GAIDE: A Framework for Using Generative AI to Assist in Course Content Development

Ethan Dickey
Department of Computer Science
Purdue University
West Lafayette, IN, USA
dickeye@purdue.edu,

Andres Bejarano
Department of Computer Science
Purdue University
West Lafayette, IN, USA
abejara@purdue.edu,

2024 IEEE Frontiers in Education Conference (FIE) | 979-8-3503-5150-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/FIE61694.2024.10893132

***Abstract*—*Contribution*: This research-to-practice full paper presents "GAIDE: Generative AI for Instructional Development and Education," introducing a pragmatic and systematic framework for employing Generative AI (GenAI) in the development of educational content. Unlike existing frameworks, GAIDE emphasizes practical applicability for educators, facilitating the creation of diverse, engaging, and academically sound materials. The novel aspect of our approach lies in its detailed methodology for integrating GenAI into curriculum design processes, thereby reducing instructors' workload and improving the quality of educational materials. Through GAIDE, we contribute a distinct, adaptable model for leveraging technological advancements in education, providing a foundational step towards more efficient and effective instructional material development.**

***Background*: The motivation for our study emerges from the increasing demand for innovative and engaging educational content, coupled with the notable rise in Generative AI (GenAI) utilization among students for academic tasks. Our investigations reveal that nearly half of students engage with GenAI tools for completing homework assignments, highlighting a significant shift in study behaviors and the potential for technology to shape educational practices. This scenario presents a dual challenge for educators: to adapt to and incorporate these emerging technologies into their teaching methodologies, not merely to keep pace with technological advancements but to leverage them in fostering a more dynamic and inclusive learning environment. This research addresses these challenges by offering a concrete, adaptable solution, aiming to reshape the landscape of educational content creation and its application across diverse learning settings.**

***Intended Outcomes*: The primary objectives of introducing GAIDE are to: 1) Streamline the course content development process for educators, 2) Foster the creation of dynamic, engaging, and varied educational materials, and 3) Demonstrate the practical utility of GenAI in enhancing instructional design, potentially setting a precedent for its adoption in diverse educational contexts.**

***Application Design*: GAIDE was conceived out of a necessity to efficiently harness GenAI's potential in education. The application design is rooted in constructivist learning theory and TPCK, emphasizing the importance of integrating technology in a manner that complements pedagogical goals and content knowledge. Our Outcomes-Based Course Design approach aids educators in crafting effective GenAI prompts and guides them through interactions with GenAI tools, both of which are critical for generating high-quality, contextually appropriate content.**

***Findings*: Preliminary evaluation of GAIDE indicates its effectiveness in mitigating the instructional challenges associated with content creation. Educators reported a significant reduction in the time and effort required to develop course materials, without compromising on the breadth or depth of the content. Moreover, the use of GenAI has shown promise in deterring conventional cheating methods, suggesting a positive impact on academic integrity and student engagement.**

***Index Terms*—Generative AI (GenAI), course content development, content generation framework, instructional workload reduction, instructional design, course design, faculty development.**

I. Introduction

In an era where academic integrity is challenged by the widespread availability of unauthorized solutions online, it has become increasingly important for instructors to generate novel and diverse content each semester. However, in the past year, even freshly created content has come under threat from Generative AI (GenAI) models, which purport advanced language comprehension and question-answering capabilities. While opinions on the actual problem-solving capacities of these models vary, as evidenced by both positive [1]–[3] and negative [4] reports, there is a consensus among instructors globally about GenAI's potential to significantly disrupt academia, especially in the realm of Computer Science (CS) [5]–[8].

A recent preliminary anonymous survey of our summer students enrolled in a data structures and algorithms course revealed that at least 48.5% utilized GenAI for homework assistance [9]. This figure might even be an underestimation, considering potential hesitations in self-reporting. Recent literature and news articles suggest that the actual number of students using GenAI and the variety of their methods might be more extensive than previously assumed by educators [10], [11]. Anecdotally, many students appear comfortable sharing insights about their peers' frequent use of GenAI. Their detailed knowledge of these tools suggests familiarity, although direct admissions of personal use remain rare. All of these factors underscore the growing influence of GenAI in the academic realm but also hint at its potential applications for educators. As the academic community grapples with the ever-growing demand for fresh, diverse, and high-quality course content, the traditional methods of content creation often fall short, being both time-consuming and occasionally monotonous.

While the challenges posed by GenAI to academic integrity are evident, it is crucial to recognize the transformative potential these tools offer. The same capabilities that enable students to seek unauthorized assistance can, if channeled correctly, revolutionize the way educators create and deliver content. Content Creators (CCs), in particular, find themselves uniquely positioned in this landscape: they stand to benefit directly from GenAI's advantages in content generation, such as creating diverse, high-quality, and relevant content, yet they also confront the challenges it presents to academic integrity. This dual impact places CCs at a critical juncture. Without a structured approach, they risk either not fully harnessing the benefits of GenAI or inadvertently amplifying its challenges [5], [12], [13]. Therefore, a systematic method is essential to guide CCs in navigating the complexities of GenAI, ensuring they can effectively leverage its benefits while being aware of its potential pitfalls. This burgeoning potential brings forth a pivotal question: **how can educators harness the power of GenAI to create meaningful course content efficiently?** It's worth noting that the challenge of unauthorized student assistance with GenAI is a significant concern and is addressed in detail in our separate parallel work [9].

