ment learning)

Human as Collaborator

Instead of a curator, a human could be part of a creation process as a collaborator[1]. Specifically, Fig. 1E illustrates humans to base their work on, perform, and process the model outputs. We are witnessing *human-as-collaborator* applications in a wide range of art, ranging from a painting to a film. Collaborative drawing and painting workflows utilize AI and robotic arms to co-create with humans (**?**). In AIBO, the opera performer's spoken words were used to generate texts from GPT-2 (**?**). During a live performance, Time Waves, human artists synthesized and sampled audio and visual elements according to directions from GPT-3 (**?**). Although not real-time like aforementioned examples, Sunspring (**?**) is a short film based on a script written by LSTM (**?**). This LSTM-generated script provided the setting and dialog, which were performed by the director and actors. This process can be seen as general modal translation (**?**), such as from a text to a multimedia or from a musical note to a sound wave. This type of HITL for creative techniques will increase in use, as humans fill in certain roles that AI models are not yet suitable for. In a long term, this flips the conventional roles of humans and computers in which humans dictate and computers support. The next level of *human-as-collaborator* approach would include humans' interpretative and performative works in further training (Fig. 1F). In aforementioned examples, human artists have played a significant role in creating a final output based on what the computing system has produced. Viewing humans as processing units, we are interested in better understanding and encoding how humans integrate certain inputs and generate artworks. We, as humans, use a variety of personal and cultural contexts – even subconsciously. While this information is difficult to quantify or summarize, HITL provides a framework to work with our innate creative and emotional responses. For example, with multi-modal architectures such as DALL-E (**?**) and CLIP (**?**) which are trained on a large corpus of images and texts, a human can paint according to a AI-generated text, and that human painting is fed back into the system. We may create more nuanced optimization problems, such as balancing of HITL paintings and initial training set of paintings.

These types of interactions require new user interfaces, network architectures, training schemes, and more. Innovative technical developments will aid in how we create AI art, what it means to encode creativity, and how we conceptualize machine creativity in general. For example, GPT-3 is not open-sourced and only accessed via texts; advanced computer vision requires powerful GPUs for real-time applications. A robust speech interface for GPT-3 based on speech recognition and synthesis would not only result in interesting experiments but also change our perspective In future HITL for machine creativity, human emotions and environmental vibes would be approximated, represented, and fed back into the AI model. This will aid in teaching AI how to better process and represent our creative and emotional responses in generative and expressive manners. The AI models would learn more varied, nuanced, and multi-modal understanding; e.g., texts like 'sadness' can not only link to 'blue' or 'minor chords', but also unexpected sonic textures based on more advanced concepts of harmony and composition. The proposed HITL that relies on more fluid and real-time inputs and outputs would push AI models to discover subconscious relationships and create novel artworks. Furthermore, this process will help us rethink what it means for machines to be creative and for our emotional responses to be encoded.

[1]Roles of curators and collaborators might be overlapping, along a spectrum of engagement in art making.

Conclusion

Creative expressions derived from AI systems, particularly using neural networks, are flourishing. Expensive sales, prestigious exhibitions, persistent coverage by the media, and longterm interest in academia shows that the rise of AI is a pivotal moment in the history of art. A majority of recent AI art relies on computer vision, audio processing, language models, and other algorithms that are excellent at mimicking, combining, and compositing styles learned from training. When human interactions are needed, they are often limited to initial training or curatorship afterwards. Future HITL for machine creativity will extend this interaction, in order to link different modalities, to interpret high-level concepts, and to mirror emotional responses. Humans can work as curators and collaborators to provide our creative and emotional feedback into creating engaging pieces. For machine creativity to be appreciated and valued, we need to think about how such outputs can relate to aesthetics and invoke emotions. Even in contemporary art, where aesthetics and emotional responses may be unintentionally or intentionally neglected, it is the temporal and societal contexts that make such pieces interesting. Such contexts would include, but not limited to, histories, traditions, heritages, and technological developments, as well as subconscious and individualistic connections. Teaching an algorithm to take into account the vastness of human experience and diversity of culture is likely impossible. Thus, we propose how expressive and creative training using HITL may help teach AI to understand and mimic certain complex associations and

Our proposal for present and future HITL for machine creativity attempts to better reflect this reality. Particularly, human technologists and artists are deeply coupled with final outputs. Imparting authorship or responsibility to AI systems has been dubious and challenging. It would be more fitting to describe, for example, how a human acted as a highly important curator that has chosen one digital While we may initially focus on creative applications of AI, these proposed approaches can be applicable in wider domains. Expertises of financial advisors or radiologists may not be easy to encode, yet highly important for profitable selection of investments or prognosis for malignant tumors. Nonetheless, how and when to incorporate a narrow and personal set of domain knowledge

Discussion on 'creativity' has a long history in cognitive science, computer science, and art (e.g., **????**). As intelligence can neither be easily defined nor be singularly opti-

mized, creativity is a broad and far-flung goal which can't be quantified, even for humans. What is clear is that state of