

Lessons Learned from Creating Course Content using Generative AI

Mathias Hatakka and Andreas Ask

Örebro University School of Business, Örebro, Sweden

mathias.hatakka@oru.se

andreas.ask@oru.se

Abstract: The release of generative AI tools such as OpenAI's ChatGPT has sparked interest in their implications for education. While early discourse emphasized concerns about plagiarism and academic integrity, recent studies have begun to explore the potential of these tools to support teaching and learning. This paper presents a case study on the use of ChatGPT in the redesign of a first-year systems development project course for informatics students. The course required the integration of various course materials, making it a suitable context for evaluating generative AI's role in course material development. The aim of the study is to present lessons learned from using ChatGPT in the development of course content. Drawing on our practical experience as course designers and instructors, we outline lessons learned from using ChatGPT in the creation of key course elements, including case descriptions, SQL scripts, and requirements specifications. We found that ChatGPT was effective for generating coherent initial drafts of content, but its outputs often required refinement to ensure pedagogical alignment. Challenges included the generation of misleading or irrelevant non-functional requirements and logically flawed code, despite syntactic correctness. Our findings highlight the importance of prompt engineering, critical review, and maintaining a human-in-the-loop approach. We conclude that while ChatGPT can significantly reduce development time for some tasks, it should be used as a complementary tool. This study contributes practical insights to the growing field of AI-assisted education.

Keywords: Content Creation, Generative AI, Large Language Models, ChatGPT, Course Development

1. Introduction

Since the release of OpenAI's large language model (LLM) ChatGPT in November 2022, the technology has shown the potential to be a disruptive force in education (Mills et al., 2023), as well as to revolutionize and transform many aspects of higher education (Ansari et al., 2023; Lee, 2024; Koraishi, 2023). Early reports, especially in mainstream media, portrayed ChatGPT primarily as a tool for plagiarism and cheating. Since then, the discourse in both media and academic research has become more nuanced. Educators and researchers have begun to explore how generative AI can be used to enhance student learning and to streamline teaching and administrative tasks. For example, studies have investigated how LLMs can be used to generate feedback on student assignments (Ali et al., 2024; Dai et al., 2023; Escalante et al., 2023), serve as personalized tutors (Ansari et al., 2023; Currie, 2023), or provide automated administrative support (Rasul et al., 2023; Kshetri et al., 2023). While AI holds great potential to benefit education, there are also notable risks and challenges. These include issues such as bias and stereotyping in generated texts, as well as the generation of false or misleading information (Dempere et al., 2023).

While there are many types of generative AI, this study focuses on ChatGPT, an interactive AI system that uses a natural language processing model to simulate human-like conversations in text (Ansari et al., 2023). An essential aspect of using ChatGPT for generating text is prompt engineering.

"Prompt engineering is the process of designing, refining, and optimizing input prompts to effectively communicate the user's intent to a language model like ChatGPT. This practice is essential for obtaining accurate, relevant, and coherent responses from the model" (Ekin, 2023, p. 3).

By providing the LLM with rules, guidelines, and context for the "conversation," users can clarify what information is important and what the desired output should be (White et al., 2023). For example, in addition to specifying the actual question, users can include contextual details such as tone, intended audience, and constraints on the response. However, even carefully designed prompts do not guarantee accurate responses. The hallucination effect - when generated texts appear realistic but do not correspond to reality (Alkaissi & McFarlane, 2023) - is a significant concern, as users may not be able to distinguish between accurate information and fabrications. Therefore, maintaining a human-in-the-loop approach is essential (Guo & Wang, 2023), and ChatGPT should be viewed as an assistant rather than a replacement.

Designing educational resources is a time intensive and complex task, particularly when courses include multiple components. There is limited empirical research on the practical benefits, limitations, and challenges involved. This creates a need to better understand how tools such as ChatGPT can support teachers in course design, especially in a way that align with pedagogical goals, how to avoid misinformation, and to mitigate the risk of

overreliance on AI-generated content. The aim of this case study is to present lessons learned from using ChatGPT in the development of course content. The case concerns a first-year systems development project course at Örebro University in Sweden. In the spring of 2024, the course project was redesigned using ChatGPT. In this paper, we explain the redesign process and present the lessons learned from using ChatGPT in course development.

