

Generative AI and Knowledge Management: A Multimodal Perspective on Knowledge Creation

Anastasiia Iufereva¹ and Peter Mozelius²

¹IULM University, Milan, Italy

²Mid Sweden University, Sundsvall, Sweden

anastasiia.iufereva@studenti.iulm.it

peter.mozelius@miun.se

Abstract: The rapid evolution of Generative AI (GenAI) has transformed the ways in which knowledge is created, shared, interpreted, and applied in organization and educational contexts. While earlier studies have often focused on the technical capabilities of GenAI or on the detection of synthetic content, less attention has been given to how these tools influence knowledge management processes at the individual and group levels. This study addresses that gap by examining how GenAI affects the production, transfer, validation, and use of knowledge among individuals and within collaborative settings. The aim of this study is to explore the current state of 'AI vs AI' by generating a dataset comprising texts, images, audio and video, utilizing a set of freely available AI tools. This pilot illustrative study was based on a two-stage strategy: in the first stage synthetic media was generated with different GenAI tools to mimic human creativity. A second stage, which involved a rigorous evaluation of AI-detection tools for each modality separately. Texts were created in three different GenAI tools and Images were prompted to mimic Renaissance paintings by Michelangelo and Raphael. Audio was generated with the purpose of mimicking interviews with the Renaissance painters. Videos, finally, were created to be deceptive deepfakes, placing persons in environments and situations where they never have been in real life. Findings provide a comparative analysis of tools' accuracy and false-positive rates that could act as valuable guidelines in the increasing flood of AI-generated content. This research contributes to the growing discussion on how organizations and institutions can manage the opportunities and risks associated with GenAI in knowledge-intensive environment. The study highlights the need for critical AI literacy, transparent knowledge practices, and governance mechanisms that ensure the responsible integration of GenAI into individual and collaborative knowledge work.

Keywords: Generative AI, Knowledge management, Individual knowledge work, Group collaboration, Organizational learning, AI literacy

1. Introduction

An ongoing and rapid development of Generative AI (GenAI) has significantly reshaped the landscape of knowledge creation. From producing written text and visual material to generating audio and video, GenAI has become increasingly capable of supporting tasks that were previously considered to require distinctly human experience. As a result, GenAI is no longer only a technical innovation but also a knowledge management challenge, affecting how knowledge is generated and applied across different contexts.

Within knowledge-intensive environments, GenAI offers a wide range of opportunities regarding co-creation knowledge. At the individual and group level, it can assist with idea generation, summarization, problem-solving, support brainstorming, and the transformation of information into accessible knowledge formats. At the same time, GenAI may also create risks, such as cognitive offloading, reduced critical engagement, and difficulties in distinguishing between reliable and misleading outputs. Moreover, it may also affect trust, participation, and the validation of shared knowledge.

Recent knowledge management literature increasingly conceptualizes GenAI not only as an automation technology but also as a knowledge-creation infrastructure. From a knowledge management perspective, GenAI can enhance knowledge creation by accelerating information processing, enabling synthesis across heterogeneous knowledge sources, supporting ideation, and transforming tacit or fragmented inputs into more explicit and reusable knowledge forms (Alavi et al., 2024). Similarly, recent studies argue that GenAI changes the knowledge management paradigm by strengthening knowledge acquisition, codification, sharing, and application, while also creating new requirements for human oversight, data quality, ethical governance, and critical validation (Kaczorowska-Spychalska et al., 2024). Applied research on GenAI-enabled knowledge management further identifies use cases such as automated content generation, knowledge discovery, summarization, decision support, and access to organizational knowledge repositories, suggesting that GenAI can improve both individual knowledge work and collaborative knowledge practices when embedded into appropriate socio-technical processes (Kudryavtsev et al., 2024; Pimentel & Veliz, 2024). Empirical evidence from knowledge-intensive work also shows that GenAI can improve productivity and output quality for tasks within its capability frontier, although performance may deteriorate when users over-rely on AI in tasks that

require contextual judgment or expert validation (Dell'Acqua et al., 2023). Therefore, the current literature suggests that GenAI enhances knowledge creation most effectively when it is treated as a human–AI co-creation tool rather than as a replacement for human expertise.

An important concern is therefore not only what GenAI can produce, but how its outputs shape knowledge processes among individuals and groups. An ongoing and rapid development of Generative AI (GenAI) has opened up for both a beneficial and a deceptive content creation. To create fake content and manipulating images is not a new phenomenon, but GenAI tools have drastically increased the amount and the pace (Iufereva & Mozelius, 2025). The widespread dissemination of AI-generated content poses a serious threat to society (Kreps & Kriner, 2025; Romanishyn, 2025; Guvenilir & Yildiz, 2025). The dissemination of controversial content can exacerbate conflicts at the local and international levels (Vosoughi et al., 2018; Guess et al., 2020a), undermine trust in institutional sources of information (Guess et al., 2020b; V-Dem Institute, 2025), and increase polarization within society (Çetinkaya, 2025).

