Conceptual Patent Outline: Generative Cadence Canvas

Disclaimer: This document is a conceptual outline for informational purposes only and does not constitute legal advice. It is not a substitute for consultation with a qualified patent attorney. Patent law is complex and requires specialized legal expertise for proper application and prosecution.

1. Title of the Invention

System and Method for Multi-Modal, Iterative Human-AI Collaborative Art Creation via a Generative Cadence Canvas

2. Inventors

3. Assignee

Conexus Global Arts Media [or your legal entity name]

4. Abstract

A system and method for facilitating collaborative art creation through a multi-modal, iterative human-AI interaction process, herein termed the "Generative Cadence Canvas." The invention enables users to sequentially contribute to evolving artworks across various modalities (e.g., visual, textual, audio, video) within defined time limits. The system leverages generative AI models to interpret human inputs, offer diverse creative transformations, and manage the sequential flow with controlled visibility, thereby fostering unique and unpredictable artistic outcomes. All resulting collaborative artworks are dedicated to the public domain, emphasizing the process of co-creation over individual ownership.

5. Background of the Invention

5.1. Field of the Invention

The present invention relates generally to artificial intelligence (AI) systems, human-computer interaction, digital art creation, collaborative software platforms, and generative media. More specifically, it pertains to novel methods for orchestrating human and AI collaboration in a sequential, multi-modal creative process.

5.2. Description of Related Art

The digital creative landscape has seen significant advancements in both collaborative tools and generative AI.

- **Traditional Collaborative Art Platforms:** Existing platforms facilitate human-to-human collaboration, often through shared digital canvases (e.g., Drawpile , Miro Lite) or sequential "telephone game" / "exquisite corpse" formats (e.g., Skwiz , MoMA Exquisite Corpse , Tel-AI-phone). While engaging, these often lack integrated, intelligent AI responsiveness across multiple modalities.
- **Generative AI Tools:** Modern generative AI models can produce high-quality content from text prompts across images (e.g., DALL-E 3, Midjourney, Stable Diffusion, Adobe Firefly), music (e.g., Suno AI, Udio, Mubert, Google Lyria), and video (e.g., Sora, Runway, Synthesia).
- **Human-AI Co-creation:** Research indicates that effective human-AI co-creation involves a "dance" where AI augments human creativity, helping overcome blocks and explore ideas, with humans providing "judgment" and "taste". However, a significant limitation is AI's tendency to produce "homogenous" or "generic" outputs when left to its own devices, lacking true originality or nuanced context.
- **Limitations in Current Integration:** A discernible gap exists in platforms that can truly facilitate a fluid, real-time, multimodal interaction within existing digital co-creation environments. Current AI tools often function as "black boxes" lacking granular human control or explicit collaborative features. Furthermore, the legal and ethical frameworks for AI-generated content, particularly concerning authorship and ownership, remain ambiguous, with purely AI-generated works generally not being copyrightable.

There is a need for a system that systematically integrates multimodal AI, structured sequential evolution with controlled visibility, and dynamic human-AI iteration to consistently produce unique, collaborative artistic outputs while proactively addressing ethical and ownership complexities.

6. Summary of the Invention

The present invention provides a novel system and method for multi-modal, iterative human-AI collaborative art creation, referred to as the "Generative Cadence Canvas." The invention addresses the limitations of existing creative tools by providing an intuitive platform where users engage in a structured, sequential "telephone game" of art, with generative AI acting as a transformative co-creator at each step.

The core of the invention lies in its unique process:

1. **Sequential, Time-Boxed Contributions:** Users contribute in defined turns, each with a strict time limit (e.g., 30 seconds), fostering rapid, intuitive creativity.
2. **Controlled Visibility:** Projects are initially obscured (e.g., blurred) to maintain an element of surprise, revealing full content only upon a user's commitment to contribute.
3. **Multimodal AI Orchestration:** At each turn, a human provides input (text, sketch, audio, or video clip), and generative AI (e.g., Google Gemini via Vertex AI Live API) processes this input to generate multiple, diverse multimodal outputs.
4. **Human-in-the-Loop Curation & Refinement:** The human user selects their preferred AI-generated output and can further refine it (e.g., drawing overlays, text edits, masking) before passing it to the next participant. This ensures human judgment and "taste" guide the creative evolution, counteracting AI's tendency towards homogeneity.
5. **Public Domain Output:** All resulting collaborative artworks are explicitly dedicated to the public domain, simplifying intellectual property ambiguities and emphasizing the shared joy of creation over individual ownership.

