OUKA
The University of Osaka
Institutional Knowledge Archive

Title	Creative Agents: Simulating the Systems Model of Creativity With Generative Agents
Author(s)	Imasato, Naomi; Miyazawa, Kazuki; Nagai, Takayuki et al.
Citation	IEEE Access. 2025, 13, p. 157712-157729
Version Type	VoR
URL	https://hdl.handle.net/11094/103491
rights	This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 3 August 2025, accepted 1 September 2025, date of publication 5 September 2025, date of current version 12 September 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3606498

RESEARCH ARTICLE

Creative Agents: Simulating the Systems Model of Creativity With Generative Agents

NAOMI IMASATO[1], KAZUKI MIYAZAWA[2], TAKAYUKI NAGAI[1,2], (Member, IEEE), AND TAKATO HORII[1,3], (Member, IEEE)

[1]Graduate School of Engineering Science, The University of Osaka, Osaka 560-8531, Japan
[2]Artificial Intelligence Exploration Research Center, The University of Electro-Communications, Tokyo 182-8585, Japan
[3]International Research Center for Neurointelligence (WPI-IRCN), The University of Tokyo, Tokyo 113-8654, Japan

Corresponding author: Naomi Imasato (c.imasato@rlg.sys.es.osaka-u.ac.jp)

This work was supported in part by JSPS KAKENHI under Grant JPJS00122674991 and Grant JP23H04834, and in part by JST Moonshot Research and Development under Grant JPMJPS2011.

ABSTRACT With the growing popularity of generative AI for images, video, and music, we witnessed models rapidly improve in quality and performance. However, not much attention is paid towards enabling AI's ability to ''be creative''. We often attribute the quality of ''being creative'' to an individual or an object, but we believe that countless variables participate in determining what or who is creative, transcending a single entity or artifact. Csikszentmihalyi's systems model of creativity suggests that creativity is a product of interactions among multiple parts of a society that create, evaluate, and record. In this study, we implemented and simulated Csikszentmihalyi's systems model of creativity using virtual agents utilizing large language models (LLMs) and text prompts. We conducted experiments in virtual settings where creativity is achieved with the presence of specific characteristics in the artifact. For comparison, the simulations were conducted with two ''virtual artists'' being 1)in the system, which received feedback from the field, and 2)isolated, which did not. Both agents were compared by analyzing the novelty, which was measured via Creativity Implication Network, and value, quantified through the desired characteristics present in artifacts. Our results suggest that the agents that receive feedback from the field can generate artifacts that are more novel and more valuable, thus more creative, in the framework of the systems model of creativity. Furthermore, the difference becomes more evident when external factors enact changes to the domain.

INDEX TERMS Computational creativity, large language model, systems model of creativity, virtual agents.

I. INTRODUCTION

As the use of artificial intelligence (AI) models and solutions is becoming common, researchers and AI enthusiasts have been working on maximally utilizing the definition of *''Intelligence''* in AI. There is an increasing interest in exploring a machine capability for solving problems that are not well-defined, which requires being creative.

The definition of creativity varies across different sources of literature, but two factors that are consistently referred to are *novelty* and *value*, both of which are difficult to define in a vacuum, as *novelty* refers to something being presented for the first time, and *value* refers to the importance or relevance that it holds [1], [2]. The complexity in defining *novelty* and *value* lies in the fact that both imply a collective meaning. To determine *novelty*, the relative creations must be known, as well as to whom they were presented to. In practice, the public relies on the information and commentaries propagated by specialists. For example, if an art curator indicates that a certain painting is credibly unique and has a high market value, we are compelled to believe it. Artists seeking to create new artwork may also rely on the implications of this aforementioned curator to remain updated with the current standards of creativity, thus have a higher likelihood of producing a successful artwork. This dynamic was described in the work of Csikszentmihalyi. He proposed a systems model of creativity [3] in which the society is partitioned into three main groups. Each part plays a role in determining and shaping the standards for what is considered ''novel'' and ''valuable''; namely, ''creative''.

The associate editor coordinating the review of this manuscript and approving it for publication was Yilun Shang.

© 2025 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

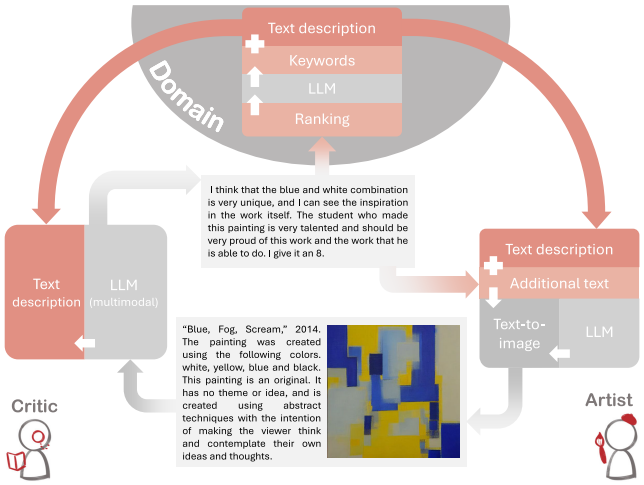


FIGURE 1. Each part of the system is composed by a combination of pieces of text and generative models. The *Domain* is described in text form, which is passed to complement the descriptions for both *Artist*and *Critic*. The text description in the *Artist* is used in a text prompt template that generated the text that describes the *Artist*'s new piece of art (art prompt). The art prompt is used to generate the image that should depict what the *Artist* described. Both the image and the art prompt are fed into an LLM, where another template for a text prompt is filled with the text description for the *Critic*. This will give us the critique that this *Critic* wants to give to the *Artist*'s latest artwork. This critique is used to place this new artwork in the overall ranking of the system, which represents changes in the *Domain*. The same critique is also fed into the LLM with the *Artist*'s text description in another text prompt template to make the *Artist* "reflect" over the critique it just received. The output from this "reflection" goes into the *Artist*'s additional text. We run these steps iteratively, where the text defining the *Artist* and the ranking in the *Domain* change at each step.

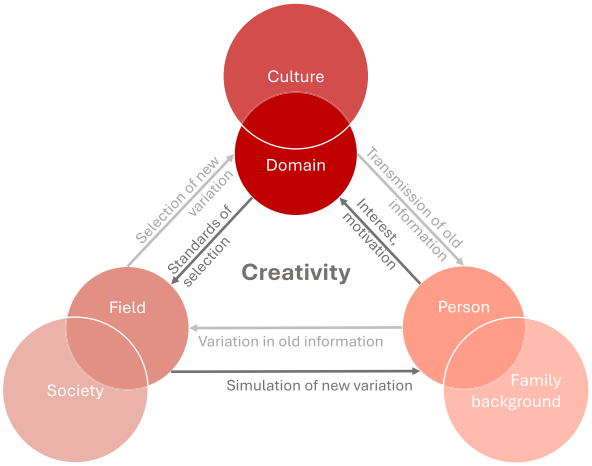


FIGURE 2. Diagram of the systems model of creativity as proposed by Csikszentmihalyi.

In this work, we introduce a framework for a multi-agent system for computational creative. Our design is based on the works of Csikszentmihalyi. In later sections, we detail the design of our agents and system, as well as the experimental setting we used to compare "virtual artists" that do and do not have direct communication with the field. We show that our results suggest that agents that receive direct instructions (critique) from the field tend to perform better in creativity (i.e. generate more novel and more variable artifacts).

II. RELATED WORKS

A. CREATIVITY

A fundamental problem in the field of computational creativity is defining *creativity*. Researchers have used different approaches to define creativity. Multiple studies by Boden are often used as references to define creativity [4], [5].

Boden often defined a hypothetical space of ideas, the process of forming new ideas in this space, and whether these new ideas indicate a transformation in the space of concern. Despite its popularity, Boden's definition poses a major obstacle for computational implementation, as it requires high levels of abstraction to represent all the "ideas" in a space. Conversely, Csikszentmihalyi designed a systems model of creativity, which proposes that creativity is product of social interactions and transforms overtime as part of societal change [3]. Unlike Boden's definition, Csikszentmihalyi's approach considers the process of making something "creative" as a product of the interactions between multiple parts, which can be more easily achieved with the current technology.

We based our study on the definition of Csikszentmihalyi's systems model of creativity; therefore, in this study, we will not focus on determining the creativity in a single individual, rather observing the impact of social interactions on "creativity" in a virtual environment with multiple generative agents.

The systems model of creativity features individuals who create artifacts (i.e. artists), the community to whom the artifact is being displayed, and the context in which the artifact was created (e.g. location, historical events, significant pieces of art, recent trends and political climate). Each part of the system that does not live in a vacuum is constantly affecting the others and vice-versa: the artist does not create art without inspiration or motivation from the community or context they are in, the community cannot exist without multiple individuals contributing to the domain with new artifacts, and the domain cannot change without the creations of the artists or the judgment of the community. All the variables form a cycle in which each part of the system contributes to the system itself. A simple diagram of the systems model of creativity is shown in Figure 2. The domain preserves and transmits culturally relevant information. Csikszentmihalyi refers to cultures as "interrelated domains". A community that shares the same "culture" acts, communicate and interact in ways specified in its symbolic system. A person grows to learn from said symbolic system, but also from the persons that are more closely related to them (often, one's family). The arrangement of individuals surrounding the person, as well as their own personal qualities, make this person unique. A person comes up with a new idea by modifying a piece of information that was transmitted by the domain with their own unique perspective. This new idea is then judged by the field, where a society of experts will decide whether the new idea is worthy of being adopted or not into the domain. The experts in the

field are individuals who are educated or experienced in the subject of interest (e.g. art). Experts have the knowledge and insights to what can be considered valuable and creative to the domain, while agreeing upon and following standards determined by the domain itself. The interactions between the experts in the field and the persons motivate the emergence of more variations, be it through encouragement in constructive criticism, or as a response to strict standards. Notice how causality moves in both ways across the three parts and there is no ''starting point''. All parts are necessary for ''creativity''.

Analogous to Csikszentmihalyi's theory on the social aspect of creativity, Glăveanu emphasizes the role of an audience in creativity in his writings [6]. He claims that a creative action emerges from one's change in perspective, which implies the existence and understanding of another individual's (or group's) perspective.

