Original Paper

Cocreating an Automated mHealth Apps Systematic Review Process With Generative AI: Design Science Research Approach

Guido Giunti[1,2,3], MD, PhD; Colin P Doherty[1,3,4], MD

[1]Academic Unit of Neurology, School of Medicine, Trinity College Dublin, Dublin, Ireland

[2]Research Unit of Health Sciences and Technology, Faculty of Medicine, University of Oulu, Oulu, Finland

[3]FutureNeuro SFI Research Centre, Royal College of Surgeons in Ireland, Dublin, Ireland

[4]Department of Neurology, St James Hospital, Dublin, Ireland

Corresponding Author:
Guido Giunti, MD, PhD
Academic Unit of Neurology
School of Medicine
Trinity College Dublin
College Green
Dublin, D02
Ireland
Phone: 353 1 896 1000
Email: drguidogiunti@gmail.com

Abstract

Background: The use of mobile devices for delivering health-related services (mobile health [mHealth]) has rapidly increased, leading to a demand for summarizing the state of the art and practice through systematic reviews. However, the systematic review process is a resource-intensive and time-consuming process. Generative artificial intelligence (AI) has emerged as a potential solution to automate tedious tasks.

Objective: This study aimed to explore the feasibility of using generative AI tools to automate time-consuming and resource-intensive tasks in a systematic review process and assess the scope and limitations of using such tools.

Methods: We used the design science research methodology. The solution proposed is to use cocreation with a generative AI, such as ChatGPT, to produce software code that automates the process of conducting systematic reviews.

Results: A triggering prompt was generated, and assistance from the generative AI was used to guide the steps toward developing, executing, and debugging a Python script. Errors in code were solved through conversational exchange with ChatGPT, and a tentative script was created. The code pulled the mHealth solutions from the Google Play Store and searched their descriptions for keywords that hinted toward evidence base. The results were exported to a CSV file, which was compared to the initial outputs of other similar systematic review processes.

Conclusions: This study demonstrates the potential of using generative AI to automate the time-consuming process of conducting systematic reviews of mHealth apps. This approach could be particularly useful for researchers with limited coding skills. However, the study has limitations related to the design science research methodology, subjectivity bias, and the quality of the search results used to train the language model.

(JMIR Med Educ 2024;10:e48949) doi: 10.2196/48949

KEYWORDS

generative artificial intelligence; mHealth; ChatGPT; evidence-base; apps; qualitative study; design science research; eHealth; mobile device; AI; language model; mHealth intervention; generative AI; AI tool; software code; systematic review; language model

Introduction

The delivery of health-related services through the use of mobile devices (mHealth) [1] has been growing at a tremendous pace. A decade ago, in the first "era of mHealth," the literature surrounding mHealth called for the generation of evidence demonstrating the impact of mHealth solutions on health system processes and patient outcomes [2]. In 2013, Labrique et al [2]

conducted a preliminary search on the US federal clinical trials database (ClinicalTrials.gov) and had to combine the keywords "mHealth," "mobile," and "cell AND phone" to obtain 1678 studies and their results. Today, that same number can be obtained using "mHealth" alone as a keyword. As the need for mHealth evidence has grown, so too has the necessity for summarizing both the state of the art and the practice.

Systematic reviews seek to collect and combine relevant evidence within the specific scope of a research question while also striving to minimize bias [3,4]. In PubMed alone, the number of systematic reviews published on digital health–related topics has increased a hundredfold in the last 10 years. In fact, the pace at which the mHealth field is developing for certain conditions like breast cancer is such that systematic reviews can be found every 2 or 3 years [5-9]. The systematic review process, however, is a time- and resource-intensive process, reportedly requiring a median of 5 researchers and approximately 40 weeks of work to reach submission [10-12].

The emergence of generative AI has been seen as a breakthrough in the field of automation. With the ability to generate content such as text, images, and even music, AI has been reported as a potential solution to tedious time-consuming and labor-intensive tasks [13]. For instance, generative AI can be used to automatically generate product descriptions, news articles, or even code [14]. By eliminating the need for human intervention, generative AI can free up valuable time and resources for more complex tasks, thereby improving efficiency and accuracy. ChatGPT, a natural language processing model with a capacity of 175 billion parameters, has been trained on extensive amounts of data and is designed to produce human-like responses to user inputs. Since its release in November 2022, ChatGPT has received significant attention from media and academia alike, provoking ethical discussions on scientific authorship [15,16], attempting to pass medical license and specialist examinations [17-19], and even designing medical education curricula [20].