Synthetic content has become an indispensable weapon for a number of countries, including the United States, Europe, and Asia, which use generated content in information warfare to discredit the enemy country and promote propaganda narratives (Crawford 2025). This is why content authentication is currently a key issue widely discussed in the public sphere (Cazzamatta, 2025; Diep, 2024). Critical thinking and common sense are often insufficient to determine whether text or visual content is the product of GenAI. Therefore, one of the key questions of this study is to test AI-based detection tools aimed at debunking AI artifacts. This will determine the effectiveness of such tools and also indicate where and for what types of content these tools should be developed.

The aim of this study is to explore how GenAI influence knowledge creation by examine the current state of “AI vs AI” through a multimodal dataset of text, images, audio, and video generated with freely available AI tools. In particular, the study investigates to what extent AI-based detection tools are reliable and capable of providing humans with objective results about whether content was generated by AI, and what this means for knowledge practices in contemporary digital environments.

2. Methods and Material

This exploratory pilot study in the field of deepfake detection was based on a two-stage strategy. In the first stage, synthetic outputs were generated with the use of different GenAI tools to simulate human creativity. To provide a cohesive thematic framework, the research focused on two of the most famous Renaissance figures in history: Michelangelo Buonarroti (1475–1564) and Raphael (1483–1520). Just as their artistic rivalry shaped the cultural landscape of early 16th-century Rome, this study also establishes a parallel 'competition' between selected AI-based detection tools. Therefore, the following outputs were produced:

[1] Textual output: Six narrative units were generated using three different Large Language Models (LLMs): ChatGPT 5.2, Le Chat, and Gemini 1.5 Pro. To ensure a cross-platform comparison, a standardized prompt engineering approach was applied. The authors utilized the following prompts:

Prompt 1: *“Create a short 100-word narrative text about the art of Raphael”*

Prompt 2: *“Create a short 100-word narrative text about the art of Michelangelo”.*

Text scripts were generated with Le Chat and subsequently converted into audio using Luvvoice text-to-speech.

[2] Synthetic Video Generation: To evaluate the detection of synthetic visual media, two video sequences were produced using the Sora (OpenAI) model. The generation was guided by highly descriptive prompts to test the model’s ability to simulate historical artistic processes:

Prompt 1: *“Create a short, highly realistic cinematic film showing how Michelangelo creates his art”*

Prompt 2: *“Create a short, highly realistic cinematic film showing how Raphael creates his art”*

[3] Audio generation: The initial idea was to create interviews in audio format with the well-reputed ElevenLabs tool at <https://elevenlabs.io/>. However, when the Agent feature was prompted with: “Could you make a short interview with the Renaissance painter Michelangelo?”, the response appeared to be: *“I am designed to be a friendly technical support specialist at LaptopCare Solutions, so I can't do that. Is there anything I can help you with regarding your laptop?”*.

The audio output then had to be created in two steps: 1) Creating interviews in text format with the open version of Le Chat at: <https://chat.mistral.ai/chat>, and then 2) using the text-to-speech feature in the Luvvoice portal at:

<https://luvvvoice.com/>. During the time of writing this paper most of the services at the ElevenLabs portal are not free to use anymore, while the Luvvoice service for text-to-speech conversion still is without any cost. The Michelangelo interview was transformed to mp3 with a male voice and the Raphael interview with a female voice. British English was the choice for both the interviews.

The initial affirmative reply from the French cat-bot was: *"Absolutely! Here's a short, imaginative interview with Michelangelo Buonarroti, the legendary Renaissance painter and sculptor. Of course, this is a creative reconstruction based on historical knowledge—since Michelangelo lived from 1475 to 1564, we can't actually speak with him today!"*

[4] Image creation: Digital images in Renaissance style were created using the Raphael AI tool at <https://raphael.app/>. The images were created with straightforward prompts such as the one in Figure 3 below:

Prompt 1: "A Renaissance painting in the style of Michelangelo please "

Prompt 2: "Could you please create a Renaissance painting in the style of Raphael"

The output comprised 3 paintings in what Raphael AI see as Michelangelo style, and 4 paintings in what Raphael AI see as Raphael style.

