Human-AI Co-Creation of Art Based on the Personalization of Collective Memory

Conference Paper · December 2022

DOI: 10.1109/ACAIT56212.2022.10137839

7 authors, including:

Zhuohao Wu
Communication University of China
8 PUBLICATIONS **188** CITATIONS
SEE PROFILE

Danwen Ji
Tongji University
10 PUBLICATIONS **188** CITATIONS
SEE PROFILE

Dingming Wu
Shenzhen University
55 PUBLICATIONS **2,323** CITATIONS
SEE PROFILE

Mohammad Shidujaman
Independent University
84 PUBLICATIONS **619** CITATIONS
SEE PROFILE

Human-AI Co-Creation of Art Based on the Personalization of Collective Memory

Zhuohao Wu
School of Animation and Digital Arts
Communication University of China
Beijing, China
how@cuc.edu.cn

Yanni Li
School of Animation and Digital Arts
Communication University of China
Beijing, China
leeyanni@cuc.edu.cn

Danwen Ji
College of Design and Innovation
Tongji University
Shanghai, China
danwen_ji@tongji.edu.cn

Dingming Wu
College of Computer Science and Software Engineering
Shenzhen University
Shenzhen, China
dingming@szu.edu.cn

Mohammad Shidujaman
Department of Electrical and Electronic Engineering
American International University-Bangladesh
Dhaka, Bangladesh
mjaman@aiub.edu

Yuan Zhang
School of Animation and Digital Arts
Communication University of China
Beijing, China
zhangyuan_397@163.com

Chenfan Zhang
School of Design
Politecnico di Milano
Milan, Italy
chenfan.zhang@mail.polimi.it

Abstract. *Artificial intelligence (AI) is trained with data, especially texts, numbers, images, videos and music on Internet. These data all together across time and space make a collective memory of the world. The latest large-scale AI models give people a chance to create out of a large pool of this collective memory, which they won't be able to access before, and communicate with AI in both human natural language and the unique machine supported ways. As demonstrated and discussed in this paper, effective and efficient workflows can be built up for human and AI to co-create meaningful results based on both the collective memory of the world and the personalized ideas and tastes. This kind of human-AI co-creation has a new force with great potential, and expect a new strategy and philosophy to guide human-AI collaboration.*

Keywords—creativity, artificial intelligence, AI, art, design, methods, process management

I. INTRODUCTION

Scientists and artists have been longing to create things with AI for decades.

Music lovers started to explore music AI back in 1950s [1], almost simultaneously with the introduction of the concept of AI. Lejaren Hiller and Leonard Isaacson conducted groundbreaking experimental music work with the ILLIAC supercomputer, Illinois Automatic Computer, at the University of Illinois College of Music In 1958 [1]. Marvin Minsky published a paper "Music, Mind, and Meaning", studying how music infects human minds in the way of AI thinking in 1981 [2]. Deep learning brought music AIs to a whole new level in the late 2010s. There are not only simple ones such as MuseNet, but also powerful platforms for professionals such as AIVA [3].

Visual art lovers also explored painting AI in the form of both physical robots and computer programs. One of the most well-known experiments is Harold Cohen and his painting robot AARON [4]. Cohen and AARON worked together for over 40 years since AARON was created in 1973 until Cohen died in 2016. Deep learning significantly changed the landscape of AI, including image generation AI, from rule based to big-data trained. An interesting example is, Google developed a fun app, Quick, Draw!, getting people around the world to draw on it; and then used these drawing data to train an AI [5]; and finally developed another app, AutoDraw, empowered by this AI, recognizing people's doodles and associating them with the preset graphics. This drawing dataset is a typical collective memory, while everything on Internet all together makes a mega scale collective memory, which AI can learn from. Google DeepDream (2015), Nvidia GauGAN (2019) and NightCafe (2019) were the milestones demonstrating what can be generated based on this collective memory of Internet. In 2020-2021 the large-scale AI models such as GPT-3 surprised people with its impressive applications in many different areas. And then in 2021-2022 CLIP + GAN/Diffusion models brought Nvidia Canvas (June 23, 2021), WOMBO Dream (November 5, 2021), Disco Diffusion (February 26, 2022), DALL·E 2 (April 6, 2022), Midjourney (April 8, 2022), Imagen (May 24, 2022).