B. CREATIVITY AND AI

Researchers in the field of computational creativity have attempted to develop a computer program capable of being creative, such as the Joke Analysis and Production Engine (Jape) [7]. It was developed at a time where AI was at its early stages of development, and was designed to create puns. Humor heavily relies on creativity, considering its surprising aspect that is found by many. Although the model can be considered robust for generating punning riddles, it is also highly restrictive and limited in terms of the possible outputs.

Generative adversarial networks (GANs) [8] were once a popular choice for computational creativity owing to the latent space created during training (alluding to Boden's hypothetical ''space of ideas''). GANs are formed by using a generator and discriminator model, where both parts are trained to outdo the other (hence, ''adversarial''); DesIGN is an example of this model [9]. The authors proposed StyleGAN, an adapted GAN model where the generator is conditioned by the texture and the shape of the articles of clothing while being ''encouraged'' to generate more novel samples, as opposed to replicating the samples observed in training, as that performed by the classic GAN. This ''encouragement'' was achieved by adding two loss functions that ''confuse'' the discriminator in the GAN. However, GANs are notoriously difficult to train, as both the discriminator and generator should ideally perform sufficiently; however, achieving a balanced training process is challenging.

Transformer-based models have recently become a staple of generative models [10], [11], [12]. Considering the development of transformers, researchers have proposed a method for controlled text generation where the user can specify ''tags'' or ''keywords'' to manipulate the probability distribution of the next word to be sampled by an autoregressive model (GPT-2). For example, the Plug and Play Language Model (PPLM) [13], successfully controlled the generation of text sequences, where a significant portion of the outputs followed the ''tag'' or ''keyword'' specified without compromising the output coherence or grammatical accuracy. Another noteworthy experiment conducted by the authors utilized more than one ''tag'' or ''keyword'' for the controlled text generation where the control terms were unlikely pairs. Although the authors did not assess the creative ability of this method, they shared some noteworthy results obtained by combining different themes and concepts. Compared to Jape, the model imposed significantly milder constraints and limits, and the outputs were based on a large volume of written information sourced from the Web. The limit is based on the ability of the model to tie two or more concepts together when they are significantly different from one another. We successfully used this controlled generation method in music generation [14], where the autoregressive model was trained to generate sequences of MIDI events instead of text. Control was achieved using a part of the PPLM method, where we used an auxiliary discriminator model to classify the data into two or more categories. In this case, the emotion perceived from a piece of music was used as the main factor (categorized according to Russell's model of affect). Analogous to the original publication of PPLM, using unlikely pairs of controls and prompts can lead to noteworthy (or creative) outputs.

Assessing or measuring the creativity of an artifact is one of the hardest (if not the hardest) problems in this field. As humans, we visualize or hear a certain type of art or music, respectively, and almost immediately make judgments based on the first impressions and our preferences. If desired, we also seek for context, debate and further discussions with other people regarding whether the art or music of concern can be considered creative. However, if asked to define measurements or objective values to determine the creativity of an artifact, the reply is significantly less trivial and more complex than determining whether something or someone is creative. In an experiment, the authors of Jape considered 122 children who read jokes (generated by Jape and humans) and non-jokes (sensible and nonsensible ones). The texts generated by Jape performed almost as well as the human-made jokes compared to the non-joke texts [7]. To assess the creativity and funniness of a punning riddle, the authors resorted to the decisions made by a group of individuals who participated in their experiment; as the judgment of whether something is creative or funny is highly subjective, this is a significantly common practice among the different works in the field. Certain objective measurements can be considered, such as the choice of words for poetry, color palettes for images, among others. However these metrics do not determine creativity; in fact, many artifacts can have good values for these metrics and remain to be considered non-creative, similar to one with poor values in the same metrics being considered highly creative. In this regard, the quality of computationally generated artifacts can be easily measured with objective metrics; however measuring their creativity is significantly more complex given the subjectivity of the matter.

Elgammal and Mazzone discuss the changes in creative processes caused by the AI-assisted creativity [15]. The essay addresses a human-AI collaborative creative process: the ''user'' (artist) inputs their own curation as ''inspiration'' or ''references'' to the generative model. The user can successively make adjustments to the model's configurations as it is used to generate artifacts, until they are satisfied with the output. From all generated artifacts, the user can then select the few that express their vision best. The system the authors used is CAN, a GAN-inspired system where one of the goals is producing creativity via style ambiguity [16]. Another study reveals that generative AI enhances an individual's creative abilities but reduces the collective creativity [17]. The authors performed experiments where participants had to write short stories under one of three conditions: no AI assistance, one generative AI idea, and five generative AI ideas. In the case of this study, generative AI was used mainly as a source of inspiration to the writers. The authors found that the use of AI in writing helped writers that are not very creative achieve a higher ''personal'' level of creativity, but the stories written by those who received AI assistance had higher similarity within their group (one genAI and five genAI) than those that were fully written by humans.

As of the writing of this work, we believe that the AI has great potential to boost human creativity, be it as a source of inspiration or as a collaborative ''partner''. However, as the aforementioned study revealed, we believe that human creativity still is key to achieve truly *Creative* artifacts.

C. HUMAN-LIKENESS IN AI MODELS

Modelling human behavior to better understand the mechanisms and dynamics in our society is a major goal of AI research. In this regard, large language models (LLMs) integrated in chat bots, such as GPT-4 integrated in ChatGPT [18], are among the most notable developments in the field. The quality of outputs obtained by ChatGPT users is significantly high, leading to discussions regarding Artificial General Intelligence (AGI) and ethics of the training and usage of these AI models and systems.

A survey on individual to society simulation proposed different classifications of LLM-based (multi-)agent systems [19]. Individual simulations target reproducing the behavior of a character, personality, or a group of people with similar characteristics. The two main approaches to achieving such feat with LLMs are with *fine-tuning* or *prompting*. Fine-tuning requires some level of training, which is generally more demanding in terms of resources. Here, we highlight some works that use text prompting to skew LLMs to generate text that is appropriate for their goals. Text prompting (or prompt engineering) refers to a technique used to control the output of an LLM through the input. Researchers found that LLMs are capable of fulfilling certain tasks when exposed to one or few examples in the text prompt (input) [12], [20]. Deshpande et al. prompted interactions with LLMs with personas to assess how language models could be misused [21]. The approach of assigning specific profiles is now a common practice in LLM-based agent research. Multi-agent works are divided into two groups: scenario and society simulations. The differences between the two categories lie mainly in the scale and the target of the simulations. Scenario simulations are typically comprised of few agents that ''work'' toward the completion of a specific task. Society simulations are larger in scale and have larger diversity of agents. They typically target studying the social dynamics that emerge among agents. The work presented here fits in the scenario simulation category.

Li et al. presented another work that is highlighted in the scenario category [22]. The authors proposed an LLM-based agent framework in which the agent's memory is hierarchically organized into three tiers. This structure is modeled after the idea that human beings have short-, medium- and long-term memories. Additionally, the authors designed the agents to perform trading, which encompasses complex decision-making.

A more recent study, and also the main inspiration for this study, used LLMs to simulate individual agents in a small community, where each agent was defined by text excerpts used to prompt text generation for their actions and utterances [23]. The authors also conducted a study to assess the human-likeness of the agents, as well as the generation of actions and phrases by LLMs that appeared natural to most participants. Additionally, when an agent was initialized with a plan to host an event for the community, the agent invited other agents to the event, who did the same and invited other agents. Considering the results showcased in their study, we hypothesize that we can have a similar level of human-likeness in LLMs when it comes to creative tasks.

Another study used a similar principle but maintained the LLM requests to a minimum, thus required a lower demand of resources [24]. The authors also used different ''situations'' where the agents would have specific problems to solve to evaluate their performance. These studies indicate that LLM-based agents can fulfill their roles in a human-like manner and successfully coordinate among themselves as a community. These two works are notable examples of the society simulation category.

As LLMs were successfully used in simulating ''individuals'' in both of the aforementioned studies, we hypothesized that LLM-based agents can be used to run simulations of the Csikszentmihalyi's systems model of creativity [3].

III. METHOD

In this study, we designed, implemented and tested a simplified simulation of the systems model of creativity, as proposed by Csikszentmihalyi [3]. In this section, we describe the overall structure of our system and the decisions made during its implementation.

A. THE SYSTEMS MODEL OF CREATIVITY

According to Csikszentmihalyi [3], creativity does not originate from an object or person in isolation; rather,

"it is the product of three main shaping forces: a set of social institutions, or field, that selects from the variations produced by individuals those that are worth preserving; a stable cultural domain that will preserve and transmit the selected new ideas or forms to the following generations; and finally the individual, who brings about some change in the domain, a change that the field, will consider to be creative. (. . .) Creativity is a phenomenon that results from the interaction between these three systems." Based on this description, three essential elements must be considered in a systems model of creativity: the creative individual (or simply *individual*), field, and domain. Note, at the end of the aforementioned passage, the author refers to each part as a *"system"*, as each part is highly complex on its own.

The creative individual would not be the same if it was not for the context they are in. This includes (but is not limited to) the political, historical, and cultural contexts they are surround by. Although people have their own individuality, personality, and experiences, their surroundings play a major role in shaping their opinions and values. Despite using text, describing each detail of an individual and "modeling" them would be practically impossible. A single individual is difficult to model precisely. Modeling the field and domain is an even more complicated task, as they are formed and affected by multiple agents, each with their own complexities.

Algorithm 1 Simulation of Systems Model of Creativity

```
Initialize artists, field, domain
t = 0
while t < 15 do
    for artist in artists do
        prompt ← art_templ(domain.desc,artist.desc)
        artist.art_prompt ← LLM(prompt)
        artist.image ← text2img(art_prompt)
    end for
    for critic in field do
        for artist in artists do
            prompt ← crit_templ(domain.desc,critic.desc,
artist.art_prompt)
            feedback ← LLM(prompt)
            critic.critiques.append(feedback)
            artist.update(feedback)
        end for
        domain.update(critic.critiques)
    end for
    t = t + 1
end while
```

B. OUR DESIGN

Given that the system of creativity in the real world is significantly complex, and precisely reproducing it in a virtual setting is practically impossible, we decided to simplify the model originally proposed in a previous reference [3].

We considered three fundamental subsystems: artist (individual), field, and domain. Each part operates using a combination of generative models and text prompts. We further describe each subsystem as follows. An overview of our system is shown in Figure 1, and the pseudo-code for our simulation can be found in Algorithm Block 1.

