an excellent opportunity for the Inno Wing to provide introductory experiences and make these techniques accessible to students. The focus of this stage is to help students become familiar with the capabilities of existing technology and develop skills in becoming proficient users of these technologies, such as composing textual prompts for image generation using the stable diffusion model [21].

In the 2023/24 academic year, the team initiated the design of a generative AI walkthrough experience for students (Figure 1), accessible at [22]. The initial offering aims to create an AI-generated virtual avatar that can synthesize the user's speech. This would be of great interest to students as it allows them to create short presentation clips, which are useful for various project presentations. Therefore, it would be best to host this walkthrough in the podcast studio of the Inno Wing, which consists of multimedia equipment for students to create podcasts and present materials. The podcast studio is upgraded with high-performance computing facilities to support the computationally intensive gen-AI software. The learning materials are carefully designed to include various latest software, such as stable diffusion

models for avatar image generation, Elevenlabs for voice generation, and D-ID for voice-avatar synchronization .

In terms of ethical considerations, this stage offers an authentic opportunity for students to engage in ethical review and fact validation when working with crafted content. For example, when students witness the capabilities of AI-generated images, the importance of "validating facts" becomes evident, as it highlights the potential for image manipulation or falsification. Additionally, when students experience the capability of voice synthesis, it raises concerns about the potential misuse of downloaded voices from online sources to create a realistic voice model of someone else. This raises awareness among students regarding privacy infringement and potential criminal offenses, prompting them to adopt strong ethical, responsible, and lawful practices when utilizing this technology.

2. Development stage: Evolve into developers of new generative AI technologies

The focus of this stage is on providing technical training to empower students to become developers of new generative AI technologies. We place emphasis on their ability to utilize the latest libraries and tools, including application programming interfaces (APIs) offered on various cloud platforms. For the 2023/24 academic year, the team has defined the core technical competencies as language completion, text-to-speech conversion, semantic search, and computer vision.

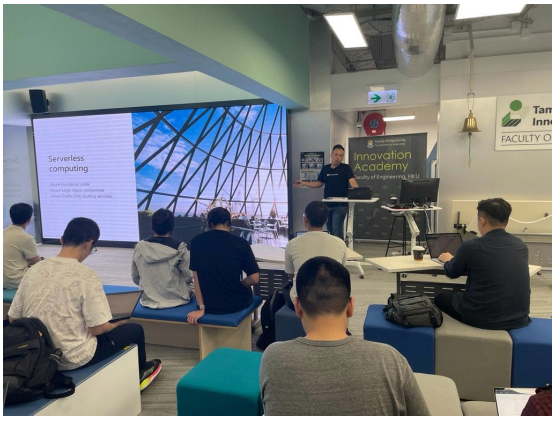

Figure 2. The workshops "*Building AI Solutions on Microsoft Azure Cloud*" organized by industrial experts from Microsoft Azure.

Figure 3. Student interns visiting and exchanging ideas with professional developers from Google Cloud.

To acquire this knowledge, the team explores various open-source large language models such as Llama2 [23] and utilizes cloud services like Microsoft Azure OpenAI and Google GCP Vertex AI. Additionally, the team organizes hands-on workshops with industrial experts (Figure 2) and arranges visits to interact with professional developers (Figure 3). This process enhances the technical prowess of the team, positioning them to lead other students in acquiring these technical skills on a larger scale.

In order to translate the acquired knowledge into learning materials for other students, the team designs task-based workshops that incorporate the learning of core competencies. In the

2023/24 academic year, the team developed a workshop where students can design and develop a user-customized chatbot that leverages APIs to implement large language model optimization techniques, such as Retrieval-Augmented Generation (RAG) [24], as depicted in Figure 4. Ongoing development extends these techniques to include fine-tuning a language model and integrating deep learning-based computer vision techniques. These workshops, hosted at the Inno Wing, are open to all students, enabling them to acquire fundamental competencies in generative AI application development. This equips them to apply their technical knowledge to solve real-life problems in the subsequent stage.

Figure 4. The workshops "*Unleash Creativity with Generative AI through Open AI Engine and ChatGPT Build Your Personalized Chatbot*" organized by student interns for other students.

