ISSN: 2473-4901
v12 n6705

Teaching Case

Creating videos for class at scale using digital avatar and voice cloning technology

David R. Firth
david.firth@umontana.edu
University of Montana
Missoula, MT 59812, USA

Bryan.Hammer
bryan.hammer@umontana.edu
University of Montana
Missoula, MT 59812, USA

HOOK

This teaching case describes how to create a fully AI-generated digital twin, digital voice clone, highly authentic video of "you" discussing the slides that you use for your in-person class. The video can then be used for teaching online course content, and also to support in-class discussions around AI ethics and responsibility.

ABSTRACT

The increasing demand for online, hybrid, and flexible learning environments has amplified the need for scalable approaches to creating high-quality instructional content. While video-based instruction has become a standard component of contemporary higher education, producing and maintaining instructional videos remains a time-intensive process for faculty. Recent advances in generative artificial intelligence (AI), digital avatar platforms, and voice-cloning technologies offer new opportunities to streamline content creation while preserving instructor presence and engagement. This paper describes how to generate an AI-generated digital avatar with a highly-realistic synthetic voice using this technology to create instructional videos at scale for Management Information Systems (MIS) courses. The resulting approach enables instructors to rapidly generate, revise, and deploy instructional videos with significantly reduced production effort compared to traditional recording and editing methods. The paper discusses the content development process, and practical considerations associated with deploying avatar-based instructional media in higher education settings. These include using the AI generated videos to stimulate in-class discussions around the ethics of AI and the responsible use of AI.

Keywords: Generative AI, Video production, Ethics of AI, Responsible use of AI, Online education

Creating videos for class at scale using digital avatar and voice cloning technology

David R. Firth and Bryan.Hammer

1. INTRODUCTION

The rapid advancement of generative artificial intelligence (AI), digital avatar systems, and voice cloning technologies is reshaping the landscape of higher education content delivery. In particular, these technologies offer new opportunities for Management Information Systems (MIS) educators to address longstanding challenges associated with online and hybrid instruction, including scalability, consistency, production time, and student engagement. While video-based instruction has become a central component of contemporary pedagogy, the creation of high-quality instructional videos remains both labor-intensive and technically demanding for faculty members (Lowenthal & Dunlap, 2018; Martin, et al., 2019). As institutions continue to expand online offerings and emphasize flexible learning environments, educators increasingly seek efficient methods for producing engaging instructional media at scale.

Recent developments in AI-driven avatar generation and synthetic voice technologies provide a potential solution to these challenges. Platforms such as Synthesia, HeyGen, and ElevenLabs enable instructors to create realistic digital representations of themselves capable of delivering instructional content with minimal incremental production effort. These systems allow faculty to generate narrated videos from text scripts without repeated recording sessions, professional video editing, or extensive multimedia expertise. Consequently, instructors can rapidly update course materials, personalize communication, and create large volumes of instructional content while maintaining consistency in presentation and branding. Prior studies published in the *Journal of Information Systems Education (JISE)* and *Information Systems Education Journal (ISEDJ)* have emphasized the growing importance of multimedia learning environments, instructor presence, and technology-supported engagement in MIS education (Everett, 2015; Carraher-Wolverton, et al. 2025).

This technology and approach described here is not to be confused with deepfakes. Deepfakes are AI-generated synthetic media that realistically depict individuals saying or doing things they never actually said or did (Westerlund, 2019). However, unlike malicious deepfakes intended to deceive, educational avatars are typically developed with the knowledge and consent of the represented individual. These educational video avatars can be used in-class to stimulate discussions about AI ethics and responsible use of AI. (Westerlund, 2019). However, unlike malicious deepfakes intended to deceive, educational avatars are typically developed with the knowledge and consent of the represented individual. Usefully, though, these educational video avatars can be used in-class to stimulate discussions about AI ethics and responsible use of AI.