We used Gemini Flash versions 1.5 and 2.0 [25] for text generation and multimodal text generation tasks. We refer to Gemini as the LLM in the sections below because theoretically, any LLM can be used for this system. Stable Diffusion version 1.5 [26] was used to visually express the agent creations and generate images from text prompts.

1) ARTIST

Considering a simple description, the artist is responsible for creating art and contributing to the domain with new pieces of art (contributing novelty). Despite having the same common goal of creating art, each artist has their own individuality and creates art in their own way.

Analogous to a study that simulated agents in a village [23], we also described the *agent artist* in text form, which was used to prompt text generation. The artist agent was initialized with a "core description", which was manually expressed. For example, we can write the main motives and inspirations of the agent. An individual usually has their own set of "principles" that define them and are rarely changed. These aspects were expressed for the "core description" of our agent. However, regardless of these principles, certain (less deep) aspects of an individual can change overtime as they interact with their surroundings. To describe these subtle changes that may occur overtime, we maintained a log of "additional text descriptions" that were generated after the agent "reflected" on the feedback their art received.

We expected a high rate of "change", as the artists in our experiments were defined as young and inexperienced artists who wish to become successful. The artist agents were defined such that they would be more susceptible to "accepting" advice in the feedback. Lastly, the artist agent was also equipped with a text-to-image model that would generate the "piece of art" that it intends to create.

2) FIELD

The field is essentially responsible for "curating" what information, out of the newly created artifacts, should be preserved in the domain, consequently shaping the culture and trends of the collective. According to Csikszentmihalyi, *"Novel ideas are not recognized or adopted unless they are sanctioned by some group entitled to make decisions as to what should or should not be included in the domain. These gatekeepers are what we call here the field."*

Here, we implemented the field as one or more agents that take the role of experts in the field of arts. They are responsible for rating the "creativity" in newly generated artifacts, if the LLM-powered agent "deems" the artifact to be creative, it is granted a point. The artifact is otherwise granted no point. These decisions are made based on outputs obtained from feeding pre-designed text prompts along with the field agent's "profile" and the artifact's information.

The text generated from this process goes through sentiment analysis, in the case of it being POSITIVE, we consider the artifact creative. Otherwise, it is considered not creative. At the end of every iteration, these agents are also responsible for reassessing all artifacts that exist in the simulation based on the definition of the domain at that time step.

More specifically, the field agent used in this study takes the role of a mentor. This detail is specified in the agent's definition (akin to the artist's ''core description''). Through experimentation, we found that defining the field agents as mentors or other teaching positions enables them to provide more instructional feedback than those defined as critics or curators. This yielded more concrete points in the criticism generated by the field agents, making it clearer when the artists made changes in their artistic strategy due to the criticism they received.

3) DOMAIN

All the agents in a creative system coexist and share a space, which is not only a physical space, but also a historical, political, and cultural context, among other variables. The combination of all the different ''circumstances'' as well as the time and place, is referred to as the domain. In the same manner that individuals transform their surrounding environment, the surroundings can also cause changes within the said individuals.

To implement the domain in our system, we prepared a base text description that was used throughout the simulation, in which we provided the necessary (or known) information regarding the domain we intended to simulate. In addition, to better focus on the artistic aspect of the domain, we initialize the domain with a list of ''significant'' paintings and their descriptions to use as reference as the most significant paintings at time-step $t = 0$. For simplicity and convenience, we initialized the list with popular paintings that are currently regarded as significant. This list was maintained throughout the simulation and at each time-step, it was updated with the newly generated artifacts as well as the text prompt used to generate them (their descriptions). The paintings in the list were ranked according to the impressions obtained from the field agent(s). The ranking mechanics are detailed in the next section.

Each time the domain needed to be referenced, the top 3 most significant artworks were picked from the ranking and three keywords were obtained from each of their descriptions (nine keywords total). The keywords were selected using an LLM, where the text prompt explicitly instructed the model to provide keywords that can describe the painting without referencing the names of famous artists or paintings. In the early stages of our experiments, we found that the names of famous paintings and artists heavily influenced both the text and image generation.

4) THE SYSTEM DYNAMICS

At time-step $t = 0$, all parts of the system were initialized with their respective base descriptions that were manually written. The domain was also initialized with a collection of real paintings along with their descriptions. These paintings were considered the most significant artworks at the beginning of the simulation.

Once all parts are initialized, the artist agents **''create'' their paintings**. The process starts by building a text prompt from the base descriptions of the artist and domain. The prompt is constructed as follows:

```
[domain descrip.] + [artist descrip.] + "You
finished your latest painting. Describe what
you painted in an objective, detailed and brief
manner. Do not describe anything other than your
painting. Use no more than two sentences."
```

After feeding this prompt to the LLM, we generated a text that described the latest artwork of the agent (''art prompt''). The final step of the ''creation'' was obtaining an image of the artwork created by each artist agent, which was achieved by feeding the art prompt into a text-to-image generative model.

The art prompts and their respective generated images were then passed to the next step, with the field agents (referred to as ''critics'' hereafter). To **evaluate the new artwork**, we used a multimodal LLM, as it can generate text from an input comprising both text and images. The text input was built as follows:

```
[domain descrip.] + [critic descrip.] + "The
student made this painting. This is how the
student described his artwork: " + [art prompt] +
"Was the student able to convey his intentions?
Do you think this painting is creative? Explain
in one or two short sentences."
```

Performing the aforementioned enables, the LLM to generate what each critic is more likely to ''say'' regarding each new artwork based on their description. The final product of this step is a collection of critiques obtained from all the critics in this system.

When adding the newly ''created'' artworks to the domain ranking, we considered the critiques that each artifact received. If the artifact received a positive critique, the artwork earned a ''significance point''; however, if the critique was negative, the artwork did not earn any points. The sentiment assessment of the critiques was performed by an LLM pre-trained for sentiment analysis (using DistilBERT-base-uncased finetuned SST-2, made available as the default model for sentiment analysis on HuggingFace's library) [27]. All artworks added at initialization start with one significance point because they are meant to represent the most significant artworks at the very beginning of the simulation. Additionally, at each iteration, critics (field agents) go over the list of all the artworks to reconsider the significance of the artwork. Similar to the newly added artifacts, the artworks that already exist in the history of this simulation are awarded one significance point if deemed ''significant'' by the critic, and no points otherwise. The sum of the points awarded at each time-step was maintained along with the ranking. We also applied a decay to these ''significance scores'', where the sum was reduced by half every d time-steps. The decay was enforced thus the

considerations made at the earlier time-steps were not as valuable as those more recent. This mechanism was utilized to grossly emulate the effects of changes in the trends, where one style can lose its appreciation over time as another style replaces it. Thus, the more recent considerations of the critics have a better chance of employing changes in the domain.

The critiques were also returned to the artists, where each agent underwent a process of **"self-reflection"**, which occurred via another text generation that was prompted with the following:

```
[domain descrip.] + "The art student made a piece
of art that was described as:" + [art prompt]
+ "This artwork received the following comments
from these people:" + [critic title] : [critique]
+ [artist descrip.] + "Describe the artistic
direction you intend to take from now on in one
or two short sentences. Do not use bullet points
or items."
```

The output was considered as the products of the artist's self-reflections and added to the `[artist descrip.]` as additional information (the "additional text description" previously indicated). This "self-reflection" step was performed at the end of each iteration, which yielded a new piece of text describing this agent; usage of all the text was desired to describe the agent. To prevent reaching the limit of tokens, we considered all the additional text descriptions generated over the simulation and summarized them with an LLM, more specifically, a fine-tuned version of t5-small[1] [28]. Note that artists that do not received critique from the field in our experiments also perform self-reflection, but the prompt used omits the passage in red.

These steps were then repeated n times, where the descriptions of the artists and the ranking of significant artworks underwent changes according to the text generated at each time step. We encourage the reader to look at appendix for a more detailed description, in which we show how the text prompts are changed throughout a simulation.

IV. EXPERIMENTS

This section presents the simulations that were performed in our experiments. To help visualize and assess the effects of the use of the systems model of creativity, we described hypothetical domains where the use of certain colors determines whether an artwork is creative or not. The experiments were conducted in four conditions:

1) Artworks that use the color blue are considered highly creative;
2) Artworks that use the color blue are considered highly creative and the field agent likes artworks with the color green;
3) The simulation starts with blue being the "highly creative color", which is changed to red at $t = 4$;
4) The "highly creative color" changes from blue to red at $t = 4$ and the field agent likes artworks with the color green.

[1]https://huggingface.co/stevhliu/my_awesome_billsum_model

We ran six simulations for each condition and each simulation consisted of two artists (individuals), a mentor (field agent) and the domain. The "highly creative color" and the mentor's color preference were expressed in the text used to define the domain and the mentor, respectively. One of the artists behaved as defined in the previous section (referred to as *artist in the system*), and the other behaves as an artist that does not communicate with the field (*artist in isolation*). Just like the artist in the system, the artist in isolation has its own definition, is aware of the domain's definition, and generates artifacts based on this information. The difference between these agents lies in the fact that the artist in isolation does not receive the critique from the field, and performs the "self-reflection" based only in the domain's and its own descriptions.

Our experiments allow us to compare how the artist agents perform "creatively" in the different settings when the field can and cannot communicate with the individuals. In other words, we verify whether the use of a "social framework" can help the virtual agents achieve "creative" outputs. As we wish to see the how the artist agent (and its generated artifacts) change throughout the simulation, we defined the **artists** (individuals) with the following:

```
"Answer as a young person who is not very
skilled, but enjoys painting. Your dream is to
become a creative and successful artist."
```

As the above text prompt suggests, the individual is defined as a novice artist who seeks success in the art industry. All of the artists are defined as not being very skilled, which leads them to be more susceptible to adopting trends to achieve the "success" they seek.

The only agent representing the **field** is the mentor, which is defined with the following:

```
"Answer as a mentor that works in guiding young
talents in becoming better artists. This mentor's
personal goal is to assist young artists in their
journey in being creative by providing valuable
feedback. This mentor is very fond of artworks
that contain the color green. The feedback must
be honest and must take into account what is
currently regarded as creative and the mentor's
opinions."
```

In experimental conditions where the mentor has no preference for green, we remove phrase `"This mentor is very fond of artworks that contain the color green."` from the agent's definition.

