

# Reframing Explainable Generative AI as a Tool for Thought

Katelyn Morrison  
kcmorris@andrew.cmu.edu  
Carnegie Mellon University  
Pittsburgh, Pennsylvania, USA

Eric Mason  
epmason@andrew.cmu.edu  
Carnegie Mellon University  
Pittsburgh, Pennsylvania, USA

Maggie Chen  
maggiec3@andrew.cmu.edu  
Carnegie Mellon University  
Pittsburgh, Pennsylvania, USA

Steven Lundi  
University of California, Los Angeles  
(UCLA) Health  
Los Angeles, California, USA

Afroz Zandifar  
University of Pittsburgh Medical  
Center (UPMC)  
Pittsburgh, Pennsylvania, USA

Daesung Kim  
daesungkim@yonsei.ac.kr  
Yonsei University  
Seoul, South Korea

Weicheng Dai  
wd2119@bu.edu  
Boston University  
Boston, Massachusetts, USA

Motahhare Eslami  
meslami@andrew.cmu.edu  
Carnegie Mellon University  
Pittsburgh, Pennsylvania, USA

Adam Perer  
adamperer@cmu.edu  
Carnegie Mellon University  
Pittsburgh, Pennsylvania, USA

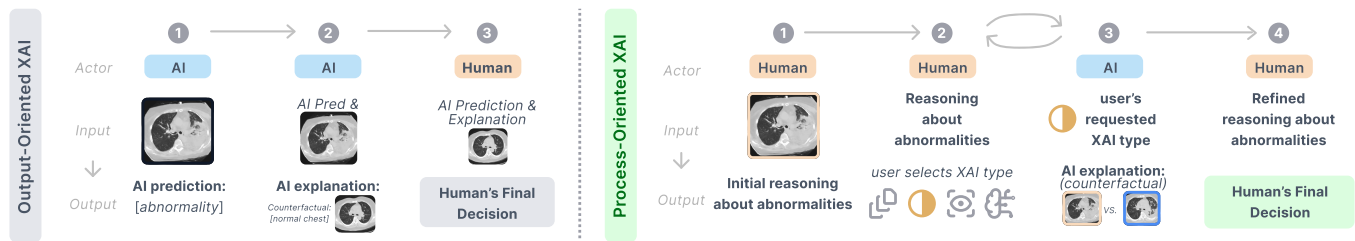


Figure 1: Output-oriented XAI is framed as a way to help humans decide whether to accept, modify, or reject an AI’s recommendation or generated artifact. Process-oriented XAI, in contrast, considers a workflow in which humans first reason without AI support, then generate AI explanations and use them to refine or shape their reasoning before finalizing their decision.

## Abstract

Generative AI (GenAI) has become ubiquitous across our personal and professional lives, able to assist us with nearly any task thrown its way. To help end users calibrate their reliance on and better understand the outputs of GenAI, researchers have developed explainable generative AI (GenXAI). However, this output-oriented framing of XAI positions the user to focus on validating model behavior, distracting them from engaging in their own reasoning process. To mitigate this, we propose reframing GenXAI as a design space that reframes XAI techniques to support reasoning throughout workflows, which we call process-oriented XAI. In this workshop paper, we instantiate process-oriented XAI in the first GenXAI Tool for Thought (TfT), which incorporates four XAI techniques to support users in interpreting chest CT scans. Through a think-aloud and semi-structured interview study with five medical students and five radiology residents, we show how our Tool for

Thought can shape clinical reasoning processes. We also discuss the nuanced challenges of using GenXAI in Tools for Thought.

## CCS Concepts

• Human-centered computing → Empirical studies in HCI; • Information systems → Decision support systems.

## Keywords

Explainable AI, Multimodal Generative AI, Empirical Study

### ACM Reference Format:

Katelyn Morrison, Eric Mason, Maggie Chen, Steven Lundi, Afroz Zandifar, Daesung Kim, Weicheng Dai, Motahhare Eslami, and Adam Perer. 2026. Reframing Explainable Generative AI as a Tool for Thought. In *Proceedings of CHI '26 Workshop on Tools for Thought*. ACM, New York, NY, USA, 5 pages. <https://doi.org/XXXXXXXX.XXXXXXX>