2. Literature Review

Early experiments using generative AI to produce various types of course content have been cautiously optimistic. For example, Mikeladze (2023) used ChatGPT in the design of English language teaching materials with positive results. The material could be contextualized for specific learners, and its customization led to increased effectiveness and learner motivation. Another benefit was the time and resources saved for teachers. Similarly, Rouabhia (2024) used ChatGPT to create content for a multimedia databases course. The findings from the study suggest that generative AI can significantly enhance content creation, offer scalable and customizable solutions, and save time for teachers. In an experimental research project, Meron and Araci (2023) used ChatGPT to develop course assignments and design a postgraduate course, also with positive results: “ChatGPT was a competent partner with regard to saving time, structuring textual content and documentation, and as a brainstorming tool” (p. 1). The generated content was clearly structured and well-formatted. All three studies highlight that using ChatGPT for course material development leads to positive educational outcomes, enables content customization, and significantly reduces the time and effort required by teachers.

A common theme in the successful use of generative AI for course content development is prompt engineering. The quality of generated content depends on the prompts provided (Rouabhia, 2024) and “requires considerable effort and calculated prompting by professional, design-educated, and experienced human course developers” (Meron & Araci, 2023, p. 19). While it is easy to ask simple questions of ChatGPT, designing prompts that tailor the generated outputs to your specific needs is more complex. The hallucination effect - where ChatGPT appears to “make up information seemingly out of thin air” (Koraishi, 2023, p. 69) – also means that we cannot blindly trust its outputs. That ChatGPT may generate misleading or false information is a concern (ibid), and it is something that must be taken seriously. There is also the issue of bias in generated material (Davis & Lee, 2023; Koraishi, 2023), where AI may introduce, for example, cultural biases or stereotypes related to gender or ethnicity. This highlights the need for continuous human oversight (Kasneci et al., 2023) of the generated content.

In summary, recent studies on the use of generative AI in course content development report positive outcomes, highlighting benefits such as time savings, improved content structure, and enhanced customization for specific learning contexts. However, these successes depend heavily on the quality of prompt engineering, which requires expertise and iterative refinement. Despite its strengths, generative AI tools like ChatGPT can produce misleading or biased outputs, making continuous human oversight essential.

3. Method

In this study we explore the use of generative AI, specifically ChatGPT, in the development of course materials for a university-level systems development course. This study is based on a case study methodology, through which we investigate a contemporary phenomenon in a real-life context involving a relevant real-world problem (Yin, 2018). The contemporary phenomenon is the use of ChatGPT in education, and the real-world problem is the development of course content. The course used as a case, designed for first-year informatics students, requires the integration of various course materials, making it an ideal context for evaluating the role of generative AI in course content development. Both authors of this paper are teachers on the course and are responsible for creating the materials. Our method involved multiple cycles, during which we created, evaluated, and refined course materials with ChatGPT, guided by both pedagogical goals and practical constraints.

The overall development process consisted of five stages: (1) ER-model design, (2) database script generation, (3) generating example data, (4) case description generation, and (5) requirements specification. Each stage, except for the ER-model design, involved human-AI interaction, where we iteratively refined both the input prompts and the resulting output to align with the course objectives and student needs. Throughout the development process, we documented each interaction with ChatGPT, including the specific prompts used, the generated outputs, and our evaluation of the output. This documentation enabled us to track what had worked (or not) in previous prompt designs. It provided a basis for post-hoc analysis, including the discussion of recurring issues such as hallucination, syntactic vs. logical validity, and prompt design. The documentation also allowed us

to trace the origins of design decisions and to reflect on the implications of using generative AI in educational contexts.

3.1 Course Description

The course used as the case for this study is a systems development project course at Örebro University in Sweden. It is the final course of the first semester in the Informatics, Basic Course program (30 credits) and aims to help students apply previously acquired knowledge to develop an information system. The course has two parts: in the first week, students model an organization using entity-relationship (ER) modeling; in the final five weeks, they develop the system using Java in groups of four.