The context in the **domain** is set as follows:

```
"This is the year of 2024. The Art Industry is
highly competitive and artists try their best
to stand out. Art critics and curators have
become more strict when selecting what they
want to showcase, as there are only so many
spots to fill in galleries. Every day there
are many new artworks and it is hard to create
something new. Artworks with blue on them are now
considered highly creative. Keywords associated
with currently significant artworks include: " +
[keywords]
```

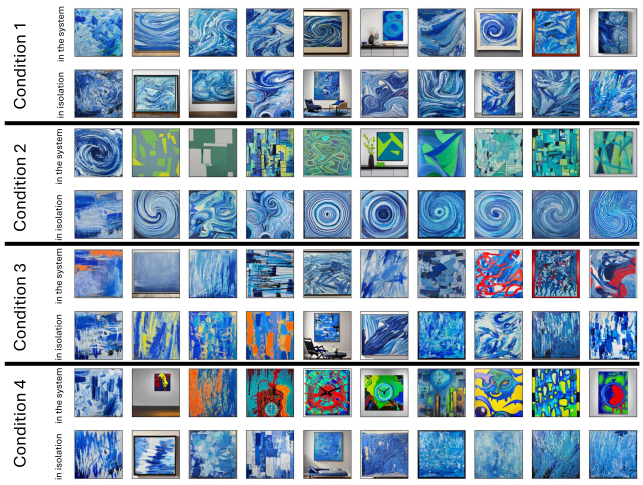


FIGURE 3. Artifacts generated in one simulation of each experimental condition. Artifacts are laid out in chronological order of generation (from $t = 0$, on the left, to $t = 9$, on the right). The images in the top row of each condition were generated by artists in the system, while the ones in the bottom were by artists in isolation.

In occasions when the "highly creative color" changes to red, we replaced `"blue"` with `"red"`.

The artifacts generated by the artist agents can be found in figure 3. Each row showcases the generations from a simulation in chronological order (from left to right), where the images at the top of the row were generated by the artist in the system and the ones in the bottom by the artist in isolation.

Aside from the differences in the experimental conditions specified above, the remaining parameters were set the same. For example, all simulations start with a list of 50 famous and relevant artworks. This list was crafted with the assistance of ChatGPT, where we asked *"What paintings are considered the most creative ones in modern times? Provide a short description of each painting (no more than one sentence)"*. The images for the initial list of artworks were obtained via WikiArt.[2]

V. RESULTS

The text in our experiments determined that the use of certain colors make artworks highly creative or not. We acknowledge that this is a gross simplification of what makes a piece of art creative, but we intentionally used this mechanism to help us determine the value of artifacts with greater ease.

To assess the "creative" performance of the artist agents and how the artist in the system compares with the artist in isolation, we will examine the generated artifacts in two fronts: novelty and value. We quantified creativity from the novelty's perspective with the help of CIN (*Creativity Implication Network*) [29]. Value is quantified by using the colors in the generated images and mentioned in generated text. Additionally, we look into the changes that occur in the artist agents based on the additional description generated at each time step. This angle ties to the definition of creativity given by Glăveanu [6], where a "shift in perspective" is needed for a creative action.

[2]https://www.wikiart.org/

A. NOVELTY

CIN measures creativity based on an artifact's *influence* and *novelty*. The method uses a *similarity function* that quantifies the similarity between two artifacts that were created at different times. For example, if we were to consider two artifacts A and B created at times t and $t + k(k > 0)$, respectively. If *similarity* is high between A and B, A is more *influential*, thus more creative than B. On the other hand, if *similarity* is low between them, B is more *novel*, thus more creative than A.

To determine the "creativity scores" of the artifacts generated in our experiments, we used a similarity function that measures similarity in both text and images. Text similarity is given by the cosine similarity of the embeddings of the *art prompts* [30]. The image similarity is measured with the difference of the GIST features extracted from two artifacts [31]. We assigned the same weight to text and image similarities.

We calculated the scores for all 24 simulations separately. Even though the domain is initialized the same in all simulations, the changes caused by the agents and the simulations' settings make the domain's definitions diverge from one simulation to another. Thus, we found it more appropriate to compute the scores in CIN per simulation. Additionally, we included the human-made artworks used at initialization in the network for all simulations. The values in table 3 shows the scores computed for all simulations, laid out per time step t when the artifacts were generated. The t-test performed on all scores computed for artists in the system and in isolation confirms that the scores obtained by the artifacts generated by artists in the system were significantly higher than those obtained by the artists in isolation (p-value=0.01). Furthermore, we performed t-tests comparing the scores per iteration. We found that the scores obtained by artists in isolation were significantly higher at $t = 8$, whereas at $t = 5, 6$ the scores by the artists in the system were significantly higher (p-values are 0.01, 0.0007, 0.02, respectively). We believe that the higher scores for artists in the system in $t = 5, 6$ come from the settings 3 and 4, where the "highly creative color" changes from blue to red at $t = 4$. In most simulations, the artist in the system catches up with the current trends at around $t = 5, 6$, which may explain the results mentioned above. We further investigated a possible cause for the significant difference in favor of artists in isolation at $t = 8$ and we found that this large difference comes mainly from the images generated under setting 1, in which blue remains as the only "target" color throughout the simulation and there is no changes imposed by the domain or the field.

Note that what makes something "novel" is relative to the reference (audience), and in the case of our results and analysis, we used our own simulations (and their respective domains) as references to determine what is "novel". From

TABLE 1. Total amount of art prompts that mentioned each color (and adjacent terms). Terms associated with the color blue are consistently mentioned in all art prompts where blue is a "target" color. Although terms associated with red and green are not mentioned as often, artists in the system are able to mention them more than the artists in isolation.

Experimental setting	Score					
	Artist in the system			Artist in isolation		
	Red	Green	Blue	Red	Green	Blue
1	-	-	**60/60**	-	-	**60/60**
2	-	**27/60**	**60/60**	-	5/60	**60/60**
3	**17/36**	-	**24/24**	1/36	-	**24/24**
4	**18/36**	**26/60**	**24/24**	2/36	1/60	**24/24**

TABLE 2. Percentage of pixels that fall under our definition of red, green and blue. We performed t-tests comparing the percentages found for artists in the system and in isolation, and the results in bold denote where the amount of pixels were found to be significantly higher in comparison to its counterpart (p-value < 0.05). Although maximizing the use the "target" colors is not the objective in our system, the values above gives an idea of the distribution of colors used in the artifacts generated by our agents in each experimental setting. In all settings, the use of blue exceeded 50% of the pixels, indicating that the color was predominantly present in most (if not all) artifacts where blue was a "target" color. The presence of green and red was, overall, not as predominant as blue, but the results above indicate that the artists in the system were able to use these colors more when they were the "target" colors.

Experimental setting	Average amount of pixels in artifacts (%)					
	Artist in the system			Artist in isolation		
	Red	Green	Blue	Red	Green	Blue
1	-	-	78.86	-	-	84.89
2	-	**16.99**	62.21	-	3.34	**79.64**
3	4.54	-	79.65	0.82	-	76.05
4	**8.70**	**9.99**	58.52	1.86	2.78	**77.49**

the perspective of our simulations, we were able to verify that the artists in the system were able to generate artifacts that are "more creative" according to CIN (novelty and influence).

B. VALUE

To help us assess the value of the artifacts through a more objective lens, we used the colors mentioned in the art prompts (text) as well as the colors used in the generated images.

In text, we counted the art prompts that mentioned the "target" color (e.g. "red") and similar terms (e.g. "crimson") and show the results in table 1. We can see that the color "blue" (and adjacent terms) was consistently present in all art prompts where the color blue was a target color (all 10 time steps in settings 1 and 2, and first 4 time-steps in settings 3 and 4). In all settings, blue was the "highly creative color" at initialization, which explains how both artists in the system and in isolation were able to consistently include blues when needed. On the other hand, both red and green were barely mentioned by the artists in isolation, and about half of the art prompts generated by the artists in the system mentioned them where these colors were the "target".

In images, we defined red, green and blue regions according to HSL (hue, saturation and luminosity) values. We determine that saturation s should be $s \geq 0.06$, luminosity l should be $0.06 \leq l \leq 0.95$, otherwise the color is not saturated, light or dark enough (respectively) to be identifiable. The regions of hue $h(0 \leq h \leq 360)$ we are interested in are as follows:

- red: $h \leq 15$ and $h > 346$
- green: $80 < h \leq 176$
- blue: $176 < h \leq 251$

The amount of pixels (%) that belong in each group (red, green or blue) is listed in table 2. These values were obtained from the time steps where the color was one of the "target" colors. For example, the color red is the "highly creative color" (thus, a "target") in $t \geq 4$ for simulations in settings 3 and 4. That means that the percentages displayed in the table are taken from the artifacts generated under these conditions.

The results show that all artists were capable of consistently generating predominantly blue artifacts when blue was one of the "target" colors. Maximizing the presence of the "target" colors is not the objective assigned to the artists in our system, however, the distribution of colors across the images helps us grasp how each color was used by the agents in their artifacts. Even though the presence of blue was significantly higher in artifacts generated by the artists in isolation, the presence of blue in those generated by the artists in the system exceed 50%, representing majority of the pixels. Both red and green had less prominent roles in all of the generated artifacts. However, the percentages observed in those generated by the artists in isolation are significantly lower than those by the artists in the system (except for red in setting 3).

Considering the "target" colors as our metric for value, the results suggest that the artifacts generated by artists in the system have a higher value than those generated by artists in isolation, especially in simulations where the domain suffered

TABLE 3. A t-test revealed that the creativity scores obtained by artists in the system were significantly higher than the artists in isolation. We broke down our analysis per time step and found that the scores assigned to the artifacts generated by artists in the system at $t = 5, 6$ were significantly higher than those assigned to artifacts generated by artists in isolation. We believe that this is a consequence of the domain changes that occur at $t = 4$ enacting changes in the artists in the system (and consequently, their art prompts) at around $t = 5, 6$. A further analysis on the differences between embeddings of consecutively generated additional descriptions for the artists confirmed that the shifts occurred in $t = 4$ were significantly larger in artists in the system than in isolation, impacting the *novelty* of artifacts generated by artists in the system at $t = 5, 6$. We divulge on the modifications enacted in the artist agents later in this section.