This paper explores the use of digital avatar and voice cloning technology to create instructional videos for MIS courses at scale. Specifically, the paper examines the technological workflow, implementation considerations, pedagogical implications, and operational efficiencies associated with AI-generated instructional content. By documenting the design and deployment of an AI-assisted video production process, this paper contributes to the emerging conversation surrounding generative AI in IS education and offers practical guidance for educators seeking sustainable approaches to multimedia course development.

2. LITERATURE REVIEW

The use of multimedia and video-based instruction has become increasingly important in online and hybrid learning environments. Prior research demonstrates that instructor presence and professionally designed video content positively influence student engagement, satisfaction, and perceived learning outcomes in online courses (Martin, et al., 2019). These findings are particularly relevant in MIS education, where asynchronous instructional delivery continues to expand (2025 CHLOE 10 Report).

Recent advances in artificial intelligence and automation technologies have also transformed knowledge work and digital content creation. Davenport and Ronanki (2018) argue that AI-enabled systems increasingly support routine cognitive tasks and improve operational scalability across organizational settings. Applied to higher education, AI-generated avatars and synthetic voice technologies may reduce the time and technical effort required for instructional video production while maintaining consistency and quality.

In addition, emerging research on trust and human interaction with AI systems suggests that users can develop positive perceptions of synthetic agents when systems exhibit realism, consistency, and effective communication characteristics (Glikson & Woolley, 2020). This stream of research provides theoretical support for the use of digital avatars in instructional contexts, where instructor presence and communication effectiveness are critical factors in student learning experience.

3. THE STEPS

contains a table-version of this section that highlights the steps. The slides themselves are in Appendix B. The slides were created for the fully in-person version of the class. As such, they are mainly just images so that they can be used to talk about the content.The following lays out the steps taken to create a fully digital video facsimile of a course instructor "talking" about the content of the course slides. Appendix A contains a table-version of this section that highlights the steps. The slides themselves are in Appendix B. The slides were created for the fully in-person version of the class. As such, they are mainly just images so that they can be used to talk about the content.

Step 1: Create slides for class

Step 1 is no different from any other class in that PowerPoint (or something similar) slides are prepared to cover the course content. We do note that generative AI can also be used at this stage to help determine what to be covered, and even to generate images for the slides. This paper does not cover this.

Step 2: Import the slides into Google's NotebookLM

Using Google's NotebookLM product (https://notebooklm.google.com/), the slides from Step 1 are imported. This creates a Retrieval-Augmented Generation (RAG) AI. Instead of relying solely on what the generative AI platform learned during training, the AI system is given a search engine and a library of specific documents. When you ask a question, the system first looks up the relevant pages in that library, hands those pages to the AI, and says, "Read this first, then answer the user's question using only this information.

Step 3: Create a podcast

One of the powerful features of Google's NotebookLM is that in RAG mode you can create a podcast. This is done simply by hitting the "Audio overview" button. Save the audio podcast.

Step 4. Create an account in HeyGen

Create an account in HeyGen.com. HeyGen is a generative AI video platform designed to create professional, spokesperson-style videos entirely from text scripts without requiring traditional filming equipment or actors. The platform utilizes highly photorealistic "Digital Twin" avatars and voice cloning to deliver natural-looking human presentations complete with matched emotional gestures and expressions.

Step 5. Import the audio podcast from Step 3 and update the transcript

Import the audio podcast from Step 3 into HeyGen. This will create a transcribed version of the podcast. The NotebookLM podcast is a two-person podcast, and the video you are creating is a one-person video. Go through the transcript to remove or update anything that doesn't make sense for a one-person video.

Step 6. One-time create a digital twin

The creation of your digital twin is a one-time process. With your laptop camera and microphone on, you read a paragraph of words provided by HeyGen. You are prompted to add emphasis to words that make sense to be emphasized, and also to use your hands to add additional emphasis or comment. As the screenshot in Appendix A shows with the image at the left, which is taken directly from my recording, you can change the clothes you are wearing during this one-time digital twin creating process to reflect how you want your digital-twin output to be. I had not thought this through for the recording, and was wearing a hoodie with a distinctive logo on it. I was able to change this to a more business-casual look by swapping the hoodie out for a.