6. **Showroom & Community Features:** Completed "Canvas Chains" are showcased in a gallery, with clear attribution to all human and AI contributors, fostering a positive and engaged community.

The "Generative Cadence Canvas" framework ensures a dynamic "dance" between human and AI, leading to genuinely unique, unpredictable, and collaborative artistic expressions at scale.

7. Brief Description of the Drawings (Conceptual)

- **FIG. 1:** Illustrates a high-level architectural overview of the Generative Cadence Canvas system, showing user interaction, cloud infrastructure, and AI API integrations.
- **FIG. 2:** Depicts the user interface for initiating a new "Canvas Chain" project, including options for initial prompt input (text, sketch, audio).
- **FIG. 3:** Shows the user interface for a "Continue Someone Else's" project, illustrating the blurred preview of ongoing projects.
- **FIG. 4:** Details the in-turn user interface, including the 30-second timer, the AI-generated options, and the human refinement tools.
- **FIG. 5:** Illustrates the sequential flow of a "Canvas Chain" project, showing how multimodal inputs and AI transformations build upon each other through multiple turns.
- **FIG. 6:** Presents the final "Showroom" display of a completed collaborative artwork, highlighting attribution to all contributors.

8. Detailed Description of the Invention

The present invention describes a system and method for the "Generative Cadence Canvas," a novel platform for human-AI collaborative art creation. The system is implemented as a mobile application ("Conexus Canvas") and supported by a robust cloud-based backend infrastructure that orchestrates various generative AI APIs.

8.1. System Architecture (Referring to FIG. 1)

The system comprises:
- **Client-Side Application (Conexus Canvas):** A mobile application (iOS/Android) providing the user interface for interaction. Developed using modern frontend frameworks (e.g., Svelte, Next.js) for optimal performance and user experience.
- **Backend Server:** A scalable cloud-based backend (e.g., Google Cloud Platform) responsible for:
  - User authentication and profile management.
  - Managing project state, turns, and queues.
  - Orchestrating calls to various generative AI APIs.
  - Storing project history and final outputs.
  - Implementing content moderation filters.
- **Database:** A scalable database (e.g., Supabase, PlanetScale, Weaviate for vector embeddings) for storing user data, project metadata, and references to AI-generated content.
- **Generative AI API Integrations:** Direct API integrations with leading multimodal generative AI services. This is crucial for cost-effectiveness and leveraging state-of-the-art models without custom AI model development from scratch.

- ○ **Google Gemini API (via Vertex AI Live API):** Primary for real-time, low-latency, bidirectional text, audio, and video input/output, and session memory for fluid co-creation.
  - ○ **Image Generation:** OpenAI DALL-E 3, Stability AI (SDXL), Adobe Firefly, Google Gemini (Imagen).
  - ○ **Music/Audio Generation:** Google Lyria (via Vertex AI), Suno AI, Udio, Riffusion.
  - ○ **Video Generation:** Runway Gen-2, Sora (part of ChatGPT Plus), Fliki.
- ● **MLOps Infrastructure:** Tools for experiment tracking (MLflow, Weights & Biases), data and model versioning (DVC), and continuous integration/deployment (CI/CD) for AI models, ensuring reproducibility and continuous improvement.

8.2. Method of Operation (Referring to FIGS. 2-6)

The "Generative Cadence Canvas" operates as follows:
1. **User Onboarding & Choice (FIG. 2, FIG. 3):**
   - ○ A user signs up/logs in.
   - ○ They are presented with two primary options: "Start New Creation" or "Continue Someone Else's."
   - ○ If "Continue Someone Else's" is selected, a list of ongoing projects is displayed. These projects are initially blurred or abstractly represented to maintain an element of surprise, revealing full content only upon selection.
2. **Project Initiation (FIG. 2):**
   - ○ If "Start New Creation" is chosen, the user is presented with a blank canvas.
   - ○ The user provides the initial creative input, which can be:
     - ■ A text prompt (e.g., "A serene forest with glowing mushrooms").
     - ■ A simple sketch or image upload.
     - ■ A short audio recording (e.g., a sound effect or melody).
   - ○ The initiator also defines game parameters, such as the number of turns (e.g., 4-6) and privacy (public or private).
3. **AI Generation & Human Refinement Loop (FIG. 4):**
   - ○ **Turn Allocation:** Only one user can contribute to a project at a time. If multiple users attempt to join, they are placed in a queue with an estimated wait time.
   - ○ **Time Limit:** Once a user's turn begins, a strict 30-second timer (or a creator-defined 2-minute limit) starts, encouraging rapid, intuitive creativity.
   - ○ **AI Interpretation & Generation:** The AI processes the previous contribution's output (or the initial prompt) and the current user's input (if any). It then generates multiple, diverse multimodal options (e.g., 3-4 images based on text, 3-4 text descriptions based on an image, 3-4 audio clips based on text).
   - ○ **Human Selection & Refinement:** The user reviews the AI-generated options and selects their preferred output. They can then use basic refinement tools (e.g., drawing overlays, text edits, masking) or provide further text prompts for targeted AI regeneration, ensuring human control and artistic direction. Once selected/refined, the user "commits" their contribution.
4. **Sequential Chain Continuation (FIG. 5):**
   - ○ The system ensures that the modality of input/output changes or evolves with each turn to encourage diverse creative interpretations (e.g., text to image, image to audio, audio to text).
   - ○ Steps 3 and 4 repeat with subsequent users until the predefined number of turns is