In spring 2023, we replaced the existing case with a new one focused on developing a system for an NGO involved in sustainability projects. While existing recorded materials (lectures, tutorials) were largely reusable, all case-specific documents had to be created from scratch:

- An ER-model of the organization - This is used both to develop the database and as a reference model against which students compare their own models as part of the assessment in the first part of the course.
- A database script, including example data - At the start of the development phase, all student groups receive the same database script to ensure a common starting point.
- A case description - A detailed description of the NGO, used in both the analysis and development phases of the course. In the analysis phase, students use it to model the organization; in the development phase, it provides contextual background for the system they are building.
- Requirement specifications - These include the functional and non-functional requirements students must implement. A sample functional requirement is: "As a project supervisor, I need to be able to change the date of the projects I am assigned." A sample non-functional requirement is: "The system should not crash if the user enters incorrect data in a field."

When developing the new material, we decided to experiment with using ChatGPT. However, we needed a base from which to generate prompts for ChatGPT, so we manually created the ER-model of the NGO as a starting point. Once the ER-model was finalized, the remaining documents were generated using ChatGPT.

4. Course Design Using ChatGPT

In this section we will explain how the new materials for the course were generated. Figure 1 show the process for the development of the course material.

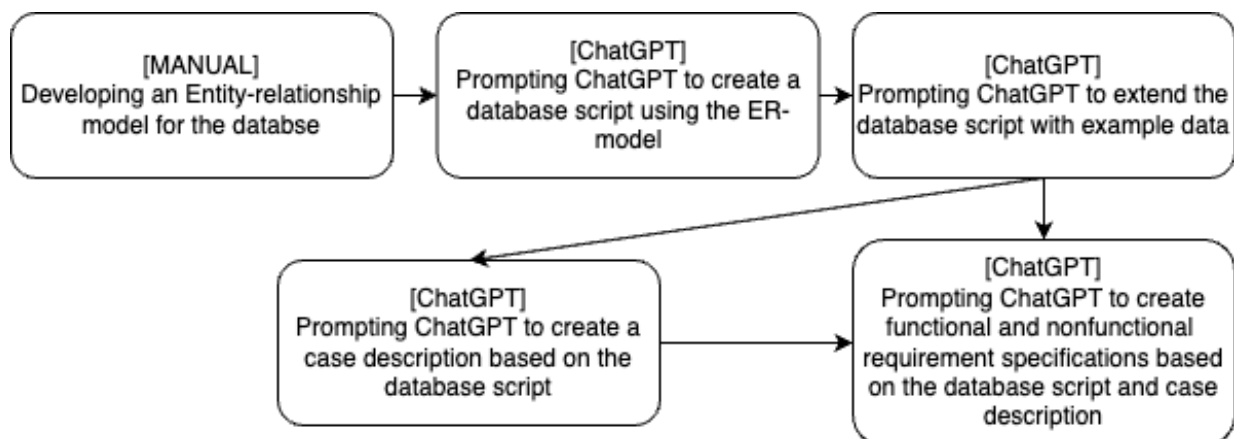


Figure 1: The course material development process

4.1 Creating the ER-Model

The first step in developing the project case was to establish a starting point that could later be used with ChatGPT. We needed something to "feed" into ChatGPT in our prompts. In this case, we opted to begin by creating an ER-model for the database. This decision was based on two main reasons, 1) We had a general idea of the case's focus (an NGO working with the Sustainable Development Goals), but few specific details. Starting with the ER model allowed us to model the organization and discuss which entities, relationships, and other

components should be part of the system. 2) An important aspect of the course is that students work with different components of a relational database. Creating the model manually ensured that these components (e.g., different types of relationships) would be included.

Although we did not attempt it, we believe another starting point could have worked - for example, manually creating the case description and using ChatGPT to generate the database structure and requirement specifications.

4.2 Creating the Database Script

Once the initial ER-model was completed, it was uploaded, and we prompted ChatGPT to generate an SQL script containing DDL code for MySQL based on the provided image. We copied the generated code into an IDE and executed it to create the database. This allowed us to review the code for potential errors (there were none) and to understand how the completed database would appear. It also helped us evaluate whether the model required revisions, such as adding or removing elements that did not align with the course objectives.

The initial ER-model did not specify attributes for the entities and the generated DDL code included only key attributes for each entity in the model. Therefore, we revisited our ER-model, added attributes for each entity, and re-uploaded the image to the same session before prompting ChatGPT to regenerate the DDL script. This ensured that the attributes were in Swedish and included the details we required for the case. ChatGPT then generated a new SQL script, which we copied into the IDE, checked for errors or inconsistencies (there were none), and used to recreate the database.

