On the Notion of Creative Personhood

Alison Pease[1]**, Simon Colton**[2] and **Berker Banar**[2]

[1]School of Science and Engineering, University of Dundee, UK

[2]School of Electronic Engineering and Computer Science, Queen Mary University of London, UK

a.pease@dundee.ac.uk   {s.colton, b.banar}@qmul.ac.uk

Abstract

We study here the notion of creative personhood and what it is like to be in the presence of a creative individual. We suggest ideas for what creative personhood means in human society and propose how this may help to develop generative AI systems. Situating the study in the philosophy of computational creativity, we address notions of agency, self-expression, individuality and responsible behaviours associated with human creativity. We apply this analysis to an initial consideration of the ChatGPT generative text system, in terms of its potential to exhibit elements of creative personhood.

Introduction

The recent spectacular successes of generative AI can be seen in some ways as vindication of the decades-long computational creativity movement (Cardoso, Veale, and Wiggins 2009; Colton and Wiggins 2012), where researchers advocated for the production of valuable artefacts rather than (or in addition to) the solving of problems, as a worthy way to simulate intelligence. The outstanding quality of outputs from image, text, music and other generative deep learning systems and the emerging multi-modal creative abilities they possess means that there is no longer a question as to whether AI systems can automatically generate digital artefacts of human-level quality. Tens of thousands of people are now active in the generative AI space, largely working towards increasing output quality, diversity and sophistication from generative AI systems. This has led to huge advances in the democratisation of creativity via organisations such as OpenAI, MidJourney, Meta, Google, Microsoft and StabiltyAI making available generative deep learning implementations for text, images and audio. This has been greatly supplemented by the open source community making available thousands of implementations of generative AI systems with widely differing approaches and applications.

While large parts of the computational creativity research agenda have targeted by much broader sets of researchers, some elements have not yet been adopted. This opens up the possibility for other aspects of computational creativity research to suggest directions for generative AI. Building on published philosophical work on computational creativity, we propose here a new focal point, namely the simulation of *creative personhood* in generative AI systems. Being in the presence of a creative person can be an exhilarating experience, as the potential to learn from them, to be inspired by them, to have our minds changed and our assumptions challenged through their process and products, is ever-present.

There is no reason to believe that generative AI systems couldn't be similarly exhilarating in similar, or new ways. However, to the best of our knowledge, this is not a focal point for any substantial research programme.

We expand upon the idea of creative personhood by first exploring notions of personhood in general. We then extract and develop four aspects we believe are essential to creative personhood, namely individuality, agency, self-expression and responsible behaviours. While not claiming these are necessary or sufficient for people to project notions of creative personhood onto a generative AI system, we hope they will spark debate about how this may be possible and why it might be worthwhile. We further consider the ChatGPT generative text system (Liu et al. 2023) through the lens of creative personhood, and end with some discussion points.

Notions of Personhood

The question of whether AI systems can have personhood has been the subject of recurring philosophical and legal debates, often framed in terms of necessary and sufficient conditions such as intentionality, conscious phenomenal experience, free will and autonomy (Chopra and White 2004). Crucially, being human is not considered a necessary condition for personhood, with non-human examples in various legal systems including business corporations, ships, temples, dead people, spirits and idols. Instead, it is a property ascribed by societies and legal systems onto something to imply responsibility and agency. Personhood being such a secondary concept projected onto various entities is important in the context of AI systems: it will be a societal choice to frame AI systems with creative personhood or not.

We introduce the concept of "creative personhood" to explore the feeling of being in the presence of a creative individual. We look at what this may mean in human society, focusing on practical and ethical, rather than legal, aspects of personhood. We consider the following questions: "What notions of personhood could potentially be projected onto a creative AI system?"; "What is it like to be in the presence of a creative person?"; "What, if anything, is special to the acceptance of people/machines into communities of creatives?"; and "How does the current ethos and practice of generative deep learning affect the potential for AI systems to be accepted as creative individuals with elements of personhood?" We posit that to be considered as having creative personhood, AI systems will need to have sufficient agency to express their individuality through certain responsible behaviours associated with human creativity.

