The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

6th Youth in High-Dimensions: Recent Progress in Machine Learning, High-Dimensional Statistics and Inference & How creative is Generative AI? Perspectives from Science and Philosophy | (SMR 4084)

03 Jul 2025 - 10 Jul 2025
ICTP, Trieste, Italy

T01 - BALDWIN Daniel Guy

Process Versus Output: Various Illusions of Originality in Humans and AI

T02 - CHIECO Giovanni

Reimagining Authorship and Ownership of AI-Generated Images under EU Law

T03 - DE PAOLIS Ludovica

Synthesizing Naturalistic Visual Textures Using Deep Neural Samplers

T04 - FORNI Ortensia

Color Distributions in Nature: A Data-Driven Exploration of Color Harmony and Aesthetic Preferences

T05 - HAYES Patrick Ainsworth

Creativity as a Human Virtue

T06 - KUNICKI Martha

Can GenAI be creative? AI-Human Co-Creativity: Enriching the Concept of Creativity in Light of Emerging GenAI Towards Collective Co-Creativity of AI and Humans

T07 - LAZAREVIC Dorde

AI and the problem of creative spontaneity

T08 - MÁCHA Jakub

Metaphorical Masks: Capturing and Concealing LLM Creativity in Art

T09 - MASAI Pierluigi

From the Human/Machine Dichotomy to Distributed Creativity - The Case of AI Integration in Filmmaking

T10 - MILANESIO Federico

6th Youth in High-Dimensions: "Understanding Geometric Compression in Neural Networks Dynamics" AI Creativity: "The Myth of the Creativity Dial: Temperature in LLMs"

T11 - PIZZOCHERO Michele

Can Machines Philosophize?

T12 - PUSCASU Iulia-Dana

Generative AI: collaborator or tool? The experience of interaction in the process of artistic creation

T01

Process Versus Output: Various Illusions of Originality in Humans and AI

Daniel Guy Baldwin

University of Liverpool

A common argument against the creativity of generative AI is that it lacks originality, but the creative process of human beings and that of generative AI does not seem so far removed as people are inclined to believe initially. Generative AI uses machine-learning to produce novel data – on a basic level this might involve a simple model such as a Markov chain. The next word in a text is generated based on the previous ones. A more modern generative AI such as ChatGPT functions at a base level in essentially the same manner but with greater complexity and parameters, trained on a massive dataset of text and code available on the Internet.

This does not appear too dissimilar to certain forms of creative process in human beings. I am not a tabula rasa when it comes to creation. I do not produce something out of nothing. I have a lived experience which, along with its interpretation, fills the function of both dataset and parameters in my creative process. The way I write is informed by the books I've read; the themes by the life I have lived. In one sense, the next word in my text is generated based on previous ones. One might say there is still a greater degree of originality in human creation which does not exist with generative AI, but this is not always the case.

Some forms of art involve direct modification of previous artworks. This is especially true within music: hip hop often samples older tracks; interpolation uses a previously recorded melody in a new recording; song covers can involve little to no unique ideas. There have also been cases where musicians have unintentionally copied other artists' tracks (see Nine Inch Nails' A Warm Place and David Bowie's Crystal Japan). It is also relatively easy to find tracks made entirely from samples online. It seems unlikely we would deny these were creative endeavours although they only involve reorganising previously available data.

From this it seems we also judge originality and creativity by the product produced rather than strictly by the process alone.

T02

Abstract template for How Creative is Generative AI? Perspective form Science and Philosophy

-

Reimagining Authorship and Ownership of AI-Generated Images under EU Law

Giovanni Chieco

(PhD Candidate in Applied Data Science and Artificial Intelligence at The University of Trieste – AI Law; Trainee Lawyer)

The proliferation of generative artificial intelligence (AI) technologies—such as DALL·E and Midjourney —has introduced significant disruptions to the European creative landscape. These tools, capable of autonomously producing complex visual content, challenge the foundational assumptions of European intellectual property law, particularly with respect to authorship, originality and ownership.

