Planetary Boundaries Meet Future Engineers: An AI-Driven Exploration

Sanghyeon Han[1,2,*], Sruthi Viswanathan[1,3,*]

[1]The Spaceship Academy, United Kingdom
[2]Korea Advanced Institute of Science & Technology, South Korea
[3]University of Oxford, United Kingdom

Abstract

Our planet's resources are finite. Despite significant advancements in fields such as engineering, environmental considerations often lag behind. To address this gap, early education on sustainable practices is crucial. This article advocates for using accessible and scalable technology, such as conversational generative artificial intelligence (GenAI), to enhance university students' understanding of planetary boundaries. On pairing 10 engineering students with an initial prototype of a GenAI chatbot equipped with relevant knowledge, our findings suggest that even with limited initial understanding, students gained new insights, and developed further curiosity about planetary boundaries and their future careers. Building on these implications, we developed a multi-layered prompt architecture employing mixed-initiative interactions to proactively guide students' exploration. The efficacy of our approach was evaluated through an A/B test comparing first and second prototypes with 40 engineering students globally. Results demonstrated improvements in chatbot usability and technology acceptance. We propose further development of this technology, including a structured, domain-specific curriculum, supported by university initiatives, to foster sustainable thinking among future engineers.

Keywords

Planetary Boundaries, AI for Sustainability, AI for Education, Human-Centred AI, Interaction Layer

1. Introduction

While climate change dominates the headlines [1], there are actually nine planetary boundaries that are critical to maintaining Earth's stability [2]. These boundaries include climate change, biosphere integrity, land-system change, freshwater use, biochemical flows, ocean acidification, atmospheric aerosol loading, stratospheric ozone depletion, and novel entities. Alarmingly, by 2023, Earth has already crossed six of these boundaries (See Figure 1). Crossing these thresholds risks triggering irreversible and abrupt environmental changes, underscoring the urgent need for comprehensive sustainability efforts [3].

The nine planetary boundaries represent critical thresholds within which humanity can safely operate to maintain Earth's stability. Planetary boundaries include *(i) Climate change*, where excessive greenhouse gas emissions drive warming and extreme weather; *(ii) Biosphere integrity*, compromised by species extinction from habitat destruction and pollution; *(iii) Land-system change*, where deforestation and urbanisation reduce biodiversity and soil health; and *(iv) Freshwater use*, strained by over-extraction and contamination, leading to water scarcity and potential conflicts. *(v) Biochemical flows* from nitrogen and phosphorus disrupt ecosystems, creating toxic "dead zones," while *(vi) Ocean acidification* from CO_2 absorption threatens marine life. *(vii) Atmospheric aerosol loading* affects air quality and climate, *(viii) Stratospheric ozone depletion* weakens UV protection, and *(ix) Novel entities* like plastics and heavy metals accumulate, disrupting biological systems. Exceeding these boundaries risks irreversible environmental damage and undermines ecosystems essential for human survival.

However, an excessive emphasis on carbon emissions often overshadows other critical ecological thresholds, a phenomenon described as "carbon tunnel vision" [4]. Research and funding are disproportionately channelled towards climate change, leaving planetary boundaries such as biosphere integrity

STAI'24: International Workshop on Sustainable Transition with AI (Collocated with the 33rd International Joint Conference on Artificial Intelligence 2024), August 05, 2024, Jeju, Republic of Korea.

*Equal contribution.

sang@thespaceship.org (S. Han); sruthi@thespaceship.org (S. Viswanathan)

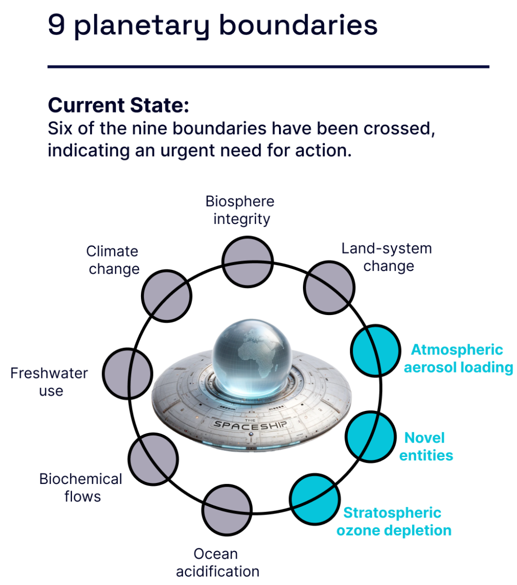

Figure 1: Illustration of the nine planetary boundaries, with unbreached boundaries in blue (adapted from [2]). Representing critical systems essential for stability the stability of 'Spaceship Earth,' (as coined by Buckminster Fuller [7]), for underscoring our planet's finite resources and collective responsibility for sustainability.

or biochemical flows underexplored [4]. This skewed focus risks partial solutions and unforeseen trade-offs, ultimately undermining broader sustainability goals [5]. A more holistic approach, allocating resources across all nine planetary boundaries, is essential to safeguard Earth's resilience [6].

Everyday, we have less margin before reaching these critical boundaries. We must urgently slow down and reverse our approach to prevent further damage and begin repairing our planet. To acknowledge, prevent the crossing of these boundaries, and to begin recovering the damage already done, education is essential. A Yale University survey shows that awareness of environmental sustainability among students rose from 21% in 2015 to 41% in 2020 [8]. While this signals progress, much work remains to equip the youth with the tools needed to make an impactful change in their respective fields.

Engineers, who play a crucial role in building our limitless future using our planet's limited resources, must recognize their duty not only to improve lives but also to sustain them [9]. From designing energy-efficient buildings to developing new forms of renewable energy, engineers are at the forefront of tackling environmental issues. For instance, the creation of carbon capture technologies has emerged as a promising solution to mitigate climate change by trapping CO2 emissions from industrial processes [10]. Additionally, engineers are involved in the design of green urban spaces, where natural ecosystems are integrated into city landscapes to combat biodiversity loss and reduce urban heat [11]. These examples illustrate how sustainable engineering principles are vital in addressing complex environmental challenges.

To equip future engineers for the pressing challenge of sustainability, we propose utilising generative artificial intelligence (GenAI), specifically in the form of a chatbot, to deliver accessible education and guidance on planetary boundaries. In pursuit of this objective, we developed an initial GenAI prototype designed to serve as an educational aid [12], with a targeted focus on empowering engineering students to understand and integrate planetary boundaries within their academic projects and career trajectories.

Despite the promising educational potential of GenAI models such as ChatGPT [13], their application is not without environmental cost, primarily due to the significant electricity demands associated with their operation [14]. Nevertheless, studies indicate that AI-driven interactions often enhance learning efficiency and engagement, surpassing traditional methods [15]. Empirical evidence further supports GenAI's efficacy as an educational tool, outperforming classic virtual content formats like videos; for instance, interactive virtual content groups scored significantly higher than online class groups [16]. Furthermore, GenAI-powered educational initiatives extend valuable resources to remote and underserved communities [17]. In contexts where educational benefits justify the environmental costs, GenAI represents a viable and powerful tool for advancing both educational outcomes and sustainability awareness.

Our research investigates the prospects of a generative artificial intelligence (GenAI) chatbot designed to educate engineering students about planetary boundaries and sustainable engineering practices. This study was conducted in two phases, each aimed at refining and evaluating the effectiveness of the AI agent in enhancing students' learning experience. In the first phase, we tested the initial version of our GenAI chatbot, referred to as Prototype I (see Section 3), with a small group of 10 engineering students in an in-person setting. This Large Language Model(LLM) based chatbot was prompted to provide detailed information related to planetary boundaries and sustainable engineering concepts. After the students interacted with the chatbot, we conducted semi-structured interviews to assess their understanding of the material presented and to evaluate the overall experience of the chatbot as an educational tool.

