Workshop: Generative AI and the Future of Creation

Introduction

Workshop Objectives

In this workshop, we will explore a wide range of topics from the basics of Generative AI to its applications, examining the possibilities of AI in creation. Through practical examples of text, image, music, video, and code generation, participants will experience how AI can be utilized in creative work. We will explore the transformations AI brings to art and design, while also critically considering its limitations and challenges.

Stance on Generative AI

The emergence of Generative AI is similar to the advent of internet search engines.

- **Powerful Convenience**: A shift from "searching" to "consulting." It can generate new things from vast amounts of data.
- **Problems**: Issues like hallucinations (falsehoods), copyright, information leakage, and ethical challenges also exist.
- **This Workshop's Position**:
  - We do not blindly trust or uncritically accept AI.
  - However, rather than rejecting it, it is important to enjoy its convenience while **thinking critically** about its problems and structures.
  - First, we will actively engage with and try out the technology to gain an accurate understanding of it.

Fundamentals of Generative AI - Transformers and LLMs

The Relationship Between AI / ML / DL / GenAI

- **Artificial Intelligence (AI)**: Computer systems that mimic human intelligence. (The broadest concept)

- **Machine Learning (ML)**: A field of AI where systems learn patterns from data.

- **Deep Learning (DL)**: A method of ML that uses neural networks modeled after the human brain.

- **Generative AI (GenAI)**: A subset of DL that generates new content (text, images, music, etc.).

- (Reference Video: Introduction to Generative AI)

The 30-Year History Leading to ChatGPT

- **Early Days (1980s)**: Recurrent Neural Networks (RNNs) introduced the concept of "memory," and research on sequential learning (e.g., word prediction) began.

- **Turning Point (2017)**: The **Transformer** architecture was developed.

- **Evolution (2018-2020)**:
  - **GPT-1**: Applied the Transformer to "next word prediction."
  - **GPT-2**: Trained on large-scale data collected from the web.
  - **GPT-3**: 175 billion parameters. Became capable of zero-shot learning (performing tasks from instructions alone).

- (Reference Video: ChatGPT: 30 Year History | How AI Learned to Talk)

How LLMs and Transformers Work

- **The Essence of LLMs**: They are mathematical functions that **probabilistically predict** the next word.

- **Transformer**:
  - Previous models (RNNs) processed one word at a time.
  - Transformers can process the entire text in **parallel**.

- **Attention Mechanism (Attention)**:
  - This is the core technology of the Transformer.
  - It calculates how strongly each word in a sentence is related to other words.
  - This allows it to understand context-appropriate meaning (e.g., determining if "bank" means a financial institution or a riverbank).

- **References**:
  - Video: How LLMs Work (Simple Version)
  - Illustrated Guide: The Illustrated Transformer
  - Interactive: TRANSFORMER EXPLAINER

History of Generative Art and AI Art Examples

What is Generative Art?

- **Generative Art**: Art in which the artist does not directly create the work itself, but instead defines the "**rules**," "**processes**," or "**systems**" that generate the work.

- Algorithms or systems autonomously generate the artwork.

- Current "**Generative AI**" can be seen as the latest form of this generative art.

- (Reference Video: What is Generative Art: TouchDesigner's Role)

History of Generative Art (Excerpts)

(Reference: Generative Art Timeline)

- **1950s (Analog Era)**: Ben Laposky created "Electronic Abstractions" using an oscilloscope.

- **1960s (Digital Era)**: Artists like Frieder Nake created algorithmic drawings using early computers and plotters.

- **1970s (Artist-Programmer Era)**: Harold Cohen developed "AARON," a pioneer in AI art.

- **1990s (Net Era)**: John Maeda published "Design By Numbers."
- **2000s (Tooling Era)**: **Processing** and **openFrameworks** emerged, popularizing creative coding.
- **2010s (AI Era)**: Google's **DeepDream** (2015) and GANs appeared.
- **2020s (On-Chain / Diffusion Model Era)**: DALL-E (2021) and Stable Diffusion (2021) were introduced.

Examples of AI-Powered Artworks

- **Refik Anadol, *Unsupervised* (2023)**
  - A work that trained an AI on the entire collection of MoMA (Museum of Modern Art) and continuously generates new "machine hallucinations" in real-time.

- **Nao Tokui - AI DJ Project (2016)**
  - A project where an AI and a human DJ perform in a B2B (back-to-back) format, taking turns selecting music.

- **The Beatles, *Now And Then* (2023)**
  - Using AI technology, John Lennon's vocals and piano were separated from an old demo tape. Based on the clarified vocals, the remaining members added new performances, completing it as the final new song by The Beatles.

A Curator's Perspective: What Counts as Art?

- According to The New York Times article "Amid the A.I. Deluge, What Counts as Art?," many curators believe the value of AI art lies not in its technical novelty, but in the artist's "**intention**" and "**process**" behind it.