Step 7. Create a digital voice clone

HeyGen has several default voices that can be added to your digital twin. For my videos for class I felt it was important for the video of me to also have my voice, so I created a digital voice clone inside HeyGen. The process here is similar to the creation of the digital twin in Step 6 above, but you are talking generally rather than reading from HeyGen prompts. The idea behind this is that the voice that is being cloned doesn't sound like you are reading, and is more natural instead.

Step 8. Create the video

The video creation setup is shown in Appendix A. In the left panel you have the transcript from NotebookLM. The transcript can be edited here. In the right panel you have the options to select the avatar and the voice. Appendix A shows these as being my digital twin and my digital voice clone. The main middle panel is the video itself, and the part below it is the video timeline. This is much the same as a video editing tool, and some basic trimming of the video is an option here. Actual creation of the video once you've decided that the script, avatar and voice are what you want, is obtained by selecting the "Generate" button. This 7-minute video took some 20-minutes to render out.

Step 9. Decide on a delivery mechanism for your video

I prefer to store and share my videos from my YouTube channel. I typically set the "Visibility" to "unlisted" as this allows the video to be accessed only by someone who has the exact YouTube-sharing URL. This particular video is set to "Public" visibility so it can be seen by anyone who has access to my YouTube channel. When I post my videos in Canvas (or any other LMS set by the institution), I use the YouTube embed script and embed the video into Canvas.

4. STUDENT REACTIONS

I have used this video in several in-person classes, and several campus and community presentations. The process I use is to walk through the first four slides of my generative AI presentation using the slides in Appendix B. I do this to set a base level of what a presentation by me of the slides would look and sound like. I then talk about the Google NotebookLM podcast option, and I play the podcast from Step 3 above. This is an important step, as students get to hear what is going to be said in the video that I play next. Finally, I show the video. I do this without saying that it is fully AI-generated. I usually show about 2 minutes of it. I then talk about how it is fully AI-generated, and show it again. Once students have seen the video for the second time, and understand what is going on, I then ask students what they think.

I have consistently had very strong, very broad student participation based on this approach. This includes classes with 200-level all-major business students, 400-level management information systems

undergraduate students, as well as talks given in the community to a wide variety of business audiences. Initial comments are usually focused on how amazing the quality of the video is, and how lifelike it seems. The comments then always pivot to the ethics of doing this, including what sort of disclosure should be required by me for providing students with a completely AI-generated video lecture. There is then a discussion around what it means for me to have a digital twin and digital voice clone available to me, and HeyGen having access to that too. This is a very good opportunity to discuss AI Ethics. Table 1 gives a summary of the different concepts around ethical considerations.

- **Ethical AI:** AI systems that adhere to ethical principles and thus ensure that these systems are aligned with social norms and values.
- **Safe AI:** A key concern is that AI systems may cause physical and psychological harm to those affected by its output, either directly or indirectly.
- **Responsible AI:** While similar to ethical AI, the idea here is that AI should be used in a responsible way that does not cause any harm. While ethical AI focuses on the underlying principles, responsible AI focuses on the practices, policies, and governance of AI.
- **Explainable AI:** It is difficult, if not impossible, to understand how a certain prediction has been reached in complex AI systems. Explainable AI promotes practices that alleviate this problem. For example, using simulations, What-if scenarios, and counterfactuals.
- **Robust AI:** Robustness is a technical requirement, arguing that AI models should produce consistent output even in extreme or edge cases, reducing the risk of unintended consequences.
- **Trustworthy AI:** The concept of trustworthiness was introduced by the EU to bring together three key requirements that AI systems should fulfil:
    - Legal compliance, for example, with privacy laws such as the GDPR
    - Technical robustness, to ensure consistent performance of the model
    - Ethical soundness, to ensure adherence to social norms and values

Table 1. Extract from Oxford University Artificial Intelligence Programme, June 2026, Module 5 on AI Ethics

Typically for class, I focus on just the "ethical" and "responsible" concepts. I discuss having AI being aligned with social norms and values, and that AI should be used in a responsible way that does not cause any harm. The discussions are always lively, and purposeful.