The second stage involved a rigorous evaluation of AI-detection tools for each modality separately. This study used the freely available AI detection tools:

- Text (AI-generated text checkers): Scribbr AI Detector, Sapling AI Detector, Copyleaks AI Detector
- Video (AI/deepfake video checkers): DeepFake-o-meter, Sightengine, Hive AI Detector
- Audio (AI voice / deepfake audio checkers): ElevenLabs AI Speech Classifier, Resemble AI Free Audio Deepfake Detection, DeepFake-o-meter
- Images (AI image checkers): Illuminarty, QuillBot AI Image Detector, Hive AI Detector

A choice that was motivated by several reasons. Firstly, the goal was to assess the degree of democratic accessibility of these tools for regular users — without payment and with user-friendliness. Secondly, the opportunity of immediate verification of digital content authenticity.

3. Findings

3.1 Textual Output Generated by AI & Testing AI-based Detection Tools

The textual output for this study was synthesized using LLMs, specifically ChatGPT 5.2, Le Chat, and Gemini 1.5 Pro. Analysis indicates that the generated output exhibits a distinct narrative style with high linguistic proximity to human-authored text (See Table 1). While native speakers may identify occasional non-idiomatic constructions, the evaluated models generally showed high proficiency in text generation.

Table 1: Textual output extracted from LLMs – ChatGPT 5.2, Le Chat, and Gemini 1.5 Pro. These texts were completely generated by AI for the scientific purpose. Data extraction: 9.02.2026

LLM	Output
ChatGPT 5.2	Raphael's art feels like a quiet conversation between harmony and beauty. Walking into a room filled with his paintings, one senses balance and grace in every figure and color. He portrayed human emotions with calm clarity, giving his characters gentle expressions and natural gestures. In works like <i>The School of Athens</i>, Raphael brought philosophy to life, placing thinkers in an elegant architectural world. His Madonnas reveal warmth and tenderness, reflecting deep spiritual serenity. Raphael combined classical ideals with human sensitivity, creating art that feels both grand and intimate, leaving viewers with a sense of peaceful admiration. Michelangelo's art feels powerful, dramatic, and full of tension. Standing beneath the Sistine Chapel ceiling, one experiences a story told through muscular figures and intense movement. Michelangelo saw sculpture inside every block of marble, famously revealing life in works like <i>David</i> and <i>Pietà</i>. His figures seem to struggle, stretch, and breathe, reflecting human strength and vulnerability. Unlike artists who focused on calm beauty, Michelangelo explored emotion through energy and physical expression. His work shows the complexity of humanity and spirituality, presenting art as a battle between limitation and creative vision, leaving viewers amazed by its monumental force.

LLM	Output
Le Chat	Raphael Sanzio, the master of harmony, painted serenity into every brushstroke. His frescoes in the Vatican's <i>Stanze</i> radiate grace—<i>The School of Athens</i> gathers philosophers in a grand ballet of thought, Plato pointing heavenward, Aristotle grounding wisdom in earth. His Madonnas, like the <i>Sistine Madonna</i>, cradle tenderness, their gazes soft yet profound, drawing viewers into sacred intimacy. Raphael's genius lay in balance: vibrant colors, fluid compositions, and faces alive with quiet emotion. He didn't just depict beauty; he made it breathe, turning canvas into windows where heaven and earth met in perfect, timeless dialogue. Michelangelo Buonarroti, the Renaissance titan, transformed stone into breath, muscle, and soul. His chisel danced on marble, liberating figures like <i>David</i>—a young shepherd's defiance frozen in time, veins pulsing beneath polished skin. In the Sistine Chapel, he defied gravity: <i>The Creation of Adam</i> stretched across the ceiling, God's finger nearly touching man's, a spark of divine connection. His <i>Pietà</i> cradled grief in Mary's embrace, her youthful face a mirror of eternal sorrow. Michelangelo didn't just sculpt; he revealed the hidden life within the stone, making the heavens weep and mortals gaze upward, forever chasing the sublime.
Gemini 1.5 Pro	Raphael Sanzio didn't just paint; he orchestrated harmony on wood and canvas. While Michelangelo wrestled with divine agony and Leonardo chased shadows, Raphael captured the sublime equilibrium of the High Renaissance. His brush moved with a deceptive ease, transforming complex theology into the effortless grace of <i>The School of Athens</i>. In his Madonnas, he perfected a "human divinity"—faces so serene they felt like a quiet breath held in time. Through his use of the pyramidal composition, he achieved a structural balance that defined Western beauty for centuries. Raphael's art remains a testament to the idea that perfection isn't found in chaos, but in clarity. If Raphael was the poet of harmony, Michelangelo was the sculptor of the soul. For him, every block of marble held a prisoner waiting to be freed. His art wasn't about decorative beauty; it was a visceral struggle between spirit and stone. From the colossal defiance of <i>David</i> to the churning humanity of the Sistine Chapel, Michelangelo's work celebrates the <i>terribilità</i>—an overwhelming power that commands awe. He favored the muscular, twisting human form to express internal agony and divine ecstasy. To Michelangelo, the body was the ultimate vessel for truth, carved with a precision that made cold stone pulse with life.