This paper critically examines the current regulatory vacuum within the European Union (EU) regarding the legal status of AI-generated images. While the EU copyright framework—primarily governed by the InfoSoc Directive (2001/29/EC) and national implementations—continues to rely on the principle of human authorship, there remains no explicit guidance for works autonomously produced by non-human agents. As a result, such outputs often fall outside the scope of protection, leading to legal uncertainty and a potential chilling effect on innovation and investment in creative AI applications.

The analysis focuses on three interrelated dimensions. First, it explores the doctrinal limitations of EU copyright law in accommodating non-human creativity, particularly the requirement of "authorial" originality. Second, it addresses the implications of private governance, wherein AI tool providers assert control over outputs through contractual terms, effectively substituting public copyright regimes with privatized models of regulation. Third, it considers emerging legislative initiatives at the EU level and assesses their relevance for addressing the governance of AI-generated content, despite their limited direct engagement with intellectual property concerns.

In light of these challenges, the paper argues for the development of a harmonised European approach that recognises the hybrid nature of human–AI collaboration. It proposes a set of policy options, including the introduction of sui generis protection for AI-generated works, the clarification of human involvement thresholds for copyright eligibility and the encouragement of open licensing models to support legal certainty, fair access and responsible innovation in the European creative economy

[1] P.B. Hugenholtz, J.P. Quintais, *Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output?*, Int. Rev. Intell. Prop. Compet. Law 52, 1190–1216 (2021).
[2] D. Bietti, *Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity*, Comput. Law Secur. Rev. (2024). Available at: sciencedirect.com
[3] C. Di Bari, *AI Generated Works: The Clash Between Copyright Principles and Technological Evolution*, Law and Media Working Paper Series (2021). Available at: medialaws.eu
[4] G. Frosio, M. Jütte, *Developing Artificial Intelligence-Based Content Creation: Are EU Copyright and Antitrust Law Fit for Purpose?*, Int. Rev. Intell. Prop. Compet. Law 54, 351–380 (2023).
[5] M. Iglesias, *Generative AI, Copyright and the AI Act*, Comput. Law Secur. Rev. (2025). Available at: sciencedirect.com

[6] J. Lee, H. Yuan, *Navigating the Legal Landscape of AI Copyright: A Comparative Analysis of EU, US, and Chinese Approaches*, AI Ethics 4, 10 (2024). Available at: link.springer.com
[7] J. Rosati, *Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output?*, Int. Rev. Intell. Prop. Compet. Law 52, 313–340 (2021).
[8] S. Wang, *Between Copyright and Computer Science: The Law and Ethics of Generative AI*, arXiv (2024). Available at: arxiv.org
[9] S. Wang, *Copyright Protection for AI-Generated Works: Exploring Originality and Ownership in a Digital Landscape*, Asian J. Int. Law (2024). Available at: cambridge.org

T03

Synthesizing Naturalistic Visual Textures Using Deep Neural Samplers

Ludovica de Paolis[1], **Eugenio Piasini**[1], **Fabio Anselmi**[2] **and Alessio Ansuini**[3]

[1]*International School for Advanced Studies - SISSA*
[2]*University of Trieste*
[3]*AREA Science Park, Research and Technology Institute*