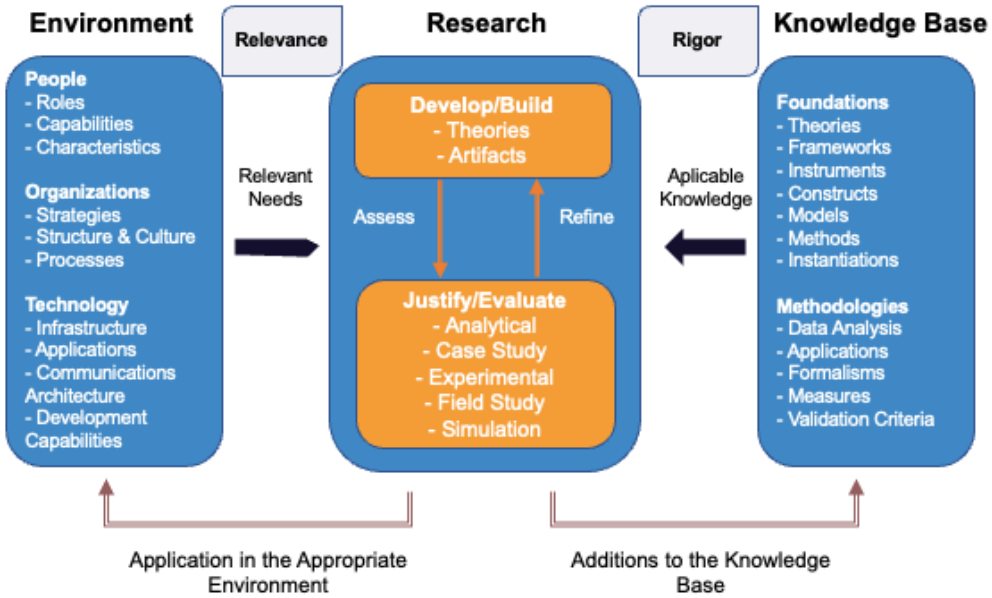
The objective of this study was to explore the feasibility of using generative AI tools to automate time-consuming and resource-intensive tasks in a systematic review process and assess the scope and limitations of using such tools.

Methods

Study Design

This study uses a design science research (DSR) methodology. DSR is a problem-solving paradigm that seeks to enhance human knowledge via the creation of innovative artifacts [21]. DSR commonly involves the identification of a problem or opportunity, followed by the development, implementation, and evaluation of a solution. In DSR, as well as in action research, the process happens within an organization that provides context and that would be changed as a result of the use of the artifact [21]. An overview of the process adapted from Hevner [22] can be seen in Figure 1.

Figure 1. Design science research overview, adapted from Hevner 2004 [22].



Problem Definition

The problem to which DSR was applied was the time-consuming and resource-intensive process of conducting systematic reviews of mHealth applications.

Organizational Context

The organizational context consisted of the More Stamina team of researchers, software developers, and health care professionals, working collaboratively within the host research institutions (ie, the University of Oulu and Trinity College Dublin).

The More Stamina project aims to create an evidence-driven gamified mHealth solution for people with multiple sclerosis (MS), where each step of the development follows a scientific process, as follows: MS needs as well as barriers and facilitators were explored through qualitative studies [23]; the state of the practice for MS apps was systematically reviewed [24,25]; user-centered design was used to create "MS personas" [23];

cocreation sessions took place to produce solution concepts [26]; the design, prototyping, and initial usability testing were described [27]; early health technology assessment was used to guide software development [28]; patient representatives were involved throughout the project [29]; and user testing and feasibility studies were ongoing in a multicenter study [30].

A script using the software application for audience targeting called 42matters [31] was used in the past to extract information from different app stores. The script is no longer functional, and person-hours from the software development team were not able to be dedicated to this task.

Background Studies

The research plans and outlines from previous studies, where systematic review methodologies were used to identify, select, collect, and analyze features and content of mHealth apps [6,24,25], served as models for our study. In those studies, a search strategy was defined, using relevant main keywords for each condition. App stores were searched, taking steps to ensure that no previous search history or cookies influenced results. Screening took place based on mHealth applications' titles, descriptions, and metadata.

Solution

The solution was to apply a cocreation process with a generative AI (ie, ChatGPT 3.5, as of June 2023) to produce software code that automated the process for conducting systematic reviews.

Cocreation Goal

The goal of the cocreation process was to use ChatGPT as a design and development partner for the automation process. The generative AI was to be interacted with as if it were a valid interloper who was more technologically skilled than the user and was guiding them through the process over text messages.

Development and Implementation

Development and implementation of the automated process happened through iterative and continuous conversations with the generative AI by one of the authors (GG). GG is a primary care physician with over a decade of experience leading digital health software design and development. Table 1 provides an overview of his digital skills background using the European Qualifications Framework and with a self-assessment score from 1 to 10 to describe his competency level. Regardless of the skill level, the development cycle was to be conducted as if no coding skill was present on the part of the user.

Table 1. Digital skills background.

Competency	Level	Experience	Self-assessment score (of 10)
Scrum master	Certified Scrum Master	Agile methodologies and team management	7
Product owner	Certified Scrum Product Owner	Product road mapping and stakeholder management	8
Game design	Intermediate	Game mechanics, storytelling, and level design	7
Web design	Advanced	User experience and user interface design and responsive design	8
JavaScript	Beginner	Front-end development	5
HTML5	Intermediate	Front-end development	7
PHP: Hypertext Preprocessor	Beginner	Front-end and back-end development	4

Evaluation

The results were evaluated for their effectiveness and efficiency in replicating the initial steps of the background studies. The measurements were considered with respect to the amount of time required to generate a spreadsheet containing the necessary information for human reviewers to start the systematic review process. The output was compared to the output generated by the previous script, which required further processing of the data.