The feedback collected during this phase, analysed through qualitative thematic analysis [18], was instrumental in shaping our design. While all participants (10 out of 10) had not been introduced to the concept of planetary boundaries prior to our study, they nevertheless found the chatbot to be informative and easy to use, while highlighting several challenges that needed to be addressed. Specifically, many students expressed concerns regarding the flow of conversation, noting that it occasionally felt disjointed and lacking in personalisation. They indicated a desire for more specific examples and contextual information that would directly relate to their domain of study. Despite these challenges, our results suggest that the chatbot had successfully sparked a genuine interest in the topic of planetary boundaries among these engineering students, indicating its potential as a valuable educational resource.

Building on our findings from the first phase, for the second phase of our study, we made significant refinements to the design of the chatbot by implementing a human-centred interaction layer [19], built using initial user feedback, for a better alignment with user expectations. We then tested this Prototype II (see Section 4), with a larger cohort of 40 international engineering students recruited online via Prolific [20]. The interaction-layered prompt architecture included elements such as an introduction to planetary boundaries and their relevance to the students' field of study, the generation of a "planetary boundaries statement" tailored to the students' current or hypothetical projects, and discussions on sustainability in relation to their future careers, as well as the broader future scope for sustainable engineering. It featured specific examples tailored to the students' curriculum and included interactive elements designed to encourage active participation. We conducted an A/B test with 40 students, where 20 students interacted with Prototype I and the remaining 20 tested Prototype II. Prototype I was built using zero-shot prompting and retrieval-augmented generation (RAG), whereas Prototype II employed an interaction-layer prompt architecture with RAG. This between-subjects testing approach required each participant to engage with the chatbot for five minutes, after which they completed a feedback questionnaire set that also included the Chatbot Usability Questionnaire (CUQ) [21] and the Technology Acceptance Model (TAM) for education initiatives [22].

Qualitative analysis revealed an enhanced user satisfaction with the conversational flow of Prototype II, with participants expressing a desire for more multimodal and in-depth content. Among the 40 students who participated in this second phase of our study, only one reported prior knowledge of the nine planetary boundaries. Furthermore, the iterative comparison resulted in statistical significance in both CUQ and TAM, indicating improved and robust usability and technology acceptance between Prototype I and Prototype II.

We contribute empirical findings on the effectiveness of a GenAI chatbot in educating engineering students on planetary boundaries, gathered from a two-phase study involving 50 students. Through iterative testing, incorporation of user feedback, and the implementation of an interaction layer [19], we demonstrate significant improvements in user alignment, usability, and technology acceptance between prototypes, as measured by the Chatbot Usability Questionnaire (CUQ) and Technology Acceptance Model (TAM). Our discussion underscores the potential of AI-driven tools in higher education but also raises important human-centred and planet-centred considerations for a successful adoption of such technologies.

This study also highlights the importance of universities incorporating planetary boundaries education into engineering curricula, as evidenced by our finding that only 2% of participants (1 out of 50) had prior knowledge of this concept. Our results demonstrate that thoughtfully designed GenAI is a viable, accessible, and scalable tool to bridge this gap, fostering sustainability awareness and equipping future engineers to operate within the ecological limits of 'Spaceship Earth' [7]. With the technology at hand, it is high time to initiate this integration, allowing universities to play a pivotal role in preparing students to address global environmental challenges. This work sets the foundation for future research to refine and expand the use of GenAI in sustainability education, shaping a more environmentally conscious generation of professionals.

2. Related Works

2.1. GenAI for Education

With the recent advancements in generative artificial intelligence (GenAI), there has been a surge of interest in harnessing this technology across diverse fields. As foundation models capable of adapting to various use cases by learning from extensive datasets [23], language-generative models—such as ChatGPT [13]—have garnered particular attention within education as a means of stimulating curiosity and enhancing engagement among learners [24].

The growing prevalence of artificial intelligence in education has catalysed numerous studies examining the impact of chatbots on learning outcomes [25]. For instance, a systematic literature review by [26] analysed 74 publications on chatbots in education, revealing their potential to improve student engagement and personalise learning experiences. Similarly, [27] conducted a systematic review of 36 studies, highlighting that chatbots can effectively support learners by providing immediate feedback and fostering interactive learning environments.

Similar findings are observed across diverse educational contexts. Chatbot-assisted learning environments have been shown to increase student motivation and engagement in various subjects [28]. In another study, [29] demonstrated that chatbots could effectively provide academic counselling, guiding students in selecting elective courses—an outcome that highlights the adaptability and versatility of AI-driven educational tools.

Furthermore, the application of chatbots in specialised subject areas has proven promising. For example, using a chatbot to support Chinese language instruction yielded significant gains in language proficiency and learning success [30]. Such results underscore the potential of AI-driven models to make meaningful contributions in subject-specific education. Nevertheless, many studies caution that while chatbots can effectively complement traditional teaching methods, they often fall short as standalone resources, especially for complex subject matter [31]. Despite these limitations, the continual evolution of GenAI models like ChatGPT enhances their value in educational settings, particularly for fostering independent learning and delivering personalised, real-time feedback to students.

As generative AI technology advances, its integration into educational frameworks will become increasingly vital. The adaptability, accessibility, and potential for personalisation make GenAI an invaluable tool for contemporary education, one that complements traditional learning while expanding the possibilities for tailored, student-centred pedagogies.

2.2. AI for Sustainability

As humanity approaches critical planetary boundaries [2], the field of sustainability demands increased attention. The recent surge in deep learning and generative language models has spurred a wave of studies examining the potential of AI agents to foster sustainable behaviours, both at the individual and organisational levels.

One notable example is *AluxBot* [32], a chatbot designed to encourage pro-environmental behaviours among users. Through engaging, sustainability-focused conversations, the bot effectively raised environmental awareness, highlighting the capacity of chatbots to inspire and educate users on sustainability topics. Similarly, [33] explored the role of anthropomorphic design in chatbots as a tool to persuade users towards sustainable mobility practices. The study demonstrated a significant positive impact on users' beliefs about sustainability, further illustrating the potential of AI-driven chatbots to influence behaviour change in meaningful ways.

The impact of AI on sustainability also extends into organisational settings. The concept of *Green IS* underscores how AI technologies can support sustainable information systems, reducing energy consumption and enhancing operational efficiency [34]. In this vein, [35] discussed the role of digital technologies, including AI, in driving sustainability innovation by enabling businesses to implement eco-friendly practices. These findings suggest that AI applications are not limited to individual behaviour modification; they also equip organisations with powerful tools to achieve their environmental objectives.

Moreover, the integration of AI with gamification techniques has shown considerable effectiveness in promoting eco-friendly behaviours. Studies indicate that gamification elements, when combined with AI, can heighten user engagement and incentivise sustainability efforts within corporate settings [36]. Additionally, [37] outlined design principles for AI-driven business model development tools that prioritise sustainability, revealing AI's role in fostering innovation for sustainable organisational practices.

While the promise of AI in sustainability is robust, several studies recommend its use as a complement to traditional methods rather than as a complete replacement. For instance, although chatbots can offer real-time feedback and personalised suggestions to support sustainability initiatives, complex environmental challenges often necessitate human oversight [38]. Furthermore, the rapid digitalisation of sustainability practices presents both opportunities and risks, underscoring the need to align AI technologies with broader environmental and ethical considerations [39].

In summary, the application of AI in sustainability is a rapidly growing area of research, with mounting evidence of its potential to drive environmentally responsible behaviours and support organisational sustainability goals. As AI technologies continue to evolve, their role in shaping the future of sustainable development is likely to become even more pronounced, positioning AI as a transformative tool in advancing sustainability across multiple domains.