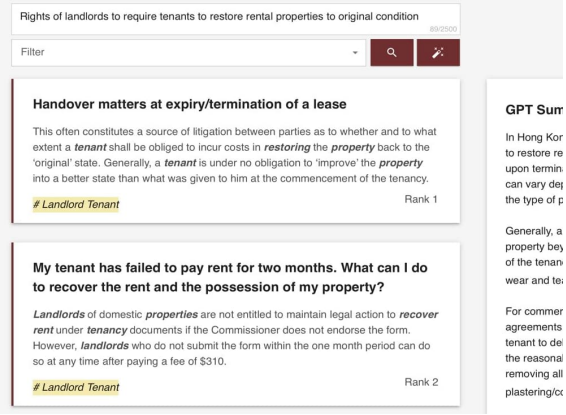

Figure 5. Students built a legal chatbot to assist underprivileged people in asking legal questions without the use of professional and complex terms [25].

In terms of ethical considerations, this stage presents an authentic scenario for practicing transparent data collection and authorized use of intellectual properties. When students prepare training data to construct a customized language model during the workshop, they are prompted to ensure compliance with regulations pertaining to the use of intellectual properties. Furthermore, the potential logging of all chatbot conversations raises awareness about user privacy. Questions arise regarding the ethical implications of utilizing conversations to enhance current dialogue or for further training of language models, which could potentially expose conversation content to other users. This situation highlights the importance of responsible, ethical, and lawful practices among generative AI developers.

3. Application stage: Apply generative AI skills to solve real-world problems

In this stage, students actively participate in designing and developing generative AI-based solutions to address real-world problems through cross-disciplinary and industrial collaborations. To ensure successful academic collaboration, our strategic focus involves exploring social good projects with specific applications in non-engineering fields.

In the 2023/24 academic year, the team established two cross-faculty collaborations. The first collaboration involved partnering with the law faculty, engaging students in developing an additional layer for a chatbot that answers legal questions. Recognizing the challenges faced

by underprivileged groups in expressing legal queries in complex terms, the team worked on a chatbot service that recommends legal topics based on a client's situation (Figure 5). This initiative aims to assist underprivileged individuals in understanding their legal protection and rights [24]. The second collaboration was with the social science faculty, exploring the use of generative AI to develop an assistive chatbot for social workers. This chatbot provides information on Social Welfare Department policies, addressing topics such as tenancy regulations and working family allowances for those in need. Additionally, the team creates simulated AI counseling agents to train professional social workers.

To ensure successful industrial collaboration, all stakeholders, including the academic staff-in-charge, industry partners, and student interns, must have a clear understanding of their roles and responsibilities. The collaboration should align industry objectives with academic learning outcomes, fostering a synergy that provides a real-life environment for practical technology application while allowing for academic exploration. For instance, during the pilot run with a railway company in the 2023/24 academic year, the theme was “AI and robotics”, aligning seamlessly with the company’s strategy. Professional engineers and student interns worked in the industrial environment to explore the adoption of generative AI in training and maintenance aspects of railway operations.

In terms of ethical considerations, the team encounters a critical challenge in data collection, emphasizing the necessity for ethical, responsible, and lawful practices. This involves employing precise data handling techniques such as anonymization and data masking to safeguard privacy before utilizing the data for training AI models. During the visit to the data studio of the railway company, both parties emphasize the importance of consensus in determining the permissible use of data, avoiding privacy-sensitive or commercially confidential information. Unlike synthetic cases encountered in coursework, the data preparation process emerges as a demanding aspect of the overall development. This hands-on experience provides students with valuable insights into the significance of professional practices, highlighting how professionals in academia and industry navigate these issues through established policies.

Methodology

We conducted a recruitment process to select non-final year undergraduate students for the pilot run of our three-stage progress framework. Through an open call for participants within our institution, we successfully recruited nine undergraduates for the pilot run. These students possess basic technical proficiency in machine learning and natural language processing acquired through computer science courses. Among them, eight students are from the engineering faculty, while one student is from the law faculty. Starting in June 2023, these students underwent the entire three-stage progress framework. To gather their feedback, we invited them to provide written reflections in response to the guiding questions presented in Table 2 upon their completion of the application stage of the progression framework. To ensure an unbiased analysis, we maintained a neutral stance without predetermined themes and utilized a coding scheme outlined in [26]. The analysis was carried out by two researchers.