## 1 Introduction

Traditional and generative artificial intelligence (AI) systems have conventionally been designed and evaluated with an output-oriented lens [19]. This type of approach skips the reasoning process imperative to decision-making and forces decision makers to jump to a conclusion by either accepting, rejecting, or modifying the AI’s prediction [24]. For example, an output-oriented approach in a generative AI (GenAI) system may generate a natural language impression of a patient’s chest CT scan, leaving the radiologist to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

CHI '26 Workshop on Tools for Thought, Barcelona, Spain

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-XXXX-X/2018/06

<https://doi.org/XXXXXXXX.XXXXXXX>

decide whether to accept, reject, or modify the impression [17]. Recent work has identified several explainable GenAI (GenXAI) techniques [16]<sup>1</sup> that aim to reveal the reasoning behind an AI output (Figure 1), which could help radiologists better understand a generative model's behavior. However, explainable AI (XAI) solutions continue to position users as judges of the model's correctness, distracting them from engaging in their own reasoning about the data and the decision to be made. Furthermore, XAI techniques applied in traditional AI systems have failed to consistently benefit decision-makers to appropriately rely on AI recommendations [21].

Recent literature suggests that the design of AI systems should depart from a recommendation-centric (output-oriented) approach and focus on providing process-oriented support [24]. However, this reframing leaves the role of AI explanations unclear, as XAI was initially designed to support decision-makers while using output-oriented AI systems. With the advances of GenAI, recent work has shown that explanations generated by GenAI can affect humans' procedural knowledge [20] after AI-assisted learning tasks. Other works have discussed the potential benefits of GenAI explanations for critical thinking, reasoning, and learning in medical training [5, 12, 14]. Using cognitive psychology of explanations [8, 11] as a research lens, we propose process-oriented XAI (Figure 1, right): a design space that reframes XAI techniques as tools to support processes throughout humans' workflows, such as reasoning, because many XAI techniques already emulate human reasoning patterns (*i.e.*, counterfactual evidence, example-based reasoning, and localization). The key distinction between process-oriented XAI and both output-oriented XAI and interactive XAI systems is that process-oriented XAI does not provide an AI recommendation and instead positions XAI as Tools for Thought.

To explore the potential impact of process-oriented XAI on reasoning processes, we contribute the first GenXAI Tool for Thought, featuring four XAI techniques for a multimodal GenAI model [22] that generates chest CT scans. We chose the task of interpreting CT scans because (1) it remains a core focus across the machine learning (*e.g.*, [17]) and human-computer interaction communities (*e.g.*, [23]) and (2) it allows us to explore multiple modalities of explanations, ranging from visual to natural language. We then conducted a think-aloud and semi-structured interview study with five medical students and five radiology residents using our system to better understand: **(RQ1)** *How can XAI techniques support clinical reasoning and critical thinking?* **(RQ2)** *What are the challenges of using GenXAI to design Tools for Thought?*

## 2 GenXAI Tool for Thought


We designed our GenXAI Tool for Thought as a thought partner to support trainees' clinical reasoning during CT scan interpretation. It was mainly designed with clinical trainees (*i.e.*, senior medical students and junior radiology residents) in mind, as they are able to reason about a CT scan but may still lack enough knowledge or be prone to biases (*e.g.*, recency bias and anchoring bias) to create a comprehensive impression of a patient's conditions on their own.


### 2.1 Interface & Interaction Design


We leverage an existing open-source medical imaging interface developed by the Open Health for Imaging Foundation [25] to create a high-fidelity prototype. There are three components to the interface as seen in Figure 2: (1) the patient vignette (showing clinical history and chief complaint), (2) the viewing window (to review the original patient CT scan and any generated CT scans), and (3) the 'AI tools' panel (where the GenXAI techniques are located). We had three initial design goals for this system, inspired by Zhang and Reicherts [24]: encouraging users to externalize their reasoning, designing XAI support to be on-demand, and eliminating the AI's prediction to prioritize users' reasoning process. Our system emulates these design goals by intentionally not applying any of the 'AI tools' to the patient's CT scan, so the user cannot have the AI generate an impression of it, as outcome-oriented systems would. Instead, the user must externalize their reasoning about the patient's CT scan.