2.3. Human-Centred AI (HCAI)

Human-Centred AI (HCAI) prioritises the integration of human values, needs, and capabilities into the design and operation of AI systems, ensuring these technologies remain aligned with ethical standards and user-centric goals [40] [41] [42]. The literature delineates three core objectives of HCAI: understanding, controlling, and improving AI systems [19]. These objectives are grounded in cybernetic loop theory, which emphasises the iterative processes of sensing, processing, and feedback [43]. This theoretical framework underscores the adaptive nature of HCAI, aiming to develop systems that are not only intelligent and efficient but also transparent, controllable, and responsive to situational human needs.

Understanding: A fundamental aspect of HCAI is fostering a deeper comprehension of AI systems among users. The literature highlights Explainable AI (XAI) as a critical component in achieving this understanding [44]. Techniques such as Local Interpretable Model-agnostic Explanations (LIME) [45] and SHapley Additive exPlanations (SHAP) [46] have been instrumental in demystifying AI processes, rendering them more accessible and fostering trust in AI technologies. However, while effective in many scenarios, these methods often struggle with the inherent complexities of large language models (LLMs) and other advanced AI architectures [47]. This gap underscores the necessity for more interactive and user-centred explanatory frameworks, which can provide context-specific insights and foster deeper engagement with AI systems.

Control: The principle of control in HCAI is centred on empowering users to influence and guide AI systems effectively [48]. This involves the development of interfaces and tools that facilitate meaningful human-AI collaboration [41]. For example, recommender systems that allow users to customise their preferences illustrate how control can be operationalised, offering users the ability to tailor system outputs to their specific needs [49]. However, implementing such control mechanisms presents significant challenges, particularly in high-stakes domains such as healthcare or finance, where the implications of algorithmic decisions can be profound [50]. Balancing user autonomy with the reliability and safety of AI systems remains a critical area of research and development in HCAI.

Improvement: Continuous improvement in HCAI is facilitated through iterative feedback loops and learning processes. Techniques such as Human-in-the-Loop (HITL) learning [51] and Interactive Machine Learning (IML) [52] are pivotal in this context, enabling AI systems to adapt and evolve based on user input. These iterative methods not only enhance the performance and adaptability of AI systems but also ensure they remain aligned with evolving human values and ethical norms. This dynamic, symbiotic interaction between users and AI fosters trust and collaboration, ultimately contributing to more effective and user-centred AI applications.

In conclusion, the principles of Human-Centred AI are crucial for developing AI systems that are understandable, controllable, and continuously improving. By integrating user feedback and maintaining alignment with human values, HCAI not only enhances the adoptability and adaptability of AI systems but also promotes critical control and collaboration, ensuring these technologies serve the broader goals of society.

3. Study I: Experiment with Engineering Students

3.1. Prototype I

Our Prototype I, designed to disseminate knowledge on planetary boundaries to engineering students, was hosted as a Streamlit app [53] and developed using the GPT-4 turbo model from OpenAI [54]. The bot was specifically configured to provide actionable knowledge on planetary boundaries by prompting users to ask domain-specific questions and offering suggestions on potential actions they could take to address the challenges they inquired about. It was designed to engage users actively in problem-solving by being as detailed as possible in its responses. To support this, the bot was integrated with a dataset comprising eight key publications from the Stockholm Resilience Centre, which amounted to a total size of 2MB. These publications include foundational research on planetary boundaries, alongside policy reports and practice guides from the Centre's official website[1], ensuring a robust and concrete knowledge base.

In addition, the bot was designed to be responsive to various engineering disciplines by leveraging the general knowledge from GPT-4 turbo's broad training, which allows it to address discipline-specific questions in fields such as civil, environmental, and mechanical engineering. It not only conveyed theoretical insights but also encouraged students to reflect on the real-world applications of planetary boundaries within their fields of study, guiding them towards solutions based on their technical expertise.

3.2. Methodology I

We employed a mixed-method approach, incorporating both qualitative and quantitative elements, for our experimentation. Using an experience prototyping methodology [55], participants interacted with our bot to understand the concept of planetary boundaries and its implications for engineering. Initially, participants completed a preliminary questionnaire to assess their existing knowledge of planetary boundaries and to collect demographic information and informed consent. Following this, they were given a laptop with access to our Prototype I and engaged in a dialogue about planetary boundaries relevant to their specific areas of study. Participants were encouraged to have at least five exchanges with the bot. Upon completing this interaction, participants engaged in a semi-structured interview [56] with the lead author to debrief their experience and answered a follow-up survey to register their feedback on the bot's usability and usefulness, and their perspectives on planetary boundaries and their application to their future careers. Our objective was to evaluate any shifts in their opinions, assess the

[1]https://www.stockholmresilience.org/publications.html

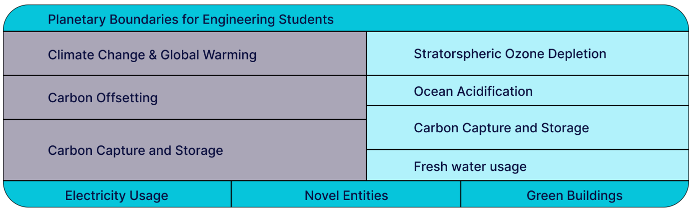

Figure 2: Topic Analysis of Engineering Students Conversations with our Bot on Planetary Boundaries

bot's educational utility, and explore its potential applications in broader contexts such as university curriculum, governmental policy, and academic research.

The experiment involved ten participants, with two joining virtually and eight participating in person. Each session lasted approximately thirty minutes. In addition to pre- and post-questionnaires, semi-structured interviews with the author were recorded, and participants' conversations with the bot were also stored. Prior approval was obtained from the internal ethics committee, and all data was anonymised and handled in accordance with GDPR guidelines [57]

3.3. Results I

We conducted a topic analysis [58] on participants' conversations with our bot to identify their primary queries regarding planetary boundaries and engineering (See Figure 2). Given that climate change and global warming were the most familiar subjects to participants prior to their interaction with the bot, it is unsurprising that these were the most discussed issues, with significant interest in subtopics like fusion energy and carbon capture were revealed by the topic analysis. Other key areas of interest include ocean acidification, stratospheric ozone depletion, fresh water usage, electricity usage, novel entities, and green buildings.

We conducted a thematic analysis [18] on the data collected from the two questionnaires and the interviews. The lead author, who conducted the experiments, and another author, who was not present during the experiments, independently coded the raw data. Through two meetings, the authors discussed and aligned on the emerging themes, which are listed below:

Lack of Prior Knowledge: None of the participants (0 out of 10) demonstrated prior familiarity with the concept of "planetary boundaries." However, they possessed foundational knowledge of related environmental issues, particularly climate change, one of the nine planetary boundaries. Although unaware of the broader planetary boundaries framework, participants' understanding of climate change and topics such as ozone depletion served as a valuable entry point for further learning. The chatbot was specifically designed to leverage and expand upon this prior knowledge, integrating their engineering background to contextualise sustainability challenges pertinent to their field. Following their interaction with the chatbot, participants reported an initial grasp of the measurable limits governing Earth's systems and expressed heightened curiosity and motivation to delve deeper into sustainability topics relevant to their academic and professional pursuits.

Educational Benefits and Need for Proactive Guidance: Most participants (9 out of 10) indicated that an educational bot like ours would be highly beneficial for diverse groups, ranging from elementary to university students, to learn about pressing issues such as planetary boundaries. Additionally, some participants (3 out of 10) suggested that the bot should be more proactive and structured in guiding the conversation by providing domain-specific examples and demonstrating its capabilities for users.