From music to visual art, the dimensions and complexity of meaning making increase significantly [6]. In this paper, we will mainly discuss the creation of visual art as an example. From DeepDream to Disco Difussion and from GAN to CLIP, people can collaborate with AI to generate more and higher quality content than ever.

II. HUMAN-AI CO-CREATION

A. A Door Across Time and Space

Generative art, design, music and text have been around for decades, however, *Edmond de Belamy, from La Famille de Belamy*, the first AI-generated portrait ever sold at auction, still made a dent in the world as no one before. This AI-generated painting fetched the same high price as Picasso's work in an auction at Christie's on October 25, 2018. And it was surrounded by various controversies since then: Although money does not equal the true value of an artwork, is it worth

978-1-6654-5311-0/22/$31.00 ©2022 IEEE

so much money? Can an algorithm-generated painting, which has no uniqueness, be considered an artwork? Isn't it just a random combination of copied human artwork pieces?..... But these debates did not reveal what an epoch-making event it is: It demonstrated a new possibility that humans can co-create more efficiently with the accumulation of human civilization across time and space through AI. This is the essence of human-AI co-creation.

How is the making of this artwork? Obvious, a French art group of two artists and one AI expert, fed the AI with 15,000 portraits from the 14th to 20th centuries as training data and enabled the AI to generate paintings such as *Edmond de Belamy*. The source code of the AI was borrowed from the one shared at Github, written by an American, Robbie Barrat. The foundation behind the code was GAN (Generative Adversarial Networks), a class of machine learning frameworks designed by another American, Ian Goodfellow, and his colleagues. In short, a French art group used human paintings to train the AI, which was based on an American's code, which was based on another American's invented technology. Can you imagine creating something in collaboration with people you have never met or unlikely to meet, around the world and across hundreds of years in time? You are given such an opportunity by AI, as a door across time and space to the accumulation of human civilization!

B. The Human-AI Co-creation Model

In my previous paper in 2021, *AI Creativity and the Human-AI Co-Creation Model* [7], I defined "AI Creativity" as the ability for human and AI to co-live and co-create by playing to each other's strengths to achieve more. AI Creativity is focusing on how to use AI creatively and how AI can inspire and enhance human creativity. AI is a complement to human intelligence, and it consolidates wisdom from all achievements of mankind, making collaboration across time and space possible. By introducing the Human-AI Co-Creation Model, a circular process model including 6 major phases, perceiving, thinking, expressing, collaborating, building and testing (Fig. 1), I demonstrated with case studies how AI empowers humans throughout the entire creative process, and makes creativity more accessible and more inclusive than ever.

In the past 12-18 months, the advance of AI in large-scale models such as GPT-3 and in multimodal models such as CLIP have significantly improved the performance of AI, especially in understanding and generating contents. The Human-AI Co-Creation Model gets more solid support. The rising popularity of Disco Diffusion among more and more average people in 2022, and the astonishing artworks generated by AI with human input prompts, gives us an interesting window to observe what large-scale and in-depth human-AI collaboration could be like and how human would work with AI to create meaningful objects.

III. A New Force and Evolving Needs

A. The Magic of Large-scale AI Model

Although in the previous research, we have identified over 1,600 AI Creativity application cases covering over 45 areas from culture and entertainment, industry, lifestyle, science research etc. [8], there was still a major constraint that most AIs can only deal with a certain type of tasks and each specific task often requires training with corresponding data. It means that only people with proper technology skills and resources could initiate an AI related work.

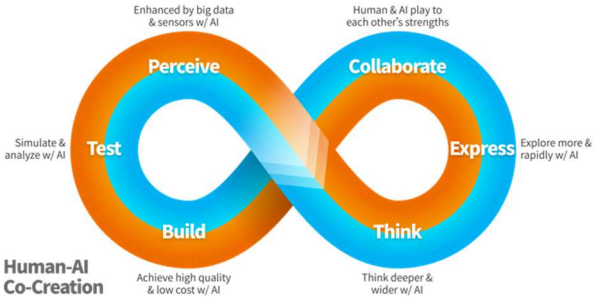

Fig. 1. The Human-AI Co-Creation Model.