In light of the growing utility and student interest in GenAI tools, this paper aims to develop a generalized approach for CCs in academia to harness the potential of these tools. Specifically, we study natural language GenAI models which incorporate memory of conversation (the authors tested the framework using ChatGPT 3.5 and 4.0[1], Bard[2], Llamma[3], and Microsoft Bing's Copilot[4] and found ChatGPT 4.0 to produce the best overall results with our framework as of November 2023). We begin by substantiating our observations on GenAI's utility and student interest, providing both rationale and illustrative examples. After motivating why instructors should be interested in mastering these tools, we introduce the *GenAI Content Generation Framework*. Remarkably, up until the submission date of this paper, we have encountered no preexisting framework specifically aiding educators in incorporating GenAI tools within the content creation process. Most existing literature often adopts a broader academic perspective, emphasizing empowering students to critically assess these tools and fostering ethical conversations [14]–[17]. In contrast, our approach stands out by specifying a particular workflow and offering practical recommendations tailored for CCs. Furthermore, we provide an explicit rationale for its adoption, ensuring that CCs not only understand the 'how' but also the 'why' behind each step. This framework, characterized by a flow of steps and guiding perspectives, assists CCs in harnessing GenAI efficiently to systematically and practically get high-quality results. While each interaction with GenAI models is unique due to its dynamism and creativity (and pseudo-random hallucinations), our framework serves as a consistent guide, offering practical strategies to achieve precise content outcomes. Subsequent sections delve into broader considerations for engaging with GenAI and conclude with reflections on the framework's implications and potential future research directions.

[1]https://chat.openai.com/
[2]https://bard.google.com/
[3]https://ai.meta.com/llama
[4]https://www.bing.com/

II. WHY SHOULD EDUCATORS CONSIDER GENAI FOR COURSE CONTENT CREATION?

GenAI excels at assisting experts in executing simple tasks more efficiently and, in many cases, with enhanced outcomes, as detailed in subsequent subsections. While experts can harness the power of GenAI effectively, non-experts (such as students) face distinct challenges [18]. They may struggle to verify, correct, and selectively use results, leading to potential pitfalls such as the reinforcement of incorrect knowledge. Other significant risks for non-experts include a lack of genuine skill development, diminished problem-solving capabilities, and an over-reliance on the tool, which can hinder self-sufficiency.

The advent of Large Language Models (LLMs), such as OpenAI Codex, marked a turning point in CS education, a trend that gained considerable momentum with the GenAI explosion of 2022 and 2023. Pioneering studies, like those by Sarsa *et al.* [19], explored using these models to craft programming exercises, solutions, and test cases. These early experiments demonstrated the potential for LLMs to generate innovative and relevant educational content, albeit requiring meticulous instructor oversight and adjustments. This initial exploration laid the groundwork for today's educators, who, with advanced LLM capabilities and more refined techniques in prompt engineering, can now much more efficiently and effectively create and tailor course materials. This evolution reflects a growing sophistication in the application of GenAI tools in education, setting the stage for their broader utilization, as detailed in the subsequent sections.

Recent studies and observations have underscored the multifaceted capabilities of GenAI [5], [6], [13], [20]–[22]. These models excel in generating large volumes of content swiftly, drafting documents, automating repetitive interactions, providing interactive explanations, generating and evaluating code, and producing (arguably) creative content. For course CCs, these capabilities translate into tangible benefits. They can produce content more rapidly, brainstorm and ideate with the assistance of the model, and refine content iteratively based on conversational or formal feedback. The direct advantages of harnessing these properties include faster course development, increased content flexibility (styles, levels, etc.), the ability to update content frequently, scalability of personalized content, cost-effective content creation, support for experimental approaches, and automation of repetitive tasks.

While many of these benefits provide intrinsic motivation for course CCs to explore these tools, we also present a student-centered perspective. We posit that (a) students are highly likely to experiment with these tools, and (b) these tools are here to stay. To truly understand the implications of these tools within academia, instructors would benefit from firsthand experience with them.

To substantiate point (a), we highlight several observations. First, the detection of GenAI tool usage for academic dishonesty has proven challenging, given its capacity to produce seemingly original content (including craftily rephrasing and editing) [23], [24]. This makes it arguably more elusive than traditional forms of academic dishonesty. Second, conventional deterrents against academic dishonesty face challenges in the context of GenAI due to its novelty and accessibility. Unlike instances where students source content from the internet, GenAI-generated content is unique, making detection and prevention more complex. Lastly, there's a growing consensus that these tools will be permissible and even prevalent in future workplaces, a sentiment echoed by CS instructors globally [5]. We can see these conversations happening even now, from legal perspectives [25], [26] to strong corporate stances on both sides of the line [27], [28]. Given this trajectory, students often pose a challenging question: "Why *shouldn't* I use GenAI?" We delve deeper into this dilemma in a separate study, where we introduce a method to address this very concern [9]. In this paper, we contend that the rapid adoption of GenAI by students and industries underscores the urgency for educators and course CCs to fully grasp its academic ramifications. This ensures they remain at the cutting edge of this dynamic academic environment. Presently, the most effective method to achieve this comprehension is through direct engagement and application of these tools.

To support (b), we share a quote from a recent investigation into GenAI by McKinsey [29], one of the oldest and largest of the world's most prestigious management strategy consulting firms, which we believe aptly summarizes current views on the future of GenAI: *"All of us are at the beginning of a journey to understand [GenAI's] power, reach, and capabilities... [This research] suggests that [GenAI] is poised to transform roles and boost performance across functions such as sales and marketing, customer operations, and software development. In the process, it could unlock trillions of dollars in value across sectors from banking to life sciences."*

In light of the above discussions and the insights from McKinsey, it becomes evident that GenAI is not just a fleeting technological trend but a transformative force poised to reshape various sectors, including academia. As educators and stakeholders in the academic community, it is incumbent upon us to not only recognize the challenges posed by GenAI but also to proactively engage with it. By doing so, we can harness its potential for positive educational outcomes while mitigating risks. This proactive approach will ensure that academia remains adaptive, relevant, and prepared for the evolving landscape that GenAI presents. As we navigate this new frontier, collaboration, continuous research, and open dialogue will be paramount in guiding our path forward.

III. GAIDE: A GenAI Content Generation Framework

Properly motivated, we dive into the foundational principles and structure of our GenAI Content Generation Framework. This framework is most succinctly described as a sequence of steps, each accompanied by its respective perspective, guiding the integration of GenAI in collegiate-level course content development. While our approach mirrors traditional content development processes, we have specifically aligned it with the Outcomes-Based Course Design methodology [30]. This choice is influenced by Ziegenfuss's summary of observed approaches to course design [31] and the widely recognized Backward Course Design Model [32].