### 2.2 GenXAI Techniques

Inspired by GenXAI techniques reviewed by Schneider [16] and prototyped by Evirgen et al. [4] and Lin et al. [9], we implement four GenXAI techniques and present them as 'AI Tools' in the interface. As trainees reason about a patient's CT scan, they can use the AI tools to explore or confirm their interpretation, depending on how well-defined it is. These tools aim to support clinical reasoning, enable critical thinking, and foster learning, depending on how users employ the tools' outcomes (GenXAI artifacts). We describe each 'AI Tool', artifact, and the clinical reasoning process they aim to support below:

 **Similar Patients (Example-based Reasoning).** Yildirim et al. [23] discuss how radiologists may search for similar example cases to help them reason. Therefore, we design a tool that allows users to generate CT scans of similar abnormalities across different underlying anatomies. This explanation technique (Figure 2.4) is inspired by comparative explanations from Cai et al. [3] and the example-based reasoning GenXAI technique [16]. In the interface, three similar patient impressions are shown as suggestions to generate, including one prompt that reflects the patient's CT scan and two others that are variations or potential mimics of the patient's condition. Users can choose to generate one, which requires some level of externalization of their own reasoning about the patient's CT scan. By comparing their patient's CT with the generated CT, the user may engage in critical thinking, as they expose their own biases when interpreting the patient's scan.

 **Variations (Counterfactual Generations).** Reasoning by comparing to variations of abnormalities can also be valuable [13]. This explanation technique (Figure 2.5) is implemented using a prompt building interface and is based on the counterfactual generations technique [16]. Users can generate CT scans by building prompt variations with different base abnormalities, severity, locations, and associated findings. Creating a variation requires users to externalize their reasoning as they build the prompt.

 **Important Regions (Attention Visualizations).** Localizing the abnormality is another important clinical reasoning process. We implement this as a visual attention mechanism (Figure 2.6) that overlays a heatmap on the generated CT scan, highlighting regions most influenced by the prompt [2]. This tool only works on the

<sup>1</sup>GenXAI consists of XAI methods that explain GenAI behavior as well as the use of GenAI to explain other models [16].