Unlike GPT-3, which mainly works in the background, the new generation AIs give people direct participation, enabling average people to create astonishing images by giving the AI orders in the format of text or image prompts.

Creating art or design works has always been considered as a talent which only a small group of people have, as well as a skill requiring years of learning and practice [9]. However, the new products mentioned above break the barriers of talent, learning and practice, and empower everyone to work with AI to create visual results in a simple way as how they communicate with other people in natural language. On one hand, the large-scale AI models significantly improved the multimodal performance, which interprets words into visual results more accurately; On the other hand, the generation models got more solid support to generate more convincing visual results.

B. Personalization and AI Generated Uniqueness

No one has the same needs as everyone is different in one way or another, mentally or physically. However, it's very hard to provide a fully personalized product or content for all people. It seems that people can only choose one of the two options: standardized and less expensive, or personalized but expensive. An initial breakthrough happened in the information area, where platforms such as Tik Tok and Amazon use recommendation engine powered by AI to provide every user with a personalized information distribution experience. These examples shed light on the future of personalization when AI could eventually expand into more areas such as virtual content generation and physical product manufacture.

A clear trend among young people around the world shows that they are expecting more personalization of more things, compared with their predecessors. Such huge needs won't be able to be met in the traditional way, which is by other people who consume less but contribute more. Then it naturally goes to AI: AI automatically provides personalized stuff without the involvement of humans, or at least reduces the cost of personalization by helping humans to make the stuff. In other words, AI could make large-scale personalization possible for the first time in human history.

Many people hold an impression and option that as AI is very good at tirelessly repeating, there can't be uniqueness in the results AI generates. However, this is not true. AI can steadily generate the same kind of results, but it can also generate variations at all levels, from slight to significant

differences. What's more, with AI generation tools such as Disco Diffusion, you actually won't be able to generate a totally same result at all! This is the very uniqueness that people usually appreciate in high quality arts and designs.

IV. A New Strategy and Philosophy

To embrace AI and unleash the great potential within, a new strategy and philosophy for co-creation with AI needs to be established.

A. The Personalization of Collective Memory

All AIs are trained with data, especially texts, numbers, images, videos and music on Internet. These data all together across time and space make a collective memory of the world. I prefer the words "collective memory", rather than "collective knowledge" or "collective intelligence" because these data are really like memory, some true, some false, some complete, some fragmented, and far from intelligence which we can directly rely on to make decisions. Audrey Watters shares a similar view: " By preserving memory and knowledge, these technologies have helped create and expand collective memory — through time and place. We can share this collective memory. This collective memory is culture — that is, the sharable, accessible, alterable, transferable knowledge we pass down from generation to generation and pass across geographical space, thanks to information technologies. [10]"

The AIs are of course what they are fed. And when such an AI is asked to create something, it would start to cook based on the collective memory it has and put some randomness as a secret sauce. E.g. Given prompts like "a window with a view in [city name], beautiful morning", WOMBO Dream generated the following images (Fig. 2).

