Proceedings of the 9th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'23)
Brunel University, London, United Kingdom - August 03-05, 2023
Paper No. CIST 125
DOI: 10.11159/cist23.125

Automatically Enriching Content for a Behavioral Health Learning Management System: a First Look

Greg Barish[1], Lauren Marlotte[2], Miguel Drayton[2], Catherine Mogil[2], Patricia Lester[2]

[1]Allyance Logic, Inc.
San Ramon, California, USA
gbarish@allyance.io
[2]UCLA Semel Institute for Neuroscience and Human Behavior
University of California at Los Angeles
Los Angeles, California, USA
{lmarotte, mdrayton, cmogil, plester}@mednet.ucla.edu

Abstract **-** Deep generative AI models have been evolving rapidly and are being applied to assist users in many domains. We consider an initial case study of applying this technology to semi-automated content enrichment for an online learning management system used by behavioral health professionals. Our goal is to optimize the use of domain expertise while also scaling the production efficiency of learning assets for users. We identify the possible opportunities for its use, discuss potential challenges and concerns. Finally, we provide prompt engineering strategies and initial quantitative results towards semi-automating one type of rime consuming editorial task, to gauge the feasibility of our approach. Results show that there is significant promise in using such an approach, and suggest that a larger, more rigorous study is required.

Keywords**:** artificial intelligence, behavioral health, learning management systems, generative models, e-learning

1. Introduction

Deep generative AI models such as GPT-3 have demonstrated general applicability, propelling the industry into seeking ways to integrate such large language models (LLMs) into a variety of workflows across many domains. One such potential use case involves semi-automating the authoring of medical education materials [1], such as education related to behavioral health practices. Such models can be used to initiate content that is ultimately edited by human curators/editors. We are exploring this approach in the context of an online behavioral health learning management system (LMS). Our motivation was driven by the desire to scale content creation, as well as our informal experimentation with ChatGPT, which provides GPT-3.5 and GPT-4, finding that it facilitated the conversion of basic search results to readable, clean descriptions. However, as with many such informal experiments, we also witnessed slightly awkward or superfluous responses. Nevertheless, we wondered if such models could facilitate and enhance our current content development process.

We wish to deploy our work to systems such as the Wellbeing for LA Learning Center (the "Learning Center"), an LMS that is in wide use in Los Angeles County, California. Funded by the Los Angeles County Department of Mental Health (DMH), the Learning Center was developed to support the well-being of Los Angeles county's workforce and the clients they serve. Its trauma- and resilience-informed, evidence-based content is also applicable to under-resourced communities, for example during COVID-19 [2]. The LMS houses over 1,000 assets, serves over 34,000 users, and annually delivers tens of thousands of online trainings for the county's workforce, such as social workers, health care providers, early childhood workers, community ambassadors, public transportation workers, and library employees.

The audience and mission of the Learning Center are relevant in prompting generative AI models like GPT. For example, Los Angeles County is a multicultural region, where 56% of households speak a language other than English [3], most commonly Spanish [4]. The Learning Center audience is diverse; some are graduate-level trained psychologists; others are new community ambassadors, who may not have completed high school and whose first language may not be English. Creating content for such a diverse audience takes considerable time, effort, and sensitivity. Thus, it is important that the Learning Center not only attracts diverse learners but also provides accessible content, and an awareness of audience needs.

2. Enrichment opportunities

The Learning Center contains several types of assets: courses, trainings, pathways, articles, podcasts, collections, and so forth. Content is available both synchronously and asynchronously and can be browsed interactively or discovered from a search. A sample online training asset is shown in Figure 1 below. Such trainings are popular on the Learning Center, as they are typically delivered by experts, engage users about topics in depth, and offer continuing education credit.