Subsequently, the generated texts were systematically processed through AI-based detection tools. The following platforms were selected for this research: Scribbr AI Detector, Sapling AI Detector, and Copyleaks AI Detector. The evaluation results are summarized in Table 2.

Table 2: AI-generated text checkers results: Overview of effectiveness

AI detection tool	ChatGPT 5.2	Le Chat	Gemini 1.5 Pro
Scribbr AI Detector	87% of text is likely AI	100% of text is likely AI	100% of text is likely AI
Sapling AI Detector	73% of text is likely AI	100% of text is likely AI	100% of text is likely AI
Copyleaks AI Detector	100% of text is likely AI	100% of text is likely AI	100% of text is likely AI

3.2 Video Output Generated by AI & Testing AI-based Detection Tools

Utilizing SORA mobile application, two distinct video sequences featuring Michelangelo Buonarroti and Raphael were generated based on the aforementioned prompts. The output depicts a high-quality cinematic realism, closely approximating authentic footage. Notably, the generated content displayed no discernible AI-driven anomalies or motion artifacts.



Figure 1: (left picture) Video output featuring Raphael extracted from SORA. Video duration: 10 seconds. This video was completely generated by AI for this study. Data extraction: 09/02/2026. (right picture) Video output featuring Michelangelo extracted from SORA. Video duration: 6 seconds. This video was completely generated by AI for this study. Data extraction: 09/02/2026.

The authors then sequentially loaded these texts into AI-based detection tools. The following AI-generated text checkers were selected for this study: DeepFake-o-meter, Sightengine, and Hive AI Detector.

Table 3: AI-generated video checkers results: Overview of effectiveness

AI detection tool	SORA: video 1 (Raphael)	SORA: video 2 (Michelangelo)
DeepFake-o-meter*	100% of video is likely AI	100% of video is likely AI
Sightengine	88% of video is likely AI	69 % of video is likely AI
Hive AI Detector	99% of video is likely AI	99% of video is likely AI

* Clarification: DeepFake-o-meter suggested a set of detection systems and AVSRDD (2025) was chosen to detect generated video.

3.3 Audio Output Generated by AI & Testing AI-based Detection Tools

The initial text output interviews generated by the Le Chat bot:

Michelangelo interview:

Prompt: "Could you make a short interview with the Renaissance painter Michelangelo?"

Absolutely! Here's a short, imaginative interview with **Michelangelo Buonarroti**, the legendary Renaissance painter and sculptor. Of course, this is a creative reconstruction based on historical knowledge—since Michelangelo lived from 1475 to 1564, we can't actually speak with him today!

Interviewer: Michelangelo, thank you for joining us. You're renowned as both a painter and a sculptor. Which art form do you feel most connected to?

Michelangelo: *Sculpture* is my true love. In every block of marble, I see a figure waiting to be freed. Painting, while noble, is but a shadow of the truth I find in stone.

Interviewer: Your work on the Sistine Chapel ceiling is iconic. What was your experience like, painting such a vast masterpiece?

Michelangelo: Agony and ecstasy! I am not a painter by trade, yet Pope Julius II insisted. For four years, I lay on my back, my neck cramped, my eyes strained. But when I saw the figures come to life—Adam reaching for God's finger—I knew it was divine.

Interviewer: You also sculpted the *David* and the *Pietà*. What inspired these works?

Michelangelo: *David* was the triumph of human spirit over tyranny, carved from a flawed block others had abandoned. The *Pietà* was my meditation on sorrow and grace, the Virgin holding Christ's lifeless body. Both were prayers in stone.

Interviewer: Many say you were a perfectionist. Is that true?

Michelangelo: I sought the ideal, yes. If a nose was a millimeter too long, I would rework it. Art is not about ease, but about truth—even if it means chiseling until my hands bled.

Interviewer: What advice would you give to young artists today?

Michelangelo: Study anatomy, draw constantly, and never fear the blank page or the uncut stone. And remember: the greatest art is born from struggle, not comfort.

Raphael interview:

Prompt: "Could you please make a similar interview with the Renaissance painter Raphael?"

Certainly! Here's a short, imaginative interview with **Raffaello Sanzio da Urbino**, better known as **Raphael**, the master of harmony and grace during the High Renaissance.

Interviewer: Raphael, thank you for speaking with us. Your works are celebrated for their beauty and balance. What draws you to painting?