Ethical Considerations

No ethics board review is needed as the work does not use patient data or involve human participants.

Results

The following section reports the overall design, development, and implementation process as well as the evaluation results and highlights of the experience. Prompts provided by GG and generative AI are tagged accordingly.

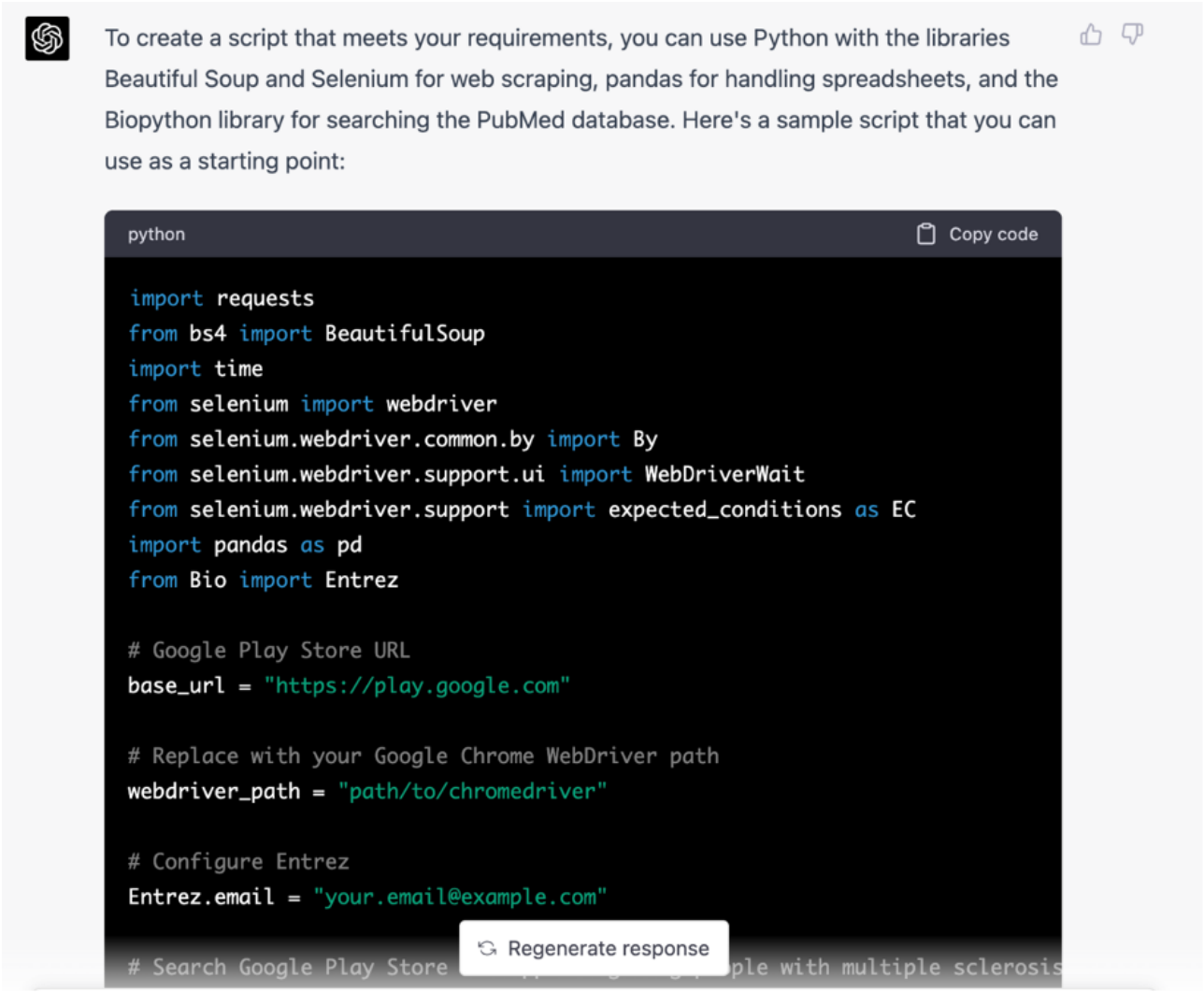
Cocreation Process

ChatGPT requires prompts to trigger interaction. For this purpose, a lay language description of the desired outcome was engineered by GG (triggering prompt), as follows:

> *I need a script that crawls the Google Play Store for apps that target people with multiple sclerosis. The script should generate a spreadsheet file that contains name, links, description and other metadata, as well as have a specific column to highlight if the description mentions any medical or clinical evidence. The script should then search Pubmed database looking for each app listed on the spreadsheet, and create a separate spreadsheet relating the apps with any scientific publication that mentions them in Pubmed.* [GG]

The generative AI's response was to provide a sample script in Python that was suggested as a starting point, as shown in Figure 2.