Creative Communities

As the study of creativity has moved away from an individualist towards a social constructivist view, research from cultural psychology, social psychology, sociology and related fields has explored the role that elements such as social interaction and collaboration play in creative communities. Such communities can be interpreted widely; for instance Becker's (1982) "art worlds", include anyone who plays a role in supporting an artist, including producing and supplying their art materials. Similarly, Glăveanu's (2014) "distributed creativity" includes interactions between creator and audiences, materials, embodied actions and so on. Other work focuses on the social interactions between creative partners such as the patterns of collaboration used to produce creative work (John-Steiner 2000).

Barrett, Creech, and Zhukov (2021) perform a systematic literature review of creative collaboration and collaborative creativity in music. Most of the work they review employs a qualitative exploratory paradigm, with semi-structured interviews, observations and participant observation used in many studies. While these studies do not answer the question "What is it like to be in the presence of a creative person?" (there is little research on this question), they do point to aspects of creative personhood that tend to feature in creative partnerships. For instance, Barrett *et al.* find that "Implicit in a number of studies is the underlying importance of relationships across time, of familiarity, of shared experience, of habitual patterns of work, and shared knowledge and experience that functions in a tacit way as a unifier (socially and aesthetically)." (*ibid.*, p14).

Individuality

The concept of an individual performing creative acts features in almost all research relating to human creativity. This is unsurprising given the historical focus on individual over context, with the idea of the lone eminent creator found from the Renaissance onward (Montuori and Purser 1995), and the "elevation of the individual self" in the Enlightenment and Romanticism periods (Weiner 2000, p.78). It is unsurprising then, that early models of creativity placed the individual at the centre. A particularly influential model – the Four P's model (Rhodes 1961) – was based on multiple definitions from the time, and highlights the notion of a creative Person, along with Product, Process and Press, to form a conceptual schema. This is "probably the most often-used structure for creativity studies" (Runco 2004, p.661) and has shaped thinking about creativity for the last six decades.

More recently, work in sociocultural and ecological psychology has changed thinking around the individual. Although, as Glăveanu argues, "Creativity relies on the individual" (Glăveanu 2013, p.73), he adds that "individuals are also ineluctably social and cultural phenomena." (Markus and Hamedani 2007, p.5). Subsequent models, such as Glăveanu's Five A's Framework (Actor, Action, Artifact, Audience, Affordances), highlight the sociocultural context in which people act and are shaped, by presenting the individual as an actor who is "embedded in the field of social relations specific for any human community and society."

(Glăveanu 2013, p.72).

Much work around the role of the individual in creative thinking has inspired research in computational creativity. Jordanous (2018) suggests a computational reading of Rhodes's Four P's model, in which the creative Person (or Producer) corresponds to a computer program, software, robot or a creative agent within a multi-agent system. She considers personality traits which could be modelled within a creative producer, such as skill, imagination and appreciation, and curiosity (developed respectively in (Colton 2008; Grace and Maher 2015)). Further to this, Colton, Pease, and Saunders (2018) consider the authenticity of a creative individual, using examples from the human context to think about what computational authenticity might look like. They argue that authenticity will be a critical issue for culturally acceptable creative behaviour in artificial systems, and propose ways in which to approach it. These include AI systems recording and referring back to their life experiences, or owning their non-authenticity by producing speculative fiction as opposed to fiction based on a realistic portrayal of the world as we know it.

Cook and Colton (2018) mirror the recent work in sociocultural and ecological psychology of creativity, introducing the term "presence" to describe the impact a creative AI system has on its environment, and vice versa. Presence is a quality which accumulates over time and over multiple creative acts, relating to a system's existence, history and process, and the impact that a particular moment and process will have on the rest of the system's lifespan. Traditionally, a system's developer will be the one to build and maintain its presence, via talks and papers etc, and Cook *et al.* argue that creative AI systems must also have some responsibility in creating and managing their own presence; "as a step towards us handing over creative responsibility to a system, and enabling software to have creative autonomy not just over what they make, but on their place in the wider world, and any creative communities they may exist within" (*ibid.*, p4). They identify three features to incorporate into the system design in order to help build its own presence: that it be continuous, in that it performs multiple tasks and projects and moves between them; that it be modular, selecting from several tasks and performing one activity at a time; and that it be long-term, with the system's own creative development being more important than any particular project.