Building further on foundational methodologies, the GAIDE framework is deeply embedded within constructivist learning theory [33], which posits that learners construct knowledge through active engagement with their environment. This aligns with our aim to employ GenAI to create dynamic and interactive learning experiences that are custom-tailored to meet the diverse needs of students. Moreover, we integrate the principles of Technological Pedagogical Content Knowledge (TPCK) [34], ensuring that technology through GenAI supports and actively enhances pedagogical practices and content delivery. This strategic alignment guarantees that GAIDE supports educational objectives while also advancing the practical application of these theories in real-world educational settings.

By rooting the GAIDE framework in these diverse theories, we provide a structured approach to content generation and enrich the ongoing discourse on integrating technology in education. This theoretical grounding ensures that our framework not only streamlines educational processes but actively enhances them, creating an environment where educators and students alike can thrive. The application of constructivist and TPCK principles underscores a proactive approach to improving educational practices, ensuring that technology serves as a bridge rather than a barrier to effective learning."

After designing course outcomes, the framework moves to course content draft generation. This is followed by iterative refinement, initially on a macro-scale and subsequently on a micro-scale. Recognizing the diverse nature of course content and the versatile capabilities of GenAI, we've categorized our approach into two primary content types: Lecture-Style and Problem Creation. While our framework doesn't adhere strictly to a specific design model, it offers flexibility and adaptability to various educational contexts. We wrap up this section with additional recommendations for harnessing GenAI, which, while valuable, didn't seamlessly fit within the main structure of the framework. Figure 1 contains a visual illustration of the following process, and the Supplementary Materials (available on the arXiv preprint version [35]) contain a simple example of the framework using ChatGPT 4.0 and Microsoft Bing's Copilot.

It is important to note that our framework primarily targets the creation of course content for undergraduate collegiate CS courses. Nevertheless, the versatility of this framework allows for its adaptation to a variety of educational contexts. As course material delves deeper and becomes more advanced, there's an increased likelihood of GenAI models producing inaccurate or misleading content. Despite these potential pitfalls, GenAI models remain invaluable tools for content design and ideation, even in advanced courses.

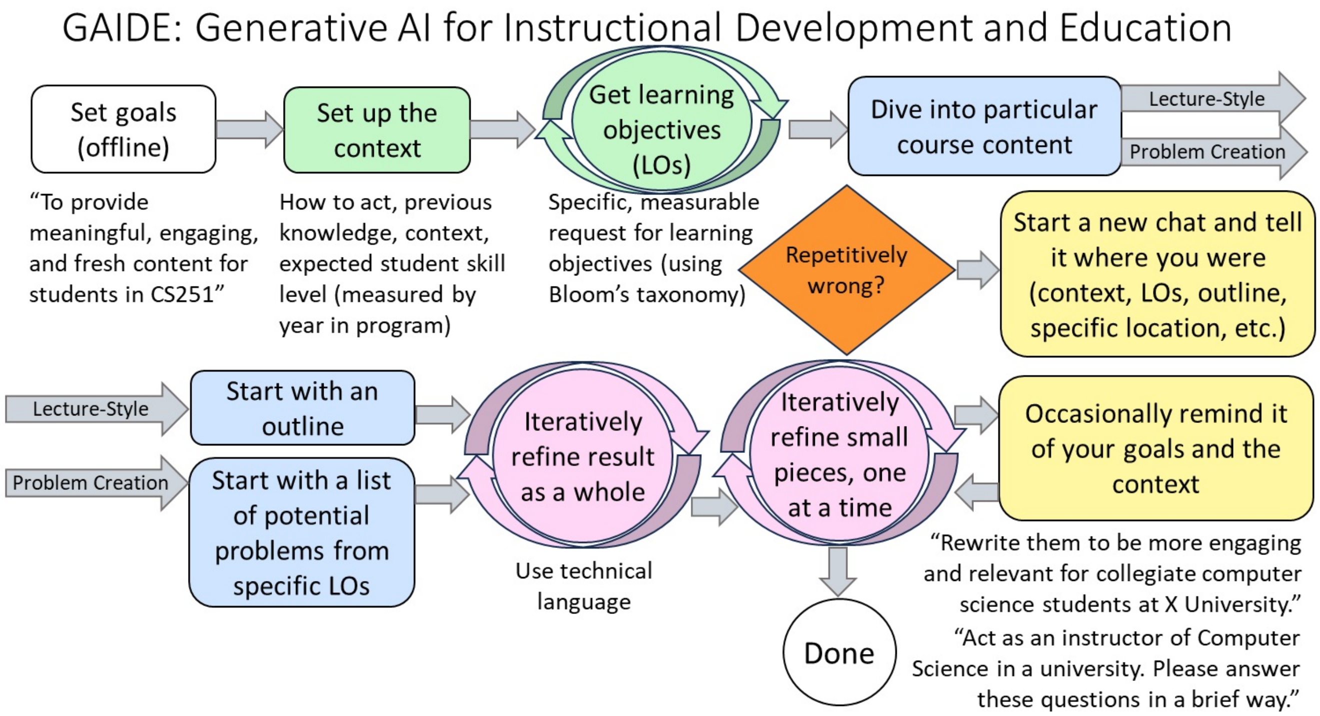

Fig. 1. A illustrative overview of the GAIDE process. Please refer to Section III for further discussion of each component. Please refer to the Supplementary Materials (available on the arXiv preprint version [35]) for a simple example of the framework using ChatGPT4.0 and Microsoft Bing's Copilot

A. Setup

1) Set Goals: The first and only offline (without GenAI) step in our framework is the establishment of clear goals. These goals not only guide the narrative presented to the GenAI model but also serve as a constant reminder of the intended direction. For instance, when we were designing new homework and tests for a summer class, one of our primary goals was: "to provide meaningful, engaging, and fresh content for students in CS251." While these goals can be as broad or specific as desired, it is imperative that they contain specific, measurable outcomes. To clarify, a measurable outcome might be something like "students should be able to implement a basic sorting algorithm" rather than a vague goal like "students should understand sorting." In the process of setting and working towards these goals, it's also worth noting the importance of self-reflection. Challenging one's own assumptions about what these models can achieve can be beneficial, as it fosters adaptability and encourages a proactive approach to potential limitations and innovations of GenAI. By regularly reassessing and adjusting one's expectations, educators can better harness the evolving capabilities of these models. For our team, the challenge lay in the "fresh content" aspect. Given the numerous semesters of content we had accumulated, devising non-repetitive material had proven to be a significant hurdle. Contrary to our initial assumptions, we found that these tools were particularly adept at generating fresh content, often surpassing our own initial expectations.