During the initial coding phase, the researchers engaged in an open exploration of the data, allowing themes to naturally emerge. They categorized words and sentences according to each criterion and assigned to mutually exclusive categories [27]. Through subsequent readings, they gained a deeper understanding of the reflections and identified overarching themes. This iterative coding and categorization process followed a systematic framework, ensuring a comprehensive and unbiased analysis. The resulting themes, free from any preconceived notions, provide valuable insights into the diverse perspectives and experiences of the participants in the pilot run of our framework. The themes presented in the next section are aligned with the overall educational framework of the project.

Table 2. Guiding questions for written reflection.

	Guiding questions for written reflection
1	What are the technical strengths and limitations of generative AI technologies (including but not limited to Large Language Models) that you learned from your experience?
2	Are you aware of any ethical issues due to the applications of these technologies and how to practically avoid these ethical issues?
3	Can you list some of the irresponsible practices of using and developing generative AI technologies?
4	Please list three examples of the possible unlawful practices (e.g., copyrights or privacy) when applying and developing generative AI technologies.

Results and Discussions

1. The strengths and weaknesses of Generative AI

Exploring the technical strengths of generative AI reveals a shared recognition among students for its efficacy in various tasks. Large Language Models (LLMs) exhibit a notable capacity to comprehend and execute human natural language commands, streamlining repetitive work. A student highlights this, noting, "*LLMs' ability to understand human natural language commands and perform everyday tasks enhances task efficiency for repetitive work.*" Another technical strength identified by students is the swift content generation of LLMs, contributing to the acceleration of the development lifecycle. As the student mentions, "*Their high speed in generation was also impressive, enabling them to automate certain tasks and expedite research pipelines.*" Additionally, generative AI technologies excel in creative tasks such as text, image, or video generation, simplifying low-level content creation, as noted by most students.

The limitations of generative AI, especially within Large Language Models, form a consensus among students. A notable concern revolves around the struggle with intellectually demanding tasks due to the absence of suitable training data. As explained by a student, "*The lack of appropriate training data poses challenges for LLMs in tackling complex problems that require specialized domain knowledge or complicated logical reasoning.*" Moreover, the computational resource requirements for training large models tailored for specific applications present a challenge, as highlighted by another student. In their words, "*Limitation – computation power is a substantial barrier, both in terms of the training phase and the utilization (e.g., paid versions) of generative AI tools.*" The reflections also highlight

limitations in training datasets that can lead to biased output and potential societal impacts, such as racial or gender discrimination. Additionally, privacy issues during information exchange processes are brought to the forefront, with a student noting, "*Privacy issues might also arise during the information exchange process, as users inadvertently disclose personal or confidential information to the large model.*"

A special insight involves the lack of explainability in generative AI, hindering understanding of how AI arrives at conclusions. A student notes, "*Another challenge we encountered is the lack of explainability, making it difficult to understand how these AI arrive at their outputs, hindering accountability and potentially masking hidden biases.*"

To address these limitations, students have learned to leverage alternative techniques to overcome the lack of domain-specific knowledge and implement predefined rules or guidelines to control the creative abilities of LLMs, reducing the chance of misinformation. One student suggests, "*The hallucination of LLMs is an unsolved problem but to reduce its chance of misinformation, we can use multiple other LLM agents to monitor and examine the output of the original LLM.*"

2. Ethical issues in Generative AI

Ethical considerations surrounding the applications of Generative AI technologies, particularly Large Language Models (LLMs), are a focal point for students. Trustworthiness within the field emerges as a significant ethical issue, highlighted by a student "*The black-box nature of large models masked the thought process, which raises the possibility and concerns of hallucinations.*" This perspective draws attention to the challenge of interpreting and understanding the decision-making mechanisms of Generative AI, highlighting the need for transparency and explainability to build trust. Another student expands the ethical discourse by focusing on the vulnerability to manipulation, specifically through the creation of deepfakes and AI-generated propaganda. This perspective brings forth the potential malicious use of Generative AI, emphasizing the necessity for safeguards to protect information integrity and public trust. They state, "*The creation of deepfakes and AI-generated propaganda, which pose threats to information integrity and public trust.*" This empahsizes the ethical responsibility to anticipate and counteract misuse, acknowledging the societal impact of manipulated content.

