Narrative Visual Summarization for Enhancing Patient-Provider Communication Understanding

PI: Dr. Samah Jarad-Fodeh (Associate Professor, Yale University)
Co-Investigator: Dr. Yi-Chun Chen (Postdoctoral Associate, Yale University)
Department: Department of Emergency Medicine
Contact Email: samah.fodeh@yale.edu, yi-chun.chen@yale.edu

1 Project Description

Asynchronous messaging between patients and healthcare providers is increasingly common, yet patients often struggle to interpret, recall, or act on the information shared. Fragmented threads, unclear instructions, and emotionally neutral tones are especially problematic for people with limited health knowledge. These communication challenges are not only cognitive—they carry clinical consequences such as poor adherence and reduced patient trust. This project proposes a human-AI co-creative system for narrative visual summarization. We will extract meaningful communication elements such as reassurance, instructions, and decision points from patient-provider messages and generate animated visual summaries using metaphor-driven illustrations and lightweight animation. The system aims to improve comprehension, recall, and emotional engagement by combining natural language processing, visual storytelling, and prompt-driven generation. A formative user study will evaluate the system output and guide iterative design.

1.1 Motivation and Background

Patients often struggle to understand, recall, or act on shared information during medical interactions, challenges that are magnified in asynchronous digital settings such as patient portals and secure messaging systems. Research shows that miscommunication and message complexity are key contributors to poor adherence and patient dissatisfaction, especially among those with limited health knowledge [26, 12, 18, 15]. Asynchronous communication further complicates matters with fragmented threads, emotionally flat language, and delayed clarifications [13, 14, 6, 25].

Efforts to simplify communication often rely on structured templates or discharge summaries, but these solutions can still fall short in clarity and emotional tone [9, 10]. Studies suggest that visual augmentation—such as narrative framing and metaphor—can significantly improve comprehension, retention, and engagement [5, 32, 28, 29]. Related work using sentiment analysis and topic modeling has highlighted how affective and experiential dimensions of healthcare communication are often overlooked in digital records [1].

Visual summarization techniques, including animated infographics and data-driven storytelling, offer a promising strategy for enhancing understanding. Tools like *Canis*, *Cast*, and *Gemini* enable structured data to be rendered into coherent visual sequences through grammar-based or recommendation systems [8, 7, 16]. Platforms like *Narrative Player* and *Data Animator* further support expressive storytelling by linking semantic content with visual pacing and transitions [21, 30]. Design-focused research on animated explainer tools emphasizes clarity, coherence, and visual hierarchy [22].

The cognitive benefits of narrative visuals are well-established. Visual metaphors, when used in educational or health communication, improve interpretation of abstract concepts and support memory [33, 24,

29]. Public-facing science communication examples, such as *Kurzgesagt*, illustrate how minimalistic and metaphor-rich animations can make complex topics emotionally and conceptually accessible [17].

Advances in generative AI further expand the potential of visual summarization. Techniques in vision-language modeling, prompt engineering, and LLM-assisted design offer mechanisms for converting structured content into stylized visuals [11, 31, 34]. Systems that support natural language interfaces or hierarchical generation frameworks provide scaffolding for aligning narrative structure with visual flow [20, 19, 23]. Previous work has also demonstrated how narrative theory and modular storytelling pipelines can be integrated with AI models to produce personalized and stylized visual narratives [4, 2, 3]. Social media–based analyses have similarly illustrated how NLP can surface public health concerns and emotional tone from unstructured patient-facing messages [27].

This proposal builds on these foundations, extending narrative visualization methods to asynchronous health communication through a human-centered metaphor-driven AI pipeline.

2 Goals and Objectives

The overarching goal of this project is to develop and evaluate a human-AI collaborative system that generates narrative visual summaries from asynchronous patient-provider messages. The system is intended to support better comprehension, recall, and emotional engagement, particularly for patients with limited health literacy.

- **Objective 1:** Design a pipeline for extracting key communicative elements, such as medical instructions, emotional reassurance, and decision points, from multi-turn message exchanges using natural language processing techniques.
- **Objective 2:** Develop a visual summarization framework that maps the extracted content to symbolic illustrations driven by metaphors and lightweight animation templates for temporal and affective clarity.
- **Objective 3:** Conduct a formative user study to evaluate the impact of visual summaries on patient comprehension, satisfaction, and usability, comparing text-only and visual formats.

3 Methodology

This project involves three phases: (1) extracting and structuring content from patient-provider messages, (2) generating visual summaries using metaphor-driven representations, and (3) evaluating the system through a user study.