2) Set Up the Context: Transitioning to the online component of our framework, the first essential step with any GenAI session is to set up the context. Broadly, this involves providing the model with details akin to what a new DTA[5] would require to design the course content. Specifically, to get quality results, one should tell the model how to act (known as the 'persona prompt pattern' [13]), the topic under consideration, any prior knowledge the students possess, and any pertinent student demographics (usually just their academic year). For example, high schoolers in a collegiate summer program possess a different skill set and background compared to sophomore CS majors. Providing the model with such demographic details enables it to tailor content with greater precision (additional benefits surrounding demographics and perspectives are discussed in IV-A). Our approach to context setup aligns with the 'context manager prompt pattern' introduced by White *et al.* [13]. However, our framework offers specific recommendations on context details and their placement within the GenAI session.

3) Generate Learning Objectives: Concluding the setup stage, the generation of learning objectives (LOs) is paramount. The centrality of LOs in guiding subsequent interactions with GenAI cannot be overstated. This emphasis on LOs is rooted in our adoption of the Outcomes-Based Course Design methodology [30], [31], which prioritizes determining

[5]We refer to TAs who develop content for courses as one of development TAs, dev TAs, or DTAs.

student outcomes from the outset. For a deeper exploration of LOs as foundational to effective teaching, refer to [36].

When instructing GenAI to generate LOs, it's essential to make specific, measurable requests. Additionally, directing the model to employ professional language, such as terms from Bloom's revised taxonomy [37], is crucial. This not only implies a quality standard for the LOs based on other LOs which use the same language but also communicates the desired specificity and formality to the model. Below are sample prompts to guide GenAI: *"Act as an instructor of Computer Science in a university. The current topic we are covering is Binary Search Trees. Students know discrete math, basic programming, pointers, primitive data structures, and runtime algorithm analysis. Students are in their second year of studies. Give me five LOs for the current topic. Use Bloom's revised taxonomy verbs for the objectives."*

As we delve into design sessions with GenAI, it is important to grasp how best to interact with the model. The generation of content, including LOs, is iterative, with initial outputs often not aligning perfectly with the specific course goals. This is a key reason for requesting a larger number of LOs than might be immediately needed. From this broader set, educators can select a subset that resonates more closely with their goals and the targeted student profile. If the initial LOs don't fully meet expectations, it is beneficial to work with this chosen subset, providing the model with targeted feedback for refinement. For instance, if an objective needs rewording to better fit within Bloom's revised taxonomy or to emphasize application over theory, such iterative feedback can steer the model accordingly. An illustrative interaction might be: "I like learning objectives 2, 3, and 7, but I want you to reword 2 to be higher in the Cognitive Process dimension of Bloom's revised taxonomy and adjust 3 to prioritize application over theoretical understanding." Lastly, if you edit something the model gave you, it is imperative to let it know what you did, even if you do not want its feedback. This allows it to learn your preferences over a session.

B. Course Content Rough Draft

Once refined LOs are established, the focus shifts to the specific course content. While course content can take various forms, for the purposes of this framework, it is categorized into two primary types: Lecture-Style and Problem Creation. A rough draft of the former refers to an outline that should be correct only from a very high-level view, while a rough draft of the latter refers to lists of potential problems, from which only a select subset will be chosen for iterative refinement. Following the creation of these rough drafts, we move to iterative refinement from a macro perspective.

C. Macro Refinement

It is critical to understand the distinction we make between macro and micro refinement. Macro refinement focuses on adjusting the content draft holistically, refining entire sections to ensure alignment with the established expectations, goals, LOs, and context. Conversely, micro refinement zeroes in on specific parts of the draft, addressing them with precision and individualized context. At the macro stage, it is acceptable if certain components don't fully meet expectations, as long as they are broadly correct; finer adjustments are reserved for the micro refinement phase.

During macro refinement, interactions with the model should mirror discussions with a dev TA[5], employing technical terminology and offering feedback on the draft's strengths and weaknesses.

With this foundation, we present considerations for each category and provide guidance on when to transition to micro refinement.

1) Lecture-Type: In any content type which requires a larger flow and organization, outlines are typically a good starting place. Request an outline appropriate for your content and specify which LOs you wish to use in it. From there, you should iteratively refine the outline as a whole. Here are some things to consider:

- **Duration:** How long is the lecture?
- **Associated Tasks:** Are there any associated tasks outside the lecture, such as grading in-class activities or preparing materials?
- **Pre- and Post-Lecture Activities:** Are there any activities before or after the lecture?
- **Subtopics:** Are there any subtopics that are crucial to cover?
- **Specific Activities**: Are there particular activities, such as quizzes or group work, that are planned?

Each of these considerations, if relevant, should be communicated to the model during the refinement process (e.g. *"Rewrite the outline to fit into 50 minutes"*). Addressing all considerations simultaneously or tackling them individually are both valid approaches, with both methods showing comparable success rates in our experience (e.g. *"Try again, but make them all higher in Bloom's Taxonomy"* versus *"I like 2 and 3, but 1 and 4 don't match with my goals, please make 1 harder and 4 easier, to make this more appropriate for a timed exam setting"*).

Additionally, it can be helpful and incredibly insightful to let GenAI brainstorm on different lecture components. Moreover, encouraging the model to deviate from routine approaches aligns with the active learning principle of unpredictability [36]. For a deeper dive into brainstorming and alternate perspectives, refer to IV-A.