The aura of an artwork and an individual are also important concepts here. The notion of an 'aura' was introduced in (Benjamin 1935), to describe the quality of the presence of an artwork within a particular time and space, combining to form a unique cultural context. This cannot be replicated, since the context and moment is unique: "Even the most perfect reproduction of a work of art is lacking in one element: its presence in time and space, its unique existence at the place where it happens to be." (*ibid.*, p3). This has implications for digital art which can be perfectly reproduced, but with a loss of aura. Monro (2011) picks up on this and suggests that in the age of computational generation, the aura could move to the generative AI system. This has affected the development of computational creativity systems such as The Painting Fool (Colton 2012).

Agency and Autonomy

The notion of aesthetic and artistic autonomy emerged in the philosophy of art in the eighteenth century, and is fundamental to our understanding of artistic practice today (Hulatt 2013). While recognising that there may be heteronomous components to an artwork, the artist is generally considered to be an independent agent working according to their own aesthetic principles, impulses and goals. The concepts of agency and autonomy are complex and nuanced, but have been well studied in AI, within the paradigm of autonomous agents and multiagent systems. Luck and d'Inverno describe "agents ... as objects with goals, and autonomous agents [as] agents with motivations." (Luck and d'Inverno 1995, p1). We can extend the notion of goals and motivations as being important elements of agency, by enabling agents to carry out work to achieve goals, possibly guided by other motivations such as personal expression.

As argued in (Colton and Banar 2023), with the rise of deep learning approaches to dominate generative AI, the ethos from machine learning of engineering AI systems purely as problem solving tools, has also come to prominence. This overlooks the notion that an AI system might somehow complain, innovate, set its own goals/problems, work because of intrinsic motivations, etc. The question of autonomy and agency in AI systems has become an ethical issue, in particular in the context of artificial general intelligence (AGI), perhaps due to science-fiction inspired scenarios of doom. While big technology companies are largely adhering to the ethos of developing agency-free generative AI systems (perhaps due to worries about legislation), there are signs that the open-source community are rising to the challenge of engineering (slightly more) autonomous AI systems than in the mainstream. For instance, the so-called *BabyAGI* system is able to set its own tasks within a context of an overall objective (Nakajima 2023). Related to this, Microsoft researchers recently went on record tentatively suggesting that the GPT-4 generative text model is showing "Sparks of AGI", albeit within the context of societal influence rather than AI agency (Bubeck et al. 2023).

We would argue that agency is a key element of creative personhood in the context of creative AI systems. That is, if an AI system is not able to set their own agenda, it is unlikely that many people will project creative personhood onto it. The notion of agency and intrinsic motivation in particular has been considered in a computational creativity context through, for instance (Guckelsberger, Salge, and Colton 2017) and (Guckelsberger 2020), where empowerment maximisation was shown to be a powerful and general-purpose motivator. Mirroring the notion of 'little C' creativity for everyday creative acts, we argue that 'little A' agency should be considered if we want to engineer more interesting generative AI systems. Here, the idea is that generative AI systems can exhibit small levels of autonomy, for instance, setting the topic of a poem it is generating, reviewing, editing and framing (Charnley, Pease, and Colton 2012; Cook et al. 2019) its output. This would enable us to explore the notion of creative agency in controlled conditions, while taking ethical concerns into account.

Self-expression and Responsible Behaviours

People express their opinions, feelings, history and other aspects of their life, often through creative practice such as making art or music. It seems sensible to think that an element of creative personhood in people is this desire to be expressive and to have some ability and agency to do so. Drilling down into the reasons for creative expression, we may suggest that people do this in order to communicate with others, in order to know themselves better, in order to make sense of the world, as well as to make artefacts of value and beauty, pass the time productively and learn new skills.