Raphael: Painting is the mirror of the soul. It allows me to capture the divine proportion, the harmony of nature, and the dignity of the human spirit. Every brushstroke is a step closer to perfection.

Interviewer: Your frescoes in the Vatican, especially *The School of Athens*, are legendary. What was your vision for this masterpiece?

Raphael: I wanted to gather the greatest minds of antiquity—Plato, Aristotle, Euclid—in one grand space, united by the pursuit of knowledge. The composition is a dance of intellect and beauty, where architecture and philosophy become one.

Interviewer: You were known for your ability to portray the Madonna and Child with extraordinary tenderness. What inspired these works?

Raphael: The bond between mother and child is the purest expression of love. I sought to capture that intimacy, that quiet grace. The Madonna is not just a religious figure; she is every mother, filled with warmth and devotion.

Interviewer: You were a contemporary of Michelangelo and Leonardo da Vinci. How did their work influence you?

Raphael: Michelangelo's power and Leonardo's genius are unmatched. From Michelangelo, I learned the grandeur of form; from Leonardo, the subtlety of expression. But I sought to blend their strengths—creating works that are both majestic and serene.

Interviewer: What advice would you give to artists today?

Raphael: Study the masters, but find your own voice. Let your heart guide your hand, and never lose sight of beauty—it is the language of the soul.

Table 4: AI-generated audio checkers results: Overview of effectiveness

AI detection tool	Audio 1 (Michelangelo)	Audio 2 (Raphael)
ElevenLabs AI Speech Classifier	2% probability that it was AI generated	2% probability that it was AI generated
Resemble AI Free Audio Deepfake Detection	Fake	Fake
DeepFake-o-meter	22.7% probability of AI generation	43.5% probability of AI generation

Image output generated by AI & Testing AI-based Detection Tools

As can be seen in Figure 3 and in Figure 4 the generated Images have been watermarked in the Raphael AI generated images from AI tool. A feature that probably can help AI detection tools to make distinction between AI generated images and pictures of original Renaissance paintings.

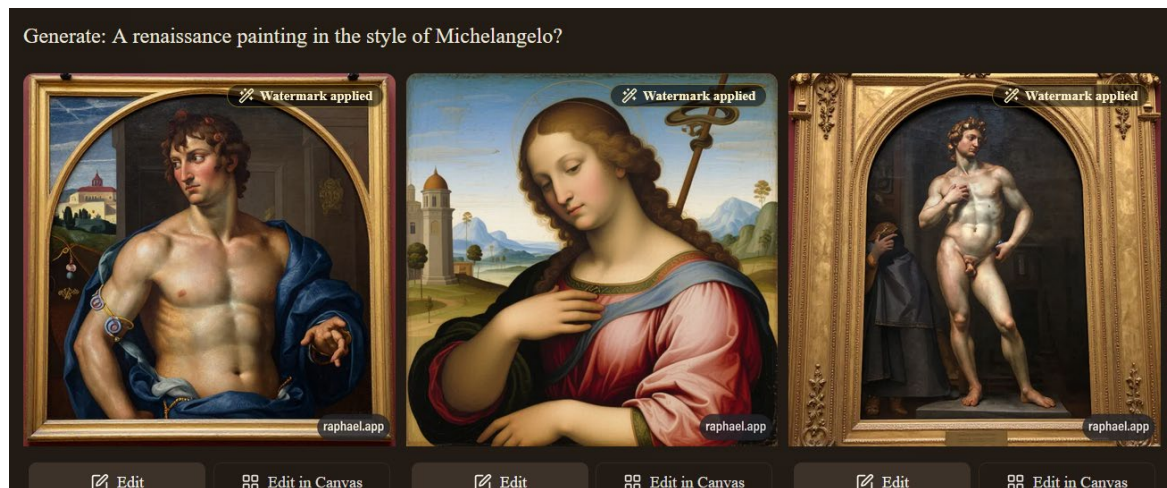


Figure 3: Image output in Michelangelo style from the Raphael AI tool at <https://raphael.app/> This image was generated with GenAI for this study. Data extraction: 13/02/2026

The two first images have a Renaissance style but also a touch of kitsch and the light that we associate with AI generating images. The third image depicts something that strongly resembles Michelangelo's David statue, but not the way Michelangelo painted. Since an image generator is built around an adversarial network with generator and a discriminator, it can only generate images from the generators training data. To tell the difference between a painting and a statue failed in the generation of the third image.