3.1 Message Analysis and Narrative Structuring

We will apply natural language processing (NLP) to de-identified patient-provider message logs to extract:

- Key communicative intents (e.g., instructions, reassurance, decision points)
- Emotional tone, urgency, and sequential order
- Narrative groupings across multi-turn threads

These elements will form structured representations for downstream visual generation.

3.2 Visual Summarization and Metaphor Mapping

Extracted content will be mapped to a curated library of visual metaphors and animation templates:

- Symbolic visuals will represent common medical intents (e.g., roadmaps for plans, icons for symptoms).
- Lightweight animations will provide temporal cues (e.g., fade-ins for sequence, zooms for emphasis).
- Prompt-based generation via LLMs will assist in selecting appropriate visuals and styles.

The pipeline will be prototyped using Python for NLP, and Unity or WebGL for visual rendering.

3.3 Evaluation and Iteration

A small-scale user study will compare visual and text-only summaries of simulated or real message threads. Evaluation metrics include:

- **Comprehension:** Accuracy in recalling instructions or key takeaways.
- **Engagement and satisfaction:** Ratings on emotional clarity, helpfulness, and trust.
- **Usability:** Participant feedback on layout, pacing, and visual interpretation.

Study results will guide refinements to content extraction and visual design. All artifacts will be iteratively improved based on user insights.

4 Contribution to the Yale AI Initiative

This project aligns closely with the Yale AI Initiative's mission to advance responsible, interdisciplinary, and human-centered AI research. By applying AI techniques to improve healthcare communication, our work exemplifies how technical innovation can address real-world challenges in an ethically grounded and socially impactful manner.

Specifically, this project contributes to the Yale AI Initiative in the following ways:

- **Human-Centered AI:** Our system is designed around patient needs, focusing on accessibility, empathy, and health literacy. The integration of narrative modeling, metaphor design, and user feedback reflects a commitment to AI systems that are not only functional but deeply usable and meaningful.
- **Interdisciplinary Collaboration:** The project brings together expertise in computer science, emergency medicine, digital health communication, and visual storytelling. It embodies the kind of cross-domain synergy that the Yale AI Initiative seeks to foster.
- **Responsible and Transparent AI:** By grounding the visual summarization system in explainable symbolic elements and conducting human evaluation throughout the development process, the project addresses transparency, safety, and accountability in AI-assisted decision support.
- **Platform for Future Research:** The resulting system, annotated datasets, and visual design components will serve as a foundation for future research in narrative AI, multimodal communication tools, and healthcare education. We anticipate open-sourcing components to support the wider Yale AI community.

In summary, this proposal supports the Yale AI Initiative's priorities by advancing technical innovation alongside human-centered impact, demonstrating how generative AI can enhance communication equity and healthcare outcomes.

References

[1] Sharon Chekijian, Huan Li, and Samah Fodeh. "Emergency care and the patient experience: using sentiment analysis and topic modeling to understand the impact of the COVID-19 pandemic". In: *Health and Technology* 11.5 (2021), pp. 1073–1082.

[2] Yi-Chun Chen and Arnav Jhala. "A customizable generator for comic-style visual narrative". In: *arXiv preprint arXiv:2401.02863* (2023).

[3] Yi-Chun Chen and Arnav Jhala. "Collaborative Comic Generation: Integrating Visual Narrative Theories with AI Models for Enhanced Creativity". In: *arXiv preprint arXiv:2409.17263* (2024).

[4] Yi-Chun Chen and Arnav Jhala. "Panel Transitions for Genre Analysis in Visual Narratives". In: *arXiv preprint arXiv:2312.08720* (2023).

[5] Amita Chugh et al. "Better transitions: improving comprehension of discharge instructions". In: *Frontiers of health services management* 25.3 (2009), pp. 11–32.

[6] Artin Entezarjou et al. "Experiences of digital communication with automated patient interviews and asynchronous chat in Swedish primary care: a qualitative study". In: *BMJ open* 10.7 (2020), e036585.

[7] Tong Ge, Bongshin Lee, and Yunhai Wang. "Cast: Authoring data-driven chart animations". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–15.

[8] Tong Ge et al. "Canis: A High-Level Language for Data-Driven Chart Animations". In: *Computer Graphics Forum*. Vol. 39. 3. Wiley Online Library. 2020, pp. 607–617.

[9] Alexander F Glick et al. "Accuracy of parent perception of comprehension of discharge instructions: role of plan complexity and health literacy". In: *Academic pediatrics* 20.4 (2020), pp. 516–523.

[10] Richard T Griffey et al. "The impact of teach-back on comprehension of discharge instructions and satisfaction among emergency patients with limited health literacy: A randomized, controlled study". In: *Journal of communication in healthcare* 8.1 (2015), pp. 10–21.