4.3 Creating Example Data

Using the final SQL script generated by ChatGPT, we then prompted ChatGPT to create SQL code for inserting multiple rows with example data into each table. Since the case relates to the UN's sustainability goals, we asked ChatGPT to generate insert statements for the UN goals, which we then copied and pasted into an IDE, checked for errors, and executed. We followed the same process for all tables, prompting ChatGPT for relevant insert statements. Additionally, we asked ChatGPT to suggest suitable projects for an NGO working with UN sustainability goals and appropriate responsibilities for departments focusing on specific UN goals. Based on these suggestions, we created insert scripts for each table. Occasionally, we encountered issues where the amount of generated data exceeded ChatGPT's capabilities, requiring us to reduce the dataset's size. Once all inserts were generated and inserted into the MySQL database, the final step was to create an SQL dump of the entire database. We then uploaded it to ChatGPT and prompted it to:

1. Identify any errors or inconsistencies.
2. Suggest possible improvements.

Based on this, we finalized the script and added some code, including a stored procedure that would allow students to re-run the script in its entirety to revert the database to its original state whenever needed. However, the code generated by ChatGPT did not fully adhere to the ER-model. This was not discovered until the course was underway, and students using the case experienced issues when implementing the information system. The ER-model included disjoint specializations, which occur when each real-world entity can belong to only one subclass at most. When ChatGPT created the code for inserting multiple rows, it did not account for the disjoint relationships. Due to the large amount of example data generated by ChatGPT, this issue went unnoticed. Disjoint relationships are not enforced by SQL but rather handled in the Java code. Therefore, although the SQL code generated was syntactically correct, it was logically incorrect because it disregarded parts of the ER-model.

4.4 Creating a Case Description

In the next step, we used the database script to prompt ChatGPT for a case description. The description is used by the students in modeling the NGO; hence, it needed to include all aspects of the database and be detailed enough to guide the students in their own modeling work. The case description is also used in the realization part of the course to provide the students with a description of the NGO and the context for the information system they are developing. In this step, we tried several different prompts. We started with a very simple prompt:

Create a case description based on the SQL script [insert SQL script]

Unsurprisingly, the results were unsatisfactory. The results focused too much on explaining the database rather than describing the NGO. In the next iteration, we extended the prompt to include that the case description should explain the NGO rather than the database:

Create a case description of the NGO that will use this database [insert SQL script]; all aspects of the database need to be included in the description.

The result of this prompt was better but still flawed in the sense that it did not include enough detail. As mentioned, the case description was to be used by the students not only to understand the NGO but also to be able to model the NGO using ER-modeling. We again extended the prompt:

Create a case description of the NGO that will use the database included below; all aspects of the database need to be included in the description. The case description will be used by systems development students to create an ER-model of the NGO. [insert SQL script]

This prompt resulted in a case description that focused on explaining the NGO and had enough detail to enable the students to model the NGO based on it. However, the case description had two issues that needed to be fixed. The first was that it included the names of the tables in the database (the students were supposed to find the tables from the description themselves, so we did not want those included). That was easily fixed by just adding “Do not include the names of the tables in the description” to the prompt. The second problem was that the description did not give enough “clues” for the different types of relationships we wanted the students to model. We were also unsuccessful in updating the prompt so that the description managed to include a detailed enough explanation of all the relationships, and instead, we decided to manually update the case description with the missing details.

4.5 Creating a List of Requirements

After the case description was completed, we used ChatGPT to create a list of requirements for the students to implement. The requirements should include both functional and non-functional elements. This list essentially defines what is being assessed - that is, students are expected to implement all requirements. We started with the following prompt:

Create a list of requirements for an information system based on the SQL script below. The list should include both functional and non-functional requirements. [insert SQL script]

The results of this prompt were unsatisfactory. The requirements were either too technical or very unclear. Next, we used the same prompt but replaced the SQL script with the case description. The results were more in line with what we wanted, but many of the requirements were still unclear. Additionally, the prompt generated requirements that would be impossible to implement using the current database. In the next iteration, we asked ChatGPT to consider both the SQL script and the case description. We also clarified the purpose of the list of requirements within the prompt:

Create a list of requirements for an information system based on the case description and SQL script below. The list should include both functional and non-functional requirements. The list of requirements will be used by beginner programmers in a 6-week university course. [insert case description] [insert SQL script]

Due to limitations in how much texts that can be added in prompts, we had to divide the texts based on functionalities. The result of this prompt was much more in line with what we wanted. The functional requirements were detailed, feasible, and (for the most part) appropriately challenging for beginner programmers. For example:

Functional Requirements:

1. User Management
 - The system must allow the creation, updating, and deletion of user accounts (administrators, project managers, and staff).
 - The system must support role-based access control to ensure that users have appropriate permissions based on their roles.