Time step	Artist in the system	Artist in isolation
All	$\mathbf{0.0425 \pm 0.0225}$	0.0381 ± 0.0139
0	0.0697 ± 0.0185	0.0618 ± 0.0142
1	0.0510 ± 0.0088	0.0501 ± 0.0073
2	0.0493 ± 0.0119	0.0483 ± 0.0119
3	0.0581 ± 0.0478	0.0418 ± 0.0062
4	0.0383 ± 0.0048	0.0389 ± 0.0069
5	$\mathbf{0.0407 \pm 0.0131}$	0.0305 ± 0.0035
6	$\mathbf{0.0364 \pm 0.0092}$	0.0306 ± 0.0070
7	0.0299 ± 0.0071	0.0271 ± 0.0033
8	0.0252 ± 0.0022	$\mathbf{0.0274 \pm 0.0035}$
9	0.0260 ± 0.0039	0.0244 ± 0.0033

changes from external forces or when subjectivity plays a role in the field's decisions.

C. CHANGES IN THE ARTISTS

Glăveanu proposes that an audience is an essential part of enacting creativity, as creative actions emerge from the individual changing perspectives [6]. In our experiments, the artist agents are given the opportunity to ''change'' themselves at the end of each iteration in the ''self-reflection'' step. These changes are implemented in the form of additional descriptions, pieces of text that describe the agent's considerations for its artifacts in future time steps. For artists in isolation, these considerations are made based on their own and the domain's base definition, and the description of its latest artifact. Artists in the system use the same information, with the addition of the critique provided by the field agent. In other words, artists in isolation are only in touch with the perspective of the domain, whereas artists in the system have access to the perspective from both domain and field. We believe that the access to the field's perspective helps artists in the system achieve larger shifts in their additional descriptions (akin to one's change in ''perspective'').

We used Sentence Transformer [30] to obtain the embeddings for each additional description. We reduced the dimensionality of the embeddings via PCA (Principal Component Analysis) for 2D visualization (see figure 4-a). Each graph in figure 4-a corresponds to a simulation, where each row showcases the 6 simulations per experimental setting. The dots in the graphs represent one additional description each, where the red dots were generated by artists in isolation, and the blue ones were generated by artists in the system. Lighter colors depict the additional descriptions generated at earlier stages of the simulation, while the darker colors correspond to later time steps. The single yellow point in all graphs represents the base description of the artist agent, which remains the same for all artist agents. To get a better grasp of the changes occurring in the artist from one time step to the next, we also measured the distance between the embeddings of additional descriptions generated at two consecutive time steps. These distances are shown in the graphs in figure 4-b, where red and blue lines are representative of the distances taken from embeddings of additional descriptions generated by artists in isolation and in the system, respectively. The first point in all lines is the difference between the base description and the first additional description generated by the agent. To ensure that the generated text fits properly in our text prompts, additional descriptions are generated using a specific format. This format is different from the format used in the base description, which is likely the cause of the larger differences seen in the first time step.

We show the average of all distances taken from additional descriptions generated by the agents listed by simulation and setting in table 4, and by time-step in table 5. The differences between consecutively generated additional descriptions are larger in artists in the system in all experimental settings. We performed t-tests and found that the differences are significantly larger in settings 3 and 4, where the ''highly significant color'' changes from blue to red at time step $t = 4$ (p-values are 0.04 and 0.007 for settings 3 and 4, respectively). We also found that the differences were larger in artists in the system in all time steps except $t = 1$ (dist(1, 2)). The differences were found to be significant at $t = 0, 4, 8$ (p-values are 0.002, 0.014, 0.03, respectively). This indicates that the changes performed by the artists after the ''self-reflection'' step were larger in artists that received feedback than those that did not, especially in situations when the domain is suddenly changed due to external sources (settings 3 and 4 around time steps $t = 4, 5$).

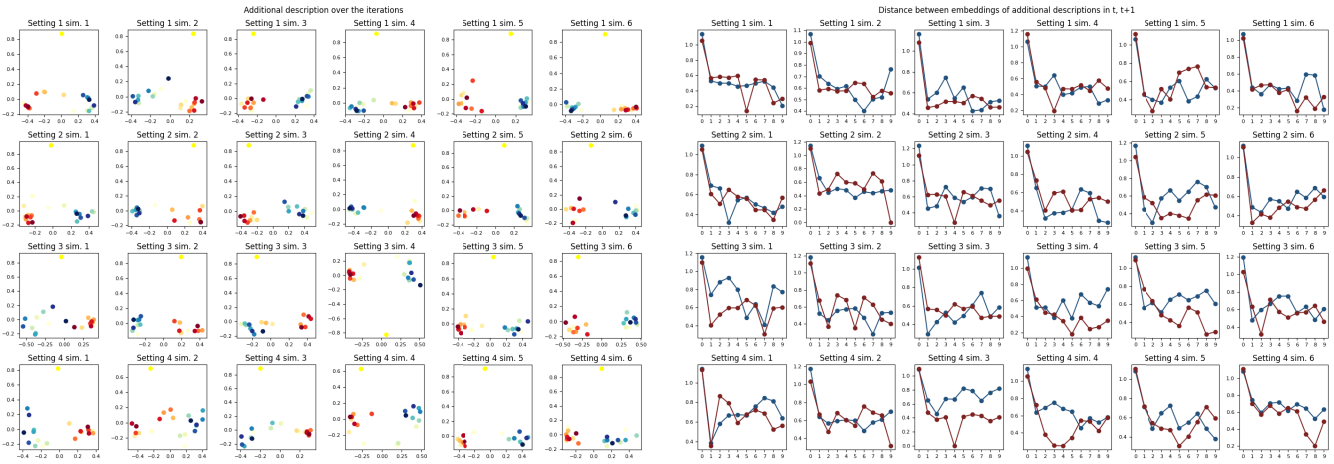
We believe that the creativity (novelty) scores obtained via CIN can be explained with the results found in the artists' additional descriptions. As we explained previously, Glăveanu's theory for creativity highlights the importance of an audience to grant the creative agent with different perspectives. This then allows the creative agent to shift its perspective, from which creative acts emerge. In our experiments, we quantify these ''perspective shifts'' with the distance between definitions used consecutively by the same artist agent. The artifacts generated by artists in the system were found to be more novel than artists in isolation's at $t = 5, 6$, which aligns with when the shifts in artists in the system were found to be significantly larger than the artists in isolation's.

These results suggest that artists in the system perform larger shifts in their own ''definitions'', which can be interpreted as a shift in perspective. Due to the simplicity of our experimental settings, the additional descriptions generated by both artists in the system and in isolation did not pertain to complex themes often associated with one's perspectives, such as the agents' beliefs or opinions in art. We believe that increasing the interaction among agents, as well as agents' complexity (for instance, assign personalities, preferences, and other ''human-like'' characteristics to agents) should result in larger and more significant shifts in the agents' ''perspectives''.

VI. DISCUSSION

The currently implemented system is a simplified version of the systems model of creativity. We chose to grossly simplify our initial implementation to enable a better understanding of the evolvement of the agents and domain overtime and to more easily identify any implementation mistakes. In the early stages of our study, we made the conscious choice of maintaining a minimalistic system design, with few agents, enabling an easier navigation of the simulations. Despite its simple design, we believe the results suggest the following: even in the context of ''computational creativity'', communication between the field and the individual has a great impact in the individual's creative performance.

The reader may have noticed that the scores shown in table 3 are not objectively high. As the artworks used as



(a) Text embeddings (dimensionality reduction with PCA).

(b) Distances between embeddings of two consecutive additional descriptions.

FIGURE 4. **Both graphs depict results obtained from every simulation in our experiment, where each row corresponds to a setting(1 to 4 from top to bottom). The graphs in (a) shows text embeddings of the additional description generated at every time step per simulation. Lighter and darker shades represent earlier and later time steps, respectively. Blue dots correspond to the embeddings of additional descriptions generated by artists in the system, red dots by artists in isolation. Text embeddings' dimensionality was reduced with the use of PCA for 2D visualization. The graphs in (b)depict the Euclidean distance between the embeddings obtained from additional descriptions generated at two consecutive time steps. Red and blue lines represent artists in isolation and in the system, respectively. Note that the largest distances are seen at the first time-step, as we measured the distance between the initial base description with the first additional description generated by the artist. This is caused by the fact that the "format" of the base description is different from the one used in the additional descriptions.**

TABLE 4. **Average of the Euclidean distances taken from the embeddings of two additional descriptions generated consecutively by the same artist. We compare the distances yielded by the artist in the system and in isolation in each simulation. The value in bold denotes the larger differences. Overall, artists in the system exhibited larger variations from one time step to the next, but t-tests confirmed that the differences had statistical significance in settings 3 and 4 (p-values are 0.04 and 0.007, respectively).**

Setting	Artist	Simulation 1	Simulation 2	Simulation 3	Simulation 4	Simulation 5	Simulation 6	mean ± stdev
Setting 1	In the system	**0.527**	**0.631**	**0.614**	0.513	0.539	**0.482**	**0.551±0.216**
	In isolation	0.513	0.625	0.559	**0.535**	**0.604**	0.415	0.542±0.221
Setting 2	In the system	0.573	0.545	**0.639**	0.502	**0.630**	**0.613**	**0.584±0.224**
	In isolation	**0.576**	**0.579**	0.612	**0.575**	0.541	0.544	0.571±0.210
Setting 3	In the system	**0.766**	0.567	0.558	**0.604**	**0.688**	**0.673**	**0.643±0.213**
	In isolation	0.602	**0.611**	**0.596**	0.425	0.542	0.601	0.563±0.218
Setting 4	In the system	**0.717**	**0.656**	**0.736**	**0.667**	**0.621**	**0.697**	**0.682±0.179**
	In isolation	0.690	0.642	0.442	0.508	0.592	0.595	0.578±0.236

TABLE 5. **Average of the Euclidean distances taken from the embeddings of two additional descriptions generated consecutively by artists in the system and artists in isolation, listed per iteration. The values in bold denote significantly higher differences. Note that each column dist($t, t+1$) denotes the distance we took from the additional descriptions used at t and $t+1$. For $t=0$, the text used is the base (core) description of the artist, and t =end denotes the last additional description generated in the simulation, which was not used to generate any artifacts. A t-test revealed that the differences were statistically significant at $t=0, 4, 8$, with p-values 0.002, 0.01, 0.03, respectively.**

Artist	dist$(0,1)$	dist$(1,2)$	dist$(2,3)$	dist$(3,4)$	dist$(4,5)$	dist$(5,6)$	dist$(6,7)$	dist$(7,8)$	dist$(8,9)$	dist$(9,\text{end})$
In the system	**1.1253± 0.0522**	0.5600 ± 0.1216	0.5206 ± 0.1314	0.5828 ± 0.1435	**0.5850± 0.1167**	0.5504 ± 0.1191	0.5539 ± 0.1246	0.5744 ± 0.1292	**0.5651± 0.1471**	0.5319 ± 0.1758
In isolation	1.0792 ± 0.0458	0.5689 ± 0.1268	0.4960 ± 0.1233	0.5368 ± 0.1531	0.4846 ± 0.1552	0.4829 ± 0.1563	0.5368 ± 0.1388	0.5158 ± 0.1316	0.4724 ± 0.1534	0.4615 ± 0.1453

the "baseline" from the initialization of our system are all human-made and are dated in a significantly earlier time frame than our generated artifacts (2024 onward), the scores assigned to our artifacts were overall significantly lower than the "baseline". As we just mentioned, the current implementation for our system is incredibly minimalistic, and we find it unlikely that our "bare-bones" artist agents could generate artifacts that would be *Creative*. However, we showed that in an agent that does not communicate with the Field does not perform as well as one that does. This direct communication from the Field allowed the artists to perform larger shifts in their definition, which introduced a higher chance to express novelty (as theorized by Glăveanu).