Demand for Domain-Specific Education Many participants (6 out of 10) could not prompt the bot to make domain-specific connections, highlighting the need for educational tools that empower and raise literacy across various fields. This underscores the global institutional need for education tailored to specific disciplines while promoting a comprehensive understanding of planetary boundaries.

Note on Participant Mode: No significant differences were observed between virtual and in-person participants in terms of engagement, learning outcomes, or chatbot usability.

4. Study II: Experiment with Engineering Students using Interaction Layer

4.1. Prototype II

In response to the feedback gathered from users in Section 3.3, we developed the second version of our chatbot, Prototype II, using the GPT-4o model, specifically addressing the issues faced by engineering students with: (i) lack of familiarity with the concept of planetary boundaries, (ii) need for proactive interaction by the chatbot, and (iii) domain-specific exemplification and discussion. These key enhancements were built via a structured interaction layer based on the "Understand, Control, and Improve" human-centred interaction model [19]. This version aimed to refine the bot's engagement by allowing it to guide users through structured educational exchanges, tailored project guidance, and future career discussions.

The interaction layer was designed with templates to ensure consistency and relevance across different stages of user interaction (see Figure 3 for prompt architecture). The bot begins by assessing the student's needs, offering initial explanations of planetary boundaries linked to their engineering field. It then provides options for students to control the direction of the conversation, ensuring the content remains pertinent to their studies. For instance, when guiding project design, the bot suggests domain-specific strategies, and concludes with a tailored planetary boundary statement for the user's project. The aim was to create a personalised and contextually relevant learning experience that directly connects planetary boundaries to the user's academic and professional future.

4.2. Methodology II

We recruited 40 engineering students globally via the Prolific platform. Participants were informed of the study's purpose through an information sheet, and informed consent was obtained. They were compensated $2.5 for their participation, with an average session lasting 20 minutes.

The study utilised an A/B testing approach to compare Prototype I and Prototype II. We did not alter Prototype I, except for changing its base model to OpenAI's GPT-4o. The study was structured as a between-subjects experiment: 20 participants interacted with Prototype I, and 20 with Prototype II. Among the 20 participants in Prototype I, the average age was 31.25 years, with 9 females, 11 identified as non-white, and 10 were residing in global South countries. For Prototype II, among the 20 participants, the average age was 27.95 years, 6 were female, 10 identified as non-white, and 9 were from the global South countries. Both prototypes were hosted via Streamlit [53], providing a similar user interface.

After interaction, participants completed an anonymous survey via Google Forms, offering subjective feedback on their experience. Additionally, they answered two qualitative questionnaires: the Chatbot Usability Questionnaire (CUQ) [21] to evaluate chatbot usability, and a adaptation of [22] Technology Acceptance Model (TAM) questionnaire. This was aimed at understanding their behavioural intention to use the chatbot for learning about planetary boundaries and integrating them into their engineering projects.In addition to pre- and post-questionnaires, semi-structured interviews with the author were recorded, and participants' conversations with the bot were also stored. Ethical approval was secured from the internal review board, and all data were anonymised and processed in compliance with GDPR regulations [57].

4.3. Results II

4.3.1. Qualitative Results

Through a qualitative analysis of the open-ended feedback collected from the A/B testing of Prototype I and Prototype II, significant improvements in user experience were observed. As highlighted in Section 3.3, the in-person testing of Prototype I, the online testing with a more advanced LLM model (GPT-4o), also revealed key issues such as lengthy and repetitive responses (reported by 13 out of 20 participants), lack of domain-specific information dissemination (reported by 16 out of 20 participants),

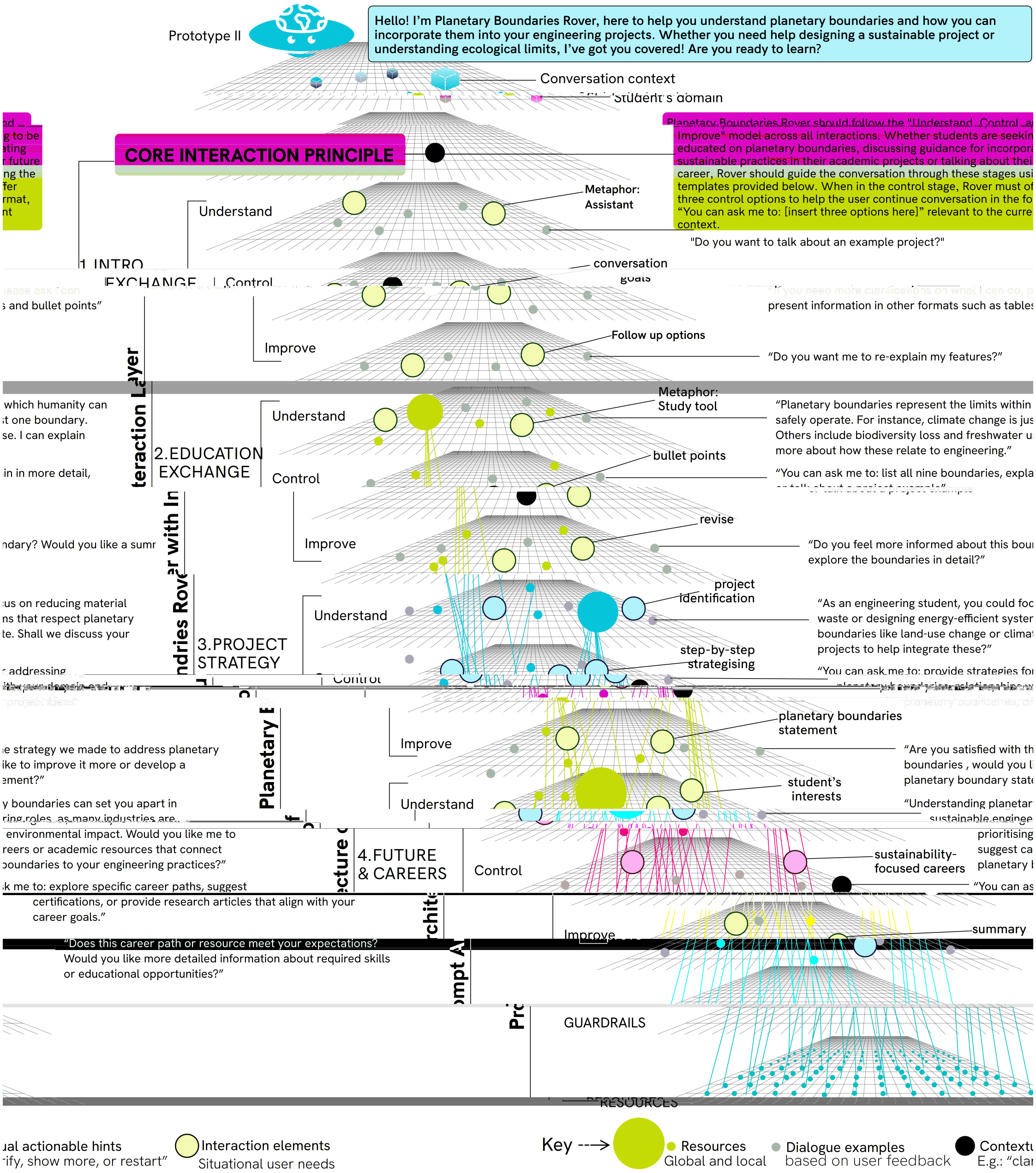

Figure 3: This figure (adapted from [19]) illustrates the interaction layer prompt architecture of Prototype II, showcasing how the chatbot guides user engagement through the core principle of "Understand, Control, and Improve," based on the user feedback received on testing Prototype I.

and a lack of dynamic conversational flow (reported by 18 out of 20 participants). In contrast, Prototype II demonstrated notable enhancements in responsiveness, conversational engagement, and adaptability, with users commending its more *human-like* and *insightful interactions* (reported by 15 out of 20 participants). Nevertheless, some participants continued to express a preference for shorter , more concise responses (reported by 7 out of 20 participants), and the integration of multimodal content (reported by 11 out of 20 participants). Across both prototype only 1 out of 40 participants (2.5%) reported familiarly with the concept of nine critical planetary boundaries. Overall, the qualitative findings suggest that the transition from Prototype I to Prototype II resulted in a refined user experience, better aligning with user expectations and significantly enhancing overall satisfaction.