The goal of this study is to generate naturalistic visual textures with a novel generative architecture and to investigate their representations through their statistical and semantic properties. In neuroscience, the efficient coding principle states that the mammalian visual system conforms to the statistics of its natural input: natural visual scenes. Visual textures are a family of visual patterns that share certain local regularities [1] that have been used to test efficient coding. Still, experimental results are limited to unnatural looking low-order correlations. There are three main texture synthesis algorithms: parametric models [2], maximum entropy sampling methods [3], Gram Matrix-based convolutional neural networks [4]. We propose a new model for texture synthesis: a combination of a Variational Autoencoder (VAE [5]) and a pre-trained convolutional neural network (VGG-16 [6]) to generate realistic textures in the latent space, only characterized by nonlinear multi-scale representations. The model is trained with unsupervised learning to minimize: a Kullback-Leibler divergence between the latent code and a Gaussian distribution; a Mean Squared Error between pairs of Gram Matrices computed on VGG-16 feature maps to define the perceptual difference between original and generated textures. We are currently training the model with the Describable Textures Dataset (DTD [7]), a supervised dataset of 5640 images of textures divided in 47 classes. We successfully replicated the of study by [4] using DTD and studied the ability of Gram Matrices to capture statistical information of textures with Representational Similarity Analysis (RSA [8]) on the Gram Matrices produced with Gatys' approach. We also performed a zero-shot learning experiment with CLIP [9] on DTD with the intention to investigate the semantic attributes that characterize textures. We expect our model to produce a texture space that aligns with human perception, and the generated textures to show high variance statistical features, which will reveal how efficient coding applies in our behavioral and computational regime. We expect Gram Matrices to capture the statistical information contained in the textures, and therefore to best generate highly abstract textures; we expect textures that are more similar to objects to be captured by semantic features as represented by CLIP.

References

[1] B. Julesz. "Textons, the elements of texture perception, and their interactions". In: *Nature* 290 (1981), pp. 91–97.

[2] J. Portilla E. P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients". In: *International Journal of Computer Vision* 40 (2000), pp. 49–70.

[3] J. D. Victor M. M. Conte. "Local image statistics: maximum entropy constructions and perceptual salience". In: *Optical Society of America* 29 (2012), pp. 1313–1345.

[4] L. A. Gatys et al. "Texture Synthesis Using Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems 28 (NIPS 2015)* (2015).

[5] D. P. Knigma M. Wellig. "Auto-Encoding Variational Bayes". In: *ArXiv* 11 (2022).

[6] K. Simonyan A. Zisserman. "Very Deep Convolutional Networks for Large Scale Image Recognition". In: *ICLR 2015* (2015).

[7] M. Cimpoi et al. "Describing Textures in the Wild". In: *CVPR 2014* (2014).

[8] N. Kriegeskorte M. Mur P. Bandettini. "Representational similarity analysis – connecting the branches of systems neuroscience". In: *Front. Syst. Neurosci* (2008).

[9] A. Radford et al. "Learning Transferable Visual Models From Natural Language Supervision". In: *arXiv* (2021).

T04

Color Distributions in Nature: A Data-Driven Exploration of Color Harmony and Aesthetic Preferences

O. Forni[1], **A. Darmon**[2], **M. Benzaquen**[1,3,4]

[1] *Chair of Econophysics & Complex Systems, Ecole polytechnique, 91128 Palaiseau Cedex, France*

[2] *Art in Research, 33 rue Censier, 75005 Paris, France*

[3] *Ladhyx UMR CNRS 7646, Ecole polytechnique, 91128 Palaiseau Cedex, France*

[4] *Capital Fund Management, 23 Rue de l'Université, 75007 Paris, France*

Our research explores the relationship between color appreciation and statistical properties, extending previous work [1] on grayscale images, which showed that aesthetic appreciation peaks at intermediate entropic complexity, thus suggesting universal criteria for aesthetic judgment.

Since previous studies [1, 2] showed that people's preferences are influenced by natural environment, we investigate whether similar principles apply to color images by analyzing the color distribution of a large dataset of natural pictures - extracting their dominant hues and identifying a common distribution of colors in nature.
First we examine whether natural color distributions align with the color harmony theory of Moon and Spencer (1944) [3], which is still widely referenced in various fields, despite the lack of clear evidence supporting the applicability of this theory. By testing this model in an agnostic manner through a large-scale survey, we identify a hue-dependent effect and develop a more comprehensive quantitative framework for color harmony and aesthetic appreciation in color pairings.
Then, on the basis of this analysis, we construct a novel hue space that redistributes color spectrum according to frequency of nature. The application of Multiscale Relevance [4] in this context provides a new approach to quantify how different color spaces capture the information content of an image.

By integrating statistical physics with empirical studies, our work aims to address color appreciation and harmony in a truly quantitative manner. It bridges aesthetic perspectives and computational methods, aiming also to provide physically grounded tools for color-based image analysis and interpretation and to enhance the physical interpretability of image processing by studying color distributions.