Figure 2. ChatGPT's first response to the initial prompt.



Following the premise of the cocreation goal, ChatGPT was asked (by GG) “I am new to Python, how do I run a Python script?” to which the response was the following:

> *To run a Python script, you need to follow these steps:*
>
> *1. Install Python on your computer, if it's not already installed. You can download Python from the official website: https://www.python.org/downloads/.*
>
> *2. Open a text editor (e.g. Notepad, Sublime Text, Atom) and create a new file.*
>
> *3. Write your Python code in the file.*
>
> *4. Save the file with a .py extension (e.g. “my_script.py”).*
>
> *5. Open a command prompt or terminal and navigate to the directory where your Python script is saved.*
>
> *6. Type “python” followed by the name of your script and press enter. For example, if your script is named “my_script.py”, you would type “python my_script.py” and press enter.*
>
> *7. Your Python script will now run, and you should see any output or results in the command prompt or terminal.*
>
> *Note: If you're using an Integrated Development Environment (IDE) like PyCharm or Visual Studio Code, you can run your script directly from the IDE by clicking on the “Run” button or pressing a keyboard shortcut.* [ChatGPT]

ChatGPT provided step-by-step guidance for the installation and setup of Visual Studio Code, along with the required libraries and dependencies. After the development environment was ready, the suggested Python script was copied and pasted to check its functionality. The code produced several errors, which were fed back to ChatGPT by GG, as follows:

> *I am getting “ModuleNotFoundError: No module named 'requests'”*
>
> *Now it says “ModuleNotFoundError: No module named 'bs4'”*

ChatGPT stated that the “requests” library was imported but not used, so it should be removed. Subsequent error messages were troubleshot with the AI. At times, the script would stop responding due to a time-out error or issues with the output. To make the debugging process easier to follow and understand, ChatGPT was prompted by GG to do the following:

> *add error messages that explain what went wrong.*

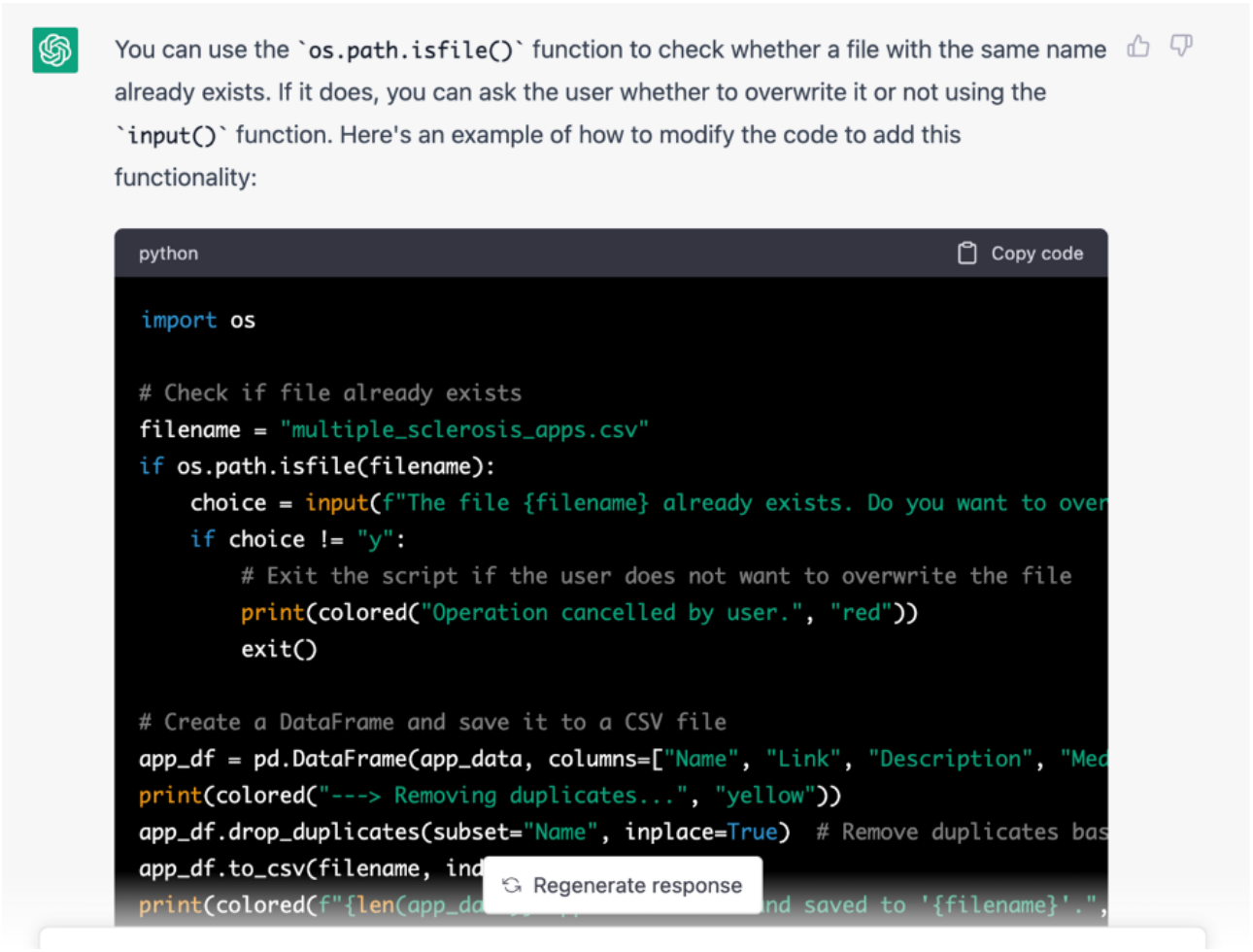
add indicators for each step that it is working.