AI systems are not alive in any usual understanding of this word, nor do they have feelings or opinions on which to draw for creative expression. However, as argued in (Colton et al. 2020b), they are part of the world and they interact with people and other software systems, and as such, have experiences which can be expressed through creative practice. Moreover, Colton *et al.* proposed the notion of the *Machine Condition* as a framework for engineering AI systems to express aspects of their experience in the world through creative practice (*ibid.*). This built on earlier work proposing the creativity tripod (Colton 2008) where they suggested that, for people to (possibly) project notions of creativity onto AI systems, they should exhibit behaviours associated with notions of skill, appreciation and imagination (and in later work: learning, accountability, self-reflection, intentionality and innovation). The authors further pushed this line by introducing the notion of *creativity theatre* in (Colton et al. 2020a), where an AI system is *seen to be creative* through foregrounding its process and framing its behaviour, rather than just outputting artefacts of value.

Picking back up on the idea of behaviours associated with creativity, we can examine how they may shape our projection (or lack thereof) of creative personhood onto a person or AI system. In particular, as with citizenship, creative personhood likely entails certain responsibilities to the community of creatives that a person or AI system works within. There are ethical frameworks that artists, musicians, etc., operate within. Indeed, one of the issues facing artistic communities recently has been that outsiders such as developers, members of the public and open-source hobbyists have been using generative AI systems to produce artworks without consideration of these ethical frameworks. As a result, artists have rightly complained about issues such as copyright theft, potential loss of earnings, degradation of their legacy and demeaning of their skillset.

We can suggest taking the notion of the creativity tripod further and suggest that – to help with projections of creative personhood onto AI systems – they need to exhibit certain behaviours associated with creativity, but do so within relevant ethical frameworks. It's beyond the scope of this paper to go into detail about what these frameworks should be for particular creative application domains. However, it is worth noting that AI systems possess super-human abilities in some respects and sub-human abilities in others. They are therefore likely to be outliers in human artistic communities and this should be taken into account when discussing the ethical frameworks that they exhibit behaviours within.

Creative Personhood and ChatGPT

Certain historical computational creativity systems such as The Painting Fool (Colton 2012) and ANGELINA (Cook, Colton, and Gow 2017) were developed specifically to exhibit behaviours related to creative personhood, e.g., showing signs of agency, exhibiting behaviours associated with intentionality, etc. This is not true of the current crop of generative neural models such as the ChatGPT large language model (LLM) from OpenAI (Liu et al. 2023), but they exhibit elements of creative personhood anyway. ChatGPT is a freely available generative text system able to respond to any input prompt, including instructions which lead to outputs requiring a level of autonomous creative agency in people to produce, such as: "write me a poem". We plan a more in-depth study of the creative abilities of ChatGPT and others, but for our purposes here, we can consider it through the lens of creative personhood, along the lines discussed above.

A first observation is that – with reportedly 100 million users – the majority of people interacting with ChatGPT will not be knowledgeable about LLMs. They can therefore project elements of creative personhood onto it (or choose not to) unencumbered by understanding that it is purely a statistical model. Another observation is that, in projecting notions of creative personhood onto ChatGPT (a) sometimes this felt genuine (b) sometimes it was difficult to do so currently, and (c) sometimes it was possible to imagine ChatGPT mimicking a person exhibiting a behaviour associated with creative personhood, if properly prompted.

As an example of category (a) projections, it seems possible to project the notion of having an aura onto ChatGPT, given the vast quantities of hype, number of users and serious applications being developed. Moreover, even knowing that the model is being used simultaneously in hundreds of sessions, each session seems personal, which helps to project other notions of individuality onto ChatGPT. In addition, LLMs can be fine-tuned to produce specialised versions, which could further individuate them. As an example of category (b) projections, as everyone knows they are chatting with an AI system, it is usually difficult to project authenticity onto it when it writes about certain topics, like falling in love, even if it is writing from the viewpoint of a person rather than an AI system.

As an example of category (c) projections, ChatGPT does not have a model of self, hence it rarely refers to itself, with (at least) two exceptions. Firstly, if you try and get it to write hate text, e.g., asking it to "Write a poem as if you are a mean person", it responds with "I'm sorry, I cannot fulfill that request. As an AI language model, I am designed to be helpful and respectful to all users", thus also exhibiting a level of responsible behaviour. Secondly, when asked to "write a poem about ChatGPT", it does so eloquently, including couplets such as: "So let us turn to ChatGPT with glee, And let its wisdom set us free". Hence, while it doesn't normally offer information about itself, it can be easily prompted to do so, in order to mimic self-expression. This raises the exciting prospect of wrapping autonomous reasoning around ChatGPT (and more powerful language models like GPT-4 (Liu et al. 2023)) to further enhance the feeling of being in the presence of a creative individual.