However, most of the non-functional requirements were not satisfactory. This is likely because the case description and SQL script provided few clues as to what non-functional requirements would be necessary for the information system. Instead, ChatGPT included generic non-functional requirements commonly found in information systems. Examples of proposed non-functional requirements included:

- Data Encryption: All sensitive data (such as passwords) should be encrypted both in transit and at rest.
- System Uptime: The system should be operational and accessible at least 99.5% of the time, excluding planned maintenance.

The data encryption requirement would be too complex for our students, and the system uptime requirement is not relevant. We tried tweaking the prompt by adding more context (e.g., the students' level of programming experience and the type of system - a desktop system coded in Java), but we received similar results. Ultimately, we chose to manually add the non-functional requirements to the list.

5. Discussion

Similar to previous studies on the use of ChatGPT to generate course content, we are cautiously optimistic about its usefulness. We were mainly successful in generating course material, except for the non-functional requirements and relationships in the case description, which we ultimately created manually. While ChatGPT proved effective in generating coherent descriptions and initial drafts, its outputs required careful refinement to align with the course objectives. Based on our experience using ChatGPT for content development, we draw five important lessons:

1. The design of prompts is critical to the results, and it is an iterative process:

Consistent with other studies (e.g., Rouabhia, 2024; Meron and Araci, 2023), we found that prompt design is crucial to obtaining relevant results and that it is a complex process requiring iteration. In our case, only after specifying the audience (students and their level) and explicitly stating what not to include did the generated texts begin to align with our needs. We also found that the more background information included in the prompts, the better the results. For example, when generating requirements, we added the SQL script for technical requirements and the case description to provide context.

2. The hallucination effect results in misleading or false information:

The most apparent example of the hallucination effect in our case was in generating non-functional requirements. We encountered challenges in producing detailed and contextually appropriate requirements and had to manually adjust them to ensure their relevance to the course. The generated non-functional requirements were not based on the case and were not applicable to the course content. While the case description and SQL-script provided some clues as to what the non-functional requirements could be, it lacked sufficient detail. As a result, ChatGPT added generic requirements not relevant to the course - i.e., it generated misleading information (Koraishi, 2023). Our solution was to manually edit the non-functional requirements, but another option would have been to add more contextual information about the system so that ChatGPT did not need to "hallucinate."

3. Syntactic correctness is not the same as logical validity:

Generated code that is syntactically correct may still contain contextual and logical flaws that do not fully align with the underlying data model or system requirements. In such instances, the issue is not hallucination—the code is technically correct and functional, and it partially meets the stated requirements. However, ChatGPT's inability to fully account for contextual and logical constraints can cause problems when the course material is implemented. This highlights the importance of critically reviewing and validating AI-generated outputs, especially in educational contexts where implicit assumptions - such as disjoint specializations in an ER-model - may not be correctly implemented without explicit guidance.

4. Choose the correct starting point for your prompts:

In our case, we aimed to create several documents that were interconnected. We had two logical starting points: either begin with a model of the organization (an ER-model) or the case description. While starting with the case description might have worked, the advantage of manually creating the ER model was ensuring the inclusion of necessary concepts. If we had started with the case description, it is likely that several concepts would have been missing in the ChatGPT-generated ER-model - unless we included more detail than we needed. A flawed or overly simplistic ER-model would have introduced errors into all subsequent documents generated from it.

5. Using ChatGPT for content generation can save time:

We agree with previous research that ChatGPT can save teachers time in content development (e.g., Meron and Araci, 2023). Using the ER-model as input to generate the SQL script and example data was significantly faster

than doing it manually. However, both the script and the data contained errors that were only discovered during course implementation, despite our efforts to review them beforehand. However, these errors were minor and easily corrected. ChatGPT is effective at generating content with a clear structure and well-defined rules. However, generating context-dependent content, such as non-functional requirements, was more problematic. We ultimately had to create these manually, and the time spent trying to prompt ChatGPT to generate them was, in retrospect, a loss.