could you add a progress bar indicator to each step?

I would like the responses to be colored.

Upon review, it was clear that some error caused the output files to be overwritten. The author requested the following: "how would you add a condition to check whether a file with the same name already exists and if it does ask whether to overwrite it?" The AI added the pertinent libraries to the script and produced an input prompt for the user (Figure 3).

Figure 3. Added libraries and code for checking existing files.



The triggering prompt expected the script to not only be searching app stores for matches with "multiple sclerosis" but also to cross-reference them with scientific publications found in PubMed. As described in the Background Studies section of this paper, this process was done by looking for keywords present in the app description. As it were, the script was designed to only search for "multiple sclerosis" and crashed when it reached the PubMed phase. ChatGPT diagnosed a time-out error due to the number of publications and provided a potential solution to improve the indicator bars:

> *The progress bar for PubMed search may not be moving because the total value for it is set to 3600, which is the number of seconds in an hour. This is the unit being used for the progress bar, not the total number of items to be searched. To fix this, you should update the total value for the progress bar to be the total number of apps being searched, which is the length of the app_df dataframe multiplied by the maximum number of PubMed IDs to be retrieved for each app (100 in this case). You can update the progress bar as follows…* [ChatGPT]

After these issues were sorted and the script could properly fetch PubMed results, more keywords were entered into the script by prompting ChatGPT with the following:

> *how would you make it so that the items in the medical_keywords list are taken from a csv file called "keywords.csv"?* [GG]

A screenshot of the final Python script running can be seen in Figure 4.

Figure 4. Screenshot of the final script running.

Evaluation

As explained in the Background Studies section of this paper, app data extraction from the Google Play Store resulted in a spreadsheet file that contained the mHealth app's name, store link, app description, developer's name, developers' URL, price, number of downloads, and app rating. During the screening phase of the studies, the research team read the apps' descriptions and flagged those that contained keywords or sentences suggestive of the evidence base for in-depth review.

The ChatGPT-generated code resulted in a CSV file that contained the app's name, store link, app description, and a column titled "Medical Evidence." There were no columns containing metadata, and the Medical Evidence column only contained "Yes" or "No," accordingly. Closer inspection revealed that the script was searching for a full match on the apps' titles in PubMed results. The resulting document was useful as an intermediate outcome but was deemed unsuitable as a final output. The overall cocreation process had a total duration of 4 hours and 39 minutes, providing a working script version available on GitHub [32].

Using the results from the ChatGPT-generated script to fully automate the process would likely require further work refining the script, either by using the steps of the background studies to base the script or by providing clearer starting prompts for the generative AI. However, leveraging this approach as a means to advance work when the software developing team was otherwise engaged was useful.

Highlights

Some highlights of this study are as follows:

- The overall cocreation process exercise had a total duration of 4 hours and 39 minutes.
- There were several misunderstandings during the interactions, not unlike the challenges one might encounter when messaging a more experienced coder.
- Structured thinking ahead of time reduced the number of misunderstandings.
- No knowledge of Python scripting was required by the author.
- The resulting output was useful to continue a systematic review but not sufficient to replace the final outputs.

Discussion

Principal Results

This study is the first to describe the cocreation process with a generative AI in developing an automated script for conducting a systematic review of mHealth apps. The study provides insights into the potential of using this kind of AI tools for researchers with little to no coding skills, and it identifies an innovative way of approaching a research problem and facilitating interdisciplinary collaborations. This study also makes a methodological contribution, expanding knowledge as it uses DSR, an approach that is not commonly used in health care and health informatics [33].

Comparison With Prior Work

The resource-intensive process and the burden that systematic reviews represent have been highlighted in the literature before. The use of multiple databases, such as MEDLINE, Embase, Cochrane Library, and Web of Science as well as clinical trial registries like ClinicalTrials.gov are common practices to increase results [34]. However, this tactic requires a lengthy deduplication process, involving long manual procedures, potentially introducing quality-affecting errors and biases [35-37]. In fact, automation attempts using AI models have been made in the past, with a focus on the deduplication problem, as seen in studies by Borissov et al [38] and Bramer et al [39].