completed.
5. **Final Reveal & Showroom (FIG. 6):**
   ○ Once all turns are complete, the entire "Canvas Chain" is revealed to all participants, showcasing the progression of the artwork from the initial prompt through each human and AI transformation.
   ○ Completed projects are displayed in a public showroom/gallery.
   ○ Clear attribution is provided for all human and AI contributors within the final display.
   ○ Users can download the final composite artwork and share it directly to social media platforms.

8.3. Core Principles Embodied in the Invention

The "Generative Cadence Canvas" embodies several novel principles:
- **Human-Centric Control:** The system ensures human artists remain the ultimate directors, providing "judgment" and "taste" at critical junctures, with AI as a powerful co-creator and assistant.
- **Dynamic Iteration:** The creative process is an ongoing "dance" between human and AI, emphasizing continuous feedback loops and successive refinement within each "cadence."
- **Multimodal Fusion:** The architecture seamlessly integrates and transitions between different creative forms (text, images, audio, video) within a single project, leveraging AI to facilitate complex transformations.
- **Controlled Serendipity:** The invention intentionally introduces elements of chance or AI-driven "misinterpretation" in a structured manner (e.g., multiple AI options, sequential hidden contributions) to break homogeneity and drive unexpected, unique outcomes.
- **Public Domain by Design:** All resulting artworks are explicitly dedicated to the public domain, simplifying intellectual property ambiguities and fostering a truly open, collaborative community.

9. Claims (Example Claims - Subject to Legal Drafting)

The following are example claims that a patent attorney would draft to define the scope of the invention. These are illustrative and would need precise legal wording.

Claim 1: A method for facilitating collaborative art creation, comprising: a. receiving, from a first user, an initial creative input for a project, said initial creative input comprising at least one of a text prompt, a sketch, an image, or an audio recording; b. processing said initial creative input using a multi-modal generative artificial intelligence (AI) model to generate a plurality of initial multi-modal outputs; c. presenting said plurality of initial multi-modal outputs to said first user; d. receiving, from said first user, a selection of a preferred initial multi-modal output from said plurality, and optionally, a refinement input; e. storing said selected and optionally refined initial multi-modal output as a first contribution to said project; f. presenting a representation of said first contribution to a subsequent user, wherein said representation is partially obscured to maintain an element of surprise; g. receiving, from said subsequent user, a second creative input in a different modality than said first contribution, said second creative input provided within a predefined time limit; h. processing said first contribution and said second creative input using said multi-modal generative AI model to generate a plurality of second multi-modal outputs; i. receiving, from said subsequent user, a selection of a preferred second multi-modal

output from said plurality, and optionally, a refinement input; j. storing said selected and optionally refined second multi-modal output as a second contribution to said project; and k. iteratively repeating steps (f) through (j) with additional users until a predefined number of contributions are received, thereby forming a collaborative multi-modal artwork.

Claim 2: The method of Claim 1, further comprising: a. displaying said collaborative multi-modal artwork in a public gallery upon completion of said predefined number of contributions; and b. explicitly attributing each human and AI contribution within said displayed artwork.

Claim 3: The method of Claim 1, wherein said predefined time limit for each contribution is between 30 seconds and 2 minutes.

Claim 4: The method of Claim 1, wherein said multi-modal generative AI model comprises an integration of at least two distinct AI-as-a-Service (AIaaS) APIs, each specializing in a different modality.