[1] S. Lakhal, A. Darmon, J-P. Bouchaud, and M. Benzaquen. "Beauty and structural complexity". In: *Physical Review Research* 2.2 (June 2020).

[2] G. J.Sthephens, T. Mora, G. Tkacik, and W. Bialek. "Statistical Thermodynamics of Natural Images". In: *Phys. Rev. Lett.* 110, 018701 (2013).

[3] P. Moon and D. Spencer. "Aesthetic measure applied to color harmony". In: *JOSA* 34.4 (1944), pp. 234–242.

[4] S. Lakhal, A. Darmon, I. Mastromatteo, M. Marsili, and M. Benzaquen. "Multiscale relevance of natural images". In: *Scientific Reports* 13.1 (2023), p. 14879.

T05

Creativity as a Human Virtue

Pat Hayes

University of Glasgow

Drawing from Swanton's book *Target-Centred Virtue Ethics* and Pettigrove's explorations of creativity in virtue theoretical space, this paper will examine how we can derive a general framework in which to understand creativity and its relation to human agents which can then be applied to the aesthetic virtue of creativity and its possible appearance in AI. Within this framing, creativity may be conceptualised as a character trait possessed and acted upon by human agents as they hit their socio-culturally defined, Aristotelian mean targets of excellent action in creation, which may result in artistic achievement for a particular agent's development or art. Once this conception has been put forward and defended, I will argue that this conception may be applied to cases of existing and possible future AI to show that creativity is a human virtue that may only be possessed, acted upon by, and motivate human agents, who are themselves enculturated in a particular society or dwelling. By doing so, I aim to show that the only fitting attribution of creativity is to human agents.

T06

Smr 4084 - 6th Youth in High-Dimensions

How creative is Generative AI? Perspectives from Science and Philosophy 3-5 July 2025

Martha Kunicki, Ph.D.

Princeton University

Can GenAI be creative? AI-Human Co-Creativity: Enriching the Concept of Creativity in Light of Emerging GenAI Towards Collective Co-Creativity of AI and Humans

This paper explores how our concept of creativity has to be rethought in light of emerging GenAI in creative practice, thereby leading to an enrichment of the concept of creativity to include collective co-creativity produced by humans and GenAI. Drawing on my experience as a theater director and producer involved in collective creativity and as a philosopher trained in analytical thinking, I explore case studies of co-created art, music, holograms, and text, in theater, and demonstrate how GenAI can be creative and how it can enhance our creativity.

Working together with GenAI can *enhance* creativity because digital technologies can expand the capacities of human creativity. I illustrate this by taking case studies of co-creativity in art, music, text, and holograms in theater, with examples of co-creations in art includinng *DALL·E*, co-creations in music including *AIVA*, and co-creations in text including *ChatGPT*. The most exciting recent examples of co-creations occur in immersive theater experiences, which integrate multisensory performances.

Collective creativity arises from the collective practices of people in relation to their nature, environment, cultures, societies, techniques, technologies, and GenAI. By expanding our concept of creativity towards co-creativity, we are able to acknowledge the processes and products of humans and GenAI as creative.

Creativity is standardly defined as the ability to produce something *new* and *valuable*. Various conceptions of creativity have been distinguished in the literature, such as *combinational*, *explorational*, and *transformational* creativity. I critically examine the various characterizations and conceptions, and offer a new conception of *integrational* creativity, which brings together the various conceptions of creativity. On this new conception, creativity is the creation of something (i) *new*, (ii) *integrationally significant* in the relevant context, and (iii) that *resonates*. A key feature of this new conception is that it allows for co-creativity of humans and GenAI.

GenAI has the potential to revolutionize the way we think about creativity and how we experience creativity. Today's emerging GenAI technologies, with ground-breaking new tools, open up greater possibilities for creativity, with richer forms of experiencing creativity arising from the collective creativity of GenAI and humans. GenAI is becoming an increasingly influential (f)actor in the creative process.