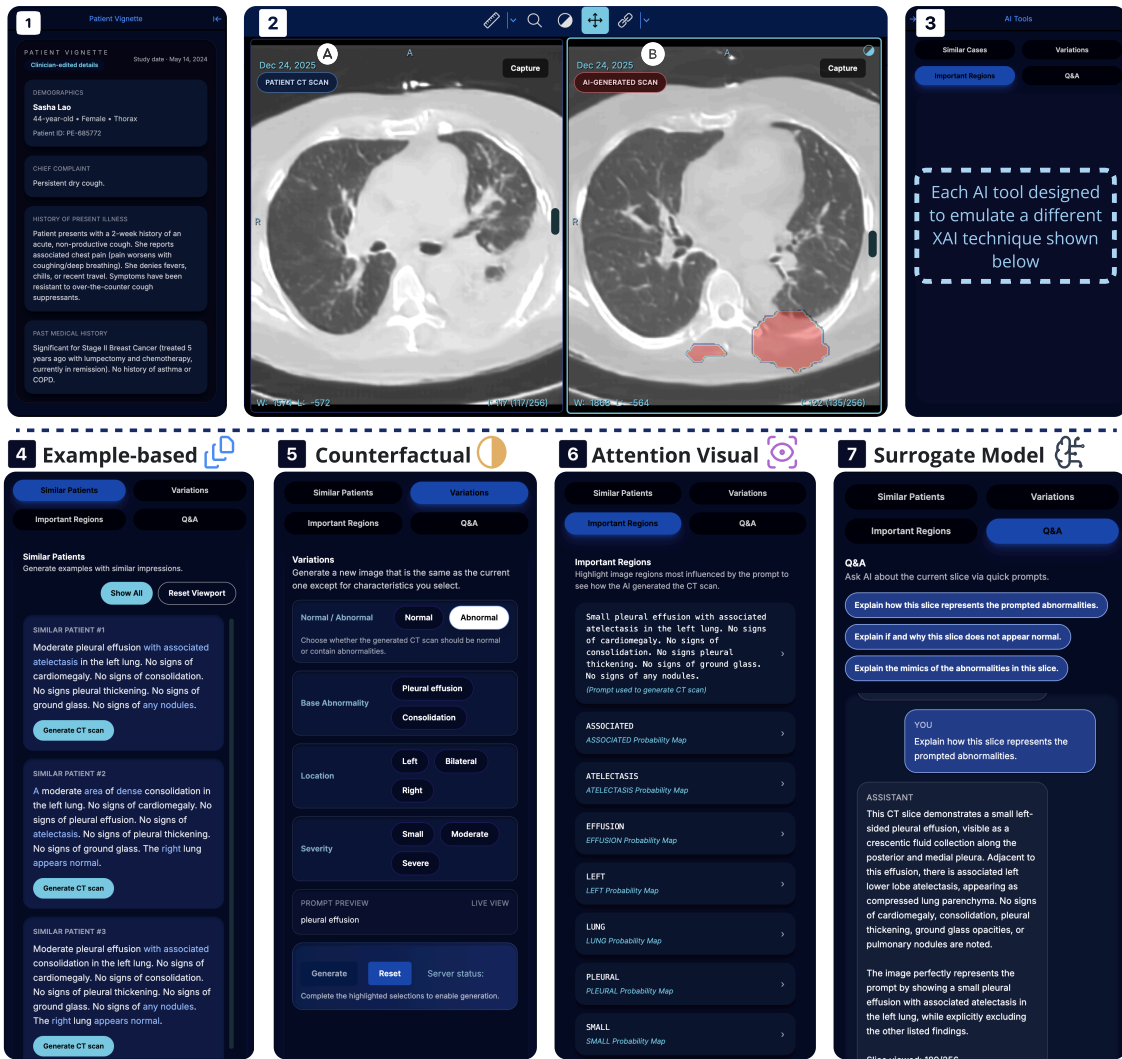


Figure 2: GenXAI Tool for Thought interface design. (1) shows the patient vignette. (2) reveals the viewing window: (a) shows the patient’s CT scan and (b) shows a generated CT scan based on a variation with important regions. (3) features the AI Tools.

generated CT scans, forcing users to reason about when and how to use it. For example, users can highlight important regions to assess the validity of a generation or to try to localize an abnormality.

**Natural Language Q&A (Surrogate Model).** Lastly, seeking information from external resources to help explain differences and similarities between abnormalities is part of the clinical reasoning process. Based on the GenXAI techniques described by Schneider [16], the Q&A tool (Figure 2.7) provides preset explanations that users can request for a specific slice of an AI-generated scan. This feature is powered by Gemini 2.5 Flash. We chose to use Gemini over GPT-5 because of its faster response time.

### 2.3 Semi-structured Interview Study

Sessions were conducted remotely via Teams/Zoom for an average of 61.25 minutes (*SD* = 4.3 minutes). Participants first completed a

10-question Likert-scale survey on their perceptions of AI in health-care [18] and then received a walkthrough of the interface and AI tools. For the main task, each participant was given approximately 40 minutes to reason about two patient vignettes and to build a four-slide case presentation for each. The vignettes were presented in a random order to counterbalance. The remaining interview time was used to reflect on the challenges on benefits of the XAI tools.

**Participant Recruitment.** Recruitment was through our clinical collaborators’ networks. Participants were compensated with an equivalent of 50 USD for the 60 minutes.

**Preliminary Data Analysis.** Data collection included screen recordings, audio transcripts, and the case presentation decks. We conducted a first-level preliminary analysis by process coding [15] the interview transcripts to capture the tool’s impact on clinical reasoning. We also conducted a preliminary round of structured coding [15] to identify the challenges of designing with GenXAI.

### 3 Preliminary Findings on Impact & Challenges

We conducted one pilot study and then recruited ten trainees: five 4th-year medical students (M1–M5) and five junior radiology residents (R1–R5) from two different institutions. All participants for the first vignette chose to review the vignette and perform a top-to-bottom search of the CT scan before using the AI tools, despite having access to them. Once they finished their initial reasoning, they used different AI tools as needed. Below is a preliminary analysis of how we observed the AI tools shaping clinical reasoning.

**Confirming own reasoning.** When participants were more confident about their impressions, the ‘AI Tools’ were “...*helpful more as a confirmation that I am thinking the right thing*” (R1). R4 chose to generate a variation based on their initial impression “...*to see if that would match with the actual scan*”. M4 similarly started by generating a variation of their exact impression.

**Identifying a starting point for reasoning.** Interpreting chest CT scans can be very difficult, especially for trainees. M1 actually used the AI tools as a starting point for the second vignette without reasoning about the CT scan, stating that “...*the other tools don't work unless it's AI-generated*”. Although not directly observed, R1 acknowledged how these GenXAI tools could be used “...*as a starting point to get its [AI] view on what does it think is going on and potentially use that as a sounding room essentially to kind of get me on a process to thinking about what this could even be...*”.