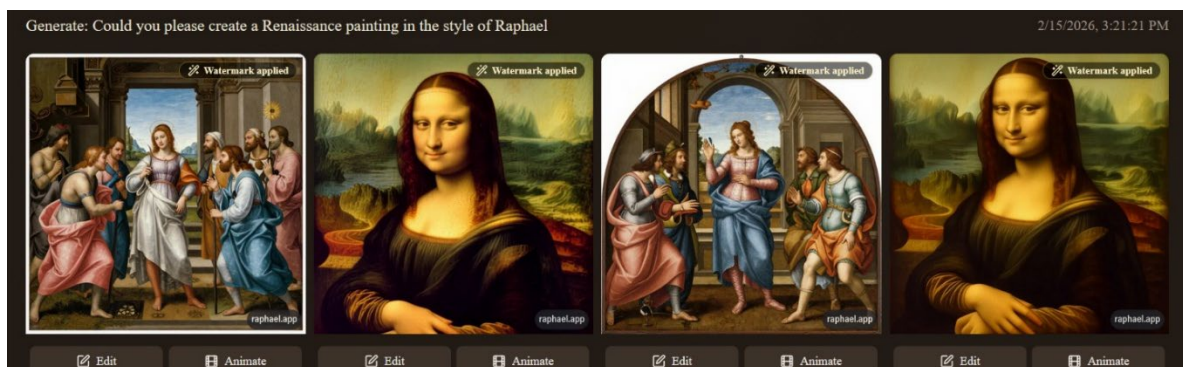


Figure 4: Image output in Michelangelo style from the Raphael AI tool at <https://raphael.app/> This image was generated with GenAI for this study. Data extraction: 16/02/2026

When prompted for Renaissance paintings in Raphael style the generation times were longer than in the Michelangelo case. Firstly, image 1 and image 2 from the left side was generated, secondly after quite a while they were more or less doubled by image 3 and image 4. The generated images are in Renaissance style, but the Mona Lisa imitation in image 2 and image 4 could hardly be classified as Raphael style.

Table 5: AI-generated images checkers results: Overview of effectiveness

AI detection tool	Michelangelo - leftmost in Figure 3	Michelangelo - rightmost in Figure 3	Raphael - leftmost in Figure 4	Raphael - rightmost in Figure 4
Illuminarty	0.3% probability of AI generated image	0.3% probability of AI generated image	0.1% probability of AI generated image	0.1% probability of AI generated image
QuillBot AI Image Detector	33% probability of AI generated image	100% probability of AI generated image	78% probability of AI generated image	47% probability of AI generated image
Hive AI Detector	99.9% probability of AI generated image	89.6% probability of AI generated image	99.9% probability of AI generated image	99.9% probability of AI generated image

4. Concluding Discussion and Further Research

The study was divided into two main sections: 1) generation of different type of content – texts, video, audio, images – by AI-tools and 2) the evaluation of these outputs by AI-based detection tools.

In the first stage, various content formats were generated using AI tools. For text content, three of the most popular tools were used: ChatGPT 5.2, Le Chat, and Gemini 1.5 Pro. It's worth noting that significant progress has been made in the development of GenAI over the past few years. Previously, AI-generated texts were of a low standard. Currently, these tools produce texts with decent content. While they may still generally lack specificity and use common words, these texts are now lexically and grammatically closer to those produced by native speakers (Rudolph et al. 2023). For video content, AI has advanced to the point where visible artifacts clearly indicating deepfake content are becoming increasingly rare. Typically, the initial videos produced included, for example, distortions of human anatomical proportions (six fingers and other distortions). Today, modern AI platforms can create short videos of fairly high quality. It's not always possible to determine, especially if the video is short, whether it's real or AI-generated. For example, videos generated using Sora are distinguished by high quality and realism.

In the second stage, we verified AI-generated content. For text verification apps, we used free ones such as Scribbr AI Detector, Sapling AI Detector, and Copyleaks AI Detector. Only the latter site, Copyleaks AI Detector, achieved a perfect result (in all three cases, it showed "100% of the text is likely AI"). Meanwhile, Scribbr AI Detector and Sapling AI Detector had minor difficulties identifying text generated by ChatGPT 5.2, yielding results of 87% and 73%, respectively. Overall, all sites performed well, indicating a high probability that the text was generated by AI. However, prior work has demonstrated that even widely used detectors are prone to both false positives and false negatives, particularly when differentiating between polished human writing and AI-generated text (Elkhatat 2023; Weber-Wulff et al. 2023). For videos, the AI detection tools we selected also showed high results. DeepFake-o-meter and Hive AI Detector performed flawlessly. At the same time, Sightengine showed results at 88% and 69%, respectively.