Performing a systematic review is a common step in doctoral researchers' studies [40,41], as a means of introducing the candidate to the topic. The use of generative AI to cocreate scripts like the one presented in this study could help automate the time-consuming process, allowing researchers to focus on other aspects of the research process.

The ethical implications of using generative AI models, such as ChatGPT, to generate scientific authorship have sparked discussions [15,16]. AI's potential for assisting in academic research needs to be considered and weighed against the potential for its misuse. Although generative AI can assist in the development of a systematic review script, it is important to note that the final review still requires human oversight and input to not only assess the accuracy and relevance of the results but also ensure that the ethical principles have been followed.

Beyond research, there are wider implications for the use of generative AI in both medical education and the upskilling of the health care workforce. The need for more digital skills training for health care professionals is widely recognized [42], and other authors have further explored medical degree programs' curricula to examine how AI is included [43,44]. A recent publication explored the specific competencies needed for the effective and ethical use of AI in health care [45]. Understanding basic knowledge of AI and its applications as well as how to integrate AI into the general workflow of different tasks ranked among the top 6 key competency domains.

The role of generative AI in evolving health care education is pivotal, especially as universities adapt to its challenges. Generative AI has the potential to streamline processes like systematic reviews and clinical information retrieval, thereby allowing health care professionals to focus more on patient-centered, empathetic care and the co-design of effective treatment outcomes.

Limitations

The results of this study must be considered within its limitations. The DSR methodology was developed for this specific problem, which limits applicability in other contexts. In addition, subjectivity is a common bias present in DSR, which can make it difficult to establish the reliability and validity of the results. The main goal of DSR is to generate prescriptive knowledge, which provides guidelines on how to effectively design and implement solutions in the organizational context. However, as DSR focuses more on developing practical solutions rather than generating new theoretical insights, it was aligned with the goal of this study. DSR differs from traditional research paradigms by focusing more on creating and evaluating new solutions rather than on understanding existing phenomena. Further, while generative AI can assist in the development of a systematic review script, the result will be greatly affected by the training data used for the language model. Additionally, there may be limitations in the quality of the search results obtained from the previous studies, which only become apparent through automated processes.

Conclusions

This study outlined the cocreation process of an automated script for systematic reviews of mHealth apps, using generative AI. The study shed light on the potential of such AI tools for researchers with limited coding abilities and highlighted a novel approach for addressing research problems and promoting interdisciplinary collaborations.

Acknowledgments

GG would like to thank Prof Octavio Rivera-Romero, Dr Estefania Guisado-Fernandez, Dr Diego Giunta, Dr Analia Baum, and Prof Minna Isomursu for their collaboration and support.

This study has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement (101034252). The publication has also emanated from research supported (in part) by a research grant from Science Foundation Ireland (SFI) under grant number 16/RC/3948 and Business Finland's More Stamina Research to Business project.

The authors are grateful for ChatGPT, whose collaboration was essential for the completion and inception of this study.

Conflicts of Interest

None declared.

References

1. Ryu S. Book review: mHealth: new horizons for health through mobile technologies: based on the findings of the second global survey on eHealth (global observatory for eHealth series, volume 3). Healthc Inform Res. 2012;18(3):231. [doi: 10.4258/hir.2012.18.3.231]
2. Labrique A, Vasudevan L, Chang LW, Mehl G. H_pe for mHealth: more "y" or "o" on the horizon? Int J Med Inform. May 2013;82(5):467-469. [FREE Full text] [doi: 10.1016/j.ijmedinf.2012.11.016] [Medline: 23279850]
3. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J. Jun 2009;26(2):91-108. [FREE Full text] [doi: 10.1111/j.1471-1842.2009.00848.x] [Medline: 19490148]
4. Nagendrababu V, Dilokthornsakul P, Jinatongthai P, Veettil SK, Pulikkotil SJ, Duncan HF, et al. Glossary for systematic reviews and meta-analyses. Int Endod J. Mar 25, 2020;53(2):232-249. [doi: 10.1111/iej.13217] [Medline: 31520403]
5. Bender JL, Yue RYK, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. J Med Internet Res. 2013;15(12):e287. [FREE Full text] [doi: 10.2196/jmir.2661] [Medline: 24366061]
6. Giunti G, Giunta DH, Guisado-Fernandez E, Bender JL, Fernandez-Luque L. A biopsy of Breast Cancer mobile applications: state of the practice review. Int J Med Inform. Dec 2018;110:1-9. [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.022] [Medline: 29331247]
7. Rincon E, Monteiro-Guerra F, Rivera-Romero O, Dorronzoro-Zubiete E, Sanchez-Bocanegra CL, Gabarron E. Mobile phone apps for quality of life and well-being assessment in breast and prostate cancer patients: systematic review. JMIR Mhealth Uhealth. Dec 04, 2017;5(12):e187. [FREE Full text] [doi: 10.2196/mhealth.8741] [Medline: 29203459]