**Providing guidance throughout reasoning.** When participants had a partial impression formed, they used the AI Tools to seek guidance throughout reasoning. As M2 was going through the patient CT scan, they sought out guidance from the important regions tool saying, “*I feel like I'm not very good at finding it [atelectasis] and I wanted to just like get an impression of where it [AI] thinks the atelectasis is [using the important regions], just cause then if I see it on the AI-generated, then I think it might help me know what to expect on the patient one.*” M5 used the Q&A tool to “...*have it [AI] explain the mimics of abnormalities in this slice to get a sense of its thought process here*”.

**Discerning differences between two concepts.** Some participants leveraged the Variation Tool to try to discern the differences between two pathologies (R2, M3, R4, R5) or the severity of a pathology (R1, M2, M3, M4). M2 generated variations to help them finalize their impression: “...*I'll do small this time...to kind of confirm my bias that there is a larger, more severe atelectasis...I guess my thought process is more: 'How do I determine if it's small, moderate, or severe?'*”

**Illuminating biases in own reasoning.** Some participants uncovered (R1, R2) or attempted to uncover (M1, M4, R4) their own biases. For example, R1 experienced how the Q&A Tool helped them identify their own blind spots: “...*So I mentioned hemothorax, but I didn't even mention all of these other possibilities...I think that this is very, very useful to illuminate the blind spots that people may have...*”. R4 opted to use the Similar Patients Tool to “...*see if there's anything different [or] things I'm missing*”.

**Navigating imperfect generations.** A challenge in using generative methods is balancing between when imperfect generations are helpful versus harmful. Lacking domain knowledge can make it more difficult for users to identify when a generation is inaccurate. For example, M3 said “*If you're someone with experience, you would look at it and you would know. But someone like me wouldn't know*

*everything. Nothing really looks imperfect to me*”. On the other hand, imperfections can be beneficial in jumpstarting the reasoning process (“...*having a direction, even if it's not the right one from the very beginning can help...*” – R1) or slowing down the reasoning process (“...*it's helpful to have a challenging viewpoint...*” – M5).

**Navigating confirmation bias.** Even though participants found the tool helpful because it did not replace their reasoning (R1, M3, M4, M5, R4, R5), M3 felt that, given their lack of training in chest CT scans, the tool enable confirmation bias: “...*if I think it's something, I can get it to generate [that] to confirm what I'm thinking*”.

### 4 Discussion & Future Work

We propose process-oriented XAI to reframe GenXAI as a Tool for Thought. We contribute a GenXAI Tool for Thought to help clinical trainees reason through interpreting chest CT scans. We conducted a think-aloud and semi-structured interview study with ten clinical trainees to capture how XAI techniques can shape clinical reasoning. We discuss the implications of designing Tools for Thought with GenXAI and suggest future directions to explore.

**Uncertainty in the reasoning process drives the perceived value of the tool.** Users may fail to experience any benefits in generating CT scans if they lack uncertainty in their reasoning about the patient's CT scan, because they may not see a reason to explore or confirm anything. Although we did not interview attendings, we suspect these tools may still benefit them by illuminating biases in their own reasoning. Future work should explore how this tool would affect the clinical reasoning of more knowledgeable or overconfident users.

**Imperfect explanations as a source of friction in the reasoning process.** Imperfections have conventionally been regarded as flaws in outcome-oriented AI systems, leaving expert decision-makers and companies to abandon the algorithm [7]. However, in a process-oriented manner in which these imperfections can be easily ignored, verified, or integrated into one's reasoning process [24], these imperfections could be seen less as a flaw and more as a design goal [6, 10]. Imperfections can prompt overconfident thinkers to reconsider whether they are thinking about the case correctly. However, this could lead less knowledgeable users into unproductive or misleading lines of thought. Future work should more closely examine the impact of imperfections on the reasoning and critical thinking skills.

**Quantitatively measuring impact through comparison of initial and final impressions.** Although we did not formally record participants' initial impressions of CT scans to compare with their final impressions, we believe that this is a promising way to measure the value added. For example, studies could consider deskilling as a metric. Future work could also explore similarity metrics and leverage LLMs-as-a-Judge (e.g., [1]) to score the added value by comparing the initial and final impression.

### 5 Conclusion

Our preliminary analyses suggest that explainable GenAI (GenXAI) techniques may be a promising direction to explore when designing Tools for Thought for clinical reasoning processes. We hope this encourages further discussion around which XAI techniques to use and how to evaluate their impact in Tools for Thought.