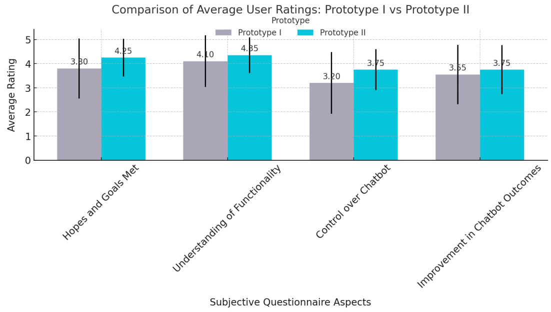

Figure 4: Comparison Of Average User Ratings: Prototype I vs Prototype II

4.3.2. Subjective Questionnaire

To evaluate the differences in user feedback between Prototype I and Prototype II, we conducted an independent sample t-test for each subjective questionnaire aspect: hopes and goals met, understanding of functionality, control over the chatbot, and improvement in chatbot outcomes. Although the average scores for Prototype II were more than that of Prototype I for all four questions (See Figure 4), results indicated that none of the differences were statistically significant. Specifically, the p-values ranged from 0.12 to 0.58 across the various aspects. Further iterations with larger sample sizes may be required to detect more subtle differences and validate the observed trends with respective to these subjective questions.

4.3.3. Chatbot Usability Questionnaire

The Chatbot Usability Questionnaire (CUQ) [21] was administered to assess and compare the usability of Prototype I and Prototype II (See Figure 5). 16 CUQ questions cover various aspects of user interaction, including the ease of use, clarity of responses, efficiency, user satisfaction, and the perceived usefulness of the chatbot in achieving its educational objectives. By systematically analysing user feedback, the CUQ provides insights into how well the chatbot facilitates learning and engagement, highlighting areas for improvement in both prototypes.

CUQ Scores Overview

- Prototype I had a mean CUQ score of 73.5 with a standard deviation of 14.7. The lowest score recorded was 37.5, while the highest was 96.9, and the median score was 76.6.
- Prototype II achieved a higher mean CUQ score of 83.4, with a reduced standard deviation of 12.8, indicating a more consistent user experience. The scores ranged from a minimum of 57.8 to a maximum of 100.0, with a median of 85.2.

T-Test Analysis An independent samples t-test was performed to determine if the difference in CUQ scores between the two prototypes was statistically significant. The t-test yielded a t-statistic of -2.26 and a p-value of 0.030, concluding that the difference in CUQ scores between Prototype I and Prototype II is statistically significant.

The significant increase in CUQ scores for Prototype II indicates that the design modifications and enhancements implemented based on feedback from Prototype I were effective in improving usability. This finding suggests that the usability improvements made in Prototype II had a meaningful impact on user experience.

4.3.4. Technology Acceptance Model

We adapted the Technology Acceptance Model (TAM) [22] to understand university students' behavioural intention towards our LLM-based chatbots to learn about planetary boundaries. It verifies various constructs such as Perceived Ease of Use (PE), Perceived Usefulness (PU), Attitude (AT), Behavioural Intention (BI), Self-Efficacy (SE), Subjective Norm (SN), and System Accessibility (SA), with

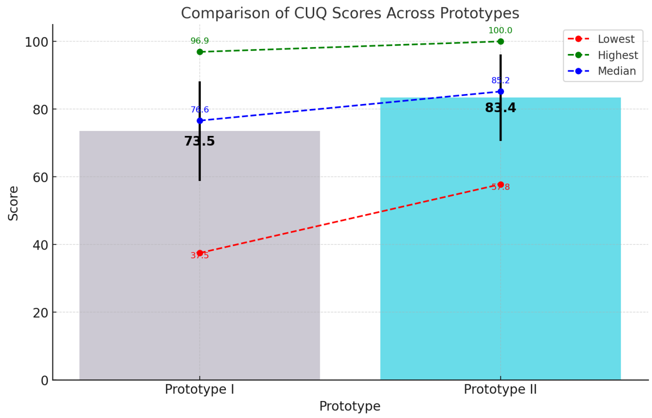

Figure 5: Comparison Of CUQ Scores Across Prototypes

17 questions answered by the users. A comprehensive statistical analysis was conducted to evaluate the differences between Prototype I and Prototype II based on our participants' responses. The methods applied included t-tests, ANOVA, and MANOVA, as visualised in Figure 6.

T-Test Analysis An Independent Sample t-Test was conducted to compare the mean scores of user responses across both prototypes. The results revealed a significant difference between the two prototypes, with a t-statistic of -4.60 and a p-value of 0.0001.

T-tests were also performed on individual questions to assess mean differences between prototypes. Statistically significant differences were observed in the following questions:

- **Q14** (Subjective Norm): $t = -2.3576$, $p = 0.0236$
- **Q15** (Subjective Norm): $t = -2.5317$, $p = 0.0156$
- **Q16** (Subjective Norm): $t = -2.6968$, $p = 0.0104$

These results indicate that user perceptions related to the chatbot's relevance and alignment with their work and societal expectations differ significantly between the prototypes.

ANOVA Analysis ANOVA was used to compare the prototypes across different sections of the questionnaire. Significant differences were observed in the following sections:

- **Behavioural Intention (BI)**: $F(2, 38) = 23.1477$, $p < 0.0001$
- **Self-Efficacy (SE)**: $F(2, 38) = 22.5068$, $p < 0.0001$

Tukey Post-Hoc Analysis confirmed these differences: BI (Mean Difference $= 0.6$, $p = 0.0212$) and SE (Mean Difference $= 0.475$, $p = 0.0267$), indicating that the prototypes differ significantly in these areas.

These findings demonstrate that the prototypes differ significantly in terms of users' perceived intention to use the chatbot in future and their belief in their ability to successfully using it.

MANOVA Analysis A MANOVA was conducted to explore the overall differences between the prototypes across all sections simultaneously:

- **Wilks' Lambda**: $\Lambda = 0.2187$, $F(17, 23) = 4.8344$, $p = 0.0003$
- **Pillai's Trace**: $V = 0.7813$, $F(17, 23) = 4.8344$, $p = 0.0003$
- **Hotelling-Lawley Trace**: $T = 3.5732$, $F(17, 23) = 4.8344$, $p = 0.0003$

These multivariate tests confirm significant overall differences between the two prototypes.