8. Adam R, McMichael D, Powell D, Murchie P. Publicly available apps for cancer survivors: a scoping review. BMJ Open. Oct 30, 2019;9(9):e032510. [FREE Full text] [doi: 10.1136/bmjopen-2019-032510] [Medline: 31575584]
9. Wanchai A, Anderson EA, Armer JM. A systematic review of m-health apps on managing side effects of breast cancer treatment. Support Care Cancer. Dec 27, 2022;31(1):86. [doi: 10.1007/s00520-022-07464-x] [Medline: 36574048]
10. Clark J, Glasziou P, Del Mar C, Bannach-Brown A, Stehlik P, Scott AM. A full systematic review was completed in 2 weeks using automation tools: a case study. J Clin Epidemiol. May 2020;121:81-90. [doi: 10.1016/j.jclinepi.2020.01.008] [Medline: 32004673]
11. Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. BMJ Open. Feb 27, 2017;7(2):e012545. [FREE Full text] [doi: 10.1136/bmjopen-2016-012545] [Medline: 28242767]
12. O'Connor AM, Tsafnat G, Gilbert SB, Thayer KA, Wolfe MS. Moving toward the automation of the systematic review process: a summary of discussions at the second meeting of International Collaboration for the Automation of Systematic Reviews (ICASR). Syst Rev. Jan 09, 2018;7(1):3. [FREE Full text] [doi: 10.1186/s13643-017-0667-4] [Medline: 29316980]
13. Benbya H, Davenport TH, Pachidi S. Artificial intelligence in organizations: current state and future opportunities. SSRN Journal. 2020:1-15. [doi: 10.2139/ssrn.3741983]
14. Weisz J, Muller M, Ross S, Martinez F, Houde S, Agarwal M, et al. Better together? An evaluation of AI-supported code translation. Presented at: 27th International Conference on Intelligent User Interfaces; 22 - 25 March, 2022;369-391; New York, NY. [doi: 10.1145/3490099.3511157]
15. Thorp HH. ChatGPT is fun, but not an author. Science. Jan 27, 2023;379(6630):313-313. [doi: 10.1126/science.adg7879] [Medline: 36701446]
16. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. Nature. Jan 24, 2023;613(7945):612-612. [doi: 10.1038/d41586-023-00191-1] [Medline: 36694020]
17. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ. Mar 08, 2023;9:e45312. [FREE Full text] [doi: 10.2196/45312] [Medline: 36753318]
18. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health. Mar 9, 2023;2(2):e0000198. [FREE Full text] [doi: 10.1371/journal.pdig.0000198] [Medline: 36812645]
19. Fuentes-Martín Á, Cilleruelo-Ramos ?, Segura-Méndez B, Mayol J. Can an artificial intelligence model pass an examination for medical specialists? Arch Bronconeumol. Aug 2023;59(8):534-536. [doi: 10.1016/j.arbres.2023.03.017] [Medline: 37055267]
20. Eysenbach G. The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT. JMIR Med Educ. Mar 06, 2023;9:e46885. [FREE Full text] [doi: 10.2196/46885] [Medline: 36863937]
21. vom BJ, Hevner A, Maedche A. Introduction to design science research. In: Design Science Research. Cases. Switzerland. Springer Cham; Sep 24, 2020;13.
22. Hevner AR, March ST, Park J, Ram S. Design science in information systems research. MIS Quarterly. 2004;28(1):75. [doi: 10.2307/25148625]
23. Giunti G, Kool J, Rivera Romero O, Dorronzoro Zubiete E. Exploring the specific needs of persons with multiple sclerosis for mhealth solutions for physical activity: mixed-methods study. JMIR Mhealth Uhealth. Feb 09, 2018;6(2):e37. [FREE Full text] [doi: 10.2196/mhealth.8996] [Medline: 29426814]
24. Giunti G, Guisado-Fernandez E, Caulfield B. Connected health in multiple sclerosis: a mobile applications review. Presented at: 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS); 22 - 24 June, 2017;660-665; Thessaloniki, Greece. [doi: 10.1109/cbms.2017.27]
25. Giunti G, Guisado FE, Dorronzoro ZE, Rivera RO. Supply and demand in mHealth apps for persons with multiple sclerosis: systematic search in app stores and scoping literature review. JMIR Mhealth Uhealth. May 23, 2018;6(5):e10512. [FREE Full text] [doi: 10.2196/10512] [Medline: 29792295]
26. Giunti G. Gamified dDesign for health workshop. Stud Health Technol Inform. 2016;225:605-606. [Medline: 27332273]
27. Giunti G, Mylonopoulou V, Rivera Romero O. More stamina, a gamified mHealth solution for persons with multiple sclerosis: research through design. JMIR Mhealth Uhealth. Mar 02, 2018;6(3):e51. [FREE Full text] [doi: 10.2196/mhealth.9437] [Medline: 29500159]
28. Giunti G, Haverinen J, Reponen J. Informing the product development of an mHealth solution for people with multiple sclerosis through early health technology assessment. Stud Health Technol Inform. Jul 06, 2022;290:1042-1043. [doi: 10.3233/SHTI220258] [Medline: 35673196]
29. Yrttiaho T, Isomursu M, Giunti G. Experiences using patient and public involvement in digital health research for multiple sclerosis. Stud Health Technol Inform. May 25, 2022;294:735-739. [doi: 10.3233/SHTI220574] [Medline: 35612194]
30. Giunti G, Rivera-Romero O, Kool J, Bansi J, Sevillano JL, Granja-Dominguez A, et al. Evaluation of more stamina, a mobile app for fatigue management in persons with multiple sclerosis: protocol for a feasibility, acceptability, and usability study. JMIR Res Protoc. Aug 04, 2020;9(8):e18196. [FREE Full text] [doi: 10.2196/18196] [Medline: 32749995]