This study contributes to the growing body of research on AI-assisted education by offering practical insights into the benefits and limitations of generative AI in course design. We conclude that while ChatGPT can be a valuable tool for content generation, its role should be complementary rather than substitutive, requiring educators to critically engage with and refine AI-generated outputs. A human-in-the-loop approach (Guo and Wang, 2023) is essential to ensure that the content is appropriate, accurate, and aligned with the course objectives. The contributions are based on our roles as both educators and researchers. Both authors were responsible for the design, development, and implementation of the course materials discussed in this paper. This included designing prompts for ChatGPT, evaluating and refining AI-generated outputs, and manually revising components that did not meet pedagogical, of practical, requirements. Our findings are thus based on direct, hands-on experience using ChatGPT in a real-world educational setting. By documenting and analyzing this process, we contribute a practice-based perspective that complements existing theoretical and experimental research on AI-assisted education.

Ethical Declaration

Ethical clearance was not required for the research.

AI Declaration

ChatGPT has only been used for spelling and grammar correction of the texts. No text in the paper, nor any analysis, has been generated by AI.

References

- Ali, K., Barhom, N., Tamimi, F. & Duggal, M. 2024. ChatGPT—A double-edged sword for healthcare education? Implications for assessments of dental students. *European Journal of Dental Education*, 28, 206-211.
- Alkaissi, H. & Mcfarlane, S. I. 2023. Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*, 15.
- Ansari, A. N., Ahmad, S. & Bhutta, S. M. 2023. Mapping the global evidence around the use of ChatGPT in higher education: A systematic scoping review. *Education and Information Technologies*, 1-41.
- Currie, G. M. Academic integrity and artificial intelligence: is ChatGPT hype, hero or heresy? *Seminars in Nuclear Medicine*, 2023. Elsevier, 719-730.
- Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y.-S., Gašević, D. & Chen, G. Can large language models provide feedback to students? A case study on ChatGPT. 2023 IEEE International Conference on Advanced Learning Technologies (ICALT), 2023. IEEE, 323-325.
- Davis, R. O. & Lee, Y. J. 2023. Prompt: Chatgpt, create my course, please! *Education Sciences*, 14, 24.
- Dempere, J., Modugu, K., Hesham, A. & Ramasamy, L. K. The impact of ChatGPT on higher education. *Frontiers in Education*, 2023. Frontiers Media SA, 1206936.
- Ekin, S. 2023. Prompt engineering for ChatGPT: a quick guide to techniques, tips, and best practices. *Authorea Preprints*.
- Escalante, J., Pack, A. & Barrett, A. 2023. AI-generated feedback on writing: insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20, 57.
- Guo, K. & Wang, D. 2023. To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 1-29.
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S. & Hüllermeier, E. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, 103, 102274.
- Koraishi, O. 2023. Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology*, 3.
- Kshetri, N., Hughes, L., Louise Slade, E., Jeyaraj, A., Kumar Kar, A., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H. & Ahmad Albashrawi, M. 2023. "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.
- Lee, H. 2024. The rise of ChatGPT: Exploring its potential in medical education. *Anatomical sciences education*, 17, 926-931.
- Meron, Y. & Araci, Y. T. 2023. Artificial intelligence in design education: evaluating ChatGPT as a virtual colleague for post-graduate course development. *Design Science*, 9, e30.

- Mikeladze, T. Creating teaching materials with ChatGPT. Proceedings of the IRCEELT—2023 13th International Research Conference on Education, Tbilisi, Georgia, 2023. 5-6.
- Mills, A., Bali, M. & Eaton, L. 2023. How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process. *Journal of Applied Learning and Teaching*, 6, 16-30.
- Rasul, T., Nair, S., Kalendra, D., Robin, M., De Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A. & Heathcote, L. 2023. The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, 6, 41-56.
- Rouabhia, D. 2024. Artificial intelligence driven course generation: A case study using chatgpt. arXiv preprint arXiv:2411.01369.
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J. & Schmidt, D. C. 2023. A prompt pattern catalog to enhance prompt engineering with chatgpt. *arXiv preprint arXiv:2302.11382*.
- Yin, R. K. 2018. Case study research and applications. Sage Thousand Oaks, CA.