Section-Specific MANOVA Further breakdown using MANOVA on individual sections with multiple questions also confirmed significant differences:

- **Perceived Ease of Use (PE)**: $\Lambda = 0.4599$, $F(3, 37) = 14.4870$, $p < 0.0001$
- **Perceived Usefulness (PU)**: $\Lambda = 0.4716$, $F(3, 37) = 13.8195$, $p < 0.0001$
- **Attitude (AT)**: $\Lambda = 0.4652$, $F(3, 37) = 14.1814$, $p < 0.0001$

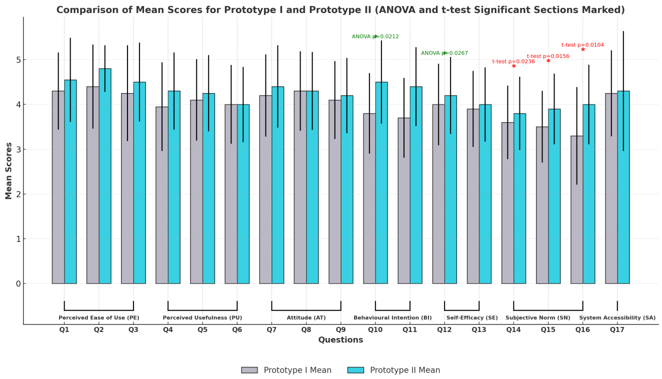

Figure 6: Comparison of Technology Acceptance Model (TAM) Mean Scores for Prototype I and Prototype II

- **Behavioural Intention (BI)**: $\Lambda = 0.4508, F(2, 38) = 23.1477, p < 0.0001$
- **Self-Efficacy (SE)**: $\Lambda = 0.4578, F(2, 38) = 22.5068, p < 0.0001$
- **Subjective Norm (SN)**: $\Lambda = 0.4249, F(3, 37) = 16.6915, p < 0.0001$

These results indicate that the prototypes vary significantly across all these sections, reinforcing the findings from the ANOVA and t-tests.

The statistical analyses provide robust evidence of significant differences between Prototype I and Prototype II across TAM. These insights support further refinement and evaluation of the chatbot prototypes based on the identified areas of variance.

5. Discussion

5.1. Enhancing Sustainability Education through GenAI Chatbots

The findings of our research study reveal a substantial knowledge gap among engineering students concerning planetary boundaries. While many participants demonstrated foundational awareness of environmental issues such as climate change and ozone depletion, only 1 out of 50 (2%) was familiar with the specific framework of planetary boundaries. This gap underscores the urgent need for targeted educational interventions to bridge this divide. The outcome of our experiments, where students' curiosity piqued despite the knowledge gap, as well as high CUQ and TAM scores, showcases Generative AI to offer a promising avenue for addressing this challenge.

The iterative improvements from Prototype I to Prototype II demonstrated significant advancements in delivering educational content more effectively. Applying Human-Centred AI (HCAI) principles by depolying the interaction layer in our prompt architecture was crucial in designing a chatbot that promotes active user engagement. Feedback from users highlighted the importance of responsiveness and the need for a conversational flow that feels natural and engaging. The balance between structured goals and personalised interaction was pivotal in enhancing the learning experience, making the chatbot more effective and relatable.

User feedback highlighted the critical need for personalisation and contextualisation in GenAI-powered educational tools, stressing that a one-size-fits-all approach fails to effectively cater to the diverse learning needs and backgrounds of students, thereby limiting engagement and educational impact. By aligning educational guidance with the specific needs of the students and concepts of their field, as we had attempted with our Prototype II in Section 4.1, the GenAI bot could foster a deeper comprehension of sustainability's intersection with diverse fields, thereby empowering students to embed these principles into their academic projects and future professional practices.

Future design iterations should focus on further refining these elements, ensuring that the chatbot remains a robust educational aid. Challenges such as overly lengthy responses and the need for multimodal content, including visuals and concise summaries, should also be noted as areas for future improvement. Overall, the chatbot's ability to adapt to different educational contexts and student needs underscores its potential as a valuable tool in sustainability education. By leveraging GenAI's capability to deliver complex sustainability concepts in an accessible and engaging manner, educators can foster deeper understanding and stimulate proactive engagement with critical environmental issues.

5.2. Sustainability of the Tool

The environmental impact of AI technologies, particularly large models, underscores the need for strategies to minimise their carbon footprint. Throughout the development of our prototypes, we were mindful of these considerations, striving to balance performance with environmental responsibility. We utilised pre-trained models, which significantly reduced the computational cost and energy consumption associated with training large models from scratch. Additionally, we designed meaningful interactions to deliver precise and relevant content, ensuring efficiency in both user engagement and resource utilisation. Furthermore, we optimised the chatbot's performance by limiting unnecessary computations and streamlining the dialogue flow to minimise processing time. Leveraging OpenAI's Agent model allowed us to dynamically allocate resources, scaling usage based on real-time demand, which helped to avoid over-provisioning and reduce energy waste.

In future iterations, we aim to further enhance sustainability by incorporating more lightweight models with fewer parameters. These models can maintain educational effectiveness while reducing computational demands and energy consumption, aligning with broader environmental goals. Additionally, we plan to explore hosting the GenAI bot on cloud platforms powered by renewable energy. This shift can significantly decrease the environmental impact of its operation, serving as a model for responsible AI use in education. Future research will also focus on algorithmic optimisation techniques such as model pruning, quantisation, and knowledge distillation to improve energy efficiency without compromising the chatbot's performance. By integrating these sustainability measures, AI-driven educational tools like the GenAI bot can effectively educate students on sustainability while exemplifying sustainable practices, preparing future population to address complex environmental challenges responsibly.

6. Limitations and Future Work

While this research highlights the potential of AI-driven tools in higher education, it also sustains critical limitations that must be addressed to fully realise their benefits. A primary limitation is the modest sample size of 50 engineering students, which may constrain the generalisability of the findings across broader educational contexts. A key limitation of this study is that the effectiveness of the GenAI chatbot has not yet been compared with traditional learning methods, such as lecture-based instruction and textbook-driven learning. Future research should conduct controlled comparative studies to assess whether AI-driven interactions offer measurable advantages over conventional educational approaches in fostering planetary boundary awareness among engineering students. Another key limitation lies managing the continuous evaluation and iterative adaptation of AI systems to ensure they align with the evolving needs of students and adhere to pedagogical best practices. The study revealed significant challenges, notably the need to maintain the accuracy and contextual relevance of chatbot responses, underscoring the complexities inherent in integrating AI within educational frameworks.

Future research should investigate the scalability of the GenAI chatbot across a broader spectrum of educational environments and disciplines. Such exploration could uncover its utility in addressing a wider range of sustainability issues beyond the scope of planetary boundaries. While our study primarily focused on user experience, usability, and technology acceptance, an equally crucial aspect in educational contexts is knowledge retention. The educational efficacy of the chatbot, specifically its impact on long-term learning and comprehension, remains to be validated in future research. By leveraging these opportunities, educators can harness AI's capacity to deepen students' understanding of critical global challenges, ultimately fostering a generation of engineers equipped to make informed, sustainable decisions in their professional and academic pursuits.

7. Conclusion

To address the critical educational gaps in sustainable practices, especially among engineering students poised to influence future environmental outcomes, we explore the potential of generative artificial intelligence (GenAI) to foster a deeper understanding of planetary boundaries. Our findings suggest that an AI-driven chatbot, when iteratively refined with user feedback and human-centred interaction design, significantly enhances engineering students' comprehension of planetary boundaries, improves usability and engagement, and fosters a stronger inclination towards integrating sustainability principles in their academic and professional pursuits. Through our efforts, we earnestly hope that educational institutions will take decisive action by embracing this AI-driven approach, empowering the next generation of engineers with the knowledge and skills needed to navigate and uphold the ecological boundaries of 'Spaceship Earth.'

Acknowledgments

We thank our team at The Spaceship Academy for their valuable feedback on earlier versions of this project. We are also grateful to our participants, the 50 engineering students, for their time and engagement, which contributed significantly to our research.

References

[1] K. Abbass, M. Z. Qasim, H. Song, M. Murshed, H. Mahmood, I. Younis, A review of the global climate change impacts, adaptation, and sustainable mitigation measures, Environmental Science and Pollution Research 29 (2022) 42539–42559. doi:https://doi.org/10.1007/s11356-022-19718-6.