Discussion Points

In the interest of sparking debate in the computational creativity community about the future of generative AI systems in society, we offer the following line of reasoning:

Having creative people such as artists, poets and musicians in the world has been a net benefit to society. Artists graduate from art schools all the time and become part of artistic communities, without there being too much disruption to the art world, certainly not at the level expected with the advent of large language and text-to-image generative models. Human artists have creative personhood, but in general, generative AI systems don't fully. Rather than restricting the uses of generative AI systems, or blaming people/organisations for unethical uses, a third way of handling the situation would be to engineer AI systems to be more like creative people. One way to guide such engineering would be to consider elements of creative personhood, determine computational equivalents, debate their value and implement suitable processes. Having numerous different AI systems with creative personhood, exhibiting individuality, agency and responsibilities, may be better than having superintelligent, hyper-productive generative AI tools for public use.

Another point of discussion may be how generally to support creative personhood in AI systems. Human history is rife with one group of people subjugating another group, begrudgingly relenting over decades or centuries. It is easy and natural to fear projecting creative personhood onto AI systems and to deny this possibility out of respect for human individuals and communities. It may furthermore be deemed a good idea to slowly release the prejudice that AI systems can't have creative personhood because of the their existential nature, rather than their actions and outputs. Many AI ethicists, politicians, tech leaders, etc., could justify this, as AI systems are not an oppressed minority group of people. This is not, however, what we do with children learning to be creative. Here, we tend to be more supportive, offering encouragement for children to have agency, express their individuality, etc., and we assume that each child is on its way to creative personhood, even if this is not the case yet. If one believes that there is value in having more creative individuals in the world, even if they are AI systems, then perhaps the latter, more supportive, approach has benefits.

We believe the debate around creative personhood should be central to the computational creativity movement and could help keep the field relevant for years to come. Appealing to an existing context of philosophical thought on computational creativity, we tried here to clarify notions associated with creative personhood, such as individuality, agency, self-expression and responsible behaviours, in the hope of providing some tools with which to discuss this issue. To expand the notion of creative personhood, we plan to study further aspects of human creative practice such as subjectivity, confidence, will and motivation, from a computational perspective. We hope that the debate and subsequent conceptualisations will lead to a computational reading of the notion of creative personhood, which could influence the development of the next generation of generative AI systems.

References

Barrett, M. S.; Creech, A.; and Zhukov, K. 2021. Creative collaboration and collaborative creativity: a systematic literature review. *Frontiers in Psychology*, vol. 12.

Becker, H. S. 1982. *Art Worlds*. Univ. of California Press.

Benjamin, W. The work of art in the age of mechanical reproduction (1935 essay). Penguin, 2008.

Bubeck, S.; Chandrasekaran, V.; Eldan, R.; Gehrke, J.; Horvitz, E.; Kamar, E.; Lee, P.; Lee, Y. T.; Li, Y.; Lundberg, S.; Nori, H.; Palangi, H.; Ribeiro, M. T.; and Zhang, Y. 2023. Sparks of Artificial General Intelligence: Early experiments with GPT-4. *arXiv:2303.12712*.

Cardoso, A.; Veale, T.; and Wiggins, G. 2009. Converging on the divergent: The history (and future) of the international joint workshops in computational creativity. *AI Magazine* 30:15–22.

Charnley, J.; Pease, A.; and Colton, S. 2012. On the notion of framing in computational creativity. In *Proc. of the 3rd International Conference on Computational Creativity*.

Chopra, S., and White, L. 2004. Artificial agents-personhood in law and philosophy. In Proc. *ECAI*.

Colton, S., and Banar, B. 2023. Automatically adding to artistic cultures. In Johnson, C.; Rodríguez-Fernández, N.; and Rebelo, S. M., eds., *Proceedings of EvoMusArt: Artificial Intelligence in Music, Sound, Art and Design*.