To address these ethical concerns, students propose solutions that span technical and architectural considerations. The suggestion of "*careful system architecture design*," as mentioned by one student, involves aligning tasks with the capability of large models to enhance overall accuracy and reliability. Additionally, the deployment of local models or isolating Generative AI components from other system parts is advocated to prevent privacy issues. This student emphasizes, "*By carefully selecting tasks that align with the capability of the large models, we can improve the overall accuracy and reliability of the system.*"

A unique perspective introduced by another student emphasizes the critical role of education in fostering responsible adoption. Beyond technical solutions, this insight recognizes the

human element in ensuring ethical practices. The call for both developers and users to "*exercise caution and critically evaluate AI outputs*" highlights the importance of informed decision-making and ethical awareness in the broader adoption and deployment of Generative AI. This student asserts, "*Education is also important to guide responsible adoption*."

3. Irresponsible practices in Generative AI

The reflections on the responsible use of generative AI showcase a shared concern among students regarding potentially irresponsible practices. A student rightly points out the risk of the "*misuse of large models*" leading to the dissemination of false information or harmful content. This emphasizes the ethical responsibility to ensure that generative AI technologies are used judiciously. The absence of disclaimers regarding possible inaccuracies before user interaction is identified as an irresponsible practice that warrants attention. Expanding the perspective, another student elaborates on the broader societal impact by addressing the manipulation of social media influence through fake accounts and content. The acknowledgment of this practice emphasizes the need for responsible development to preserve the integrity of online platforms. The quote, "*Fake accounts, contents, and comments might reduce the impact of social media content,*" highlights the potential harms associated with the misuse of generative AI.

A notable point raised by a student is the over-reliance on AI-generated content, leading to social problems such as excessive emotional attachments to AI agents or even addictiveness. This insight draws attention to the psychological impact of generative AI, and the quote, "*The over-reliance of AI-generated content might also cause social problems like excessive emotional attachments to AI agents, or even addictiveness,*" highlights the societal implications and the need for responsible deployment to avoid negative consequences.

To counteract these irresponsible practices, students advocate for the incorporation of disclaimers and content moderation. The quote, "*Presenting AI-generated contents (text, images, etc.) without clearly citing that it is AI-generated,*" highlights the concern about transparency and the importance of clearly indicating the AI origin of content. This proactive approach aligns with responsible development practices and can mitigate potential misinterpretations or misuse.

4. Unlawful practices in Generative AI

The students' reflections offer insights into the unlawful practices associated with Generative AI technologies. One student articulates the concern of copyright infringement, pointing out, "*Copyright infringement is a significant concern, as large language models can be used to replicate copyrighted works without permission*." This emphasizes the potential legal ramifications when deploying AI systems that generate content infringing upon existing copyrights. In further exploration, another student brings attention to the mishandling of data, specifically stating, "*Misusing training data (e.g., collecting and utilizing personal information without permission) can lead to privacy violations and legal consequences.*" The acknowledgment of these privacy concerns emphasizes the necessity of obtaining explicit

consent for data use, necessary for compliance with legal frameworks and ethical considerations. Turning the focus to potential misuse, students collectively emphasize the threat of generative AI for creating fake identities or producing counterfeit documents, as expressed by one student, "*The use of AI-generated content for malicious purposes like creating fake identities or producing counterfeit documents or artworks.*"

To mitigate these concerns, the students advocate for proactive measures such as obtaining consent for data use, meticulous training dataset selection, and strict adherence to legal frameworks. A student reinforces these preventative measures by enumerating unlawful examples, specifically stating, "*The collection of user data, such as their conversation with the AI tool for future training purposes without the consent of the user.*" Additionally, the student emphasizes the importance of refraining from using copyrighted material without authorization to prevent potential legal repercussions.

Common blind spots

While the insights enrich the ethical discussion, the analysis reveals significant blind spots in student reflections on generative AI. According to the academic partner with a background in social science, there's limited exploration of how students excessively rely on AI for providing factual information, despite being informed that it's not designed for that purpose. Students generally lack awareness that the information they provide to AI systems could potentially be exposed to the public. "*They are not reminded about the real 'cost' of using this 'free service'"* as warned by the academic partners. Furthermore, there is no discussion about equity and the digital divide in the realm of generative AI technology, which is regarded as another significant ethical issue that is escalating in the education system.