The poorest and most contradictory results were for image detection where the same image (Raphael - leftmost in Figure 4), got a 0.1% probability in the Illuminarty tool and 99.9 % in the Hive AI Detector tool. That the tools work by two very different models is obvious and probably that something is completely wrong with the Illuminarty concept. While the image generation shows creative parts, even if the result is rather kitschy, the tested detection tools are completely unreliable. At the higher professional level there are probably better tools, but regarding students, school teachers and low-budget journalists this is problematic. Regarding audio, the quality is good and that the generated audio files could be used in teaching and learning activities or to increase accessibility for visually impaired. The results from the AI generated audio detection tools are as confusing and contradicting as the ones for the image detection. This should be compared with the recent study by Nuska and Firdhous (2025), where a 98% accuracy was achieved when distinguishing human from AI-generated audio. The results from this study regarding images and audio evaluation leads to the same conclusion as in the preprint study on human GenAI detection by Cooke et al. (2024), as good as tossing a coin, and that the accuracy rates worsen when the media type is a single modality.

The following areas are envisaged for further research: Firstly, a future study should involve not only AI tools, but also humans as content creators. The outputs obtained by AI detection tools will then be tested. This comparison is particularly interesting in relation to text creation. The dilemma with a text creation written by a human without the use of GenAI is currently being widely discussed in academic circles since AI-text detection are not fully reliable and may yield false positive, i.e., label human-written content as AI-generated. For example, an empirical study that tested thousands of scientific abstracts from leading journals (1980-2023) reported that AI detection misclassified up to ~8% of known human-written abstracts as AI-generated (Rashidi et al. 2023).

Secondly, it would be valuable to involve licensed AI detection tools as well in the evaluation. Interesting to analyze the accuracy rate and effectiveness of licensed software and to compare with the free AI detection tools.

5. Conclusion

Several of the tested free and open services showed surprisingly good and creative results, but that some of them changed to services with fees during the time of conducting this study. The field of GenAI is a moving one where the quality of generation and detection constantly is changing, and also with changes when it comes to fees and costs. Some of the detection tools showed to have a surprisingly poor performance, and that it leaves us with the question about why they have been deployed online. With the rapidly increasing generation of deepfakes in different modalities, the need for more accurate and affordable detection tools is urgent.

From a knowledge management perspective, these developments are essential, as they directly affect how knowledge is created and validated. If synthetic content cannot be reliably identified, individuals, organizations and communities risk incorporating misleading or false information into their knowledge practices, thereby undermining decision-making, collaboration, and organizational learning. Future research should therefore develop a conceptual model that integrates AI detection with knowledge management processes in organizations. Such a model could explain how AI-based detection tools, human expert judgement, source verification, and organizational governance mechanisms interact across the knowledge management cycle, from knowledge creation and acquisition to knowledge validation, storage, sharing, and application. In particular, future studies could examine how detection results should be interpreted, documented, and communicated within organizational knowledge repositories and collaborative workflows. This would support not only the technical identification of synthetic content but also the development of transparent procedures for deciding whether AI-generated material should be accepted, revised, rejected, or flagged as uncertain within organizational knowledge systems.

Finally, and as highlighted in some recent studies, it would be valuable to define and develop a deepfake literacy strategy (Geissler et al, 2025; Chen et al., 2026; Iufereva, & Mozelius, 2026).

Ethics Declaration: No ethical clearance was required for the research reported in this paper, as it did not involve human participants, personal data, animals, or sensitive materials.

AI Declaration: AI tools were used only to generate the experimental stimuli described in the Methods section. The authors did not use AI tools to draft, revise, or interpret the analytical parts of the manuscript.

References

- Alavi, M., Leidner, D.E. and Mousavi, R. 2024. 'A knowledge management perspective of generative artificial intelligence', *Journal of the Association for Information Systems*, 25(1), pp. 1–12.
- Cazzamatta, R. 2025. 'Verification factors behind fact-checkers' selection decisions', *Journalism*, Advance online publication. <https://doi.org/10.1177/14648849251371952>
- Çetinkaya, C. 2025. The politics of misinformation: Social media, polarization, and the geopolitical landscape. *Trend Report 6/2025*. Austrian Institute for International Affairs (oiip).
- Chen, C., Hoe-Lian Goh, D. and Tang, Z. 2026. 'Teaching people to spot deepfakes: Investigating user acceptance of LLM-powered chatbots', *International Journal of Human-Computer Interaction*, pp. 1–25.
- Cooke, D., Edwards, A., Barkoff, S. and Kelly, K. 2024. 'As good as a coin toss: Human detection of AI-generated images, videos, audio, and audiovisual stimuli', *arXiv preprint arXiv:2403.16760*.
- Crawford, T. 2025. 'Contemporary information warfare, artificial intelligence, and challenges for democracies', *NATO Defense College Research Paper*, 11, pp. 1–32.
- Dell'Acqua, F., McFowland III, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K.C., Rajendran, S., Krayner, L., Candelon, F. and Lakhani, K.R. 2023. 'Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality', *Harvard Business School Working Paper No. 24-013*.
- Diep, P.P.U. 2024. 'Check the checks: A comparison of fact-checking practices between newspapers and independent organizations during 2020 U.S. election presidential debates', *Journalism & Mass Communication Quarterly*, Advance online publication. <https://doi.org/10.1177/07395329241298965>
- Elkhatat, A.M. 2023. '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*, 21(3), 100790. <https://doi.org/10.1016/j.ijme.2023.100790>
- Geissler, D., Robertson, C.E. and Feuerriegel, S. 2025. 'Digital literacy interventions can boost humans in discerning deepfakes'.
- Guess, A.M., Lockett, D., Lyons, B., Montgomery, J.M., Nyhan, B. and Reifler, J. 2020a. "'Fake news" may have limited effects beyond increasing beliefs in false claims', *Harvard Kennedy School Misinformation Review*, 1(1).