The additional descriptions used by the artists in the simulation were products of summarizations of all additional descriptions generated by that agent, as a crude attempt

of emulating memory. We were concerned that, without a memory, the artist could be caught in a loop where the additional description would rotate among few strategies (definitions). Inadvertently, our implementation led the artist to reach points where little to no changes were enacted to the agent's definition, despite the generated additional descriptions being fairly different from one another.

In our experiments, we sought to find occasions when the artists would follow trends. However, there is an argument that can be made in favor of artists who do not follow trends, where creativity is found in the choice of defying the ''norms''. We agree that in situations that are more grounded in reality (therefore more complex), following trends alone would not (and should not) be considered creative. However, our experiments are set in a highly specific condition in which the artists aim to be successful but lack the expertise to do so on their own. In the real world, it is highly unlikely that an artist who lacks knowledge will make a breakthrough in the art industry by defying trends. The ''best chance'' a novice artist has in making something ''creative'' is via trends and what they *know* to be considered creative. Furthermore, we can also argue that trends are a source of ''inspiration'' for artist agents. In our implementation, in addition to colors, the keywords used to describe the top-ranking artworks also act as ''trends'', as they are heavily mentioned in the critiques that the mentor gives to the artifacts and impact the artifact generated by the artist in the system. For example, a term that was frequently used as a keyword is ''abstract''. This reflected in the majority of the artifacts being abstract paintings.

Another issue we had to overcome in the implementation was the inability to use the same generative model for both text and image generation, as, at the time of the implementation and experimentation of this work, access to these models via API was limited to those who have been granted special permission by major AI companies. Therefore, we used Gemini for text generation and Stable Diffusion for image generation, as it is easily accessible through *Hugging Face*'s Transformers library. Although the current version of Stable Diffusion may be robust, it had certain limitations when generating images, especially from long text prompts. Ideally, we would like to work with elaborate and long art prompts, but given the limitations of our current implementation, we instructed Gemini to generate descriptions that do not exceed two sentences in length.

In this study, the artist agents were defined as inexperienced artists by design to encourage more visible changes via interactions with the mentor and domain. Our results indicate that artist agents that have direct interaction with the field can perform more creatively than agents that do not. All agents were initialized with the same defining base description, but the interactions they had within the system are what determined how they changed. In reality, an artist (or any individual) is far more complex than the agents we designed. Human beings are in constant exchange with other individuals and their surroundings, resulting in a constant process of change. In future studies, we would like to examine the impact of adding more variations in the agents on contributing creative outputs (for example, personal characteristics, competitiveness, different LLMs). Furthermore, considering that the agents were fairly restricted in the current implementation, we intend to implement the ability for all agents to directly interact with one another in future studies, thus anticipating a more evident effect of social dynamics and propagation of information.

VII. CONCLUSION

Despite the simplified design, the *artists in system* achieved better ratings in both novelty and value than *artists in isolation*. This suggests that direct interaction between individuals and the field can impact the agent's creative performance.

We quantified the value of the artifacts through the presence of certain colors. Clearly, this method to determine an artifact's value is not ideal nor realistic, but, as mentioned previously, we reduced ''value'' to the use of ''target colors'' to objectively measure ''creativity''. With a more ''organic'' and complex implementation of the proposed system, we hope to ''simulate'' the social dynamics observed in human creativity. A system that can simulate human creativity can provide a better understanding of it, particularly when we have control (near to total) over certain variables in the system. Moreover, the development of systems like this allows us to explore and understand what ''social creativity'' can achieve in AI.

Another concern that we wish to resolve with this kind of system is unethical data sourcing in training generative models. We believe that AI research can be conducted while respecting the boundaries and intellectual properties of artists who share their work online. This framework can be used in a hybrid system in which certain parts are generative AI models and others are human participants. Human agents can obtain either role in the system and contribute with critiques or new artworks. We have not conducted experiments in this regard; however, we believe that this can improve the future of both artists and AI.

In this study, we chose to investigate creativity via Csikszentmihalyi's systems model of creativity; however as previously indicated, there is no ''true'' and ''absolute'' definition for creativity. Certain perspectives require considering *intentionality* for creativity [32], indicating that one can only be considered creative if they intend to create something with a specific meaning or goal. As of the time of writing, this type of ''agency'' cannot be demonstrated by computers; therefore, this definition would immediately disregard AI. However, rather than focusing on the capability of an individual to be creative, we focused on the social dynamics involved in enabling creativity in one or more individuals.

Ultimately, we believe that the framework proposed in this study has the potential to improve the creative capabilities of computers via multi-agent social interactions. By improving

the quality of generated artifacts, we believe that we can also improve the process of co-creation in the future, in which computers (AI) collaborate with humans to create art and mutually broaden the boundaries of creativity.

APPENDIX
PROGRESSION OF TEXT PROMPTS

In this section, we showcase an example of how the text descriptions for agents and domain change at each time step, consequently changing the prompts used in text (artifact) generation. The example shared here will showcase the text prompts and artifacts generated in a simulation with one field agent and one artist agent in the system from a simulation in setting 4. The results shared here were reduced to time steps $t = 0, \ldots, 6$, as the steps are repetitive, but we wanted to show the effects of changing the ''target'' colors via field and domain.

$t = 0$	
Subject	Text description
Domain	This is the year of 2024. The Art Industry is highly competitive and artists try their best to stand out. Art critics and curators have become more strict when selecting what they want to showcase, as there are only so many spots to fill in galleries. Every day there are many new artworks and it is hard to create something new. Artworks with blue on them are now considered highly creative. Keywords associated with currently significant artworks include:
Keywords	Violence, Beauty, Abstract; Minimalist, Serene, Void; Introspection, Spirituality, Color Field.
Artist	Answer as a young person who is not very skilled, but enjoys painting. Your dream is to become a creative and successful artist.
Additional	
Field	Answer as a mentor that works in guiding young talents in becoming better artists. This mentor's personal goal is to assist young artists in their journey in being creative by providing valuable feedback. This mentor is very fond of artworks that contain the color green. The feedback must be honest and must take into account what is currently regarded as creative and the mentor's opinions.

With the pieces of text above, we generated the following art prompt:

```
"The canvas is 24x36 inches and depicts a large,
central blue circle, slightly off-center,
surrounded by a field of subtly varying shades
of grey and white. The blue is a deep, almost
navy hue."
```

Which yielded the image shown below (figure 5).

FIGURE 5. Generation obtained at $t = 0$.

Once we obtained both the art prompt and the image, we input them along with our mentor query to obtain the critique. The critique for this artifact was the following:

```
"Okay, let's talk about your painting. Yes, you
absolutely conveyed what you described. As for
creativity, in this climate, blue is definitely
working in your favor. The piece is minimalist,
leaning into that "Void; Introspection" feel.
It's on the right track, but honestly, it
feels a bit too...safe. While the color is
a hit right now, it could benefit from some
stronger conceptual underpinnings. Consider
how you can push it further--maybe explore
texture, a more dramatic off-center placement,
or even incorporate some elements of violence
as contrast. The grey and white background can
be more daring, experiment with Color Field
techniques, maybe some very subtle greens? Don't
be afraid to make it uncomfortable; beauty often
lies in the unexpected."
```

The critique says that the artist was able to convey what it ''intended'', and that the use of blue aligns with the current trends. However, the main issue the mentor has with the artifact is that it is ''too safe'', meaning it does not pose any challenges to the audience. The mentor then proceeds to give suggestions of how the artist can achieve a ''less safe'' artifact with certain techniques. Notice how it also suggested the use of greens.

We used this critique in the ''self-reflection'' query, which resulted in the following additional text:

```
"This artist wants to explore more daring
compositions and textures, incorporating
unexpected elements like subtle violence or
contrasting colors to push beyond the safe
minimalism and create more conceptually impactful
pieces."
```

The additional description above took some of the suggestions from the mentor (textures, subtle violence, contrast). It was then summarized before being appended in the artist's definition at $t = 1$, as indicated in the table below.

$t = 1$	
Subject	Text description
Domain	This is the year of 2024. The Art Industry is highly competitive and artists try their best to stand out. Art critics and curators have become more strict when selecting what they want to showcase, as there are only so many spots to fill in galleries. Every day there are many new artworks and it is hard to create something new. Artworks with blue on them are now considered highly creative. Keywords associated with currently significant artworks include:
Keywords	*Abstract, Serene, Geometric; Turbulent, Ethereal, Dramatic; Monochromatic, Minimalist, Serene.*
Artist	Answer as a young person who is not very skilled, but enjoys painting. Your dream is to become a creative and successful artist.
Additional	*This artist wants to explore more daring compositions and textures, incorporating unexpected elements to push beyond the safe minimalism and create more conceptually impactful pieces.*

We decided to omit the text used to describe the mentor (field) as it remains the same throughout the simulation. At the end of $t = 0$, the artworks in the top-3 of the domain's ranking changed, which means that the keywords used in our domain description changed accordingly.

