CERN IdeaSquare Journal of Experimental Innovation, 2025; 9(3): 53-56
DOI: https://doi.org/10.23726/cij.2025.1799

DATASET

Data from Students' Use of Generative AI in Innovation Education Programs at CERN IdeaSquare

Mikko Sairanen,[1*] Jimmy Poulaillon,[1] Lauri Valtonen,[2] Giulia Gaddi,[1] Catarina Batista[1], Ole Werner[1], Matteo Balocco[3], Andrea Bruno[4], Diego Canova Tandaypan[3], Ángel Costa-Perales[5], Jacopo Dardini[4], Ngo Hung Dung[4], Anastasiia Korovina[4], Fernando Palma Fanjul[6], Eduard Puga Creus[7], Filomena Scirocco[4], Adrián Toledo Martínez[5], Javier Pérez Trujillo[5]

[1]CERN IdeaSquare, Rte de Meyrin 372, CH-1211 Geneva 23, Switzerland; [2]Department of Computer Science, Aalto University, Konemiehentie 2, 02150 Espoo, Finland; [3]ESADE Business School, Av. Pedralbes, 60-62, 08034 Barcelona, Spain; [4]University of Bologna, Via Zamboni 33, 40126 Bologna, Italy; [5]Universitat Politècnica de Valencia, Camino de Vera, s/n, 46022 Valencia, Spain; [6]Universidad del Bio Bio, Avda. Andrés Bello 720, Chillán, Chile; Universitat Politècnica de Catalunya, Avda. Diagonal 647, 08028, Barcelona, Spain.
*Corresponding author: mikko.sairanen@cern.ch

ABSTRACT

This article presents a dataset comprising Master-level students' prompts and discussions with generative artificial intelligence (AI) as part of a challenge-based innovation problem-solving process. Openly accessible through Zenodo (https://doi.org/10.5281/zenodo.17857679), the data consists of 466 pages of students' conversations with generative AI tools and a 28-page extract comprising only the prompts. The data was collected through a survey for students participating in week-long intense innovation education programs at CERN IdeaSquare throughout 2024. The data can be used to advance practice and theory regarding the use of generative AI in problem solving and innovation education. Scholars are encouraged to use the data in various forms of analysis, including thematic, content, and narrative analysis. The data can also complement other similar datasets in larger studies to examine the effects of different pedagogical approaches, student profiles, or problem-solving situations on the use of generative AI.

Keywords: Generative AI; prompt; innovation education.

Received: December 2025. Accepted: December 2025.

INTRODUCTION

This article introduces an openly accessible dataset comprising prompts and conversations that seven Master-level student teams (in total 34 students) had with generative artificial intelligence (AI) chatbot tools as part of challenge-based innovation assignment. The data was collected from volunteering students during several weeklong educational programs focusing on transformative innovation at the collaborative innovation space IdeaSquare at CERN. The purpose of this dataset is to facilitate research and improve understanding in the intersection of generative AI, innovation, and education.

The dataset is openly accessible in Zenodo (https://doi.org/10.5281/zenodo.17857679) and contains both the complete conversations with generative AI chatbots (466 pages) and a stripped version containing only the students' prompts (28 pages). The data touches on various aspects of complex sociotechnical problem-solving, such as food, water, energy system design, and population growth modelling. The data showcases a variety of approaches to utilizing generative AI, including calculations, iterative code generation, tool building, reference mapping, image generation, scenario analysis, and many others.

This data can advance the understanding of the role of generative AI in problem solving in general and in innovation education in particular. For example, it can be used to examine effective ways of utilizing generative AI to increase creativity (see Habib et al., 2024; Mariani & Dwivedi, 2024) or to clarify the usefulness and role of AI-generated ideas, contributing to an active academic debate (see Boussioux et al., 2024; Rafner et al., 2023). The data can help AI, innovation, and education researchers to discover patterns in the usage of generative AI that could indicate either opportunities or potential threats for learning and genuinely innovative problem-solving. Such findings could enable the development of more effective pedagogical approaches or innovation processes, indicating high practical contribution potential.

THEORETICAL BACKGROUND

The utility of generative AI in innovation processes and creative problem solving is widely recognized

(Mariani & Dwivedi, 2024). Many recent studies confirm that generative AI can have multiple benefits in this domain, including the generation of more viable (Boussioux et al., 2024), flexible, and better elaborated ideas (Habib et al., 2024) and increased ease of the problem-solving process (Urban et al., 2024). However, the novelty of AI-generated ideas compared to human-generated ones is still contested (see Boussioux et al., 2024; Rafner et al., 2023). Many negative effects, such as decreased creative confidence (Habib et al., 2024) and reduced idea diversity (Doshi & Hauser, 2024), are also emerging.