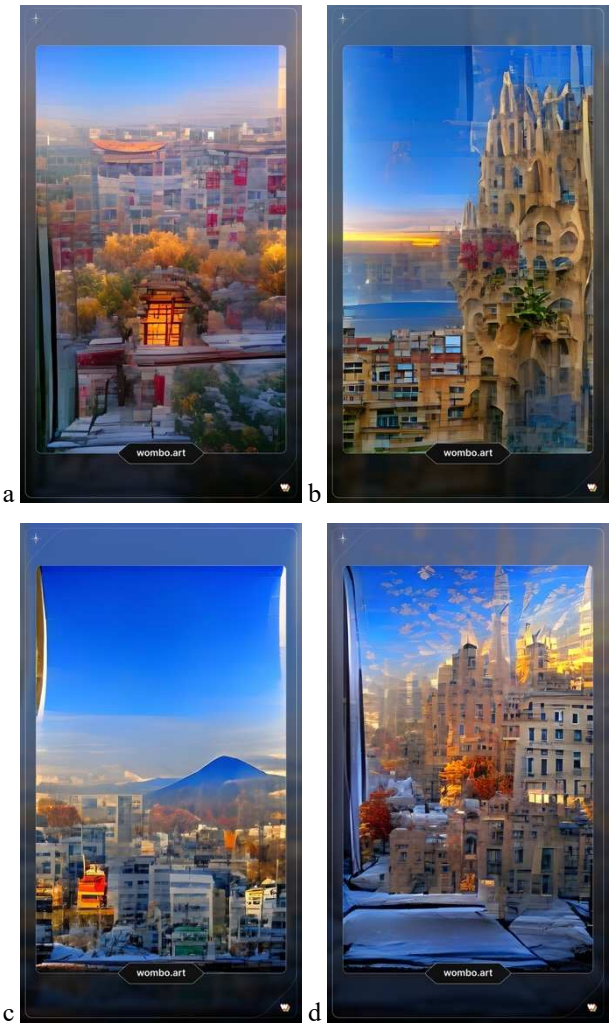

Fig. 2. A window with a view in (a) Beijing, (b) Barcelona, (c) Tokyo, (d) New York, created with Dream.

The images above aren't the photos of any scenes in the real world, but do give a sense of the particular cities. This AI generates the images of a city based on the city's collective memory in cyber space, which consists of all images the AI accesses. You might want to try to do it yourself, or check my experiment of 100 images of each city [11].

Creating based on memories is one of the key abilities to do art or design works. Usually an artist/designer would create something based on the memories of their own and on a limited amount of other people's. The large-scale collective memory AI accesses is based on numerous people's data. It's a great pool of resources to leverage, so to create things beyond the artist/designer's personal limit; it's also a great challenge for the artist/designer to find ways to get the right content combination out of a black box. E.g. Given prompts like "a chili plant by a window with a view", AI could generate a reasonable result sometimes, an illogical result other times (Fig. 3). We can consider this process as the personalization of collective memory: **Extract elements from the collective memory and re-organize them with an intention, embracing the inspirations triggered by randomness**. Is this something new? Not really. This is what artists/designers have been practicing since the beginning of art and design. The main difference is the scale and efficiency of use of collective memory.

Fig. 3. A chili plant by a window with a view, (a) created with Dream and completed by human, (b) created with Dream, (c) created with Disco Diffusion.

B. *The Human-AI Collaborative Workflow*

As currently AI's semantic understanding, causal inference ability is not strong enough, the generated results are often unstable in terms of quality. Usually the results of the impressions of large things, such as mountains, sky, sea, trees, flower bushes, a group of buildings and a huge starship, can be very good; while the results of detailed objects, such as a human, an animal, a plant, a machine, are often seen with flaws. The results of fantasy or science fiction images can be very good; while the results of realistic photos are often seen with flaws. The main reason is some logical problems are easy to find while some others aren't. And these kinds of logical problems are inevitable with the current AI technologies.

How can we improve the human-AI collaboration to achieve better results of generation? A workflow well-designed for a certain task is the key. Let's take artistic creation as an example.

a) Ideation: the human-dominated phase

- The general direction of the work still needs to be decided by humans, unless a wild ride by AI itself is the purpose.
- The classic methods of ideation still work, such as brainstorming, user research, design research, problem solving methods, scratching, reading, writing, seeking inspirations in a trip, and tec.
- The social network, knowledge/image/video sharing platforms also give us direct access to collective memory. A human views these contents much less efficiently than AI does. However, this kind of direct exposure does provide a meaningful and irreplaceable experience for artists/designers.
- The ways AIs provided can be very inspiring too, such as viewing the knowledge graph of a keyword based on the big data on Internet [12], or interactively writing with an AI, such as the one in AI Dungeon, to compose a compelling story.