Claim 5: The method of Claim 4, wherein said AIaaS APIs include at least one of Google Gemini via Vertex AI Live API, OpenAI DALL-E 3, Stability AI (SDXL), Adobe Firefly, Google Lyria, Suno AI, Udio, Runway Gen-2, or Sora.

Claim 6: The method of Claim 1, further comprising: a. providing a queue system for users waiting to contribute to an ongoing project; and b. displaying an estimated wait time for users in said queue.

Claim 7: The method of Claim 1, wherein said initial creative input or subsequent creative input is provided via a text prompt, and said multi-modal generative AI model generates a visual output based on said text prompt.

Claim 8: The method of Claim 1, wherein said initial creative input or subsequent creative input is provided via a visual input (e.g., sketch or image), and said multi-modal generative AI model generates a textual description or an audio output based on said visual input.

Claim 9: The method of Claim 1, wherein said initial creative input or subsequent creative input is provided via an audio input, and said multi-modal generative AI model generates a visual or textual output based on said audio input.

Claim 10: A system for facilitating collaborative art creation, comprising: a. a client-side application configured to: i. receive an initial creative input from a first user; ii. present a plurality of initial multi-modal outputs generated by an AI model; iii. receive a selection and optional refinement from said first user; iv. present a partially obscured representation of a previous contribution to a subsequent user; and v. receive a second creative input from said subsequent user within a predefined time limit; b. a backend server configured to: i. manage project state, turns, and user queues; ii. orchestrate calls to a plurality of multi-modal generative AI APIs; iii. store contributions for each project; and iv. manage the sequential flow of contributions; and c. a plurality of multi-modal generative AI APIs, integrated with said backend server, configured to: i. process creative inputs across various modalities; and ii. generate multi-modal outputs based on said inputs and previous contributions.

Works cited

1. Drawpile, https://drawpile.net/ 2. Music AI Sandbox, now with new features and broader access - Google DeepMind, https://deepmind.google/discover/blog/music-ai-sandbox-now-with-new-features-and-broader-ac cess/ 3. Online Whiteboard for Realtime Collaboration | Miro Lite, https://miro.com/online-whiteboard/ 4. Skwiz on the App Store, https://apps.apple.com/us/app/skwiz/id1611258643 5. What Is Multimodal AI and How It Works - IMD Business School, https://www.imd.org/blog/digital-transformation/multimodal-ai/ 6. Make a

Surrealist exquisite corpse, https://ec.moma.org/intro 7. Generative Art - Media Arts and Technology, https://www.mat.ucsb.edu/~g.legrady/academic/courses/20f594/l/generativeArt.pdf 8. Tel-AI-phone, https://telaiphone.com/ 9. AI Generated Telephone Game | Experimental Game Design - Northeastern University, https://experimentalgamedesign.sites.northeastern.edu/2023/10/24/ai-generated-telephone-game/ 10. Best AI Music Generator Software in 2025 - AudioCipher, https://www.audiocipher.com/post/ai-music-app 11. Expanding generative media for enterprise on Vertex AI | Google Cloud Blog, https://cloud.google.com/blog/products/ai-machine-learning/expanding-generative-media-for-enterprise-on-vertex-ai 12. The 11 best AI video generators in 2025 | Zapier, https://zapier.com/blog/best-ai-video-generator/ 13. Exploring Human-AI Collaboration in Creative Industries - SmythOS, https://smythos.com/ai-trends/human-ai-collaboration-in-creative-industries/ 14. Human-AI Collaboration Can Unlock New Frontiers in Creativity, https://hcii.cmu.edu/news/human-ai-creativity-tools 15. Understanding the Limitations of AI in Content Creation - Spines, https://spines.com/understanding-the-limitations-of-ai-in-content-creation/ 16. Best Approach to Use Two AI Agents for Generating PRD and Coding Plan in n8n, https://community.n8n.io/t/best-approach-to-use-two-ai-agents-for-generating-prd-and-coding-plan-in-n8n/86965 17. How Non-Technical Founders Can Build AI Startups Without a CTO I ..., https://b-works.io/en/insights/empowering-non-technical-founders-building-ai-startups-without-a-cto/ 18. Convert Text to Video with AI - Fliki, https://fliki.ai/features/text-to-video 19. Productivity Challenge Timer on the App Store - Apple, https://apps.apple.com/us/app/productivity-challenge-timer/id1117766356 20. How Multimodal Learning is Used in Generative AI - DigitalOcean, https://www.digitalocean.com/community/tutorials/multimodal-learning-generative-ai