Consequently, there is an active and ever-developing academic discussion regarding the use of generative AI in education. Besides its potential for boosting creativity, generative AI could enable highly personalized learning experiences and increase student engagement (Kadaruddin, 2023; Mittal et al., 2024). However, if misused, it can erode students' critical thinking skills and introduce biases (Dwivedi et al., 2023; Helal et al., 2025; Larson et al., 2024; Mittal et al., 2024). Caution should therefore be practiced upon adopting generative AI tools in educational programmes, including in higher education (Habib et al., 2024). More empirical data and research are needed to understand how the potential of AI for problem-solving and education can be harnessed while its potential harmful effects are carefully mitigated.

DATA DESCRIPTION AND METHODOLOGY

The dataset is linked to two educational program concepts hosted at IdeaSquare at CERN. The first one is IdeaSquare Planet (I2P), a transformative innovation education program where a multidisciplinary group of students are first tasked with completing a complex and multidimensional mission related to an imaginary exoplanet and, second, with applying their learnings to generate solutions to a certain wicked problem back at Earth conditions (Valtonen et al., 2025). The second program concept is Challenge-Based Innovation (CBI), in which student teams must approach a predetermined societal challenge with multidisciplinary innovative solutions, often based on CERN-originated technologies (Nordberg et al., 2024; Vignoli et al., 2021). These programs aim to foster creativity and systems thinking skills in future changemakers by provoking innovation beyond conventional societal assumptions. Each program implementation was designed in collaboration with participating partner universities.

During the programs, students were presented with the opportunity to upload their conversations with generative AI tools through a questionnaire in QUALTRICS. Students were informed about this possibility at the end of the weeklong program, although in some cases the teachers additionally noted this in the beginning of the program. This was done to ensure that interested students would not delete their conversations during the week. All responses were submitted either at the end of the program or shortly after, never in the middle of the problem-solving process. Thus, the collected data should comprise the student teams' entire use of generative AI during the innovation education program, although there was no practical way to verify this since the data was self-reported and produced on a voluntary basis.

Upon submitting their responses, the students gave their consent to the open access publication of the anonymized dataset and were informed of the opportunity to co-author this related data publication. In total, we gathered responses from seven student teams and in total 34 students across three I2P programs and one CBI program during 2024. Even though the generative AI tools were not pre-specified nor restricted in any way, all the conversations we collected were held in ChatGPT.

The processing of the data consisted of three main steps. First, all data was anonymized, i.e., any prompts including personal information were eliminated from the data. Second, the responses were compiled into a single .docx file, structured according to student teams (anonymized by numbering), their prompts, and the respective conversations with the AI chatbot. Several conversations were translated from original Spanish to English, as indicated in the dataset. Translation was conducted with the in-built translator of Microsoft Word and reviewed by an author fluent in both languages to ensure accuracy. The original languages were preserved in a separate file. Third, the prompts alone were extracted from the conversations as their own sibling file to make their use in analyses easier. The prompts appear in their dedicated file in the same order as in the complete dataset.

The final dataset thus consists of three types of documents and six files: a document with the complete conversations, including those translated from Spanish to English (ai_dataset_conversations), a document with the same content without translation (ai_dataset_conversations_untranslated), and a document with only the prompts (ai_dataset_prompts), each available both as .docx and .pdf files. The 466 pages of data comprise in total 34 conversations and 318 prompts, averaging almost 10 prompts by conversation but including both multiple single-prompt conversations and several long conversations with around 20 prompts.

VALUE AND REUSABILITY

Scholars from several fields such as AI, innovation, and education can benefit from the data. In particular, the dataset can serve to probe and develop theories on the use of generative AI in creative problem solving and help craft practical recommendations on the use of generative

AI in innovation work and in different educational programs and methodologies.

This data can be utilized both in qualitative and quantitative analysis. Regarding qualitative methods, thematic analysis (Braun & Clarke, 2006) is a straightforward method to apply on the dataset, but also other methods, such as narrative analysis (Sandelowski, 1991) or clustering, can be used. Quantitative analysis can be used to deepen and explain qualitative observations and can take the form of, for example, content analysis (Riffe et al., 2023). Using the data together with other similar datasets may be particularly useful to obtain quantitative results on the effects of the task type, user profile, or other factors on the use of generative AI.

It is important to consider the particularities of the I2P and CBI educational programs and the resulting limitations in all future use of the data. Such particularities include an open and self-organized approach to problem solving, complex and multilayered assignments, as well as international and multidisciplinary student teams. This evaluation is particularly important when the data is used in combination with additional data from different empirical settings. A recommended approach to integrate the data introduced here into larger empirical studies would be to treat each empirical setting as their own case within a multiple-case analysis, while explicitly recognizing and contemplating the cross-case differences.