[11] Jindong Gu et al. "A systematic survey of prompt engineering on vision-language foundation models". In: *arXiv preprint arXiv:2307.12980* (2023).

[12] Jennifer Fong Ha and Nancy Longnecker. "Doctor-patient communication: a review". In: *Ochsner journal* 10.1 (2010), pp. 38–43.

[13] Leonie Heyworth et al. "Aligning medication reconciliation and secure messaging: qualitative study of primary care providers' perspectives". In: *Journal of medical Internet research* 15.12 (2013), e264.

[14] Timothy P Hogan et al. "Patient centeredness in electronic communication: evaluation of patient-to-health care team secure messaging". In: *Journal of medical Internet research* 20.3 (2018), e82.

[15] Roy PC Kessels. "Patients' memory for medical information". In: *Journal of the royal society of medicine* 96.5 (2003), pp. 219–222.

[16] Younghoon Kim and Jeffrey Heer. "Gemini: A grammar and recommender system for animated transitions in statistical graphics". In: *IEEE Transactions on Visualization and Computer Graphics* 27.2 (2020), pp. 485–494.

[17] Kurzgesagt – In a Nutshell. *What Makes Kurzgesagt So Special?* `https://www.youtube.com/watch?v=OvKr6loCciE`. Accessed: 2025-05-23. 2025.

[18] Lucille ML Ong et al. "Doctor-patient communication: a review of the literature". In: *Social science & medicine* 40.7 (1995), pp. 903–918.

[19] Tanzila Rahman et al. “Make-a-story: Visual memory conditioned consistent story generation”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2023, pp. 2493–2502.

[20] Aditya Ramesh et al. “Hierarchical text-conditional image generation with clip latents”. In: *arXiv preprint arXiv:2204.06125* 1.2 (2022), p. 3.

[21] Zekai Shao et al. “Narrative Player: Reviving Data Narratives with Visuals”. In: *IEEE Transactions on Visualization and Computer Graphics* (2025).

[22] Leixian Shen et al. “Reflecting on Design Paradigms of Animated Data Video Tools”. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 2025, pp. 1–21.

[23] Leixian Shen et al. “Towards natural language interfaces for data visualization: A survey”. In: *IEEE transactions on visualization and computer graphics* 29.6 (2022), pp. 3121–3144.

[24] Yang Shi et al. “Supporting expressive and faithful pictorial visualization design with visual style transfer”. In: *IEEE Transactions on Visualization and Computer Graphics* 29.1 (2022), pp. 236–246.

[25] Christine A Sinsky and John W Beasley. *Texting while doctoring: a patient safety hazard*. 2013.

[26] Richard L Street Jr et al. “How does communication heal? Pathways linking clinician–patient communication to health outcomes”. In: *Patient education and counseling* 74.3 (2009), pp. 295–301.

[27] Yihua Su et al. “Twitter-based analysis reveals differential COVID-19 concerns across areas with socioeconomic disparities”. In: *Computers in Biology and Medicine* 132 (2021), p. 104336.

[28] John Sundquist. “The long and the short of it: the use of short films in the German classroom”. In: *Die Unterrichtspraxis/Teaching German* 43.2 (2010), pp. 123–132.

[29] Murray Taylor et al. “Digital storytelling and visual metaphor in lectures: a study of student engagement”. In: *Accounting Education* 27.6 (2018), pp. 552–569.

[30] John R Thompson, Zhicheng Liu, and John Stasko. “Data animator: Authoring expressive animated data graphics”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–18.

[31] Tiffany Tseng, Ruijia Cheng, and Jeffrey Nichols. “Keyframer: Empowering animation design using large language models”. In: *arXiv preprint arXiv:2402.06071* (2024).

[32] Barbara Tversky, Julie Bauer Morrison, and Mireille Betrancourt. “Animation: can it facilitate?” In: *International journal of human-computer studies* 57.4 (2002), pp. 247–262.

[33] Tz-Li Wang and Ya-Hui Bella Lien. “The power of using video data”. In: *Quality & Quantity* 47 (2013), pp. 2933–2941.

[34] Jiaqi Wu, John Joon Young Chung, and Eytan Adar. “viz2viz: Prompt-driven stylized visualization generation using a diffusion model”. In: *arXiv preprint arXiv:2304.01919* (2023).