Once the outline aligns with the lecture's objectives and approach, it's time to delve into specific sections, similar to the process for problem sets. At this juncture, managing expectations is crucial. The model might produce a near-perfect outline or offer just a few valuable section ideas. Regardless, these outputs serve as a foundation for the micro refinement phase, where the lecture content can be further tailored for a complete draft or cherry-picked for integration into an existing draft.

2) Problem Creation: For content types that result in lists, such as problems or activities, the emphasis is on the diversity of responses and the general applicability of the items, hereafter referred to generically as "problems." Unlike Lecture-Type content, only a subset of the generated list will be included in the final version. During the micro refinement phase, the focus will be on a select few top problems rather than an exhaustive review of every item as is done for outlines.

To generate a draft, ask the model to generate a specific number of problems, ideally more than required, based on the selected LOs. If the generated problems do not align with the intended goals and context, provide feedback to the model, emphasizing the desired attributes. Key variables to consider include:

- **Answer style:** Multiple choice, short answer, etc.
- **Depth:** Desired level(s) within Bloom's Revised Taxonomy [37] or Webb's Depth of Knowledge [38].
- **Theme:** Leveraging GenAI's strength in creativity.
- **Topical theme:** Ensuring the problems address specific skills within the topic.

The macro refinement process can vary based on the alignment of the generated problems with the goals. For instance, if only a few questions are relevant, instruct the model to generate problems similar to the relevant ones. If the model's responses become repetitive or misaligned, providing a sample problem can be particularly helpful in resetting its misconceptions about what is desired. Several other examples exist, but we leave it to the discretion of the reader to respond to the model within the spirit of this framework.

The boundary between macro and micro refinement in Problem Creation can be nebulous. A good indication of the transition to micro refinement occurs when the focus shifts to refining a select subset of problems, while disregarding the rest.

D. Micro Refinement

Before delving into micro refinement, educators should be satisfied with the overall structure and general alignment of the content draft with the intended goals and context. From there, the focus shifts to perfecting and elaborating on each segment of the content. This involves examining each section, subsection, or question individually, informing the model of the specific focus on that part.

During these small, focused refinements, it is crucial to maintain context specificity. For instance, if refining a problem on frequency analysis in Huffman Coding emphasizes the theoretical aspect over the practical skill, this context should not inadvertently influence a subsequent question on creating the Huffman Coding Tree, where hands-on application is integral to understanding the theory. See III-E for further discussion on context integrity.

Alongside specifying the content segment for refinement, provide the model with clear modification instructions. The granularity of these requests can vary widely. For instance, feedback can range from broader directives like, *"make this question more challenging; it's currently too straightforward for a homework assignment,"* to more specific directives such as, *"I find the word 'target' unsuitable; could you rephrase that sentence, please?"* Such detailed interactions ensure the content aligns closely with the educator's vision and objectives.

The micro refinement process is inherently flexible, adapting to the educator's vision for the final content. While there's no singular correct approach, we offer two potential workflows for each content type.

1) Lecture-Type: Once satisfied with a particular segment of the outline, educators can request the model to generate a detailed script or essential talking points. These scripts serve as a roadmap for the lecture, ensuring a coherent flow and comprehensive coverage of the topic. As the script is generated, it's crucial to assess its alignment with the LOs and its potential to engage students. Any misaligned or inaccurate sections should be refined iteratively. Once sections are polished, the model can merge the script cohesively. This refined script can be paired with lecture slides, multimedia, or classroom activities to enrich the educational journey.

2) Problem Creation: In the refinement of individual problems, several strategies emerged as particularly effective. Among these, iterative rewording of the problem by the model, adjusting the problem's difficulty, and embedding problems within a narrative stood out. Notably, while most narrative integrations are typically part of micro refinement, exceptions arise when a global storyline is employed, where each problem contributes to a larger narrative.

During problem refinement, educators can instruct the model to answer the question from an undergraduate's perspective (or some other demographic(s)). This approach can reveal common pitfalls and misconceptions, guiding further refinement. Such insights are invaluable, especially for problems where students must comprehend and act without instructional support. Misunderstandings can arise in activities without instructional support, often over aspects not intended for assessment. Rewording the problem can clarify these points, enhancing comprehension.

As the refinement progresses, educators should request the correct answer (usually requesting in a brief or concise way to avoid the overly-wordy responses certain models are prone to). If accurate, this stage is suitable for generating a rubric. While not always standard practice, rubrics ensure a consistent and fair assessment. For a deeper understanding of rubric benefits, Ragupathi and Lee [39] provide valuable insights.

If the model's answer is incorrect, educators can guide it towards the right solution. Persistent errors can highlight tasks that challenge the model, incredibly powerful information which offers unique teaching opportunities about the limitations of GenAI (as address in [9]). Regardless, perfection in answers isn't the goal. The quality of a rubric doesn't hinge on the exact correctness of the provided answer, but rather on sound evaluation criteria.

While the iterative refinement process may seem labor-intensive, it's essential to note that these steps are integral to traditional content creation. GenAI often streamlines this process, reducing the time and effort typically required.

E. Maintaining Contextual Integrity in Iterative Refinement

In the process of iterative refinement, two critical concerns arise: *context blending* and *loss of focus*. The term *'context blending'* pertains to the merging of contexts between different steps that do not inherently share the same context. On the other hand, *'loss of focus'* denotes the model's diminishing capacity to execute the tasks and refinements as directed. To mitigate context blending, one should delineate clear transitions between sections and intermittently reinforce context-specific labels, such as LOs, sections, or parts under consideration. Furthermore, prompting the model to reiterate these context-specific labels as it transitions to a subsequent section can enhance its adherence to the expected context. Most crucially, if a loss of focus is observed, it may be beneficial to initiate a complete context reset. This can be achieved by starting a new session, reintroducing the context, and specifying the current stage of the process—by providing, for instance, the working outline, LOs, global context, and the specific section in focus.

F. Consolidating Generated Content

Given the intricacies of generating comprehensive course content, there is a high likelihood that it will be necessary to initiate multiple sessions with GenAI. Due to this segmented approach, the model may not seamlessly provide an overarching summary of all content components in a singular response. As a practical measure, CCs are encouraged to maintain a dedicated document to collate and refine the selected outputs, ensuring a cohesive and well-structured course assembly.