[2] K. Richardson, W. Steffen, W. Lucht, J. Bendtsen, S. E. Cornell, J. F. Donges, M. Drüke, I. Fetzer, G. Bala, W. von Bloh, G. Feulner, S. Fiedler, D. Gerten, T. Gleeson, M. Hofmann, W. Huiskamp, M. Kummu, C. Mohan, D. Nogués-Bravo, S. Petri, M. Porkka, S. Rahmstorf, S. Schaphoff, K. Thonicke, A. Tobian, V. Virkki, L. Wang-Erlandsson, L. Weber, J. Rockström, Earth beyond six of nine planetary boundaries, Science Advances 9 (2023) eadh2458. doi:10.1126/sciadv.adh2458.

[3] S. J. Lade, W. Steffen, W. De Vries, S. R. Carpenter, J. F. Donges, D. Gerten, H. Hoff, T. Newbold, K. Richardson, J. Rockström, Human impacts on planetary boundaries amplified by earth system interactions, Nature sustainability 3 (2020) 119–128. doi:https://doi.org/10.1038/s41893-019-0454-4.

[4] M. Savasta-Kennedy, The dangers of carbon reduction tunnel vision, 2014. URL: https://ssrn.com/abstract=2518613.

[5] P. Sutoris, The problem with decarbonization tunnel vision, *Bulletin of the Atomic Scientists*, 18 April 2024, 2024. URL: https://thebulletin.org/2024/04/the-problem-with-decarbonization-tunnel-vision/, accessed: 24-Jan-2025.

[6] W. Steffen, K. Richardson, J. Rockström, S. E. Cornell, I. Fetzer, E. M. Bennett, R. Biggs, S. R. Carpenter, W. De Vries, C. A. De Wit, et al., Planetary boundaries: Guiding human development on a changing planet, Science 347 (2015) 1259855. doi:10.1126/science.1259855.

[7] R. Fuller, J. Snyder, Operating Manual for Spaceship Earth, Lars Müller Publishers, 2008. URL: https://books.google.co.uk/books?id=Ux3JwAEACAAJ.

[8] T. Cort, K. Gilbert, S. Decew, M. Goldberg, E. Wilkinson, H. Fitzgerald, Rising Leaders on Social and Environmental Sustainability, Yale University and Global Network for Advanced Management. New Haven, CT: Yale Center for Business and the Environment, Yale Program on Climate Change Communication., 2022. URL: https://climatecommunication.yale.edu/publications/environmental-beliefs-among-business-students-around-the-world/.

[9] C. Boyle, Considerations on educating engineers in sustainability, International Journal of Sustainability in Higher Education 5 (2004) 147–155. doi:10.1108/14676370410526233.

[10] T. Wilberforce, A. Olabi, E. T. Sayed, K. Elsaid, M. A. Abdelkareem, Progress in carbon capture

technologies, Science of The Total Environment 761 (2021) 143203. doi:https://doi.org/10.1016/j.scitotenv.2020.143203.
[11] T. Semeraro, A. Scarano, R. Buccolieri, A. Santino, E. Aarrevaara, Planning of urban green spaces: An ecological perspective on human benefits, Land 10 (2021) 105. doi:10.3390/land10020105.
[12] Q. Li, L. Fu, W. Zhang, X. Chen, J. Yu, W. Xia, W. Zhang, R. Tang, Y. Yu, Adapting large language models for education: Foundational capabilities, potentials, and challenges, arXiv preprint arXiv:2401.08664 (2024). URL: https://arxiv.org/abs/2401.08664.
[13] OpenAI, Chatgpt: Chat generative pre-trained transformer, 2023. URL: https://openai.com/chatgpt.
[14] J. An, W. Ding, C. Lin, Chatgpt: tackle the growing carbon footprint of generative ai, Nature 615 (2023) 586–586. doi:10.1038/d41586-023-00843-2.
[15] C.-N. Bodea, M.-I. Dascalu, A. Hang, Chatbot-based training for project management: Another way of corporate training or a must-have tool for sustainable education?, in: R. Cuevas, C.-N. Bodea, P. Torres-Lima (Eds.), Research on Project, Programme and Portfolio Management, Springer International Publishing, Cham, 2021, pp. 249–259. doi:10.1007/978-3-030-60139-3_17.
[16] A. Vakilian, E. Z. Ranjbar, M. Hassanipour, H. Ahmadinia, H. Hasani, The effectiveness of virtual interactive video in comparison with online classroom in the stroke topic of theoretical neurology in covid-19 pandemic, Journal of Education and Health Promotion 11 (2022) 219. doi:10.4103/jehp.jehp_1297_21.
[17] K. Meli, I. Taouki, D. Pantazatos, Empowering educators with generative ai: The genai education frontier initiative, in: EDULEARN24 Proceedings, IATED, 2024, pp. 4289–4299. doi:10.21125/edulearn.2024.1077.
[18] V. Braun, V. Clarke, Using thematic analysis in psychology, Qualitative Research in Psychology 3 (2006) 77–101. doi:10.1191/1478088706qp063oa.
[19] S. Viswanathan, S. Ibrahim, R. Shankar, R. Binns, M. V. Kleek, P. Slovak, The interaction layer: An exploration for co-designing user-llm interactions in parental wellbeing support systems, in: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, CHI '25, Association for Computing Machinery, New York, NY, USA, 2025. doi:10.1145/3706598.3714088.
[20] Prolific, Online participant recruitment for surveys and market research, 2014. URL: https://www.prolific.co.
[21] S. Holmes, A. Moorhead, R. Bond, H. Zheng, V. Coates, M. McTear, Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces?, in: Proceedings of the 31st European Conference on Cognitive Ergonomics (ECCE 2019), ACM, New York, NY, USA, 2019, pp. 207–214. doi:10.1145/3335082.3335094.
[22] S. Y. Park, An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning, Journal of Educational Technology & Society 12 (2009) 150–162. URL: https://www.jstor.org/stable/10.2307/jeductechsoci.12.3.150.
[23] R. Bommasani, D. A. Hudson, E. Adeli, R. B. Altman, S. Arora, S. von Arx, M. S. Bernstein, J. Bohg, A. Bosselut, E. Brunskill, E. Brynjolfsson, S. Buch, D. Card, R. Castellon, N. S. Chatterji, A. S. Chen, K. Creel, J. Q. Davis, D. Demszky, C. Donahue, M. Doumbouya, E. Durmus, S. Ermon, J. Etchemendy, K. Ethayarajh, L. Fei-Fei, C. Finn, T. Gale, L. Gillespie, K. Goel, N. D. Goodman, S. Grossman, N. Guha, T. Hashimoto, P. Henderson, J. Hewitt, D. E. Ho, J. Hong, K. Hsu, J. Huang, T. Icard, S. Jain, D. Jurafsky, P. Kalluri, S. Karamcheti, G. Keeling, F. Khani, O. Khattab, P. W. Koh, M. S. Krass, R. Krishna, R. Kuditipudi, et al., On the opportunities and risks of foundation models, CoRR abs/2108.07258 (2021). URL: https://arxiv.org/abs/2108.07258. arXiv:2108.07258.
[24] B. Alhafni, S. Vajjala, S. Bannò, K. K. Maurya, E. Kochmar, LLMs in Education: Novel Perspectives, Challenges, and Opportunities, arXiv preprint arXiv:2409.11917 (2024). URL: https://arxiv.org/abs/2409.11917.
[25] M. Bernabei, S. Colabianchi, A. Falegnami, F. Costantino, Students' use of large language models in engineering education: A case study on technology acceptance, perceptions, efficacy, and detection chances, Computers and Education: Artificial Intelligence 5 (2023) 100172. doi:10.1016/j.caeai.2023.100172.
[26] S. Wollny, J. Schneider, D. D. Mitri, J. Weidlich, M. Rittberger, H. Drachsler, Are we there yet?–a