Proposal abstract – 150-250 words direct entry

Asynchronous messaging between patients and healthcare providers is increasingly common, yet patients often struggle to interpret, recall, or act on the information shared. Fragmented threads, unclear instructions, and emotionally neutral tones are especially problematic for people with limited health knowledge. These communication challenges are not only cognitive—they carry clinical consequences such as poor adherence and reduced patient trust. This project proposes a human-AI co-creative system for narrative visual summarization. We will extract meaningful communication elements such as reassurance, instructions, and decision points from patient-provider messages and generate animated visual summaries using metaphor-driven illustrations and lightweight animation. The system aims to improve comprehension, recall, and emotional engagement by combining natural language processing, visual storytelling, and prompt-driven generation. A formative user study will evaluate the system output and guide iterative design.

Team narrative – 250-word direct entry

This project brings together complementary expertise in computer science, biomedical informatics, and healthcare communication. The project is led by Dr. Samah Jarad-Fodeh, Associate Professor in the Department of Emergency Medicine at Yale University. Dr. Jarad-Fodeh serves as the Principal Investigator, drawing on her extensive experience in digital health systems, clinical informatics, and patient-centered communication technologies. She will oversee project alignment with clinical needs, guide evaluation design, and facilitate access to patient communication data sources.

Dr. Yi-Chun Chen, Postdoctoral Associate at Yale University, serves as the technical lead and co-investigator. Dr. Chen's research focuses on human-AI collaboration, multimodal storytelling, and generative visual media. They will lead the development of the narrative summarization pipeline, including the design of the NLP system, metaphor-based visual mapping, and animation prototyping. Dr. Chen will also coordinate user study design and system iteration based on participant feedback.

Together, the team spans the full lifecycle of the project—from theoretical framing and AI development to clinical relevance and real-world applicability. The collaboration integrates methodological rigor with translational potential, positioning the project to make a meaningful contribution to healthcare communication, human-centered AI, and multimodal narrative generation.

Budget and Justification

This proposal requests a total of $99,850 to support the development and evaluation of a human-AI collaborative system for narrative visual summarization in healthcare communication. The budget spans an 12-month project period and complies with Yale AI Seed Grant guidelines.

1. Personnel – $68,000

- **Postdoctoral Associate Salary – $68,000**: Full salary support for Dr. Yi-Chun Chen for 12 months at 100% effort. Dr. Chen will serve as the technical lead, responsible for system design, NLP model development, visual narrative generation, and study coordination.

2. Cloud Computing and Software – $12,000

- **Cloud Services – $8,000**: Credits for AWS or Azure to support model fine-tuning, data storage, and deployment of generative components such as diffusion models and video synthesis pipelines.
- **Software Licenses – $4,000**: Licenses for Unity Pro, Adobe Creative Cloud, and other tools for animation, interaction design, and interface prototyping.

3. Dissemination and Outreach Materials – $7,850

- **Journal Publication Fees – $2,500**: Open-access publication costs for venues such as ACM TiiS, ACM TOMM, or IEEE TVCG.
- **Outreach Materials – $2,000**: Project website, video demos, or code documentation to support reproducibility and community engagement.
- **Workshop/Participant Supplies – $1,000**: Design kits, feedback forms, and materials for local user testing sessions.
- **Miscellaneous Research Supplies – $350**: Small equipment or technical accessories not covered by standard office supply budgets.
- **Poster Printing – $1,000**: For internal project showcases or public events such as the Yale Postdoc Symposium.

4. Travel and Dissemination – $7,000

- **AI Research Conference Travel – $3,500**: Travel and registration to present results at venues such as ACM Multimedia (ACMMM), ACM CHI, or AAAI workshops on generative or multimodal AI. These platforms align with the project's focus on narrative generation, visual AI, and human-centered design.
- **Journal Publication Fees – $2,500**: Open-access publication fees as above, supporting visibility in AI and HCI communities.
- **Outreach and Demonstration Materials – $1,000**: Visual demos and website content for broader academic and practitioner engagement.

Total Requested Budget: $99,850

Timeline and Milestones

Duration: 12 months (starting July 1, 2025)

- **Months 1–2: Project setup and planning** Finalize technical goals and system architecture; prepare de-identified message data; review literature on narrative visuals and health communication; define annotation schema.
- **Months 3–4: NLP pipeline development** Implement content extraction pipeline (e.g., message type, emotional tone); cluster and segment multi-turn dialogues; generate testable structured outputs for downstream generation.
- **Months 5–6: Visual metaphor and animation design** Build a modular library of visual elements (icons, animations, scenes); begin mapping message types to metaphor-driven templates using LLM-assisted prompts.
- **Months 7–9: System integration and user study** Integrate narrative structuring with visual rendering; conduct user study comparing visual vs. text-based summaries; gather comprehension and usability feedback.
- **Months 10–12: Refinement and dissemination** Refine outputs based on user feedback; prepare open-source demo and public materials; submit findings to AI-focused venues such as ACM Multimedia, CHI, or AAAI workshops.