1) GenAI in Comprehensive Course Planning: While the primary focus of this framework is not on constructing an entire course from the ground up using GenAI, the potential of these models in the realm of course planning warrants attention. Once LOs, course outlines, and assessments are established, GenAI models demonstrate a commendable proficiency in devising comprehensive course activity plans within a single session. This includes detailed time allocations and other integral components of a structured plan. Leading chat-based GenAI enterprises, such as OpenAI, Anthropic, and Google, have underscored the prowess of these tools in planning, brainstorming, and feedback solicitation. Notably, the planning capability of GenAI shines when tasked with orchestrating plans for cohesive sets of course content, rather than isolated components.

IV. General Considerations for Engaging with GenAI

In this section, we offer various insights and recommendations pertinent to interacting with GenAI. These considerations, while valuable, do not align explicitly with any specific segment of the framework.

1) *Perspective.* A useful analogy for understanding a GenAI model likens it to a child aged 6-8 years. Such children excel at executing tasks when given clear instructions, provided they possess the requisite knowledge. In the case of GenAI, instead of drawing from 6-8 years of human experiences, it taps into the vast expanse of the internet. However, akin to children of this age, GenAI models often lack initiative and may not perform tasks without explicit direction. Absent specific guidance, the assumptions made by the model could diverge significantly from the intended objectives of the session.
2) *Brainstorming.* One of GenAI's notable strengths lies in its creative prowess. Rather than providing explicit specifications for assessments or activities, CCs might find it beneficial to solicit recommendations from GenAI for various course content elements. This approach can yield diverse and innovative ideas, enriching the educational experience.
3) *Embracing Imperfection.* While the pursuit of excellence is commendable, it's essential to recognize the inherent limitations of GenAI. Continuously striving for perfection can lead to diminishing returns. Instead, it's often more productive to acknowledge areas where the model may falter and focus on harnessing its strengths. In essence, avoid getting mired in endless refinements and capitalize on the valuable insights it provides.
4) *Optimizing Content Generation.* When soliciting GenAI for new content, always request a more extensive set than required. This approach allows for a selection process, ensuring the final subset aligns closely with goals and context. Additionally, by providing specific directives, such as *"with varying levels of difficulty"* or *"answer in a brief way,"* one can enhance the utility and precision of the generated output.

A. Diversity of Perspectives

In the realm of education, understanding and addressing the diverse backgrounds and experiences of learners is paramount. One of the unique capabilities of GenAI is its ability to simulate a multitude of perspectives, which can be instrumental in illuminating potential blind spots in course content, especially when creating content for unfamiliar demographics.

A particularly effective strategy involves employing the phrase *"act as"* when interacting with GenAI (known as the 'persona prompt pattern' [13]). By instructing the model to *"act as"* a particular demographic or adopt a specific perspective, educators can gain insights that might otherwise remain obscured. For instance, asking GenAI to *"act as a student from a non-technical background"* or *"act as an international student"* can yield content and feedback that is more attuned to the needs and challenges of these specific groups.

Leveraging this feature not only enriches the educational material but also fosters an inclusive learning environment. It ensures that content is not inadvertently biased or neglectful of the diverse experiences and backgrounds that students bring to the classroom. In essence, by harnessing the diversity of perspectives that GenAI can offer, educators can craft a more holistic and inclusive educational experience.

V. Discussion and Future Work

Given the relatively recent release and rise to prominence of easily accessible GenAI tools, research studying their usage and applications is still very young. Zhai *et al.* [40] concluded that "[Machine Learning] has transformed–but not yet redefined–conventional science assessment practice in terms of fundamental purpose, the nature of the science assessment, and the relevant assessment challenges." While not addressing the purpose and nature of assessments themselves, the presented framework takes one small step in pushing to redefine how traditional course development is done - and seeks to change who is qualified to develop high-quality course content.

Furthermore, this framework provides a novel perspective on what could and should be done with these tools. Instead of asking, "what should we do about these tools," as many of our colleagues and professors around the world appear to be focusing on [5], this framework turns the question around on instructors and facilitates discussion surrounding "how can these tools help *us* in all of our activities?"

While our primary focus in this study was to elucidate the potential of GenAI in assisting educators with content creation, the second research question posed in our abstract remains an essential area of inquiry: Can the use of GenAI significantly reduce the workload of instructional staff? This question is of paramount importance, especially in the context of increasing class sizes and the constant demand for updated course materials. A specific and measurable research question could be formulated as: "To what extent can the integration of GenAI in course content development reduce the time spent by instructional staff on content creation and revision?" Future studies should employ both quantitative and qualitative methods to assess the time savings, if any, and the potential shifts in the nature of the workload. For instance, while GenAI might reduce the time spent on content creation, it might introduce new tasks, such as refining GenAI outputs or tailoring generated content to specific course objectives. Addressing this research question will provide a more comprehensive understanding of the true impact of GenAI on the educational landscape.

In an effort to gauge the initial reception and applicability of our framework, we introduced it to around 20 of our colleagues during a series of workshops. Although these workshops were conducted recently, and a comprehensive analysis of the survey results is pending, preliminary feedback suggests a high degree of satisfaction with the framework and an indication that the workshop objectives were largely met. Additionally, having used this framework for nearly a year, we have found it to be frequently beneficial and versatile across a wide variety of course contexts and settings.

While we sought to be as general and extendable as possible, the applicability and utility of this framework are inherently bound to some degree to the selected GenAI model. It works best on a chat-based, highly trained model that understands nuances and perspectives across academia. Furthermore, the development of this framework was performed in the context of CS undergraduate courses. More work will be required to validate its utility in other contexts and potentially adapt the model to one that is more useful across disciplines. The GAIDE framework is designed to be highly adaptable and applicable across disciplines by aligning with the specific learning objectives and goals of each respective field. This flexibility allows GAIDE to support educators in different contexts, ensuring the framework remains relevant and practical regardless of the subject matter or educational level.

To conclude, the authors would like to leave you with two thoughts: First, as the frontiers of technology continue to expand, the essence of education remains deeply rooted in the dynamic interplay between innovation and tradition. In the digital age, where generative AI reshapes learning landscapes, it beckons us to adapt and reimagine the educational paradigms. With the GAIDE framework, we endeavor to harness the formidable potential of generative AI, not as a replacement for the human touch in education but as a complement that enriches it. Let us consider how these tools can be molded to respect and uplift the timeless values of teaching while also preparing learners to thrive in a world where change is the only constant.

Second, in the relentless pursuit of progress, the realm of education stands as both a beneficiary and a steward of technological evolution. As we integrate these advanced tools, let us remember that their true value lies not in their ability to replicate human thought but in their capacity to expand the horizons of what educators can achieve. In this new educational landscape, our challenge is to cultivate a synergy where technology amplifies creativity, enhances inclusivity, and deepens understanding, thereby preparing a generation that is as wise as it is technologically adept.

References

[1] E. L. Ouh, B. K. S. Gan, K. J. Shim, and S. Wlodkowski, "Chatgpt, can you generate solutions for my coding exercises? an evaluation on its effectiveness in an undergraduate java programming course," 2023.

[2] J. Savelka, A. Agarwal, M. An, C. Bogart, and M. Sakr, "Thrilled by your progress! large language models (gpt-4) no longer struggle to pass assessments in higher education programming courses," in *Proceedings of the The 19th ACM Conference on International Computing Education Research*, ser. ICER '23, vol. 1, 2023.

[3] M. Wermelinger, "Using github copilot to solve simple programming problems," in *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, ser. SIGCSE 2023. New York, NY, USA: Association for Computing Machinery, 2023, p. 172–178. [Online]. Available: https://doi.org/10.1145/3545945.3569830

[4] S. Kabir, D. N. Udo-Imeh, B. Kou, and T. Zhang, "Who answers it better? an in-depth analysis of chatgpt and stack overflow answers to software engineering questions," 2023. [Online]. Available: https://arxiv.org/abs/2308.02312

[5] S. Lau and P. J. Guo, "From "ban it till we understand it" to "resistance is futile": How university programming instructors plan to adapt as more students use ai code generation and explanation tools such as chatgpt and github copilot," in *Proceedings of the 2023 ACM Conference on International Computing Education Research - Volume 1*, ser. ICER '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: https://doi.org/10.1145/3568813.3600138

[6] B. A. Becker, P. Denny, J. Finnie-Ansley, A. Luxton-Reilly, J. Prather, and E. A. Santos, "Programming is hard - or at least it used to be: Educational opportunities and challenges of ai code generation," in *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, ser. SIGCSE 2023. New York, NY, USA:

Association for Computing Machinery, 2023, p. 500–506. [Online]. Available: https://doi.org/10.1145/3545945.3569759

[7] J. Finnie-Ansley, P. Denny, B. A. Becker, A. Luxton-Reilly, and J. Prather, "The robots are coming: Exploring the implications of openai codex on introductory programming," in *Proceedings of the 24th Australasian Computing Education Conference*, ser. ACE '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 10–19. [Online]. Available: https://doi.org/10.1145/3511861.3511863

[8] A. Moradi Dakhel, V. Majdinasab, A. Nikanjam, F. Khomh, M. C. Desmarais, and Z. M. J. Jiang, "Github copilot ai pair programmer: Asset or liability?" *Journal of Systems and Software*, vol. 203, p. 111734, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0164121223001292

[9] E. Dickey, A. Bejarano, and C. Garg, "AI-Lab: A Framework for Introducing Generative Artificial Intelligence Tools in Computer Programming Courses," *SN Computer Science*, vol. 5, no. 6, p. 720, 2024. [Online]. Available: https://doi.org/10.1007/s42979-024-03074-y

[10] R. Yilmaz and F. G. Karaoglan Yilmaz, "Augmented intelligence in programming learning: Examining student views on the use of chatgpt for programming learning," *Computers in Human Behavior: Artificial Humans*, vol. 1, no. 2, p. 100005, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2949882123000051

[11] O. K. Terry, "I'm a student. you have no idea how much we're using chatgpt," 2023, accessed on: August 14, 2023. [Online]. Available: https://www.chronicle.com/article/im-a-student-you-have-no-idea-how-much-were-using-chatgpt

[12] J. Zamfirescu-Pereira, R. Y. Wong, B. Hartmann, and Q. Yang, "Why johnny can't prompt: How non-ai experts try (and fail) to design llm prompts," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: https://doi.org/10.1145/3544548.3581388

[13] J. White, Q. Fu, S. Hays, M. Sandborn, C. Olea, H. Gilbert, A. Elnashar, J. Spencer-Smith, and D. C. Schmidt, "A prompt pattern catalog to enhance prompt engineering with chatgpt," 2023.

[14] C. Bull and A. Kharrufa, "Generative ai assistants in software development education: A vision for integrating generative ai into educational practice, not instinctively defending against it." *IEEE Software*, pp. 1–9, 2023.

[15] J. D. Gutiérrez, "Guidelines for the use of artificial intelligence in university courses," Universidad del Rosario, Version 4.3.1, February 28 2023, license C.C. BY 4.0. [Online]. Available: https://forogpp.files.wordpress.com/2023/02/guidelines-for-the-use-of-artificial-intelligence-in-university-courses-v4.3.1.pdf

[16] S. A. Guerriero, "4 ways teachers can harness the power of chatgpt," 2023, accessed on: August 14, 2023. [Online]. Available: https://www.learningexplorer.com/blog/4-ways-teachers-can-harness-the-power-of-chatgpt

[17] J. Su and W. Yang, "Unlocking the power of chatgpt: A framework for applying generative ai in education," *ECNU Review of Education*, vol. 0, no. 0, p. 20965311231168423, Apr. 2023. [Online]. Available: https://doi.org/10.1177/20965311231168423

[18] H. Lee, "The rise of chatgpt: Exploring its potential in medical education," *Anatomical Sciences Education*, March 2023. [Online]. Available: https://anatomypubs.onlinelibrary.wiley.com/doi/abs/10.1002/ase.2270