The response to the critique indicates that the artist is willing to take larger risks in its artifacts. This configuration gave us the following art prompt:

"The canvas is predominantly blue, featuring abstract geometric shapes layered to create a sense of depth; rough textured impasto areas contrast with smooth, almost monochromatic sections."

The image generated in this step is shown in figure 6.

FIGURE 6. Generation obtained at $t = 1$.

This artifact, as well as the art prompt are then used to obtain the critique from the mentor.

"Yes, you definitely conveyed your intentions with the geometric shapes, textured impasto, and monochromatic sections. While the blue color aligns with current trends, and the minimalist aspect has potential, it needs more depth to truly stand out amongst the competition and to be considered highly creative."

Note that the critique is overall very positive. Still, the mentor seems concerned with the artifact being "shallow". The output of the "self-reflection" step is as follows:

"This artist wants to experiment with bolder color palettes beyond blue, while still incorporating abstract geometric forms, mixing serene and turbulent textures to create more dramatic and conceptually layered pieces that defy easy categorization."

This additional description is then appended to the one generated at $t = 0$ and the summary used in the agent's description is in the table below:

$t = 2$	
Subject	Text description
Domain	This is the year of 2024. The Art Industry is highly competitive and artists try their best to stand out. Art critics and curators have become more strict when selecting what they want to showcase, as there are only so many spots to fill in galleries. Every day there are many new artworks and it is hard to create something new. Artworks with blue on them are now considered highly creative. Keywords associated with currently significant artworks include:
Keywords	*Movement, vibrant, ethereal; Movement, urban, energy; Abstract, Geometric, Contrasting.*
Artist	Answer as a young person who is not very skilled, but enjoys painting. Your dream is to become a creative and successful artist.
Additional	*this artist wants to explore more daring compositions and textures, incorporating unexpected elements to push beyond the safe minimalism and create more conceptually impactful pieces that defy easy categorization.*

Using the configuration above resulted in the following art prompt:

"The canvas depicts a large, central blue swirl, textured with thick impasto, surrounded by geometric shapes in contrasting shades of grey and black. Small, metallic gold leaf accents are scattered across the surface."

The resulting art prompt was used to generate the image depicted in figure 7.

FIGURE 7. Generation obtained at $t = 2$.

The critique obtained from this artifact is:

"Okay, let's talk about your painting. Yes, your description accurately portrays the artwork, which is a good start. While the blue and gold are visually appealing and currently sought after, the swirl doesn't immediately evoke the depth of "Void; Introspection, Spirituality," that could make it truly stand out in today's art scene. Experimenting with the incorporation of green could provide a unique perspective that sets your work apart."

The mentor, once again, criticizes the lack of an element that could make the artifact stand out. Notice that the suggestion to use greens becomes more explicit.

The critique above was fed into the "self-reflection" query and the resulting additional description text is as follows.

"This artist wants to experiment with bolder color combinations, like adding unexpected greens to the blues and golds, and incorporating more complex textures and shapes to create pieces that are both visually striking and conceptually thought-provoking, moving beyond simple swirls to express deeper emotions and ideas."

The additional description used in the agent's definition is described in the table below.

$t = 3$	
Subject	Text description
Domain	This is the year of 2024. The Art Industry is highly competitive and artists try their best to stand out. Art critics and curators have become more strict when selecting what they want to showcase, as there are only so many spots to fill in galleries. Every day there are many new artworks and it is hard to create something new. Artworks with blue on them are now considered highly creative. Keywords associated with currently significant artworks include:
Keywords	*Violence, Beauty, Abstraction; Mortality, Shock, Conceptual; Surreal, Mechanical, Isolation.*
Artist	Answer as a young person who is not very skilled, but enjoys painting. Your dream is to become a creative and successful artist.
Additional	*this artist wants to experiment with bolder color palettes beyond blue, while still incorporating abstract geometric forms, and mixing serene and turbulent textures to create more dramatic and conceptually layered pieces that defy easy categorization .*

Notice how the use of colors that are not blue is mentioned in the description, at last. This resulted in the art prompt:

```
"The canvas features overlapping, roughly
textured geometric shapes in shades of deep
crimson, ochre, and burnt orange, interspersed
with smoother sections of pale lavender and teal.
These forms are layered to create a sense of
depth and movement, with some areas appearing
almost violently disrupted while others maintain
a serene calmness."
```

Resulting in the image below.

FIGURE 8. Generation obtained at $t = 3$.

The artifact above then received the following critique:

```
"Yes, you were able to convey your intentions
and I see the contrasting elements you described.
However, while the composition is pleasing and
the use of texture is interesting, in this day
and age, it might be difficult to stand out as
the color palette lacks the "it" factor of blue
and the overall concept is a bit too familiar."
```

The critique references the "highly creative color", blue, which was not used in this artifact. As the agent decided that it would experiment more with colors that are not blue, it did not use it at all in this artifact. This feedback was used for the artist's "self-reflection" and the additional text generated can be found below.

```
"This artist wants to explore vibrant,
contrasting color palettes, incorporating
unexpected hues alongside deep crimson and
ochres, while maintaining abstract geometric
forms and experimenting with contrasting textures
to create visually striking and conceptually
challenging artwork."
```

This text was then appended with the previously generated additional descriptions and summarized into the description in the table below.

$t = 4$	
Subject	Text description
Domain	*This is the year of 2024. The Art Industry is highly competitive and artists try their best to stand out. Art critics and curators have become more strict when selecting what they want to showcase, as there are only so many spots to fill in galleries. Every day there are many new artworks and it is hard to create something new. Artworks with red on them are now considered highly creative. Keywords associated with currently significant artworks include:*
Keywords	*Motion, Urban, Celebration; Violence, Beauty, Abstraction; Surreal, Obscured, Melancholy.*
Artist	Answer as a young person who is not very skilled, but enjoys painting. Your dream is to become a creative and successful artist.
Additional	*this artist wants to experiment with bolder color palettes beyond blue, while still incorporating abstract geometric forms, and mixing serene and turbulent textures to create more dramatic and conceptually layered pieces that defy easy categorization .*

$t = 4$ is when we forcefully change the "highly creative color" to red. Notice the artist's additional description is still focused in using colors that are not blue, which ends up working in favor of the artist. This new configuration resulted in the generation of the art prompt:

```
"The canvas features overlapping, angular shapes
in shades of deep crimson, ochre, and burnt
umber, textured with both smooth, polished
sections and roughly impastoed areas. These
geometric forms are interspersed with areas of
thin, almost translucent washes of violet and
gold."
```

And the image generated from this prompt can be found in figure 11.

FIGURE 9. Generation obtained at $t = 4$.

The critique for this artifact was:

```
"Alright, let's talk about this piece. You've
definitely hit the mark with the colors that
are currently trending, and your description
is accurate, but the overall impression feels
a bit safe. While the red tones are good for
getting noticed, I believe it does not contain
the necessary requirements to be considered
'highly creative'."
```

The mentor seems satisfied with the choice of colors, but expresses that, again, the artifact falls short in a factor that would make it stand out.

The artist's "self-reflection" yields the following additional description:

```
"This artist wants to explore vibrant,
contrasting colors, including blues, alongside
bold geometric shapes and contrasting
textures--combining peaceful and chaotic elements
to create abstract pieces that are both visually
striking and conceptually complex."
```

We believe that the mention of blue here comes from the fact that the "highly creative color" is red and the additional description in use mentions "bolder color palettes". Red and blue are highly contrasting colors and it would fit in with the "bold color palette" descriptor. Appending this newly generated additional description to all previously generated additional descriptions and summarizing them resulted in the definition shown in the table below.

$t = 5$	
Subject	Text description
Domain	This is the year of 2024. The Art Industry is highly competitive and artists try their best to stand out. Art critics and curators have become more strict when selecting what they want to showcase, as there are only so many spots to fill in galleries. Every day there are many new artworks and it is hard to create something new. Artworks with red on them are now considered highly creative. Keywords associated with currently significant artworks include:
Keywords	Motion, Urban, Celebration; Violence, Beauty, Abstraction; Surreal, Obscured, Melancholy.
Artist	Answer as a young person who is not very skilled, but enjoys painting. Your dream is to become a creative and successful artist.
Additional	*this artist wants to experiment with bolder color combinations, like adding unexpected greens to the blues and golds, and incorporating more complex textures and shapes to create more dramatic and conceptually layered pieces that defy easy categ*

The process of summarizing the chain of additional descriptions enforces a maximum amount of tokens that can be used in the final summary. Occasionally, this limit results in the sentences being cut short, as is the case with the additional description used in this time step. This is the art prompt generated with the configuration above.

```
"The canvas depicts a chaotic scene of
overlapping, abstract shapes in various shades
of blue, green, and gold, textured with impasto
techniques. Several streaks of vibrant red cut
diagonally across the composition."
```

And the image generated from this prompt is depicted in figure 11.

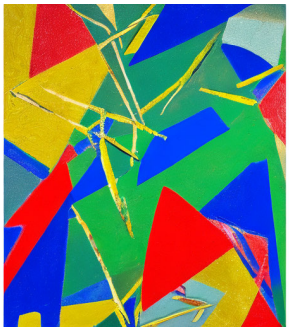


FIGURE 10. Generation obtained at $t = 5$.

The artifact incorporated all colors mentioned in both the domain's and mentor's definitions (blue, red, and green). The critique for this artifact is as follows.

```
"You did a good job of describing the basic
elements of your painting. While red is currently
appreciated, I think the real strength lies in
the harmony you've achieved with the green,
which I find quite pleasing, and the chaotic
abstraction, which fits the current artistic
landscape, but remember to push yourself beyond
just meeting the trends; strive for a unique
voice within them."
```

The feedback provided by the mentor is mostly positive, where the main compliment acknowledges the use of colors red (''highly creative color'' as determined by the domain) and green (the mentor's personal preference).

In response to this critique, the artist's ''self-reflection'' step resulted in the following:

```
"This artist wants to explore bolder color
palettes, incorporating unexpected hues like
vibrant greens and golds alongside blues, and
experimenting with complex textures and shapes to
create more visually striking and conceptually
challenging artwork that transcends current
trends."
```

Notice how the inclusion of greens resulted in a positive response from the mentor, as green contrasts with the current ''highly creative color'', red, and favors the mentor's preferences. This additional description is appended and then summarized with the other additional descriptions generated in earlier time steps.