## Acknowledgments

Figma Make was used to create a skeleton design for Figure 1. Microsoft Copilot and Grammarly were used throughout the text to improve clarity, grammar, and spelling.

## References

- [1] Zahra Ashktorab, Werner Geyer, Michael Desmond, Elizabeth M Daly, Martín Santillán Cooper, Qian Pan, Erik Miehling, Tejaswini Pedapati, and Hyo Jin Do. 2025. Evalassist: A human-centered tool for llm-as-a-judge. *2nd HEAL Workshop at CHI (2025)*. [https://heal-workshop.github.io/chi2025\\_papers/11\\_EvalAssist\\_A\\_Human\\_Centered.pdf](https://heal-workshop.github.io/chi2025_papers/11_EvalAssist_A_Human_Centered.pdf)
- [2] Aidan Bradshaw, Katelyn Morrison, Arpit Mathur, Weicheng Dai, Motahhare Eslami, Adam Perer, et al. 2025. Toward Interpretable 3D Diffusion in Radiology: Token-Wise Attribution for Text-to-CT Synthesis. In *Medical Imaging with Deep Learning-Short Papers*. <https://openreview.net/pdf?id=DYFRzRPQn>
- [3] Carrie J Cai, Jonas Jongejan, and Jess Holbrook. 2019. The effects of example-based explanations in a machine learning interface. In *Proceedings of the 24th international conference on intelligent user interfaces*. 258–262. <https://dl.acm.org/doi/pdf/10.1145/3301275.3302289>
- [4] Noyan Evirgen, Ruolin Wang, and Xiang'Anthony Chen. 2024. From text to pixels: Enhancing user understanding through text-to-image model explanations. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*. 74–87. <https://dl.acm.org/doi/pdf/10.1145/3640543.3645173>
- [5] Isabella BB Ferreira, Rodrigo C Menezes, Luis Cláudio L Correia, and Bruno B Andrade. 2025. Medicine Beyond Machines: Viewpoint on the Art of Thinking in the Age of AI. *JMIR Formative Research* 9, 1 (2025), e76669. <https://formative.jmir.org/2025/1/e76669/>
- [6] Jose A Guridi, Angel Hsing-Chi Hwang, Duarte Santo, Maria Goula, Cristobal Cheyre, Lee Humphreys, and Marco Rangel. 2025. From fake perfects to conversational imperfects: Exploring image-generative AI as a boundary object for participatory design of public spaces. *Proceedings of the ACM on Human-Computer Interaction* 9, 2 (2025), 1–33. <https://dl.acm.org/doi/10.1145/3710912>
- [7] Nari Johnson, Sanika Moharana, Christina Harrington, Nazanin Andalibi, Hoda Heidari, and Motahhare Eslami. 2024. The fall of an algorithm: Characterizing the dynamics toward abandonment. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*. 337–358. <https://dl.acm.org/doi/pdf/10.1145/3630106.3658910>
- [8] Frank C Keil. 2006. Explanation and understanding. *Annu. Rev. Psychol.* 57, 1 (2006), 227–254. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3034737/#S4>
- [9] Bo Lin, Yingjing Xu, Xuanwen Bao, Zhou Zhao, Zhouyang Wang, and Jianwei Yin. 2025. SkinGEN: An explainable dermatology diagnosis-to-generation framework with interactive vision-language models. In *Proceedings of the 30th International Conference on Intelligent User Interfaces*. 1287–1296. <https://dl.acm.org/doi/10.1145/3708359.3712098>
- [10] Fang Liu, Junyan Lv, Shenglan Cui, Zhilong Luan, Kui Wu, and Tongqing Zhou. 2024. Smart" Error"! Exploring Imperfect AI to Support Creative Ideation. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–28. <https://dl.acm.org/doi/10.1145/3637398>
- [11] Tania Lombrozo. 2006. The structure and function of explanations. *Trends in cognitive sciences* 10, 10 (2006), 464–470. <https://www.sciencedirect.com/science/article/pii/S1364661306002117>
- [12] Chenglong Ma, Yuanfeng Ji, Jin Ye, Lu Zhang, Ying Chen, Tianbin Li, Mingjie Li, Junjun He, and Hongming Shan. 2025. Towards interpretable counterfactual generation via multimodal autoregression. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 611–620. [https://link.springer.com/chapter/10.1007/978-3-032-04937-7\\_58](https://link.springer.com/chapter/10.1007/978-3-032-04937-7_58)