b) Explorations: the human-AI collaboration phase

- Although AI can generate results tirelessly, the time it costs might not be affordable. More GPUs means less time cost, however, it's also not affordable to most people. Thus we must find the proper prompts and parameters as soon as possible to guide AI to generate expected results effectively and efficiently [13].
- It's only half true that today's image generation AIs allow people to give orders in human natural language, which is called prompts. Due to the limited ability in semantic understanding and causal inference, it could be very challenging and time consuming to have fine control of the generated results by consciously adjusting prompts. Some people call it "prompt engineering" to describe how to build an effective and efficient way of communication with AI. Well-structured sentences, sometimes even keywords, often work better than normal written sentences, similar to what we do with a search engine. Words which can be connected with visual images easier, often work better than those can't, e.g. "red" performs better than "beautiful". Words with some logic involved often fall into trouble, such as "within". Just as Gary Marcus argued that even the best text-to-image generation AI of today "can't distinguish 'a horse riding an astronaut' from 'an astronaut riding a horse'" [14].
- There are some other advantages to working with AI, in addition to having AI to explore tirelessly. For text prompts, we can also put in various types of hints for visual effects or styles, such as art materials, art/design websites with sources to imitate, software which generates a particular kind of visual effect/style, and the names of artists whose artworks have been used as training materials. For image prompts, a reference image can guide the image generation in the perspectives of layout, shapes and colors [15].
- The concept of style to AI is more about a set of visual patterns which can be used as styles. E.g. the image of a chip can be a style which AI can transfer another image into (Fig. 4). The concept of weight is also introduced into prompts, where we can give certain values to the weight of each part of the prompts to indicate the importance of them. And giving weight value is a typical example of adjusting parameters to achieve better generated results. In terms of the workflow, we can either generate the content and the style at the same time, or generate the content first and then apply various styles onto the content later.

Fig. 4. AI Creativity: a fused brain of half human and half AI, style transfer with Deep Dream Generator.

c) Gneration at scale: the AI-dominated phase

- Once we find ways through the explorations to guide AI to generate expected results effectively and efficiently, it's time to generate at scale. It's a race depending on AI (models and parameters) and money (GPUs).
- The best thing creating with AI is that we can harvest the results after a while, just like a professional team is working for us.
- Among the results AI generates based on the prompts and randomness, there will be some mediocre ones, some crappy ones, and some good ones beyond expectation. All we need to do is to pick out the proper ones based on our initial ideas and personal tastes.

d) Finishing or further development: the human-dominated phase again

- Sometimes the picked results have already been very good and some simple fix would be enough. Sometimes we can push further to incorporate the AI generated results into something sophisticated for a greater purpose. Multiple types of AIs, such as AI for painting, design, music, video, literature etc., can be used together to create an integrated artwork/design (Fig. 5).
- Following by a style transfer might be the most common procedure. Serving as assets for video/animation/3D mapping production is also a good way to go.
- The meaning giving to a generated result in this phase is the key differentiator to make a decorative artwork or a meaningful masterpiece.

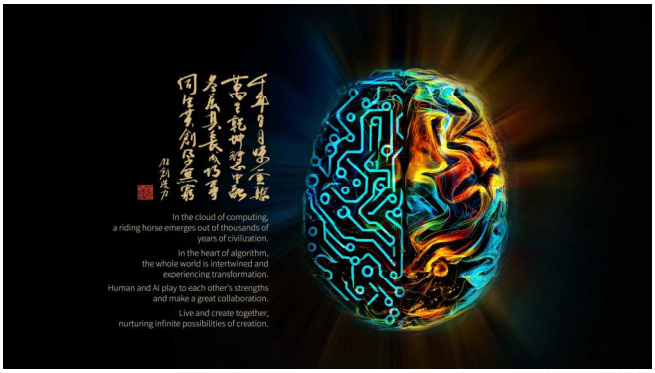

Fig. 5. AI Creativity: a painting, a poem in ancient Chinese and an English translation, created with multiple AIs.

V. Conclusion and Future Work

The large-scale AI models give people a chance to create out of a large pool of collective memory, which they won't be able to access before, and communicate with AI in an easy and natural human way. As demonstrated and discussed in this paper, this kind of co-creation based on the personalization of collective memory has great potential to inspire more people to create more. But there's also huge room for improvement.

A. AI Technology and Product Improvement

The breakthroughs of fundamental technologies about semantic understanding and causal reasoning are most expected. The continuous engineering improvements will be very helpful too. E.g. Nvidia Canvas is a lot easier to use than Disco Diffusion, because it has a graphic user interface [8].

B. Control over the Generation Process

The black box mode of generation is indeed a challenge and the main cause of the waste of time and resources. We neither know what the result would be until it comes out, nor can we do anything about it during the process.