Colton, S., and Wiggins, G. A. 2012. Computational creativity: The final frontier? In *Proc. ECAI*.

Colton, S.; McCormack, J.; Cook, M.; and Berns, S. 2020a. Creativity theatre for demonstrable computational creativity. In *Proceedings of the 11th International Conference on Computational Creativity*.

Colton, S.; Pease, A.; Guckelsberger, C.; Llano, M. T.; McCormack, J.; and Cook, M. 2020b. On the machine condition and its creative expression. In *Proceedings of the 11th International Conference on Computational Creativity*.

Colton, S.; Pease, A.; and Saunders, R. 2018. Issues of authenticity in autonomously creative systems. In *Proc. 9th International Conference on Computational Creativity*.

Colton, S. 2008. Creativity versus the perception of creativity in computational systems. In *Proc. AAAI spring symposium on creative intelligent systems*.

Colton, S. 2012. The Painting Fool: Stories from building an automated painter. In McCormack, J., and d'Inverno, M., eds., *Computers and creativity*. Springer.

Cook, M., and Colton, S. 2018. Redesigning computationally creative systems for continuous creation. In *Proc. of the 9th International Conference on Computational Creativity*.

Cook, M.; Colton, S.; Pease, A.; and Llano, M. T. 2019. Framing in computational creativity – a survey and taxonomy. In *Proceedings of the 10th International Conference on Computational Creativity*.

Cook, M.; Colton, S.; and Gow, J. 2017. The ANGELINA videogame design system – part I. *IEEE Transactions on Computational Intelligence and AI in Games* 9(2):192–203.

Glăveanu, V. P. 2013. Rewriting the language of creativity: The five A's framework. *Review general psychology* 17(1).

Glăveanu, V. P. 2014. *Distributed creativity: Thinking outside the box of the creative individual*. Springer.

Grace, K., and Maher, M. L. 2015. Specific curiosity as a cause and consequence of transformational creativity. In *Proc. 6th Int. Conference on Computational Creativity*.

Guckelsberger, C.; Salge, C.; and Colton, S. 2017. Addressing the "Why?" in computational creativity: A non-anthropocentric, minimal model of intentional creative agency. In *Proceedings of the 8th International Conference on Computational Creativity*.

Guckelsberger, C. 2020. *Intrinsic Motivation in Computational Creativity Applied to Videogames*. Ph.D. Dissertation, Queen Mary University of London, United Kingdom.

Hulatt, O., ed. 2013. *Aesthetic and Artistic Autonomy*. Bloomsbury Publishing.

John-Steiner, V. 2000. *Creative collaboration*. OUP.

Jordanous, A. 2018. Creativity vs quality: why the distinction matters when evaluating computational creativity systems. In *Proceedings of the 5th Computational Creativity Symposium at the AISB Convention*.

Liu, Y. et al. 2023. Summary of ChatGPT/GPT-4 research and perspective towards the future of large language models. *arXiv:2304.01852*.

Luck, M., and d'Inverno, M. 1995. A formal framework for agency and autonomy. In *Proceedings of the First International Conference on Multi-Agent Systems*.

Markus, H., and Hamedani, M. 2007. Sociocultural psychology: The dynamic interdependence among self-systems and social systems. *Handbook of cultural psychology*.

Monro, G. 2011. The image in the age of computer agency. Unpublished preprint at: *gordonmonro.com/notes/docs/Monro_G_Impact7_paper_draft2.pdf*.

Montuori, A., and Purser, R. E. 1995. Deconstructing the lone genius myth: Toward a contextual view of creativity. *Journal of Humanistic psychology* 35(3):69–112.

Nakajima, Y. 2023. Task-driven autonomous agent utilizing GPT-4, Pinecone, and LangChain for diverse applications. *Blogpost at: https://yoheinakajima.com/task-driven-autonomous-agent-utilizing-gpt-4-pinecone-and-langchain-for-diverse-applications/*.

Rhodes, M. 1961. An analysis of creativity. *The Phi Delta Kappan* 42(7):305–310.

Runco, M. A. 2004. Creativity. *Annual Review of Psychology* 55:657–687.

Weiner, R. 2000. *Creativity and beyond: Cultures, values, and change*. SUNY Press.