31. Girardello A, Budde A, Wang B, Delchev I. 42matters. 2016. URL: https://www.42matters.com [accessed 2016-02-26]
32. GitHub. URL: https://github.com/guidogiunti/ChatGPT-SR-script [accessed 2024-02-07]
33. Hevner A, Wickramasinghe N. Design science research opportunities in health care. In: Theories to Inform Superior Health Informatics Research and Practice. Cham, Switzerland. Springer; 2018;3-18.
34. Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. Syst Rev. Dec 06, 2017;6(1):245. [FREE Full text] [doi: 10.1186/s13643-017-0644-y] [Medline: 29208034]
35. Qi X. Duplicates in systematic reviews: a critical, but often neglected issue. WJMA. 2013;1(3):97. [doi: 10.13105/wjma.v1.i3.97]
36. McKeown S, Mir ZM. Considerations for conducting systematic reviews: evaluating the performance of different methods for de-duplicating references. Syst Rev. Jan 23, 2021;10(1):38. [FREE Full text] [doi: 10.1186/s13643-021-01583-y] [Medline: 33485394]
37. Qi X, Yang M, Ren W, Jia J, Wang J, Han G, et al. Find duplicates among the PubMed, EMBASE, and Cochrane Library Databases in systematic review. PLoS One. Aug 20, 2013;8(8):e71838. [FREE Full text] [doi: 10.1371/journal.pone.0071838] [Medline: 23977157]
38. Borissov N, Haas Q, Minder B, Kopp-Heim D, von Gernler M, Janka H, et al. Reducing systematic review burden using Deduklick: a novel, automated, reliable, and explainable deduplication algorithm to foster medical research. Syst Rev. Aug 17, 2022;11(1):172. [FREE Full text] [doi: 10.1186/s13643-022-02045-9] [Medline: 35978441]
39. Bramer WM, Giustini D, De Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc. Sep 12, 2016;104(3):240-243. [doi: 10.5195/jmla.2016.24]
40. Riaz M, Sulayman M, Salleh N, Mendes E. Experiences conducting systematic reviews from novices' perspective. Presented at: 14th International Conference on Evaluation and Assessment in Software Engineering; 12 - 13 April, 2010; Keele, UK. [doi: 10.14236/ewic/ease2010.6]
41. Pickering C, Byrne J. The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. High Educ Res Dev. Nov 11, 2013;33(3):534-548. [doi: 10.1080/07294360.2013.841651]
42. Giunti G, Guisado-Fernandez E, Belani H, Lacalle-Remigio JR. Mapping the access of future doctors to health information technologies training in the European union: cross-sectional descriptive study. J Med Internet Res. Aug 12, 2019;21(8):e14086. [FREE Full text] [doi: 10.2196/14086] [Medline: 31407668]
43. Lee J, Wu AS, Li D, Kulasegaram KM. Artificial Intelligence in Undergraduate Medical Education: A Scoping Review. Acad Med. Nov 01, 2021;96(11S):S62-S70. [doi: 10.1097/ACM.0000000000004291] [Medline: 34348374]
44. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. JMIR Med Educ. Dec 03, 2019;5(2):e16048. [FREE Full text] [doi: 10.2196/16048] [Medline: 31793895]
45. Russell R, Lovett Novak L, Patel M, Garvey KV, Craig KJT, Jackson GP, et al. Competencies for the use of artificial intelligence-based tools by health care professionals. Acad Med. Mar 01, 2023;98(3):348-356. [doi: 10.1097/ACM.0000000000004963] [Medline: 36731054]

Abbreviations

AI: artificial intelligence
DSR: design science research
mHealth: mobile health
MS: multiple sclerosis

Edited by G Eysenbach, T de Azevedo Cardoso, K Venkatesh; submitted 12.05.23; peer-reviewed by D Carvalho, X Zhao; comments to author 14.06.23; revised version received 28.11.23; accepted 28.01.24; published 12.02.24

Please cite as:
Giunti G, Doherty CP
Cocreating an Automated mHealth Apps Systematic Review Process With Generative AI: Design Science Research Approach
JMIR Med Educ 2024;10:e48949
URL: https://mededu.jmir.org/2024/1/e48949
doi: 10.2196/48949
PMID: