Generative AI in Digital Education: A Case Study of Content Co-Creation in a Design Thinking MOOC

ABSTRACT

Integrating Generative Artificial Intelligence (GenAI) into educational materials development presents opportunities and challenges in education, particularly in Massive Open Online Courses (MOOCs). This study explores the role of GenAI in developing content for MOOCs using a Design Thinking MOOC as a case study. It assesses GenAI-generated instructional materials for content accuracy, depth, and engagement potential while analyzing the level of human intervention required for pedagogical quality. Using Perplexity Pro as the GenAI tool, the study finds that GenAI efficiently generates structured drafts, fictional learning scenarios, and key takeaways. However, significant limitations emerge in GenAI's ability to differentiate complex domain-specific concepts, develop high-quality assessment items, and ensure pedagogical alignment. Human intervention remains fundamental for enhancing conceptual depth, refining instructional clarity, and fostering learner engagement. Based on these insights, the study proposes a Framework for GenAI-Assisted Content Creation in MOOC Design, outlining a structured approach to integrating GenAI while maintaining educational rigor. The framework highlights four interdependent phases: Content Planning & GenAI Preparation; GenAI-Generated Content Creation; Expert Review & Refinement; and Testing & Iterative Improvement. The study further presents Guidelines and Best Practices for MOOC Designers, providing practical recommendations for leveraging GenAI effectively without compromising instructional quality. This research contributes to the growing literature on AI-driven education, providing practical guidelines for MOOC designers seeking to optimize GenAI-driven content development.

Keywords

Design Thinking, Human-AI Collaboration; MOOCs; GenAI in Education, Perplexity Pro.

1. INTRODUCTION

The integration of GenAI in education has been considerably investigated in recent years as GenAI tools are increasingly being employed to streamline content production, reduce the workload for educators and instructional designers, and enhance scalability. For instance, previous research explored DALLE-2 and Tome.ai for their potential to accelerate MOOC development, demonstrating how automation can reduce content production time (Amado-Salvatierra et al., 2023). However, the researchers state that while AI-based tools can efficiently generate initial content drafts, human expertise remains critical in refining and contextualizing this content to meet academic objectives. Similarly, Faccia et al. (2023) investigated the role of GenAI in higher education content creation, emphasizing the importance of human oversight in maintaining accuracy and pedagogical depth. They discuss that although AI-powered models such as OpenAI's GPT and Hugging Face's Transformers can support content generation, AI-generated materials require careful curation and refinement to ensure contextual relevance, critical engagement, and alignment with pedagogical best practices.

Integrating GenAI into curriculum design has also led to the emergence of specialized frameworks for GenAI-assisted education such as GAIDE, which emphasize the collaborative role of GenAI and human creativity in content development (Dickey & Bejarano, 2024). These frameworks enable educators to leverage GenAI for dynamic content generation while ensuring that instructional materials remain pedagogically sound and academically rigorous. Moreover, Zhou et al. (2021) conducted a comprehensive analysis of AI's application in MOOCs, highlighting its effectiveness in automating assessments and teaching core algorithms. Nevertheless, they identified that pedagogical strategies such as scenario-based and project-based learning remain underutilized in AI-enhanced MOOCs, suggesting that AI tools still have significant potential to evolve in supporting interactive and experiential learning.

Beyond efficiency gains, GenAI has been recognized for its role in enhancing inclusivity and personalization in online education. Stefaniak & Moore (2024) argue that GenAI can potentially adapt learning materials to individual student needs which foster greater accessibility. However, ethical concerns such as algorithmic bias and the reinforcement of existing educational inequalities should be critically assessed. These concerns highlight the need for iterative instructional design practices that incorporate ongoing evaluation and refinement of GenAI-generated content.

The benefits of GenAI-assisted content creation in education are widely documented, with research pointing to improvements in personalized learning experiences, assessment methods, and instructional design efficiency. For instance, Liu (2024) examined the transformative potential of AI in enabling customized content development and improving student outcomes. Nonetheless, challenges such as academic integrity concerns, evolving educator roles, and data privacy issues must be addressed to ensure AI's responsible and ethical use in education.

This body of research highlights GenAI's potential as a facilitative tool rather than a replacement for educators. Ravarini et al. (2024) proposed a methodological framework positioning educators as both content creators and instructional designers while leveraging GenAI to expedite course development, improve content quality, and personalize education. Their findings align with growing perspectives that advocate for AI-human collaboration rather than full automation.

Despite GenAI's capabilities in generating instructional materials, the literature consistently emphasizes the indispensable role of human expertise in ensuring pedagogical integrity, ethical responsibility, and instructional effectiveness. While GenAI presents opportunities to improve the scalability and accessibility of MOOCs, its integration into education might be carefully approached with a balanced perspective that acknowledges its limitations and the need for human intervention.

These studies emphasize a critical need for more granular, task-level investigations of how GenAI performs across different instructional content types and how much human intervention is required to ensure pedagogical quality in MOOC design. Given GenAI's growing sophistication and its demonstrated fluency in generating text-based outputs, it was initially anticipated that many instructional tasks; especially structured and procedural ones; might require only minimal expert refinement. However, the actual degree of oversight needed across varied content formats remains empirically unclear. This study addresses this gap by asking the following research question: *To what extent can Generative AI, specifically Perplexity Pro, support instructional content development for MOOCs, and what levels of expert intervention are required to ensure pedagogical quality across diverse content types?*

2. RESEARCH METHODOLOGY

This study adopts a systematic single-case study design (Yin, 2017) to examine the use of GenAI in MOOC instructional design. The case study approach is particularly suited for investigating complex, contemporary phenomena within real-world educational settings where the boundaries between the intervention (GenAI use) and the contextual environment (course design) are blurred. The case is employed as a means to explore broader issues of human-AI collaboration in educational content development, thereby, functioning as an instrumental case (Stake, 1995). The study procedures include the identification of clear units of analysis (e.g., quizzes, scripts, exercises); systematic collection of AI prompt–output pairs; application of a structured intervention rubric (see Section 2.3) to assess revision depth and effort; and comparative analysis of GenAI performance across multiple content types.

Moreover, the study follows an exploratory, theory-building orientation, where detailed analysis of a single bounded case can inform broader conceptual insights (Eisenhardt, 1989). The theoretical framework is grounded in Activity Theory (Engeström, 2014), which enables the examination of the mediated interactions between human agents (instructional designers), technological tools (GenAI systems), and objects (instructional content). This lens supports an understanding of how GenAI tools may reshape task structures, workflows, and cognitive effort in instructional design.

2.1 Case Selection and Context

A MOOC on Design Thinking was chosen to conduct the study. Design Thinking is a "process, a mindset, and a human-centered approach to creativity, collaboration, and innovation" (Traifeh, 2023, p.4). Its purpose is to define users' needs and explore the possibilities of technology and the requirements for business success in solving complex or wicked problems, or creating innovative products (Koh et al., 2015; Razzouk & Shute, 2012). Design Thinking has been widely implemented in different industries such as business, research, education, social innovation, and other domains (e.g., Plattner et al., 2011; Kelley, T., & Kelley, 2013; Chou, 2018; Traifeh et al., 2019). Over time, several Design Thinking frameworks have been developed by individuals, universities, and organizations worldwide. While each framework employs its own terminology, they all emphasize a deep exploration of user needs to identify the core problem, ideate, prototype, and test potential solutions. The content of the MOOC at hand was developed following the d.school's framework (the Hasso Plattner Institute of Design at Stanford), which includes five phases: "Empathize, Define, Ideate, Prototype, Test" (Stanford University, n.d.).

The MOOC was structured into three primary modules, each including several learning units:

(1) Design Thinking Fundamentals: This module introduced foundational principles, historical context, and the Design Thinking process. The content units comprised video scripts, articles, case studies, self-assessment tools, and quizzes.

(2) The Problem Space: This module introduced the first two phases of the Design Thinking process- *Empathize* and *Define*. Instructional materials included practical exercises, video scripts, templates, real-world case studies, and quizzes.

(3) The Solution Space: This module explored the *Ideate*, *Prototype*, and *Test* phases. It featured video scripts, examples, case studies, and quizzes.

The course was designed to provide learners with the theoretical foundations and practical applications of Design Thinking principles, ensuring a comprehensive understanding of the methodology and its real-world implementation. Before utilizing GenAI for content generation, the lead instructional designer -who is also a Design Thinking expert- established a foundation to guide the instructional design process by developing the course outline and defining clear learning objectives. These foundational elements were critical reference points while developing course materials to ensure that GenAI-generated content aligned with pedagogical goals and maintained coherence across modules. Integrating GenAI within this structured approach supported the expert in leveraging GenAI's capabilities while exercising targeted oversight and refining the content as needed to enhance instructional quality and learner engagement.

Several methodological factors informed the selection of Design Thinking as the subject domain. First, its interdisciplinary nature and demand for both conceptual understanding and applied problem-solving make it a suitable subject for evaluating GenAI performance across varied content types. Second, the domain's fundamental emphasis on empathy and human-centered approaches offers a meaningful contrast to GenAI-generated content, revealing where human intervention is pedagogically necessary. Third, the structured five-phase model of Design Thinking provides well-defined boundaries for organizing, generating, and evaluating instructional content.

2.2 GenAI Tool Selection and Implementation

For GenAI-assisted content creation, this study employed Perplexity Pro which was selected over other large language models (LLMs), such as ChatGPT, Claude, or Gemini, due to its specific features that align with the demands of academic content creation and research (Shukla et al., 2024). First, Perplexity Pro offers automatic source attribution for generated content, enhancing transparency and enabling verification which is an essential requirement for academic content creation and fact-checking processes. Second, unlike some LLMs with fixed training cutoffs, Perplexity Pro integrates real-time web retrieval, allowing for the synthesis of up-to-date information, which is particularly valuable for constructing relevant case studies and incorporating recent examples. Third, the platform is designed to optimize academic usage by offering concise, citation-rich responses, and fact-grounded responses that support content validation. This visibility into source material improves the reliability and accountability of GenAI-generated educational materials.

Perplexity Pro was mainly employed to: (1) generate initial drafts for video scripts and articles; (2) summarize unit content into key takeaways; (3) suggest real-world case studies and fictional (hypothetical) scenarios aligned with Design Thinking principles; and (4) develop quizzes and exercises to facilitate active learning.

GenAI outputs were guided by structured prompts aligned with the course's learning objectives. For example, at the beginning of the interaction, the persona prompt pattern strategy (White et al., 2023) was applied by asking GenAI to *act as* a Design Thinking expert. Another example is when identifying relevant case studies, instead of providing a generic prompt such as "Give a real-world example of Design Thinking", a prompt was "Provide a detailed real-world example where a company successfully applied Design Thinking to solve a customer problem. Include the phases used, challenges faced, and measurable outcomes". The prompts were designed to provoke detailed explanations of Design Thinking principles, phases, and tools; generate assessment questions that target higher-order cognitive skills; and produce case studies illustrating the practical applications of Design Thinking.

2.3 Human-AI Collaboration and Intervention Classification

The integration of GenAI into the MOOC design followed an iterative workflow in which GenAI-generated content served as an initial draft, subsequently reviewed and refined by the content expert. The level of human intervention varied depending on the complexity of the task and the depth required in the content. To evaluate the effectiveness of GenAI across different content types (e.g., video scripts, case studies, quizzes), an intervention classification system was developed which is grounded in three key instructional design dimensions (Dick et al., 2001; Morrison et al., 2019), namely: content accuracy, depth of information, and engagement potential.

Content Accuracy: The accuracy of GenAI-generated content was assessed by cross-referencing facts, ensuring the correctness of terminologies, and verifying the validity of real-world case studies. For instance, when GenAI suggested the IDEO-Shimano case study for the 'Empathize' phase, the details provided were superficial and required verification through additional research. Similarly, GenAI-generated examples such as the IBM case study were fact-checked for alignment with real-world applications of Design Thinking. Another example is when a video script was generated for the 'Design Thinking Mindset & Principles", GenAI provided an incomplete list of principles, which required expert intervention to add missing principles.

Depth of Information: GenAI's ability to provide comprehensive and subtle explanations was another critical metric. Content that required surface-level explanations (e.g., summarizing key takeaways, creating empathy exercises) typically needed minimal intervention. Conversely, tasks demanding in-depth analysis (e.g., developing quizzes with critical-thinking questions or drafting video scripts) required significant refinement. For example, in the 'Brainstorming Methods' video script, GenAI struggled to differentiate 'Ideation' techniques and methods from other phases, leading to disorganized content that needed extensive human input.

Engagement Potential: The extent to which GenAI-generated content could engage learners was assessed through the clarity of examples, the practical relevance of exercises, and the alignment with learner-centered design principles.

Each piece of content was then classified into one of three categories based on the degree of human intervention required to meet instructional design standards (Table 1). The intervention levels are defined as follows: minimal (0–25% modification), moderate (26–75% modification), and extensive (76–100% modification).

Table 1. Intervention Level Rubric

Intervention Level	% of Content Modified	Indicators	Examples
Minimal	0–25%	• Accurate content • Aligned with learning objectives • Minor edits (tone/style)	Tone adjustment, formatting, rephrasing
Moderate	26–75%	• Requires additional concepts or examples • Factual fixes • Enhances existing logic	Adding missing points, extending explanations
Extensive	76–100%	• Contains major factual errors or misconceptions • Requires full restructuring	Rewriting scripts, rebuilding quiz logic, correcting concepts

2.4 Data Analysis Process

The data analysis process followed a systematic approach aligned with a qualitative case study model (Miles et al., 2014), including coding, pattern detection, triangulation, and member checking. All GenAI-generated instructional components for the course's 30 content units—such as video scripts, case studies, quizzes, and learner exercises—were coded based on the Intervention Level Rubric. The classification of each content unit into Minimal, Moderate, or Extensive intervention categories was determined using the predefined operational criteria related to content accuracy, pedagogical depth, and engagement potential (see appendix A for a coded summary of content development).

To ensure reliability, two experts participated in the review process at different stages. The lead researcher, a senior researcher with a PhD in Design Thinking and extensive instructional design experience, conducted the primary coding. A second expert, a tenured professor in digital learning and AI pedagogy with recognized expertise in design thinking methodologies, independently reviewed the intervention categorizations. Experts were identified through prior collaborations in MOOC design projects. Member checking entailed the second expert validating rubric application and coding judgments, followed by iterative discussions to resolve discrepancies. While formal inter-rater statistics were not calculated due to the qualitative orientation of the study, agreement rates exceeded 90% across content units, providing confidence in the coding reliability.

Following coding and validation, content units were analyzed to identify recurring patterns in GenAI performance across instructional formats. This process revealed which content types (e.g., summaries vs. assessments) typically required less or more human refinement, and highlighted systematic strengths (e.g., summarization capability) and recurring limitations (e.g., underdeveloped higher-order thinking questions). Triangulation was further employed by cross-verifying GenAI-generated outputs with academic literature, instructional design best practices, and expert knowledge to ensure judgments were grounded in credible reference points rather than subjective impressions.

3. FINDINGS

The systematic analysis of 30 distinct content units revealed a clear pattern in the levels of human intervention required to meet instructional quality standards. Contrary to the initial assumption that GenAI would predominantly require minimal oversight, the findings demonstrated a broader spectrum of editorial effort. Forty percent of the content units (n = 12) required only minimal human intervention. These were typically structured summaries, key takeaways, and simple learner exercises, where GenAI performed well in organizing and articulating foundational concepts with little need for revision.

The remaining units were split evenly between the moderate and extensive intervention categories, each comprising 30% (n=9) of the dataset. The moderate intervention group included video scripts, introductory explanations, and initial drafts of case studies that provided a helpful starting structure but required expert input to deepen conceptual accuracy and ensure pedagogical alignment. The extensive intervention group was dominated by assessments and conceptually complex content. In these cases, GenAI-generated content often lacked cognitive depth, contained conceptual inaccuracies, or failed to reflect domain-specific nuances which required substantial rewriting and reorganization.

This distribution aligns with Cognitive Load Theory (Sweller, 1988), which distinguishes between extraneous (organizational tasks), intrinsic cognitive load (complex domain knowledge), and germane cognitive load (meaningful learning processes). GenAI was most effective in minimizing extraneous load by generating well-structured drafts and summaries. However, it was less capable of addressing intrinsic cognitive load that is associated with complex, domain-specific reasoning, and germane cognitive load, which supports meaningful learning through instructional coherence and deep engagement. These findings emphasize the indispensable role of expert oversight in maintaining instructional quality, particularly in tasks that demand conceptual rigor and learner-centered design.

3.1 GenAI Contributions and Strengths

GenAI-assisted content demonstrated significant efficiency in generating structured drafts for instructional materials, including several video scripts, learning exercises, and key takeaways.
Overall, the GenAI tool used in this study -Perplexity Pro- successfully produced logically sequenced content, allowing the expert to focus on deepening explanations and contextual refinement rather than drafting from scratch. In the 'User Interviews' video script, for instance, GenAI effectively provided practical tips and structured guidance, requiring only minor refinements for tone and coherence. Similarly, in summarizing unit content, GenAI generated concise and well-structured 'Key Takeaways', improving content clarity and learner accessibility.

Additionally, GenAI showed strong capabilities in creating fictional learning scenarios to support experiential learning. In the 'Empathy Exercise,' for example, GenAI drafted a highly relevant and engaging scenario that required minimal expert intervention to align with learning objectives. The tool also proved beneficial in identifying real-world case studies, such as the IDEO-Shimano example for the 'Empathize' phase. Although GenAI-generated descriptions of these case studies required fact-checking and elaboration, the initial suggestions provided a helpful starting point for further refinement.

3.2 Challenges and Limitations of GenAI in MOOC Content Creation

Despite these strengths, the findings reveal several limitations that required varying degrees of human intervention, particularly in tasks requiring critical thinking, conceptual accuracy, and instructional depth. One of the most significant challenges was GenAI's inability to generate high-quality quiz questions. GenAI-produced assessment items were often surface-level, lacking the complexity to assess higher-order cognitive skills such as analysis, synthesis, and evaluation. As a result, most GenAI-generated quizzes required substantial revision or were entirely rewritten by the expert.

Another limitation was GenAI's difficulty in differentiating closely related concepts within the Design Thinking framework. For instance, in the *Brainstorming Methods* video script, GenAI misclassified ideation techniques and confused them with methods from other phases, resulting in disorganized content. Similarly, in the *Immersion* video script, it blurred the distinction between observation and immersion, requiring substantial expert correction to ensure conceptual accuracy. These findings align with previous research showing that GenAI struggles with domain-specific distinctions and requires expert guidance to maintain instructional coherence (Tuomi, 2024; Hutchins et al., 2020; Luckin & Holmes, 2016).

Where such conflations occurred, corrections were made by explicitly defining terms and illustrating them with concrete examples. *Observation* was defined as "systematically watching users interact with products or services in their natural environment without direct researcher participation, focusing on capturing authentic behaviors and usage patterns." By contrast, *Immersion* was defined as "designers placing themselves directly in the user's situation to experience challenges, emotions, and contextual factors firsthand through active participation." The corrected content included specific examples: observation might involve watching customers use a mobile banking app in a café, while immersion could mean spending a day using only public transportation to understand commuter challenges.

A similar issue arose with the Ideation phase. GenAI frequently reduced Ideation to a simple brainstorming activity. The revised content clarified that Ideation is a structured phase dedicated to generating a wide range of ideas, encompassing—but not limited to—brainstorming. Methods such as bodystorming and mind mapping were added to highlight the diversity of approaches that support creative exploration. These definitional clarifications ensured learners encountered a valid and multifaceted understanding of Design Thinking phases.

GenAI-generated content also frequently lacked contextual depth, particularly in case studies and real-world applications. While GenAI could identify relevant examples, its explanations were often superficial, requiring expert elaboration to provide deeper insights. This limitation reflects broader concerns in AI-driven education, where efficiency in content generation does not guarantee pedagogical effectiveness (Emma & Peace, 2024).

Finally, GenAI's ability to enhance learner engagement varied across content types. While it produced well-structured learning materials and effective examples, it struggled with interactive and inquiry-based elements. For instance, GenAI successfully generated structured exercises and prompts but did not effectively incorporate reflection-based learning strategies, requiring expert modifications. These observations suggest that GenAI is currently more effective in supporting content structuring rather than fully facilitating interactive and engagement-driven learning experiences.

These findings reinforce the view of GenAI as an assistive tool rather than an autonomous content creator. GenAI might effectively simplify content structuring and summarization, but expert oversight remains critical to ensure accuracy, depth, and pedagogical alignment. Its limitations in conceptual reasoning, assessment design, and contextualization indicate that its role is best framed as a content generation aid, not a stand-alone instructional designer.

3.3 Assessment Quality and Cognitive Processes

Given the centrality of quizzes to the study's findings, selected revised items were mapped to the Revised Bloom's Taxonomy (Anderson & Krathwohl, 2001) to demonstrate the intended cognitive processes. Table 2 illustrates examples of how expert refinements elevated cognitive demand beyond surface-level knowledge checks.

Table 2. Examples of AI-Generated vs Expert-Revised Quiz Items Mapped to Revised Bloom's Taxonomy

AI-Generated Quiz Item	Expert-Revised Quiz Item	Revised Bloom's Cognitive Process	Example Target
"What is empathy in Design Thinking?" (short definition	*"I actively listen to others and try to understand their perspectives."* (self-assessment on empathy practice)	Remember / Understand	Recognition and comprehension of a foundational principle

recall)			
"Brainstorming is part of which Design Thinking phase?" (single-choice)	*"Which of the following is a key rule for successful ideation sessions?"* (Solution Space quiz; correct answer: encourage wild ideas and defer judgment)	Apply	Applying rules of ideation to evaluate correct practice
"What is prototyping?" (basic recall)	*"What is the benefit of using low-fidelity prototypes in the design process?"* (Solution Space quiz; correct answer: they allow for rapid iteration and exploration of ideas and save resources)	Analyze	Differentiating between prototype types and their pedagogical purpose
"What is the purpose of the 'Define' phase in Design Thinking?" (surface level)	*"Evaluate a poorly framed problem statement: 'The app needs a better interface.' How could it be reframed to reflect the principles of the Define phase?"*(expert-added revision)	Evaluate	Critical judgment and justification beyond surface-level response

These examples illustrate the revision trajectory: while GenAI-produced quizzes tended to remain at surface-level (e.g., recall of definitions), expert revisions deliberately targeted higher-order processes. For instance, in the prototyping quiz, GenAI initially generated factual recall items, but experts reframed them to require analysis (comparing prototype fidelity) or evaluation (critiquing GenAI's draft). This ensured that assessment items addressed deeper learning goals and actively fostered the critical thinking and problem-solving skills that are at the heart of Design Thinking.

4. FRAMEWORK FOR GenAI-ASSISSTED CONTENT CREATION IN MOOC DESIGN

Based on the insights gained from this study and building upon established instructional design models, we propose a framework for GenAI-assisted content creation in MOOC Design (see Figure 1). Unlike existing frameworks such as GAIDE (Dickey & Bejarano, 2024) and IntelliFrame (Hadyaoui & Cheniti-Belcadhi , 2024) which offer generalized or assessment-focused models, this framework specifically addresses the task-level effort metrics of human-AI collaboration based on empirical evidence rather than theoretical assumptions.

The framework's structure reflects the principles of a widely adopted instructional systems model (Dick et al., 2015), which emphasizes the interdependence of instructional components from learning objectives and content materials to assessment and evaluation. Similar to Dick et al.

(2015), this framework incorporates iterative design, expert validation, and systematic feedback loops to ensure instructional coherence, especially in the context of GenAI-enhanced content development. At the same time, the framework's pedagogical scaffolding reflects principles from the Salmon's five-stage model of online learning (2013). The iterative interplay between AI-generated drafts and expert review mirrors the model's progression from early information exchange to deeper levels of knowledge construction and learner development. This structure supports both instructional scalability and pedagogical depth. The dual-theoretical grounding positions the framework as a concrete model for designing pedagogically aligned, AI-assisted instructional content in online environments. The framework consists of four interdependent phases: 1) content planning & GenAI preparation; 2) GenAI-generated content creation; 3) expert review & refinement; and 4) testing & iterative improvement.

In Phase 1, Content Planning & GenAI Preparation, instructors articulate clear learning objectives, design structured course outlines, select appropriate GenAI tools, and develop context-aware prompts before engaging in content generation (Emma & Peace, 2024). This proactive planning ensures that GenAI-generated materials are aligned with instructional goals and minimizes risks of incoherent or pedagogically misaligned output. This phase also reflects a broader trend in AI-assisted education, which emphasizes adaptive design environments guided by strong human oversight (Hadyaoui & Cheniti-Belcadhi, 2024; Amado-Salvatierra et al., 2023).

The second phase, AI-Generated Content Creation, operationalizes the GenAI's role in the instructional design process. GenAI tools can assist in drafting video scripts, developing quizzes and exercises to facilitate learner engagement, summarizing key takeaways, and suggesting real-world case studies and fictional learning scenarios. However, despite GenAI's efficiency, this study's findings show that GenAI-generated instructional materials often require refinement to enhance their depth, contextual accuracy, and engagement. This limitation emphasizes the need for the third phase, 'Expert Review and Refinement,' where human educators apply pedagogical expertise to improve accuracy, enrich conceptual depth, eliminate redundancy, assess for bias, and ensure instructional alignment. Moreover, researchers have emphasized that AI-assisted content creation must be supplemented with scenario-based learning, learner-centered adaptations, and contextual nuance (Li et al., 2024); a principle embedded throughout this phase of the framework.

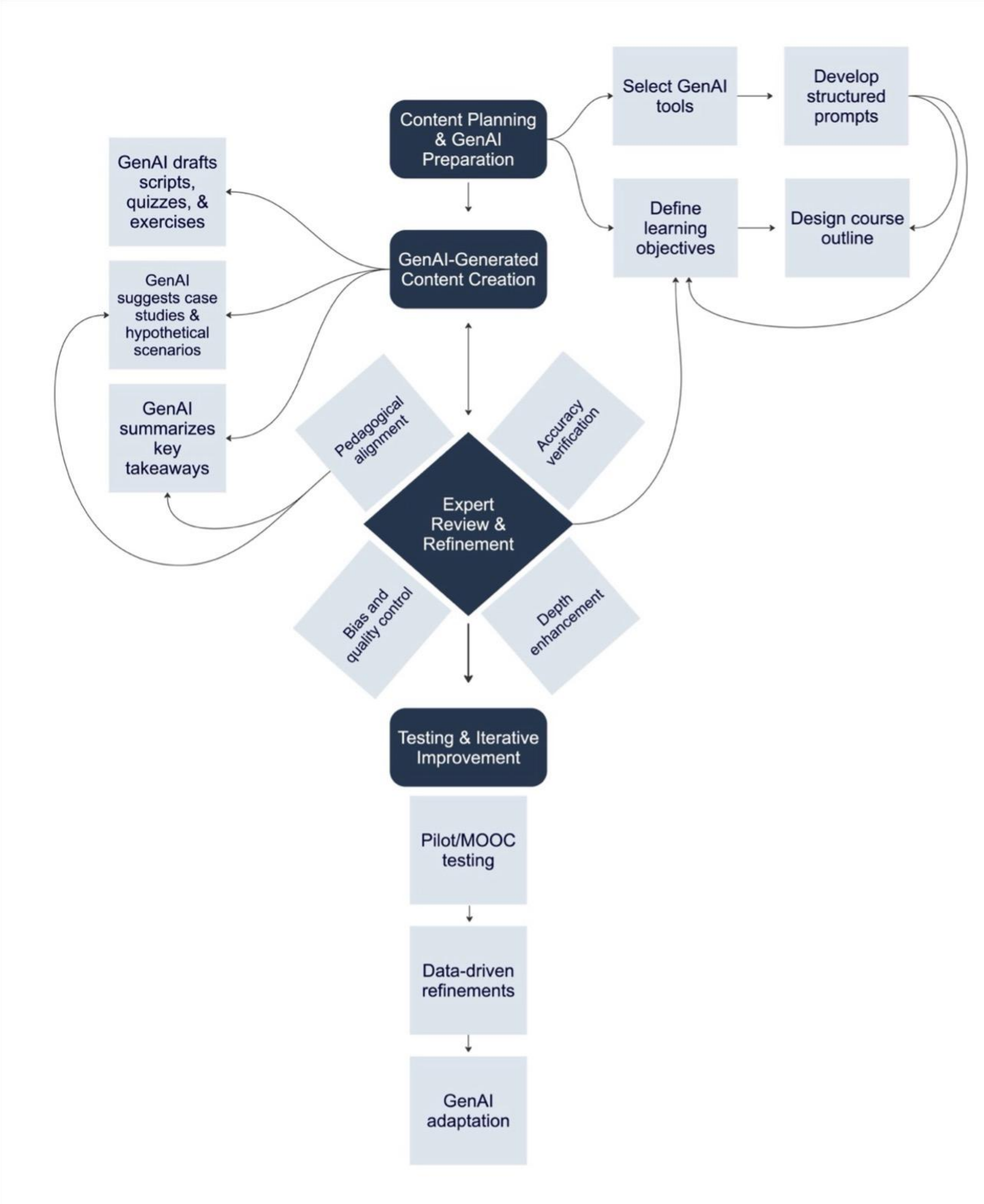

Figure 1: Framework for GenAI-Assisted Content Creation in MOOC Design

The final phase, Testing & Iterative Improvement, emphasizes real-world validation of GenAI-assisted content. This includes pilot testing, learner feedback analysis, and ongoing content revisions based on usage analytics. This phase aligns with contemporary research advocating for co-creation models in AI-assisted education (Ghariz et al., 2024), reinforcing the need for continuous improvement and contextual adaptation. Previous studies also stress that learner engagement and comprehension must be evaluated dynamically and inform content updates to ensure continued pedagogical relevance (Abbasi et al., 2024).

Although the proposed framework was developed within the context of Design Thinking education, its structure is adaptable to a broad range of disciplines. It addresses several universal challenges in AI-assisted instructional design. First, the need to manage varying levels of cognitive complexity is common across subject areas, requiring thoughtful alignment between content depth and learner capabilities. Second, the challenge of designing meaningful assessments extends beyond any single domain, making the framework's emphasis on integrated assessment development widely applicable. Third, expert validation is a critical requirement for ensuring instructional quality in any educational context, particularly when incorporating GenAI-generated materials. Finally, the framework's emphasis on iterative improvement that is driven by learner feedback and performance analytics reflects a best practice in contemporary instructional design that is relevant across all educational environments.

4.1 Comparative Positioning with Existing Frameworks

In order to position the proposed framework within the broader landscape of GenAI-supported instructional design, it is compared with existing models such as GAIDE (Dickey & Bejarano, 2024) and IntelliFrame (Hadyaoui & Cheniti-Belcadhi , 2024). Both of these frameworks provide valuable perspectives on AI integration in education, yet they differ from this study's empirically grounded approach in several important ways (Table 3).

Compared to GAIDE, which focuses on promoting creative collaboration between human experts and GenAI tools, the proposed framework distinguishes itself by offering task-level effort metrics derived from actual instructional design practice. Whereas GAIDE outlines broad phases of GenAI-supported course design, it lacks empirical data on the degree of human intervention required across content types. The current framework contributes this missing dimension by providing specific intervention percentages across 30 MOOC content units, enabling practitioners to anticipate the human effort needed for quality assurance.

Moreover, while GAIDE emphasizes planning and implementation, the proposed model expands to include ongoing refinement and formative evaluation. It also provides targeted recommendations for assessment development, an area underrepresented in the GAIDE model. Thus, the framework builds upon GAIDE’s principles but introduces an operational layer grounded in instructional design metrics.

Compared to IntelliFrame, which is primarily focused on adaptive AI-driven assessment, the proposed framework adopts a broader scope, encompassing the full range of content development tasks (e.g., video scripts, case studies, exercises, and quizzes). IntelliFrame prioritizes AI-powered personalization and learner modeling, whereas this framework emphasizes the balance of GenAI and human expertise across all phases of MOOC content creation. Additionally, by embedding formative feedback and real-world testing as core phases, this model ensures that instructional quality evolves continuously in response to learner needs.

Table 3. Comparative Framework Analysis

Framework Aspect	GAIDE (Dickey & Bejarano, 2024)	IntelliFrame (Hadyaoui & Cheniti-Belcadhi, 2024)	This Study’s Framework
Primary Scope	Broader application of GenAI in instructional design, emphasizing efficiency	AI-driven assessment in e-learning, focusing on automated evaluation and feedback	GenAI-assisted MOOC content creation with strong human oversight and iterative refinement
Empirical Foundation	Theoretical model with limited empirical validation	Technical proof-of-concept	30 Content units systematically analyzed
Phases	1) Setup, 2) Course Content Rough Draft, 3) Macro Refinement, 4) Micro Refinement, 5) Maintaining Contextual Integrity in Iterative Refinement, 6) Consolidating Generated Content	IntelliFrame is structured around key components and architectural layers rather than sequential phases: 1) Ontology-Driven Architecture, 2) Personalized AI Chatbot, 3) Adaptive Assessment Scenarios, 4) Real-Time Feedback & Monitoring, 4) LMS Integration	1) Content Generation (GenAI), 2) Human-AI Iteration, 3) Expert Review & Refinement, 4) Testing & Iterative Improvement.
Human-AI Collaboration	AI as a powerful assistant, with human designers guiding and validating.	AI primarily automates assessment, human intervention for initial setup and oversight.	Explicit, iterative human oversight at every stage, with human expertise driving pedagogical quality.

Human Intervention Metrics	Qualitative phases without specific metrics	Not specified	Quantified effort levels: 0-25% (minimal), 26-75% (moderate), 76-100% (extensive)
Outputs	Instructional materials, course outlines, learning activities.	Automated quizzes, personalized feedback, performance analytics.	Refined MOOC content, actionable guidelines for human-AI collaboration, empirically grounded framework.
Unique Contribution	Focuses on integrating GenAI into traditional instructional design models (ADDIE).	Specializes in leveraging AI for efficient and effective assessment strategies.	Emphasizes a systematic, single-case study approach to GenAI in MOOC design, providing task-level metrics of human intervention and a new framework grounded in empirical data, ongoing refinement/formative evaluation.

The *Framework for GenAI-Assisted Content Creation in MOOC Design* addresses a gap not covered by existing theoretical or assessment-focused approaches by offering a practical, evidence-based model for GenAI integration. Its contribution lies in the inclusion of task-level intervention metrics, expert validation loops, and a comprehensive approach to content development which positions it as a transferable and scalable solution for GenAI-enhanced instructional design.

4.2 Guidelines and Best Practices for MOOC Designers

Building upon the Framework for GenAI-Assisted Content Creation in MOOC Design, we provide practical guidelines to ensure effective implementation. While the framework establishes the phases of GenAI engagement, from content planning to iterative improvement, MOOC designers require actionable strategies for optimizing GenAI's use at each stage. The following Guidelines and Best Practices for MOOC Designers (Table 4) translate the framework into concrete recommendations, specifying how educators can leverage GenAI while maintaining instructional quality, pedagogical integrity, and learner engagement. These guidelines stress the collaborative role of GenAI and human expertise, ensuring that GenAI-generated content aligns with educational objectives and best practices in course development.

Table 4. Guidelines and Best Practices for MOOC Designers

Phase	**GenAI Application**	**Expert Role**	**Best Practices for MOOC Designers**

(1) **Content Planning & Preparation**	Not applicable at this stage. GenAI is not used until course learning objectives and outline are finalized. However, some educators may use it to brainstorm module topics at this stage. In this case, course structure, progression, and depth should be refined and validated to align with learning objectives.	Define learning objectives, course outline and instructional strategy before integrating GenAI. Establish the pedagogical framework and ensure alignment with learner needs.	Ensure that learning objectives and course outline are clearly defined before engaging GenAI. If GenAI is used for brainstorming, its suggestions must be critically evaluated to align with instructional goals. Establish a structured instructional design plan to guide GenAI-assisted content creation in later phases.
(2) **Drafting Learning Materials & Iterative Refinement**	GenAI generates initial drafts for video scripts and key takeaways. It can also suggest case study examples, fictional scenarios, and quizzes relevant to course topics.	Provide GenAI with structured prompts aligned with course learning objectives. Guide GenAI in generating initial drafts for video scripts, case studies, and exercises. Improve content by suggesting alternative explanations, analogies, and additional examples to enhance clarity.	Guide GenAI with specific, structured prompts to improve content relevance, depth, and completeness. Use GenAI-generated drafts as a foundation, but ensure educators iteratively refine outputs to enhance clarity, depth, engagement, and alignment with learning objectives. GenAI is most effective in structuring content but requires expert input for explanations and pedagogical depth**.**
(3) **Quiz & Assessment Design**	Suggest multiple-choice questions based on unit summaries.	Rewrite questions to enhance critical thinking and cognitive depth.	Use GenAI to generate question banks, but refine them for accuracy, cognitive complexity, and alignment with learning objectives. Complement GenAI-generated quizzes with scenario-based, open-ended, and application-based assessments designed by experts to ensure meaningful evaluation of student learning.
(4) **Case Study Development**	Recommend real-world examples relevant to the subject presented.	Fact-check, verify sources and enrich case studies with detailed analysis and application.	GenAI can suggest real-world case studies, but experts must fact-check, contextualize, and enrich them with critical analysis. Ensure that case studies align with course learning objectives and provide opportunities for deep learning and application.

(5) **Final Review & Quality Assurance**	GenAI can refine explanations, summarize key points, and offer variations of instructional content based on expert feedback.	Evaluate GenAI-generated content for accuracy, depth, and engagement potential**.** Adjust tone, complexity, and relevance as needed.	GenAI can assist in refining clarity, consistency, and grammar but should not be the sole evaluator of instructional quality. Experts must conduct a comprehensive quality review to ensure content accuracy, engagement, and alignment with learning outcomes. Final assessments should incorporate human judgment to validate pedagogical effectiveness before deployment.
(6) **Content Deployment & Iterative Improvement**	GenAI can assist in interpreting structured learner feedback and summarizing input data to support instructional refinement. It may also help generate content revisions in response to identified engagement issues when guided by expert input.	Assess learner engagement and content effectiveness to improve course materials. Use student feedback to guide improvements.	Use GenAI to summarize open-ended learner feedback and generate revision suggestions based on instructional prompts. Ensure that insights from learning analytics are interpreted by educators before making course adjustments. GenAI should support, not replace, human judgment in iterative course improvement.

5. CONCLUSION AND LIMITATIONS

This study examines the potential of Generative AI (GenAI) as a collaborative tool in MOOC content creation, illustrating its capacity to enhance instructional design efficiency while revealing several limitations that require human oversight. The findings indicate that GenAI can effectively contribute to the development of structured content, generate concise summaries, and propose fictional scenarios for learner engagement. Nevertheless, its performance diminishes in tasks requiring conceptual accuracy, contextual depth, and higher-order thinking, particularly in the generation of assessment items such as quizzes. These shortcomings require substantial expert intervention to support pedagogical quality and ensure alignment with learning objectives.

To address these dynamics, the study proposes a Framework for GenAI-Assisted Content Creation in MOOC Design, accompanied by practical Guidelines and Best Practices for MOOC Designers. Building upon existing theoretical models, this framework extends prior work by grounding its design in empirical evidence from the systematic analysis of 30 content units across a Design

Thinking MOOC. It introduces task-level intervention metrics, integrates expert validation at multiple stages, and emphasizes iterative quality improvement. The framework's utility is currently being tested in a pilot MOOC deployment, with data collection underway to assess its impact on student engagement, learning outcomes, and the perceived effectiveness of GenAI-assisted instructional content.

While the framework shows promise, several limitations must be acknowledged. First, the study is based on a single-case context (Design Thinking), which may limit generalizability to other disciplines without further adaptation. Second, the analysis was conducted by a lead instructional designer whose deep domain expertise in design thinking informed the intervention ratings. Although a second researcher independently reviewed and validated the findings to ensure methodological reliability, future studies should consider expanding the coding process to include multiple raters and inter-rater agreement metrics. Additionally, the framework does not yet account for longitudinal learner outcomes or instructor perspectives beyond content development, which may influence the broader applicability of GenAI in online learning.

Future research should explore the scalability of this framework across diverse subject domains, institutional settings, and learner populations. Empirical validation using quantitative learning analytics and mixed-method learner feedback will be essential to test the framework's effectiveness over time. Furthermore, research is needed to examine how advances in GenAI models influence the level of required human intervention and whether newer capabilities reduce or shift effort distribution across content types. Investigating how instructors adapt to GenAI collaboration and how learners perceive GenAI-generated content will also be important to guide ethical and pedagogically sound implementation.

The study concludes by reinforcing the view that GenAI should be strategically integrated within structured instructional design frameworks, rather than positioned as a replacement for human expertise. Ongoing research and practical adjustments will be necessary to maximize the effectiveness of GenAI models as they continue to evolve in education while ensuring that MOOCs maintain high academic standards, learner engagement, and contextual relevance.

Acknowledgements: We would like to thank Prof. Dr. Thomas Staubitz for the early productive discussions that helped shape this study.

References

Abbasi, K. J., Lashari, A. A., & Golo, M. A. (2024). USE OF ARTIFICIAL INTELLIGENCE AS AN ONLINE PLATFORM FOR ENGLISH LANGUAGE TEACHING AT TERTIARY LEVEL, doi: 10.5281/zenodo.14921342.

Amado-Salvatierra, H. R., Chan, M. M., & Hernandez-Rizzardini, R. (2023, October). Combining Human Creativity and AI-Based Tools in the Instructional Design of MOOCs: Benefits and Limitations. In *2023 IEEE Learning with MOOCS (LWMOOCS)* (pp. 1-6). IEEE, doi: 10.1109/LWMOOCS58322.2023.10306023.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: complete edition*. Addison Wesley Longman, Inc..

Chou, D. C. (2018). Applying design thinking method to social entrepreneurship project. *Computer Standards & Interfaces*, *55*, (pp. 73-79), 10.1016/j.csi.2017.05.001.

Dick, W., Carey, L., & Carey, J. O. (2001). The systematic design of instruction (5th). *New York: Longmann*.

Dick, W., Carey, L., & Carey, J. O. (2015). *The systematic design of instruction* (8th ed.). Pearson.

Dickey, E., & Bejarano, A. (2024, October). GAIDE: A Framework for Using Generative AI to Assist in Course Content Development. In *2024 IEEE Frontiers in Education Conference (FIE)* (pp. 1-9). IEEE, doi: 10.48550/arXiv.2308.12276.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of management review*, *14*(4), (pp. 532-550), doi: 10.2307/258557

Emma, O., & Peace, P. (2024). EVALUATING THE IMPACT OF GENERATIVE AI ON PERSONALIZED LEARNING OUTCOMES.

Engeström, Y. (2014). *Learning by expanding: An activity-theoretical approach to developmental research*. Cambridge University Press, doi: 10.1017/CBO9781139814744

Faccia, A., Ridon, M., Beebeejaun, Z., & Mosteanu, N. M. R. (2023, December). Advancements and Challenges of Generative AI in Higher Educational Content Creation A Technical Perspective. In *Proceedings of the 2023 8th International Conference on Information Systems Engineering* (pp. 48-54), doi:10.1145/3641032.3641055.

Ghariz, G., Seghir, H., Boucetta, N., Boubih, S., Janati-Idrissi, R., & El Alaoui, M. (2024). The impact of artificial intelligence on improving text in the process of conceptualization in biology: Case of education sector. *Journal of Theoretical and Applied Information Technology*, *102*(13), (pp. 5203-5214).

Hadyaoui, A., & Cheniti-Belcadhi, L. (2024). IntelliFrame: A Framework for AI-Driven, Adaptive, and Process-Oriented Student Assessments. In *Proceedings of the 20th International Conference on Web Information Systems and Technologies* (pp. 441-448), doi: 10.5220/0013070800003825

Hutchins, N.M., Biswas, G., Zhang, N. et al. (2020). Domain-Specific Modeling Languages in Computer-Based Learning Environments: a Systematic Approach to Support Science Learning through Computational Modeling. Int J Artif Intell Educ 30, (pp. 537–580), doi: 10.1007/s40593-020-00209-z

Kelley, D., & Kelley, T. (2013). Creative Confidence: Unleashing the Creative Potential Within Us All. Stanford Social Innovation Review, doi: 48558/XSA6-0F24

Koh, J. H. L., Chai, C. S., Wong, B., & Hong, H. Y. (2015). Design thinking and education. In *Design thinking for education* (pp. 1–15). Springer, Singapore, doi: 10.1007/978-981-287-444-3_1.

Li, J., Fu, Z., Wang, J., Zhu, L., & Qian, F. (2024, September). Integrating More-Than-Human Perspectives in Speculative Prototyping Based on Actor-Network Theory. In *International Conference on Entertainment Computing* (pp. 383-398). Cham: Springer Nature Switzerland, doi: 10.1007/978-3-031-74353-5_33.

Liu, D. (2024). The Impact and Prospects of AI-Generated Content in Educational Environments. *Transactions on Computer Science and Intelligent Systems Research*, *6*, (pp. 64-71), doi: 10.62051/7mewnb62.

Luckin, R; Holmes, W; (2016) Intelligence Unleashed: An argument for AI in Education. UCL Knowledge Lab: London, UK.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). SAGE Publications.

Morrison, G. R., Ross, S. J., Morrison, J. R., & Kalman, H. K. (2019). *Designing effective instruction*. John Wiley & Sons.

Plattner, H., Meinel, C. & Leifer, L. (2011). Design thinking. Understand – improve – apply. Heidelberg: Springer, doi: 10.1007/978-3-642-13757-0.

Ravarini, A.; Canavesi, A.; Passerini, K. 2024. Beyond Hype or Fear: Avoiding One-sided Perspectives on GenAI in Education. In: 10th International Conference on Higher Education Advances (HEAd'24). Valencia, 18-21 June 2024, doi: 10.4995/HEAd24.2024.17345.

Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? Review of Educational Research, 82(3), 330–348, doi: 10.3102/0034654312457429.

Salmon, G. (2013). *E-tivities: The key to active online learning* (2nd ed.). Routledge, doi: 10.4324/9780203074640.

Stake, R. (1995). *The Art of Case study research*. Cham: Springer.

Stanford University. (n.d.). *Design thinking: A quick overview*. Stanford University. https://web.stanford.edu/class/me113/d_thinking.html

Stefaniak, J. E., & Moore, S. L. (2024). The Use of Generative AI to Support Inclusivity and Design Deliberation for Online Instruction. *Online Learning*, *28*(3), (pp. 181-206), doi: 10.24059/olj.v28i3.4458.

Shukla, M., Goyal, I., Gupta, B., & Sharma, J. (2024). A comparative study of ChatGPT, Gemini, and Perplexity. Int. J. Innov. Res. Comput. Sci. Technol, 12(4), (pp.10-15), doi: 10.55524/ijircst.2024.12.4.2.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), (pp. 257-285), doi:10.1207/s15516709cog1202_4.

Traifeh, H. (2023). *Design Thinking in the Arab world: perspectives, challenges and opportunities* (Doctoral dissertation, Universität Potsdam), doi: 10.25932/publishup-59891.

Traifeh, H., Staubitz, T., & Meinel, C. (2019, October). Improving learner experience and participation in MOOCs: A design thinking approach. In *2019 IEEE Learning With MOOCS (LWMOOCS)* (pp. 165-169). IEEE. doi: 10.1109/LWMOOCS47620.2019.8939623.

Tuomi, I. (2024). Beyond mastery: toward a broader understanding of AI in education. *International Journal of Artificial Intelligence in Education*, *34*(1), (pp. 20-30), doi: 10.1007/s40593-023-00343-4.

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar. A., Spencer-Smith, J., & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with chatgpt, *arXiv preprint arXiv:2302.11382*.

Yin, R. K. (2017). *Case study research and applications: Design and methods*. Sage publications.

Zhou, X., Tong, Y., Lan, X., Zheng, K., & Zhan, Z. (2021, May). AI education in massive open online courses: A content analysis. In *2021 3rd international conference on computer science and technologies in education (CSTE)* (pp. 80-85). IEEE, doi: 10.1109/CSTE53634.2021.00023.