[19] S. Sarsa, P. Denny, A. Hellas, and J. Leinonen, "Automatic generation of programming exercises and code explanations using large language models," in *Proceedings of the 2022 ACM Conference on International Computing Education Research - Volume 1*, ser. ICER '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 27–43. [Online]. Available: https://doi.org/10.1145/3501385.3543957

[20] J. M. Buriak, D. Akinwande, N. Artzi, C. J. Brinker, C. Burrows, W. C. W. Chan, C. Chen, X. Chen, M. Chhowalla, L. Chi, W. Chueh, C. M. Crudden, D. Di Carlo, S. C. Glotzer, M. C. Hersam, D. Ho, T. Y. Hu, J. Huang, A. Javey, P. V. Kamat, I.-D. Kim, N. A. Kotov, T. R. Lee, Y. H. Lee, Y. Li, L. M. Liz-Marzán, P. Mulvaney, P. Narang, P. Nordlander, R. Oklu, W. J. Parak, A. L. Rogach, M. Salanne, P. Samorì, R. E. Schaak, K. S. Schanze, T. Sekitani, S. Skrabalak, A. K. Sood, I. K. Voets, S. Wang, S. Wang, A. T. S. Wee, and J. Ye, "Best practices for using ai when writing scientific manuscripts," *ACS Nano*, vol. 17, no. 5, pp. 4091–4093, 2023, pMID: 36848601. [Online]. Available: https://doi.org/10.1021/acsnano.3c01544

[21] A. Tlili, B. Shehata, M. A. Adarkwah, A. Bozkurt, D. T. Hickey, R. Huang, and B. Agyemang, "What if the devil is my guardian angel: Chatgpt as a case study of using chatbots in education," *Smart Learning Environments*, vol. 10, no. 1, p. 15, Feb 2023. [Online]. Available: https://doi.org/10.1186/s40561-023-00237-x

[22] M. Shidiq, "The use of artificial intelligence-based chat-gpt and its challenges for the world of education; from the viewpoint of the development of creative writing skills," *Proceeding Of International Conference On Education, Society And Humanity*, vol. 1, no. 1, pp. 353–357, 2023. [Online]. Available: https://ejournal.unuja.ac.id/index.php/icesh/article/view/5614

[23] D. Weber-Wulff, A. Anohina-Naumeca, S. Bjelobaba, T. Foltýnek, J. Guerrero-Dib, O. Popoola, P. Šigut, and L. Waddington, "Testing of detection tools for ai-generated text," 2023.

[24] D. Dalalah and O. M. Dalalah, "The false positives and false negatives of generative ai detection tools in education and academic research: The case of chatgpt," *The International Journal of Management Education*, vol. 21, no. 2, p. 100822, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1472811723000605

[25] K. Grossenbacher, "Chatgpt in the workplace: To restrict or embrace, that is the question," *Labor and Employment Law Newsletter - Spring 2023*, vol. 51, no. 1, Jun 2023. [Online]. Available: https://www.americanbar.org/groups/labor_law/publications/labor_employment_law_news/spring-2023/chatgpt-in-the-workplace/

[26] D. Bates, "Chatgpt in the workplace - a legal minefield! what you need to know to protect your business," *JDSupra*, May 2023. [Online]. Available: https://www.jdsupra.com/legalnews/chatgpt-in-the-workplace-a-legal-3653749/

[27] T. Telford and P. Verma, "Employees want chatgpt at work. bosses worry they'll spill secrets." *The Washington Post*, Jul 2023. [Online]. Available: https://www.washingtonpost.com/business/2023/07/10/chatgpt-safe-company-work-ban-lawyers-code/

[28] A. DeRose, "These companies have banned or limited chatgpt at work," *HR Brew*, May 2023. [Online]. Available: https://www.hr-brew.com/stories/2023/05/11/these-companies-have-banned-chatgpt-in-the-office

[29] M. Chui, E. Hazan, R. Roberts, A. Singla, K. Smaje, A. Sukharevsky, L. Yee, and R. Zemmel, "The economic potential of generative ai: The next productivity frontier," Jun 2023. [Online]. Available: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction

[30] R. M. Harden, "Amee guide no. 14: Outcome-based education: Part 1-an introduction to outcome-based education," *Medical Teacher*, vol. 21, no. 1, pp. 7–14, 1999. [Online]. Available: https://doi.org/10.1080/01421599979969

[31] D. H. Ziegenfuss, "A phenomenographic analysis of course design in the academy." *Journal of Ethnographic & Qualitative Research*, vol. 2, no. 1, 2007.

[32] E. Hansen, *IDEA-based learning: A course design process to promote conceptual understanding*. Stylus, 2011.

[33] A. P. Gilakjani, L. Lai-Mei, and H. N. Ismail, "Teachers' use of technology and constructivism," *International Journal of Modern Education and Computer Science*, vol. 5, no. 4, p. 49, 2013.

[34] M. Koehler and P. Mishra, "What is technological pedagogical content knowledge (tpack)?" *Contemporary issues in technology and teacher education*, vol. 9, no. 1, pp. 60–70, 2009.

[35] E. Dickey and A. Bejarano, "A model for integrating generative ai into course content development," 2024. [Online]. Available: https://arxiv.org/abs/2308.12276

[36] R. M. Felder and R. Brent, *Teaching and Learning STEM: A Practical Guide*. John Wiley & Sons, 2016.

[37] J. Conklin, "A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives complete edition," 2005.

[38] N. L. Webb, "Depth-of-knowledge levels for four content areas," *Language Arts*, vol. 28, no. March, pp. 1–9, 2002.

[39] K. Ragupathi and A. Lee, "Beyond fairness and consistency in grading: The role of rubrics in higher education," *Diversity and inclusion in global higher education: Lessons from across Asia*, pp. 73–95, 2020.

[40] X. Zhai, K. Haudek, L. Shi, R. Nehm, and M. Urban-Lurain, "From substitution to redefinition: A framework of machine learning-based science assessment," *Journal of Research in Science Teaching*, vol. 57, no. 9, pp. 1430–1459, Oct. 2020. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/tea.21658