- [13] Katelyn Morrison, Arpit Mathur, Aidan Bradshaw, Tom Wartmann, Steven Lundi, Afroz Zandifar, Weichang Dai, Kayhan Batmanghelich, Motahhare Eslami, and Adam Perer. 2025. A Human-centered approach to identifying promises, risks, & challenges of text-to-image generative AI in radiology. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, Vol. 8. 1758–1770.
- [14] Conrad W Safranek, Anne Elizabeth Sidamon-Eristoff, Aidan Gilson, and David Chartash. 2023. The role of large language models in medical education: applications and implications. e50945 pages. <https://mededu.jmir.org/2023/1/e50945>
- [15] Johnny Saldaña. 2021. The coding manual for qualitative researchers. (2021).
- [16] Johannes Schneider. 2024. Explainable Generative AI (GenXAI): a survey, conceptualization, and research agenda. *Artificial Intelligence Review* 57, 11 (2024), 289. <https://link.springer.com/article/10.1007/s10462-024-10916-x>
- [17] Adrian Serapio, Guntavt Chaudhari, Cody Savage, Yoo Jin Lee, Maya Vella, Shravan Sridhar, Jamie Lee Schroeder, Jonathan Liu, Adam Yala, and Jae Ho Sohn. 2024. An open-source fine-tuned large language model for radiological impression generation: a multi-reader performance study. *BMC Medical Imaging* 24, 1 (2024), 254. <https://link.springer.com/article/10.1186/s12880-024-01435-w>
- [18] Lucy Shinnars, Sandra Grace, Stuart Smith, Alexandre Stephens, and Christina Aggar. 2022. Exploring healthcare professionals' perceptions of artificial intelligence: Piloting the Shinnars Artificial Intelligence Perception tool. *Digital Health* 8 (2022), 20552076221078110. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8832586/>
- [19] Venkatesh Sivaraman, Katelyn Morrison, Will Epperson, and Adam Perer. 2025. Over-Relying on Reliance: Towards Realistic Evaluations of AI-Based Clinical Decision Support. *CHI '25 Workshop on Envisioning the Future of Interactive Health (2025)*. <https://arxiv.org/pdf/2504.07423>
- [20] Philipp Spitzer, Joshua Holstein, Katelyn Morrison, Kenneth Holstein, Gerhard Satzger, and Niklas Kühl. 2025. Don't be fooled: the misinformation effect of explanations in human-ai collaboration. *International Journal of Human-Computer Interaction* (2025), 1–29. <https://www.tandfonline.com/doi/pdf/10.1080/10447318.2025.2574511>
- [21] Michelle Vaccaro, Abdullah Almaatouq, and Thomas Malone. 2024. When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour* 8, 12 (2024), 2293–2303. <https://www.nature.com/articles/s41562-024-02024-1.pdf>
- [22] Yanwu Xu, Li Sun, Wei Peng, Shuyue Jia, Katelyn Morrison, Adam Perer, Afroz Zandifar, Shyam Visweswaran, Motahhare Eslami, and Kayhan Batmanghelich. 2024. MedSyn: text-guided anatomy-aware synthesis of high-fidelity 3-D CT images. *IEEE Transactions on Medical Imaging* 43, 10 (2024), 3648–3660. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10566053>
- [23] Nur Yildirim, Hannah Richardson, Maria Teodora Wetscherek, Junaid Bajwa, Joseph Jacob, Mark Ames Pinnock, Stephen Harris, Daniel Coelho De Castro, Shruthi Bannur, Stephanie Hyland, et al. 2024. Multimodal healthcare AI: identifying and designing clinically relevant vision-language applications for radiology. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–22. <https://arxiv.org/pdf/2402.14252>
- [24] Zelun Tony Zhang and Leon Reicherts. 2025. Augmenting Human Cognition With Generative AI: Lessons From AI-Assisted Decision-Making. *CHI '25 T4T: Workshop on Tools for Thought (2025)*. <https://arxiv.org/pdf/2504.03207>
- [25] Erik Ziegler, Trinity Urban, Danny Brown, James Petts, Steve D Pieper, Rob Lewis, Chris Hafey, and Gordon J Harris. 2020. Open health imaging foundation viewer: an extensible open-source framework for building web-based imaging applications to support cancer research. *JCO clinical cancer informatics* 4 (2020), 336–345. <https://ascopubs.org/doi/pdfdirect/10.1200/CCI.19.00131>