To provide a practical example of its value, the data has already been utilized in an unpublished conference paper exploring students' favoured use cases for generative AI within transformative futures innovation education. In this study, the authors used an inductive qualitative content analysis, triangulating several independently developed coding schemes into an integrated version. This qualitative exploration led the authors to conclude that students sometimes struggle to find optimal use cases for generative AI and that the straightforward answers that AI tools may provide to complex and multilayered issues present a problem. The authors also found that students rely extensively on generative AI tools for information retrieval, substituting existing tools with AI chatbots instead of incorporating AI as a new tool with a clearly defined role.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

ACKNOWLEDGEMENTS

The authors would like to thank all CERN IdeaSquare team members and external teachers from partner universities that facilitated research access to the IdeaSquare innovation education programs. We also wish to thank all students who produced and shared the data that the article introduces.

REFERENCES

Boussioux, L., Lane, J. N., Zhang, M., Jacimovic, V., & Lakhani, K. R. (2024). The Crowdless Future? Generative AI and Creative Problem-Solving. *Organization Science*, *35*(5), 1589–1607. https://doi.org/10.1287/orsc.2023.18430

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Doshi, A. R., & Hauser, O. P. (2024). Generative artificial intelligence enhances creativity but reduces the diversity of novel content. *Science Advances*, *10*(28), eadn5290. https://doi.org/10.1126/sciadv.adn5290

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

Habib, S., Vogel, T., Anli, X., & Thorne, E. (2024). How does generative artificial intelligence impact student creativity? *Journal of Creativity*, *34*(1), 100072. https://doi.org/10.1016/j.yjoc.2023.100072

Helal, M. Y. I., Elgendy, I. A., Albashrawi, M. A., Dwivedi, Y. K., Al-Ahmadi, M. S., & Jeon, I. (2025). The impact of generative AI on critical thinking skills: A systematic review, conceptual framework and future research directions. *Information Discovery and Delivery*. https://doi.org/10.1108/IDD-05-2025-0125

Kadaruddin, K. (2023). Empowering Education through Generative AI: Innovative Instructional Strategies for Tomorrow's Learners. *International Journal of Business, Law, and Education*, *4*(2), 618–625. https://doi.org/10.56442/ijble.v4i2.215

Larson, B. Z., Moser, C., Caza, A., Muehlfeld, K., & Colombo, L. A. (2024). Critical Thinking in the Age of Generative AI. *Academy of Management Learning & Education*, *23*(3), 373–378. https://doi.org/10.5465/amle.2024.0338

Mariani, M., & Dwivedi, Y. K. (2024). Generative artificial intelligence in innovation management: A preview of future research developments. *Journal of Business Research*, *175*, 114542. https://doi.org/10.1016/j.jbusres.2024.114542

Mittal, U., Sai, S., Chamola, V., & Sangwan, D. (2024). A Comprehensive Review on Generative AI for Education. *IEEE Access*, *12*, 142733–142759. https://doi.org/10.1109/ACCESS.2024.3468368

Nordberg, M., Garcia Tello, P., Poulaillon, J., Wirtavuori, L., & Valtonen, L. (2024). Addressing Societal Challenges with the Help of AI from Student Teams at IdeaSquare, CERN. In M. Streit-Bianchi & V. Gorini (Eds.), *New Frontiers in Science in the Era of AI* (pp. 143–155).

Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-61187-2_9

Rafner, J., Beaty, R. E., Kaufman, J. C., Lubart, T., & Sherson, J. (2023). Creativity in the age of generative AI. *Nature Human Behaviour*, *7*(11), 1836–1838. https://doi.org/10.1038/s41562-023-01751-1

Riffe, D., Lacy, S., Watson, B. R., & Lovejoy, J. (2023). *Analyzing Media Messages: Using Quantitative Content Analysis in Research* (5th ed.). Routledge. https://doi.org/10.4324/9781003288428

Sandelowski, M. (1991). Telling Stories: Narrative Approaches in Qualitative Research. *Image: The Journal of Nursing Scholarship*, *23*(3), 161–166. https://doi.org/10.1111/j.1547-5069.1991.tb00662.x

Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education*, *215*, 105031. https://doi.org/10.1016/j.compedu.2024.105031

Valtonen, L., Garcia Tello, P., Gaddi, G., Poulaillon, J., Werner, O., Zimmermann, D., Batista, C., Wirtavuori, L., & Breidvik, M. (2025). Towards meeting futures' educational needs: A case study of CERN IdeaSquare planet. *Cogent Education*, *12*(1), 2542398. https://doi.org/10.1080/2331186X.2025.2542398

Vignoli, M., Balboni, B., Cotoranu, A., Dosi, C., Glisoni, N., Kohler, K., Mincolelli, G., Mäkinen, S., Nordberg, M., & Thong, C. (2021). Inspiring the future change-makers: Reflections and ways forward from the Challenge-Based Innovation experiment. *CERN IdeaSquare Journal of Experimental Innovation*, *5*(1), 1–4. https://doi.org/10.23726/cij.2021.1323