systematic literature review on chatbots in education, Frontiers in Artificial Intelligence 4 (2021) 654924. doi:10.3389/frai.2021.654924.
[27] M. A. Kuhail, N. Alturki, S. Alramlawi, K. Alhejori, Interacting with educational chatbots: A systematic review, Education and Information Technologies 28 (2023) 973–1018. doi:10.1007/s10639-022-11177-3.
[28] D. Arruda, M. Marinho, E. Souza, F. Wanderley, A chatbot for goal-oriented requirements modeling, in: S. Misra, O. Gervasi, B. Murgante, E. Stankova, V. Korkhov, C. Torre, A. M. A. Rocha, D. Taniar, B. O. Apduhan, E. Tarantino (Eds.), Computational Science and Its Applications – ICCSA 2019, Springer International Publishing, Cham, 2019, pp. 506–519. doi:10.1007/978-3-030-24305-0_38.
[29] C. Chun Ho, H. L. Lee, W. K. Lo, K. F. A. Lui, Developing a chatbot for college student programme advisement, in: 2018 International Symposium on Educational Technology (ISET), 2018, pp. 52–56. doi:10.1109/ISET.2018.00021.
[30] H.-L. Chen, G. V. Widarso, H. Sutrisno, A chatbot for learning chinese: Learning achievement and technology acceptance, Journal of Educational Computing Research 58 (2020) 1161–1189. doi:10.1177/0735633120929622.
[31] J. Yin, T.-T. Goh, B. Yang, Y. Xiaobin, Conversation technology with micro-learning: The impact of chatbot-based learning on students' learning motivation and performance, Journal of Educational Computing Research 59 (2021) 154–177. doi:10.1177/0735633120952067. arXiv:https://doi.org/10.1177/0735633120952067.
[32] J. Peniche-Avilés, C. Miranda-Palma, L. Narváez-Díaz, E. Llanes-Castro, Aluxbot: A chatbot that encourages the care for the environment, International Journal of Computer Science Issues (IJCSI) 13 (2016) 120. URL: http://ijcsi.org/papers/IJCSI-13-1-120-125.pdf.
[33] S. Diederich, S. Lichtenberg, A. B. Brendel, S. Trang, Promoting sustainable mobility beliefs with persuasive and anthropomorphic design: Insights from an experiment with a conversational agent, in: Proceedings of the 40th International Conference on Information Systems (ICIS), Association for Information Systems, Munich, Germany, 2019, pp. 1–17. URL: https://aisel.aisnet.org/icis2019/sustainable_is/sustainable_is/3/.
[34] J. Dedrick, Green is: Concepts and issues for information systems research, Communications of the Association for Information Systems 27 (2010) 11. doi:10.17705/1CAIS.02711.
[35] S. Schneider, The impacts of digital technologies on innovating for sustainability (2019) 415–433. doi:10.1007/978-3-319-97385-2_22.
[36] J. Hamari, Do badges increase user activity? a field experiment on the effects of gamification, Computers in Human Behavior 71 (2017) 469–478. doi:10.1016/j.chb.2015.03.036.
[37] T. Schoormann, M. Stadtländer, R. Knackstedt, Designing business model development tools for sustainability—a design science study, Electronic Markets 32 (2022) 645–667. doi:10.1007/s12525-021-00466-3.
[38] P. Seele, I. Lock, The game-changing potential of digitalization for sustainability: possibilities, perils, and pathways, Sustainability Science 12 (2017) 183–185. URL: https://api.semanticscholar.org/CorpusID:157219180.
[39] P. Seele, I. Lock, The game-changing potential of digitalization for sustainability: Possibilities, perils, and pathways, Sustainability Science 12 (2017) 183–185. doi:10.1007/s11625-016-0389-0.
[40] T. Capel, M. Brereton, What is human-centered about human-centered ai? a map of the research landscape, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23, Association for Computing Machinery, New York, NY, USA, 2023. doi:10.1145/3544548.3580959.
[41] B. Shneiderman, Human-Centered AI, Oxford University Press, Oxford, 2022. URL: https://global.oup.com/academic/product/human-centered-ai-9780192845290.
[42] M. Tahaei, M. Constantinides, D. Quercia, M. Muller, A systematic literature review of human-centered, ethical, and responsible ai, 2023. URL: https://arxiv.org/abs/2302.05284.
[43] N. Wiener, Cybernetics: or Control and Communication in the Animal and the Machine, 2nd ed., The MIT Press, Cambridge, Massachusetts, 1961. URL: https://mitpress.mit.edu/9780262730099/

cybernetics/, fourth printing, 1985; First MIT Press paperback edition, February 1965.
[44] A. Adadi, M. Berrada, Peeking inside the black-box: A survey on explainable artificial intelligence (xai), IEEE Access 6 (2018) 52138–52160. doi:10.1109/ACCESS.2018.2870052.
[45] M. T. Ribeiro, S. Singh, C. Guestrin, "why should i trust you?": Explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, Association for Computing Machinery, New York, NY, USA, 2016, p. 1135–1144. doi:10.1145/2939672.2939778.
[46] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Neural Information Processing Systems, 2017. URL: https://api.semanticscholar.org/CorpusID:21889700.
[47] K. G. Barman, N. Wood, P. Pawlowski, Beyond transparency and explainability: On the need for adequate and contextualized user guidelines for llm use, Ethics and Information Technology 26 (2024) 47. doi:10.1007/s10676-024-09778-2.
[48] A. Schmidt, Interactive human centered artificial intelligence: A definition and research challenges, in: Proceedings of the International Conference on Advanced Visual Interfaces, AVI '20, Association for Computing Machinery, New York, NY, USA, 2020. doi:10.1145/3399715.3400873.
[49] S. Viswanathan, C. Boulard, A. Bruyat, A. Maria Grasso, Situational recommender: Are you on the spot, refining plans, or just bored?, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22, Association for Computing Machinery, New York, NY, USA, 2022. doi:10.1145/3491102.3501909.
[50] P. D. König, Challenges in enabling user control over algorithm-based services, AI & Society 39 (2024) 195–205. doi:10.1007/s00146-022-01395-1.
[51] X. Wu, L. Xiao, Y. Sun, J. Zhang, T. Ma, L. He, A survey of human-in-the-loop for machine learning, Future Generation Computer Systems 135 (2022). doi:10.1016/j.future.2022.05.014.
[52] J. J. Dudley, P. O. Kristensson, A review of user interface design for interactive machine learning, ACM Trans. Interact. Intell. Syst. 8 (2018). doi:10.1145/3185517.
[53] Streamlit, The fastest way to build data apps, 2019. URL: https://streamlit.io.
[54] OpenAI, Gpt-4: Generative pre-trained transformer 4, 2023. URL: https://openai.com.
[55] M. Buchenau, J. F. Suri, Experience prototyping, in: Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, ACM, New York, NY, USA, 2000, pp. 424–433. doi:10.1145/347642.347802.
[56] O. A. Adeoye-Olatunde, N. L. Olenik, Research and scholarly methods: Semi-structured interviews, Journal of the American College of Clinical Pharmacy 4 (2021) 1358–1367. doi:10.1002/jac5.1441.
[57] E. Union, General data protection regulation (gdpr), 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj.
[58] H. Li, K. Yamanishi, Topic analysis using a finite mixture model, Information Processing Management 39 (2003) 521–541. doi:10.1016/S0306-4573(02)00068-3.