5. CONCLUSION

The ability to generate full-length, highly realistic, videos for class from slides that are used for regular in-class presentations and discussions is revolutionary. Digital avatar and voice-cloning technologies provide a viable approach for creating instructional videos at scale while maintaining many of the elements of instructor presence valued by students. As Lowenthal and Dunlap (2018) demonstrated, students perceive instructor presence as an important component of effective online learning, making it critical that technology-enhanced instructional methods preserve meaningful connections between faculty and learners. By leveraging AI-generated avatars as a supplement to traditional teaching practices, educators may be able to achieve greater efficiency in content creation without sacrificing the social presence that contributes to student engagement and learning outcomes. The key issue here is that this approach and technology allows such videos to be generated with limited time commitment, and at scale. In addition to being able to generate these videos for online asynchronous classes, the videos can be used for in-person classes to showcase the current state of generative AI technology, and to provide a starting point for discussions about AI ethics and responsible use of AI.

REFERENCES

2025 CHLOE 10 Report: https://qualitymatters.org/qa-resources/resource-center/articles-resources/CHLOE-10-report-2025.

Brynjolfsson, E., & McAfee, A. (2017, July 18). The business of artificial intelligence. Harvard Business Review. https://hbr.org/2017/07/the-business-of-artificial-intelligence

Carraher-Wolverton, C., Lai, G., Navarre, J. T., Davis, D., & Lanier, P. (2025). Investigating the Factors That Engage Students to Be Successful in Hybrid Business Analytics Courses. Journal of Information Systems Education, 36(4), 417-431. https://doi.org/10.62273/CSVP4044

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. Harvard Business Review, 96(1), 108–116. https://hbr.org/2018/01/artificial-intelligence-for-the-real-world

Everett, D. R. (2015). Adding Value: Online Student Engagement. Information Systems Education Journal, 13(6) pp 68-76. http://isedj.org/2015-13/ ISSN: 1545-679X.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. MIS Quarterly, 27(1), 51–90. https://doi.org/10.2307/30036519

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals, 14(2), 627–660. https://doi.org/10.5465/annals.2018.0057

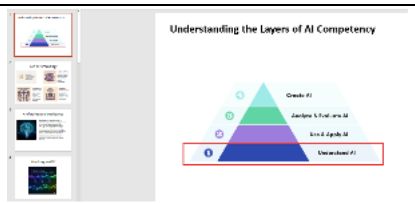
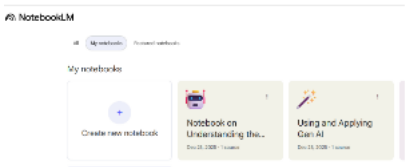
Lowenthal, P. R., & Dunlap, J. C. (2018). Investigating students' perceptions of instructional strategies to establish social presence. Distance Education, 39(3), 281–298. https://doi.org/10.1080/01587919.2018.1476844

Martin, F., Budhrani, K., Kumar, S., & Ritzhaupt, A. (2019). Award-winning faculty online teaching practices: Roles and competencies. Online Learning, 23(1), 184–205. https://doi.org/10.24059/olj.v23i1.1329

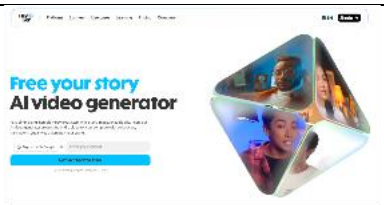
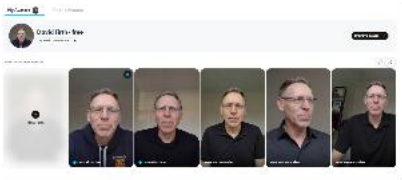
Westerlund, M. (2019). The emergence of deepfake technology: A review. Technology Innovation Management Review, 9(11), 39–52. https://doi.org/10.22215/timreview/1282