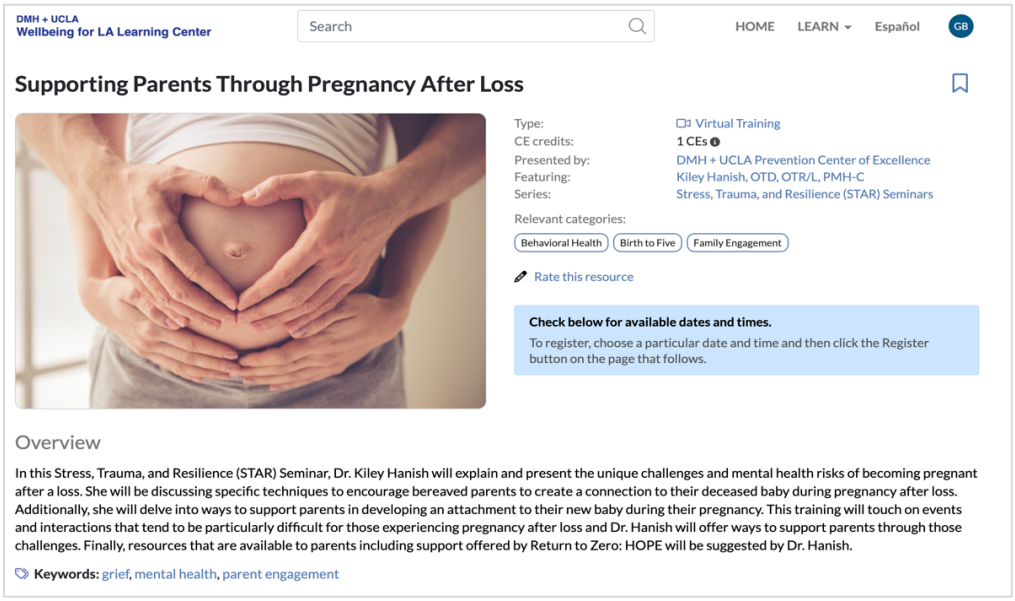

Figure 1: Detail page for an example training on the Wellbeing for LA Learning Center LMS

The figure demonstrates key elements to each page: a title, basic metadata about the type of asset it is, who is presenting it, professional continuing education credit, the content series, learning objectives (not shown), and so forth. Such metadata are important, as they set a tone and they help users understand the content and discover related assets. Much of this information is trivial for content producers to provide (i.e., no judgments/explanations needed), and some of it (e.g., duration) can be determined automatically. In contrast, there are at least four types of additional metadata that are not trivial to provide: (a) thumbnail image, (b) categories and keywords, (c) an overview, and (d) the learning objectives.

Automatically generating categories and keywords is a classic supervised machine learning problem. Content providers can mark-up several examples and models can be learned to predict which categories and keyword should be added, using the terms in the title or description as features (e.g., via a bag-of-words model). It is also possible that we can use generative AI techniques here (e.g., few shot vs traditional supervised learning) to mitigate the need for extensive markup.

Meanwhile, a more immediate use of generative AI would be to generate the thumbnail, overview, and learning objectives. The thumbnail image is a simpler task because one can choose from a set of stock photos, for example, one can imagine such images also being generated through prompted stable diffusion [6] (e.g., “illustrate a homeless encampment under a freeway overpass”).

The overview and learning objectives are the most challenging to create. They are also related; one could either generate an overview from learning objectives or one could extract objectives from an overview. It seems more realistic to assume the former, and to focus the generative AI on this task. In addition, the latter is the key text users will consider and is the most time-consuming data to generate; thus, it has the greatest impact towards scaling production efficiency.

Overviews range from a single sentence to multiple paragraphs. The ideal overview from an editorial perspective is one that is comprehensive, concise, and engaging. Typically, the content producer writes the overview. While they may be an expert on health education, they may not be skilled at promotional writing. It takes both time and some expertise to write an engaging overview; our goal is to alleviate both by using a LLM to “jump start” this process.

3. Methodology and experimental design

To begin exploring whether generative AI could help generate and/or refine asset descriptions, we constructed a simple evaluation where we asked participants to (1) write overviews for two different training topics and then (2) review and edit two separate automatically generated overviews. We asked participants to provide the final version of the overview (as written or edited) and to measure the time it took for them to complete the task and time required for background research.

Our participants were four employees of the UCLA Division of Population Behavioral Health (DPBH) who perform a range of job functions (e.g., curriculum developer, social worker, etc.) with a range of experience working at the DPBH. These employees, along with many others, help develop and deliver trauma- and resilience-informed mental health training material to the Los Angeles County workforce (e.g., DMH) and across the diverse communities they serve.

3.1 Prompt engineering

To automatically generate overviews, we prompted ChatGPT by merging metadata with pre-defined templates 1 and 2 that control for length, accessibility, and coach the LLM into how to structure a response:

1. In one paragraph of no more than 5 sentences, for reading level of 11th grade, explain why <**TITLE**> is an important topic.
2. In one paragraph of no more than 5 sentences, for reading level of 11th grade, describe an upcoming training session by <**PRESENTER**>, that starts with "Join this training" and includes the following learning objectives: <**LEARNING OBJECTIVES**>

We then concatenated the results of (1) and (2) as the generated overview for participants to review.