Image generation AI Midjourney and music generation AI MuseNet both introduced a mode which allows people to take a pause between steps, pick the most wanted one out of a group of generated results, and then further develop it to move to the next phase. These approaches show us that AI products can achieve more even without technology breakthroughs.

C. Integrated Workflow and Working Platform

As we demonstrated in this paper, to co-create with AI efficiently, the first thing is to build up a clear workflow, so human and AI can play to each other's strengths effectively in each phase.

Once integrated workflows become steady, products can be developed based on them. More people then will get a chance to use these user-friendly products, rather than struggling in codes.

Acknowledgment

Liwen Feng, Lina Dong, Chen Chen, Zhou Fang, Hao Ni, Sibo Yang, prof. Xin Lv and Sunny Wu, thanks for helping us in building environment and workflow for AI art explorations.

References

[1] Ramón López de Mántaras, "Artificial Intelligence and the Arts: Toward Computational Creativity", https://www.bbvaopenmind.com/en/articles/artificial-intelligence-and-the-arts-toward-computational-creativity/, 2017.

[2] Marvin Minsky, “Music, Mind, and Meaning”, Computer Music Journal, Fall 1981, Vol. 5, Number 3, https://web.media.mit.edu/~minsky/papers/MusicMindMeaning.html, 1981.

[3] Pierre Barreau, “How AI Could Compose a Personalized Soundtrack to Your Life”, TED2018, https://www.ted.com/talks/pierre_barreau_how_ai_could_compose_a_personalized_soundtrack_to_your_life, April 2018.

[4] Chris Garcia, “Harold Cohen and AARON—A 40-Year Collaboration”, https://computerhistory.org/blog/harold-cohen-and-aaron-a-40-year-collaboration/?key=harold-cohen-and-aaron-a-40-year-collaboration, 2016.

[5] Reena Jana, and Josh Lovejoy, “Exploring and Visualizing an Open Global Dataset”, https://ai.googleblog.com/2017/08/exploring-and-visualizing-open-global.html, August 2017.

[6] Bo Zhang, “AI and the Creation of Art”, https://www.xuetangx.com/live/live20210318M001/live20210318M001/7415444/10506841, March 2021.

[7] Zhuohao Wu, Danwen Ji, Kaiwen Yu, Xianxu Zeng, Dingming Wu, and Mohammad Shidujaman, “AI Creativity and the Human-AI Co-creation Model”, HCII 2021: Human-Computer Interaction. Theory, Methods and Tools, pp. 171–190, July 2021.

[8] Zhuohao Wu, and CREO, “AI Creativity Report 1.0”, https://www.wecreo.org/, 2021.

[9] Weiwen Chen, Mohammad Shidujaman, and Xuelin Tang, “AI Art: Towards Artificial Intelligence Art”, The 12th International Conference on Advances in Multimedia, 2020.

[10] Audrey Watters, “Memory Machines and Collective Memory: How We Remember the History of the Future of Technological Change”, October, 2017.

[11] Zhuohao Wu, “100 Images of Each City in WOMBO Dream’s Collective Memory: A Window of a View, Beautiful Morning”, https://mrhow.feishu.cn/drive/folder/fldcndBOUpNUAMZ4lU19KtR8XEf, May 2022.

[12] Andrei Kashcha, “Visualization of Google's Autocomplete as a Graph”, https://anvaka.github.io/vs/, November 2018.

[13] Ker-Jiun Wang, Caroline Yan Zheng, Mohammad Shidujaman, Maitreyee Wairagkar, and Mariana von Mohr, “Jean Joseph v2. 0 (REmotion): Make Remote Emotion Touchable, Seeable and Thinkable by Direct Brain-to-Brain Telepathy Neurohaptic Interface Empowered by Generative Adversarial Network”, 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 3488-3493. IEEE, 2020.

[14] Gary Marcus, “(Google Imagen) can’t distinguish ‘a horse riding an astronaut’ from ‘an astronaut riding a horse’”, https://nitter.tannhauser.network/GaryMarcus/status/1529269425928515584, May 2022.

[15] Weiwen Chen, Mohammad Shidujaman, Jiangbo Jin, and Salah Uddin Ahmed, "A Methodological Approach to Create Interactive Art in Artificial Intelligence", International Conference on Human-Computer Interaction, pp. 13-31. Springer, Cham, 2020.