ISSN: 2473-4901
v12 n6705

APPENDIX A: Steps in the process to create a digital twin, digital voice clone authentic-looking video of class presentation slides

Step	Screenshot of the process	Short description of the process
1		PowerPoints for class are the starting point.
2		Upload the slides from Step 1 to Google's NotebookLM Notebooklm.google.com
3		Use the "audio overview" function in NotebookLM to create a podcast

ISSN: 2473-4901
v12 n6705

4		Open an account in HeyGen.com As of May 2026 this costs $29 per month, but an additional $15 of credits was required to do what needed to be done.
5		Import the "audio overview" podcast from NotebookLM to HeyGen. HeyGen takes the audio podcast and transcribes it.
6		Record 30-seconds of yourself reading a paragraph of words to create a video digital twin. This is done just once
7		Record 30-seconds of you talking generally to create an audio voice clone
8		The creation area in HeyGen, showing the script from NotebookLM on the left, and the digital

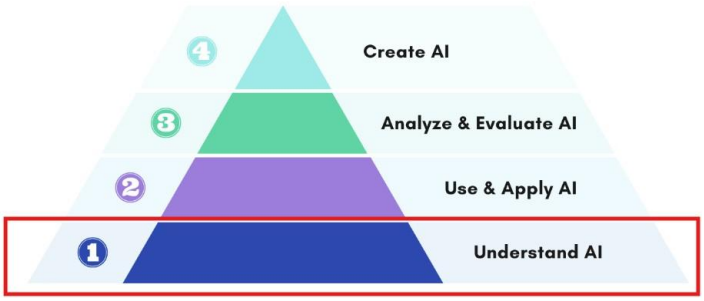
ISSN: 2473-4901
v12 n6705

		twin with voice clone in the middle. The choice of digital twin avatar and digital voice clone is made in the panel on the right.
9		Decide on where to store the video for sharing, and what sharing you will allow

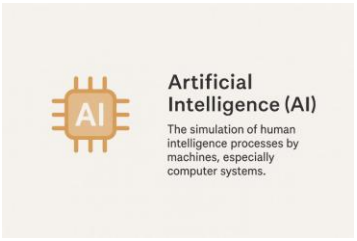
ISSN: 2473-4901
v12 n6705

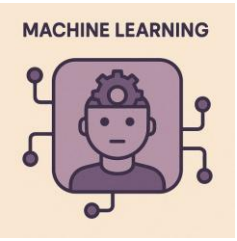
APPENDIX B: First four slides of the in-class presentation slides used as input to Google's NotebookLM

Understanding the Layers of AI Competency

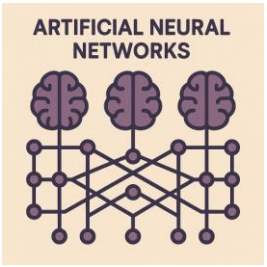


Gen AI Terminology





A subset of AI that enables systems to learn and improve from experience without being explicitly programmed



Computational models inspired by the human brain, used to recognize patterns and solve complex problems

Advanced AI models trained on vast amounts of text data to understand and generate human-like language.

ISSN: 2473-4901
v12 n6705

Artificial General Intelligence

Artificial General Intelligence (AGI) is a hypothetical type of artificial intelligence with the intellectual capability to understand, learn, and apply knowledge across a wide range of tasks, much like a human. Unlike narrow AI, which excels at specific tasks (like playing chess or recognizing faces), AGI would possess general problem-solving abilities, abstract thinking, comprehension, and creativity. It would be able to learn and perform any intellectual task that a human being can.

How long to AGI?