- A MoMA curator states, "The technology isn't the art; ultimately, it's the **individual creativity and vision** that matters."

- Works where the artist is deeply involved in building the AI model or curating the data are particularly highly valued.

Creative Workshop Using Generative AI

4-1. Image Generation (Practice)

AI that generates images from text. It can handle various styles and instructions.

- **Main Services**:
  - **Google ImageFX**: Excels at high-quality, realistic images. Available for free.
  - **DALL-E 3 (in ChatGPT)**: High fidelity to instructions, excels at rendering text.
  - **Midjourney**: Extremely high artistic and fantastical quality. Used via Discord (Paid).
  - **Stable Diffusion**: Open-source. Can be run in a local environment and offers the highest customizability.

4-1. Image Generation (Practice)

- **Practice**:
  - **Tools**: Google ImageFX or DALL-E 3 (in ChatGPT)
  - **Prompt/Task**:
    1. Try a simple prompt.
        - Example: `An oil painting of a cat in a spacesuit`
    2. Change the "expressive chips" (bolded words) in ImageFX to see the variations.
        - Example: `An **impressionist** painting of an **angry** cat in a **cyberpunk** spacesuit`
    3. (With DALL-E 3) Try modifying it conversationally.
        - Example: `Make the cat bigger and change the background to the moon`

4-2. Video Generation (Practice)

AI that generates videos from text or images. This field is evolving rapidly.

- **Technological Evolution**:
  - **GAN (～2019)**: Realistic but unstable footage (e.g., Deep Dream-style).
  - **Diffusion Models + Transformers (Present)**: Generate high-quality video from noise. Sora and Veo are based on this technology.

- **Main Services**:
  - **OpenAI Sora**: Understands physics, generates high-res video (1+ min) and audio (Limited Access).
  - **Google Veo (in Flow)**: Excels at understanding cinematic language (Limited Access).
  - **Luma AI Dream Machine**: Fast, easy to use, and high quality (Free tier available).
  - **Pika**: Good for creative effects and anime-style conversions (Free tier available).

4-2. Video Generation (Practice)

- **Practice**:
  - **Tools**: Luma AI Dream Machine, OpenAI Sora, Google Veo etc.
  - **Prompt/Task**:

    1. Generate a video from text.
       - Example: `A cinematic shot of a raccoon detective in a trench coat, walking through a rainy neon-lit city street at night.`

    2. (If you have an image) Upload an image and try to animate it.
       - Example: Try animating the "cat in a spacesuit" image you generated earlier.

4-3. Music Generation (Practice)

AI that generates full songs, including vocals and accompaniment, from text (lyrics or mood).

- **Main Services**:
  - **Suno AI**: Also supports Japanese lyrics. Generates high-quality songs (including vocals) by specifying genre and style.
  - **Google MusicFX DJ**: Allows real-time music generation and remixing by combining instruments, genres, and moods.

4-3. Music Generation (Practice)

- **Practice**:
  - **Tools**: Suno AI
  - **Prompt/Task**:

    1. Turn on "Custom Mode".

    2. Input **Lyrics**.

    3. Specify the **Style of Music**.
       - Example: `City Pop, 80s Japanese, Lo-fi`

    4. Listen to and compare the two generated song variations.

4-4. Programming (Practice)

AI assists with code generation, completion, and debugging (error fixing).

- **Subject**: **p5.js**
  - A programming language for visual arts. It can be run directly in a web browser.
- **Main AI Tools**:
  - **ChatGPT**: The easiest way. You can ask it, "Write p5.js code to do…".
  - **p5.CodingWithAI**: A dedicated p5.js editor. You can refine your code while interacting with AI.
  - **Github Copilot**: Integrated into editors like VSCode. A powerful code completion and generation tool for professionals.

4-4. Programming (Practice)

- **Practice**:
  - **Tools**: ChatGPT + **p5.js Web Editor** or p5.CodingWithAI
  - **Prompt/Task**:
    1. Have ChatGPT generate the initial code.
       - Prompt: `Using p5.js, create a program that draws a circle following the mouse position. Make the background black.`
    2. Copy and paste the generated code into the p5.js Editor and run it.
    3. Give ChatGPT additional instructions to refine it.
       - Prompt: `Change the circle's color so it changes randomly when the mouse is clicked.`
    4. Run the modified code again and check the result.

Conclusion

The Future of Generative AI and Creation

- **Is AI a Tool or a Collaborator?**:
    - As seen in today's practice, AI is a powerful "tool" for giving shape to ideas.
    - At the same time, it can also be a "collaborator" that understands human intent and makes unexpected suggestions.

- **The Role of Humans**:
  - The **vision** of "what to create" and the ability to ask questions.
  - The "editing" ability to **evaluate and select** AI's output and provide instructions for refinement.
  - The ability to design new **contexts** and **processes** that the AI has not been trained on.

- **Challenges**: We must continue to critically examine the problems behind the technology, such as data bias, copyright, and energy consumption.