As a motivating example, consider the training in Figure 1. The title of the training is "*Supporting Parents Through Pregnancy After Loss*", the presenter is "*Kiley Hanish, OTD, OTR/L, PMH-C*" and the learning objectives are as follows:

- *Identify specific techniques to encourage bereaved parents to create a connection to their deceased baby during pregnancy after loss.*
- *Describe the mental health implications during pregnancy after perinatal loss.*
- *Explain three strategies to help patients decrease stress reminders during medical appointments.*
- *Identify two suggestions to support parents to develop an attachment to their new baby during a subsequent pregnancy.*

Using the prompt templates above, we were able to use ChatGPT to automatically generate the following overview:

> Supporting parents through pregnancy after loss is a critical issue because the experience of pregnancy loss can be emotionally devastating and may affect subsequent pregnancies. Parents may experience anxiety, fear, and depression during subsequent pregnancies, and their mental health may be at risk. Adequate support and resources during this time can help parents cope with their grief and anxiety and may prevent adverse outcomes for both parents and their babies. Addressing this issue is crucial to ensuring that parents have the resources and support they need to navigate pregnancy after loss and to promote healthy outcomes for both parents and their babies.
>
> Join this training session led by Kiley Hanish, OTD, OTR/L, PMH-C, where you will learn about supporting bereaved parents during pregnancy after loss. This training will cover specific techniques to encourage parents to connect with their deceased baby during pregnancy, as well as the mental health implications of perinatal loss during subsequent pregnancies. You will also learn strategies to help patients decrease stress reminders during medical appointments and suggestions to support parents in developing an attachment to their new baby during a subsequent pregnancy. By the end of this session, you will be equipped with practical skills and knowledge to support bereaved parents during pregnancy after loss.

Figure 2: Automatically generated overview via ChatGPT, based on concatenation of results from prompts 1 and 2.

4. Initial results

Results for all participants are shown below. Three of four required less time to edit, and two showed gains of over 50%. Overall efficiency improved by 37.5%, on average. We note that in one case, no edits to a suggested overview were needed.

Table 1: Total time required to write and research two overviews vs. time to edit and research two overviews.

Participant	Minutes to write	Minutes for research (written)	Minutes to review/edit	Minutes for research (edited)
A	19	6	15	5
B	37	0	15	0
C	63	15	22	0
D	60	5	60	2
Total	179	26	112	7
Average	44.75	6.50	28.00	1.75

Figure 3 shows a marked-up fragment for one specific review task. This is a representative example that illustrates the types of changes made during the reviewing process. As shown, only a few minor edits were made to the generated text.

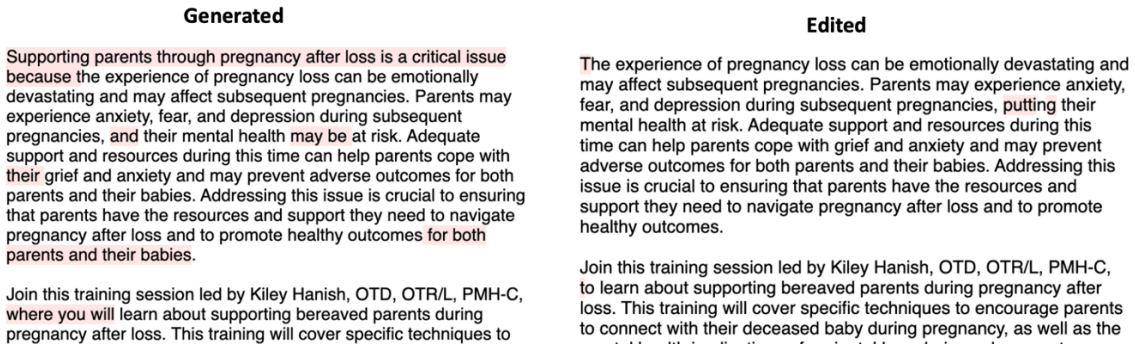

Figure 3: Comparison between generated and edited versions of text for one participant

5. Discussion

The results show that generated overviews could greatly speed up content production and suggest the need for a larger study. The results also show little/no background research was needed by the participants, possibly indicating that they were comfortable with the material and that this was likely not a confounding factor in writing or editing. Finally, we note that results from participant D are curious, and may be an outlier (without participant D, savings are more than 50%).

Acknowledgements

We acknowledge the Los Angeles County Department of Mental Health and its Public Partnership for Wellbeing mission, which directs and supports the Learning Center, and allows us to provide this content to county residents.

References

[1]. Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted Medical Education Using Large Language Models. PLOS Digital Health, 2(2). https://doi.org/10.1371/journal.pdig.0000198

[2] Ijadi-Maghsoodi, R., Harrison, D., Kelman, A., Kataoka, S., Langley, A. K., Ramos, N., Cugley, G. M., Alquijay, M. A., Tate, K., Lester, P., Mogil, C., Franke, T. M., & Bath, E. (2020). Leveraging a Public–Public Partnership in Los Angeles County to Address COVID-19 for Children, Youth, and Families in Underresourced Communities. Psychological Trauma: Theory, Research, Practice and Policy, 12(5), 457–460. https://doi.org/10.1037/tra0000880

[3] U.S. Census, 2022 [Online]. Available: https://www.census.gov/

[4] Los Angeles Almanac [Online]. Available: http://www.laalmanac.com/