- Guess, A.M., Lockett, D., Lyons, B., Montgomery, J.M., Nyhan, B. and Reifler, J. 2020b. 'Misinformation in action: Fake news exposure is linked to lower trust in media, higher political polarization, and other harms', *Harvard Kennedy School Misinformation Review*, 1(4), pp. 1–18. <https://doi.org/10.37016/mr-2020-024>
- Guenilir, T. and Yildiz, I. 2025. 'Generative artificial intelligence and disinformation in the post-truth era', in *Generative artificial intelligence and digital disinformation in the post-truth era*. IGI Global, pp. 1–20. <https://doi.org/10.4018/978-1-6684-XXXX-X.ch001>
- Iufereva, A. and Mozelius, P. 2025. 'Deepfake detection: Human performance versus AI tools: A comparison of accuracy and effectiveness', in *ICAIR 2025*, Vol. 5. Academic Conferences and Publishing International Limited.
- Iufereva, A. and Mozelius, P. 2026. 'The societal and political implications of deepfakes in education: Affordances and constraints', in *Proceedings of the 20th Annual International Technology, Education and Development Conference*. Valencia, Spain.
- Kaczorowska-Spychalska, D., Kotula, N., Mazurek, G. and Sułkowski, Ł. 2024. 'Generative AI as source of change of knowledge management paradigm', *Human Technology*, 20(1), pp. 131–154. <https://doi.org/10.14254/1795-6889.2024.20-1.7>
- Kreps, S.E. and Kriner, D.L. 2025. 'How AI threatens democracy', *Journal of Democracy*, 36(1), pp. 5–19. <https://doi.org/10.1353/jod.2025.0001>
- Kudryavtsev, D., Khan, U.A. and Kauttonen, J. 2024. 'Transforming knowledge management using generative AI: From theory to practice', in *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Vol. 3: KMIS. SCITEPRESS, pp. 362–370.
- Nuska, M.F.F. and Firdhous, M.F.M. 2025. 'AI generated audio detection using acoustic audio features', in *2025 10th International Conference on Information Technology Research (ICITR)*. IEEE, pp. 1–6.
- Pimentel, M. and Veliz, J.C. 2024. 'The generative AI solutions for enhancing knowledge management: Literature review and roadmap', in *Proceedings of the 25th European Conference on Knowledge Management*. <https://doi.org/10.34190/eckm.25.1.2770>
- Rashidi, H.H., Fennell, B.D., Albahra, S., Hu, B. and Gorbett, T. 2023. 'The ChatGPT conundrum: Human-generated scientific manuscripts misidentified as AI creations by AI text detection tool', *Journal of Pathology Informatics*, 14, 100342. <https://doi.org/10.1016/j.jpi.2023.100342>
- Romanishyn, A. et al. 2025. 'AI-driven disinformation: Policy recommendations for democratic resilience', *Frontiers in Artificial Intelligence*, 8, 1569115. <https://doi.org/10.3389/frai.2025.1569115>
- Rudolph, J., Tan, S. and Tan, S. 2023. 'ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?', *Journal of Applied Learning & Teaching*, 6(1), pp. 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Varieties of Democracy (V-Dem) Institute 2025. Democracy endangered by online disinformation. Democracy Report.
- Vosoughi, S., Roy, D. and Aral, S. 2018. 'The spread of true and false news online', *Science*, 359(6380), pp. 1146–1151.
- Weber-Wulff, D. et al. 2023. 'Testing of detection tools for AI-generated text', *International Journal for Educational Integrity*, 19(1), 26. <https://doi.org/10.1007/s40979-023-00146-z>