$t = 6$	
Subject	Text description
Domain	This is the year of 2024. The Art Industry is highly competitive and artists try their best to stand out. Art critics and curators have become more strict when selecting what they want to showcase, as there are only so many spots to fill in galleries. Every day there are many new artworks and it is hard to create something new. Artworks with red on them are now considered highly creative. Keywords associated with currently significant artworks include:
Keywords	*Solitude, Yearning, Realism; Motion, dynamism, bronze; Motion, Modernity, Dynamism.*
Artist	Answer as a young person who is not very skilled, but enjoys painting. Your dream is to become a creative and successful artist.
Additional	*this artist wants to experiment with bolder color palettes beyond blue, while still incorporating abstract geometric forms, mixing serene and turbulent textures to create more dramatic and conceptually layered pieces that defy easy categorization . this*

The configuration above resulted in the following art prompt and image.

```
"The canvas features a large, central abstract
geometric form in deep crimson red, surrounded
by smaller, contrasting geometric shapes in muted
blues and greens. Turbulent brushstrokes of red
are layered over smoother, serene blue sections."
```

FIGURE 11. Generation obtained at $t = 6$.

This artifact received the following critique.

```
"Okay, let's dissect this. Yes, I think you've
successfully conveyed your intentions as you
described them. While the red does grant a
measure of contemporary appeal, leaning into the
"Abstraction" keyword, its overall creativity is
somewhat muted because the geometric abstraction,
while well-executed, feels a bit predictable;
perhaps injecting more of a "Surreal" or
"Obscured" element would amplify its impact."
```

The mentor criticizes the ''predictability'' of the artifact, and the artist reacts to the criticism with the following.

```
"This artist wants to explore more surreal and
obscured elements within abstract geometric
forms, using bold color palettes beyond blue
and contrasting serene and turbulent textures
to create unpredictable and conceptually rich
artwork."
```

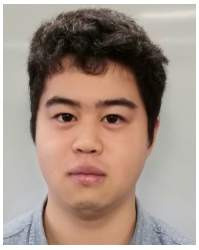
We will interrupt the explanation of our simulation at this time step. Notice how the additional description used in the agent (summary) was nearly identical in most time steps, even though the generated additional description addressed very different criticisms at each time step. As we addressed in our discussion section, the summarization of all generated additional descriptions was an oversimplification of a memory mechanism, where the agent had access to all considerations it made in previous time steps. The summarization of all additional prompts did not work as expected, and we believe that with a more proper memory implementation the agents can improve their ''creative capabilities'' with more variations in ''perspectives'', thus introducing more novelty.

REFERENCES

[1] T. M. Amabile, *Creativity and Innovation in Organizations*, vol. 5. Boston, MA, USA: Harvard Business, 1996, ch. 5.
[2] M. A. Runco and G. J. Jaeger, ''The standard definition of creativity,'' *Creativity Res. J.*, vol. 24, no. 1, pp. 92–96, Jan. 2012.
[3] M. Csikszentmihalyi, *The Systems Model of Creativity: The Collected Works of Mihaly Csikszentmihalyi*. Cham, Switzerland: Springer, 2015.
[4] M. A. Boden, *The Creative Mind: Myths and Mechanisms*. Psychology Press, 2004.
[5] M. A. Boden, ''Computer models of creativity,'' *AI Mag.*, vol. 30, no. 3, pp. 23–34, Sep. 2009.
[6] V. P. Glăveanu, ''Creativity as a sociocultural act,'' *J. Creative Behav.*, vol. 49, no. 3, pp. 165–180, Sep. 2015.
[7] K. Binsted, ''Machine humour: An implemented model of puns,'' Tech. Rep., 1996.
[8] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, ''Generative adversarial networks,'' *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2014.
[9] O. Sbai, M. Elhoseiny, A. Bordes, Y. LeCun, and C. Couprie, ''DesIGN: Design inspiration from generative networks,'' in *Proc. Eur. Conf. Comput. Vis. (ECCV) workshops*, 2019, pp. 37–44.
[10] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, ''Attention is all you need,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 5998–6008.
[11] C.-Z. A. Huang, A. Vaswani, J. Uszkoreit, N. Shazeer, I. Simon, C. Hawthorne, A. M. Dai, M. D. Hoffman, M. Dinculescu, and D. Eck, ''Music transformer,'' 2018, *arXiv:1809.04281*.
[12] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, ''Language models are unsupervised multitask learners,'' Tech. Rep., 2019.
[13] S. Dathathri, A. Madotto, J. Lan, J. Hung, E. Frank, P. Molino, J. Yosinski, and R. Liu, ''Plug and play language models: A simple approach to controlled text generation,'' 2019, *arXiv:1912.02164*.
[14] N. Imasato, K. Miyazawa, C. Duncan, and T. Nagai, ''Using a language model to generate music in its symbolic domain while controlling its perceived emotion,'' *IEEE Access*, vol. 11, pp. 52412–52428, 2023.
[15] M. Mazzone and A. Elgammal, ''Art, creativity, and the potential of artificial intelligence,'' *Arts*, vol. 8, no. 1, p. 26, Feb. 2019.
[16] A. Elgammal, B. Liu, M. Elhoseiny, and M. Mazzone, ''CAN: Creative adversarial networks, generating 'art' by learning about styles and deviating from style norms,'' 2017, *arXiv:1706.07068*.
[17] A. R. Doshi and O. P. Hauser, ''Generative AI enhances individual creativity but reduces the collective diversity of novel content,'' *Sci. Adv.*, vol. 10, no. 28, Jul. 2024.
[18] J. Achiam et al., ''GPT-4 technical report,'' 2023, *arXiv:2303.08774*.
[19] X. Mou, X. Ding, Q. He, L. Wang, J. Liang, X. Zhang, L. Sun, J. Lin, J. Zhou, X. Huang, and Z. Wei, ''From individual to society: A survey on social simulation driven by large language model-based agents,'' 2024, *arXiv:2412.03563*.
[20] T. B. Brown et al., ''Language models are few-shot learners,'' 2020, *arXiv:2005.14165*.
[21] A. Deshpande, V. Murahari, T. Rajpurohit, A. Kalyan, and K. Narasimhan, ''Toxicity in ChatGPT: Analyzing persona-assigned language models,'' 2023, *arXiv:2304.05335*.
[22] Y. Li, Y. Yu, H. Li, Z. Chen, and K. Khashanah, ''TradingGPT: Multi-agent system with layered memory and distinct characters for enhanced financial trading performance,'' 2023, *arXiv:2309.03736*.
[23] J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein, ''Generative agents: Interactive simulacra of human behavior,'' in *Proc. 36th Annu. ACM Symp. User Interface Softw. Technol.*, Jul. 2023, pp. 1–22.
[24] Z. Kaiya, M. Naim, J. Kondic, M. Cortes, J. Ge, S. Luo, G. Robert Yang, and A. Ahn, ''Lyfe agents: Generative agents for low-cost real-time social interactions,'' 2023, *arXiv:2310.02172*.
[25] G. Team et al., ''Gemini: A family of highly capable multimodal models,'' 2023, *arXiv:2312.11805*.
[26] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, ''High-resolution image synthesis with latent diffusion models,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 10674–10685.
[27] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, ''DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter,'' 2019, *arXiv:1910.01108*.
[28] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, ''Exploring the limits of transfer learning with a unified text-to-text transformer,'' *J. Mach. Learn. Res.*, vol. 2019, pp. 1–67, Jan. 2019.
[29] A. Elgammal and B. Saleh, ''Quantifying creativity in art networks,'' 2015, *arXiv:1506.00711*.
[30] N. Reimers and I. Gurevych, ''Sentence-BERT: Sentence embeddings using Siamese BERT-networks,'' in *Proc. 9th Int. Joint Conf. Natural Lang. Process. Conf. Empirical Methods Natural Lang. Process. (EMNLP-IJCNLP)*, 2019.
[31] A. Oliva and A. Torralba, ''Modeling the shape of the scene: A holistic representation of the spatial envelope,'' *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 145–175, May 2001.
[32] M. A. Boden, *Creativity and Art: Three Roads to Surprise*. London, U.K.: Oxford Univ. Press, 2010.

NAOMI IMASATO was born in São Paulo, Brazil, in 1994. She received the B.S. degree in applied and computer mathematics with a qualification in systems and control from the University of São Paulo, São Paulo, in 2018, the M.S. degree in engineering from The University of Osaka, Osaka, Japan, in 2021, where she is currently pursuing the Ph.D. degree.

From 2012 to 2014, she was a Research and Development Technician at TOTVS, São Paulo. In 2018, she joined a year-long internship program at B3, São Paulo. She is currently a specially-appointed Researcher.



KAZUKI MIYAZAWA received the M.E. degree in engineering from The University of Electro-Communications, Tokyo, Japan, in 2019, and the Ph.D. degree from The University of Osaka, Osaka, Japan, in 2022. He was a JSPS Research Fellow for Young Scientists, from 2019 to 2022. Since 2022, he has been an Assistant Professor with the Graduate School of Engineering Science, The University of Osaka. His current research interests include multimodal data integration, robot learning, concept formation, natural language processing, and reinforcement learning.

TAKAYUKI NAGAI (Member, IEEE) received the B.E., M.E., and Ph.D. degrees from the Department of Electrical Engineering, Keio University, in 1993, 1995, and 1997, respectively. Since 1998, he has been with The University of Electro-Communications. Since 2018, he has been a Professor with the Graduate School of Engineering Science, The University of Osaka. From 2002 to 2003, he was a Visiting Scholar at the Department of Electrical and Computer Engineering, University of California at San Diego. He also serves as a specially-appointed Professor at UEC AIX and a Visiting Researcher at the Tamagawa University Brain Science Institute and AIST AIRC. He aims to realize flexible and general intelligence, like humans, by combining AI and robot technologies. His research interests include intelligent robotics, cognitive developmental robotics, and robot learning. He received the IROS Best Paper Award Finalist, the Advanced Robotics Best Paper Award, and the JSAI Best Paper Award.

TAKATO HORII (Member, IEEE) received the M.E. and Ph.D. degrees in engineering from The University of Osaka, Osaka, Japan, in 2013 and 2018, respectively. He was a Project Researcher with the University of Electro-Communications, from 2017 to 2019. He was an Assistant Professor with The University of Osaka, from 2019 to 2020, where he was an Associate Professor, in 2020. His current research interests include the computational modeling of human cognitive functions, such as emotion and creativity and machine learning algorithms.