Regarding responsible AI, there is limited discussion on the role of common AI platforms and policymakers in regulating and monitoring the use of generative AI. While individual developers play an essential role, the absence of comprehensive systemic measures is highlighted as a common oversight. Broader regulatory frameworks and monitoring mechanisms are deemed necessary to address the societal impact of irresponsible AI use. Additionally, there is a lack of discussion regarding service providers' role in safeguarding users' data and the unethical practice of selling such data for commercial purposes.

On the legal aspect, students have limited exploration of international legal implications and variations in regulations concerning generative AI. The need for developers to comprehend and navigate the global legal landscape is highlighted as essential for ensuring compliance and responsible deployment of generative AI technologies on a global scale.

Addressing the blind spots in the continued development of the framework

To address the identified blind spots, the Inno Wing will collaborate with professional AI practitioners in the industry and legal experts to enhance students' understanding in three key areas. Firstly, hands-on modules will be incorporated into the framework, focusing on the limitations of AI in providing factual information and emphasizing the risks of excessive

reliance on AI systems. Students will receive education regarding the phenomenon of LLM hallucination and the potential public exposure of their information when interacting with AI.

Secondly, expert seminars will be organized to explore the role of common AI platforms, policymakers, and regulatory frameworks. Students will gain insights into international legal implications and variations in regulations related to generative AI. This enable students to navigate the global legal landscape and ensure responsible deployment and compliance. The interdisciplinary nature of the center will be leveraged to incorporate perspectives from various faculties, including law and social science, in addition to engineering.

In addition to the mentioned measures, it is important to establish a solid formal review panel with a standardized procedure to evaluate the ethical practices outlined in Table 1. This panel should actively involve students, providing them with the opportunity to critically analyze and challenge ethical practices not only in their own AI project but also in the projects of their peers. Through peer review, students can learn from each other, share insights, and collectively promote responsible AI development.

Conclusion

In this practice paper, we detail our pilot run of a three-stage progression framework aimed at enhancing the practical skills of undergraduate engineering students in generative AI technologies. The framework incorporates ethical awareness, responsible practices, and legal considerations into its development. A working guideline for using Generative AI tools and building applications is proposed and enforced in the Inno Wing in the 2023/24 academic year. Initial analysis of students' learning within the pilot implementation of the framework indicates a progression from proficient users to professional developers, developing strong core technical competencies in generative AI. Student reflections not only demonstrate an understanding of the strengths and limitations of the technology but also foster awareness of ethical, responsible, and legal practices. These insights offer valuable guidance for refining measures in guiding the pedagogical development in the realm of generative AI for both students and the institute.

However, the analysis also reveals significant blind spots in student reflections on generative AI. Notably, there is a lack of exploration regarding international legal implications and variations in regulations, emphasizing the need for developers to understand global legal landscapes for compliance and responsible deployment. Another key observation is the limited discussion on the role of common AI platforms and policymakers in regulating and monitoring generative AI, indicating a need for more comprehensive systemic measures beyond individual developers. Additionally, there's a potential blind spot concerning the psychological impact on users forming excessive emotional attachments to AI agents, prompting a call for a more holistic examination of emotional and psychological dimensions in the ethical considerations. These findings represent a work-in-progress, contributing to the ongoing engineering pedagogical development in response to the emergent and dynamic nature of generative AI technologies.

Acknowledgement

We would like to acknowledge the support provided by the Tam Wing Fan Innovation Wing, Faculty of Engineering, University of Hong Kong. We are grateful for the support received from the Faculty of Engineering, HKU, as well as the Tam Wing Fan Innovation Fund and the Philomathia Foundation Innovation Fund. Additionally, we extend our heartfelt gratitude to the founding staff of Inno Wing, including Professor Norman C. Tien, Mrs. Wendy Ma, Professor Leslie George Tham and Dr. C.K. Chui, as well as the staff and students involved in the Innovation Academy. Their dedication and contributions have been the key driving force behind the remarkable success of the project.

Ethical approval

Ethical approval for this study was obtained from the University of Hong Kong, with HREC reference number EA230632.

References

[1] Chun Kit Chui, Norman C. Tien. "The Journey of Establishing and Operating an Innovation Center to Nurture Future Engineering Innovators," *2024 ASEE Annual Conference & Exposition*. 2024.

[2] Vöneky, S., Kellmeyer, P., Mueller, O., & Burgard, W. (Eds.). (2022). *The Cambridge handbook of responsible artificial intelligence: interdisciplinary perspectives*. Cambridge University Press.

[3] UNESCO, C. "Recommendation on the ethics of artificial intelligence." (2021).

[4] Hurlburt, G. (2023). What if ethics got in the way of generative ai?. *IT Professional*, *25*(2), 4-6.

[5] Kajtazi, M., Holmberg, N., & Sarker, S. (2023). The changing nature of teaching future IS professionals in the era of generative AI. *Journal of Information Technology Case and Application Research*, 415-422.

[6] Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, *13*(9), 856.

[7] Alasadi, E. A., & Baiz, C. R. (2023). Generative AI in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education*, *100*(8), 2965-2971.

[8] Schlagwein, D., & Willcocks, L. (2023). "'ChatGPT et al.': The ethics of using (generative) artificial intelligence in research and science". *Journal of Information Technology*, *38*(3), 232-238.

[9] Lin, Z. (2024). Building ethical guidelines for generative AI in scientific research. *arXiv preprint arXiv:2401.15284*.

[10] Faculty of Information (iSchool), University of Toronto, "Practical Elements of Responsible AI Development." https://ischool.utoronto.ca/course/practical-elements-of-responsible-ai-development/

[11] Kavita Bala, Alex Colvin, (2023), Cornell University, "Generative Artificial Intelligence for Education and Pedagogy" https://teaching.cornell.edu/sites/default/files/2023-08/Cornell-GenerativeAIForEducation-Report_2.pdf

[12] MIT Professional Education (2024), "Applied Generative AI for Digital Transformation: Professional Education."https://professional.mit.edu/course-catalog/applied-generative-ai-digital-transformation

[13] UC Berkeley School of Information. "Data Science 290. Generative AI: Foundations, Techniques, Challenges, and Opportunities." https://www.ischool.berkeley.edu/courses/datasci/290/genai

[14] Brown, N., Xie, B., Sarder, E., Fiesler, C., & Wiese, E. S. (2024). Teaching Ethics in Computing: A Systematic Literature Review of ACM Computer Science Education Publications. *ACM Transactions on Computing Education*, *24*(1), 1-36.

[15] Goetze, T. S. (2023, March). Integrating ethics into computer science education: Multi-, inter-, and transdisciplinary approaches. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1* (pp. 645-651).

[16] Petelka, J., Finn, M., Roesner, F., & Shilton, K. (2022, February). Principles matter: integrating an ethics intervention into a computer security course. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education-Volume 1* (pp. 474-480).

[17] Han, B., Nawaz, S., Buchanan, G., & McKay, D. (2023, June). Ethical and pedagogical impacts of AI in education. In *International Conference on Artificial Intelligence in Education* (pp. 667-673). Cham: Springer Nature Switzerland.

[18] Sætra, H. S. (2023). Generative AI: Here to stay, but for good?. *Technology in Society*, *75*, 102372.

[19] Tam Wing Fan Innovation Wing, the University of Hong Kong "Guidance for Using Generative AI Tools and Building Applications." https://innowings.engg.hku.hk/aiguide/

[20] Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), 1-38.

[21] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10684-10695).

[22] Tam Wing Fan Innovation Wing, the University of Hong Kong, "Generative AI Walk-through Series #1 My Avatar (Oct 2023)." https://innowings.engg.hku.hk/genai_walkthrough/

[23] Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., ... & Scialom, T. (2023). Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

[24] Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., ... & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*.

[25] Huang, L., Hicks, D., Chen, X., Kao, B., & Wu, T-H. (2023, November 7). On the Use of Large Language Model and Legal Knowledge Base in Community Legal Education. *The Law via the Internet 2023 Conference*, Vienna, Austria.

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

[27] Schamber, L. (2000). Time-line interviews and inductive content analysis: their effectiveness for exploring cognitive behaviors. *Journal of the American society for Information Science*, *51*(8), 734-744.

* URL are retrieved on April 25, 2024