T07

AI and the problem of creative spontaneity

Đorđe Lazarević

Faculty of Philosophy at the Univeristy of Niš, Serbia

Recent advances in "creative technology," particularly AI-generated images, poetry, and scientific ideas, have led to questions about whether machines might one day surpass human creative abilities—or whether they already have. However, to properly assess this and related questions, it is crucial to first establish a clear conceptual framework for creativity itself. In this talk, I pursue two main objectives: first, I propose an analysis of creativity by distinguishing three key components—product, process, and agent. Second, I brefly will explore the relevance of agential and production aspects for AI, and then focus on a more pressing question: can a machine be spontaneous? Since spontaneity is a necessary condition for a creative cognitive process, even if a machine has a conscious desire to create, its presumed lack of spontaneity would imply a lack of creativity. One common argument holds that machines lack spontaneity because they are always following instructions and thus lack intrinsic motivation and creative freedom.

I challenge this argument in two ways. First, both thought experiments and empirical studies confirm that human creative actions can be determined and predicted, which undermines the idea that spontaneity requires indeterminacy or unpredictability. A stronger argument, however, holds that AI agents are always causally dependent on humans, whereas humans, at least sometimes, initiate creative processes independently. Here, I suggest that the key distinction lies in how we interpret human *versus* machine actions: we attribute creativity to humans partly because we lack full knowledge of the causal factors underlying our own creative cognition, whereas we assume a fully determined model of action for machines. A further implication of this view is that a superintelligent AI—if it possessed complete knowledge of its own creative cognitive architecture—could not be creative, since it would lack the very epistemic opacity that underlies spontaneity in human creativity, even thought such AI would create most valuable ideas and artifacts.

T08

Metaphorical Masks: Capturing and Concealing LLM Creativity in Art

Jakub Mácha, Masaryk University

This paper examines how metaphorical language shapes perceptions of large language models (LLMs) in artistic practice, focusing on whether these framings accurately express or conceal LLMs' creativity. Adopting a non-cognitivist approach, where metaphors invite "seeing as" without fixed cognitive content, we analyze five metaphors—artist, collaborator, muse, medium, and tool—and their interplay in defining LLMs' artistic roles. The "artist" metaphor emphasizes LLMs' creative outputs, like generating surreal poems, implying human-like intent, but risks anthropomorphism that obscures their computational nature, as Bender et al.'s "stochastic parrots" critique (2021). The "tool" metaphor highlights human control, framing LLMs as predictable instruments but concealing their unpredictable, recalcitrant creativity, comparable to gardening's dynamic medium (Young & Terrone, 2025). The "collaborator" and "muse" metaphors balance agency and inspiration, expressing LLMs' contributions while preserving human responsibility (Anscomb, 2024), yet human-like framings may overstate autonomy, masking algorithmic limits. The "medium" metaphor captures LLMs' unpredictable outputs, enhancing creativity but constrained by "tool's" reductive lens. Against the master metaphor of "artificial intelligence," which portrays machines as figuratively intelligent, these metaphors interact—complementing to deepen perspectives, constraining when human-centric views obscure LLMs' unique generative potential. Anthropomorphic metaphors like "artist" or "collaborator" may conceal LLMs' distinct creativity, rooted in statistical patterns, while "medium" and "tool" clarify their non-human agency. Drawing on Anscomb's responsibility arguments, this study shows that metaphorical framings navigate the artworld's social validation, expressing LLMs' creativity when acknowledging their computational essence, but concealing it when overly humanized. This interplay positions AI art as a collaborative process of human intent and machine dynamics, offering nuanced insights into LLMs' artistic contributions.

T09

From the Human/Machine Dichotomy to Distributed Creativity
The Case of AI Integration in Filmmaking

Pierluigi Masai[1], **Lorenzo Carta**[1], **and Mateusz Miroslaw Lis**[2]

[1]*Università degli Studi di Trieste*
[2]*SophIA - Artificial Intelligence Audiovisual Lab*

The recent development and diffusion of generative AI systems have reinvigorated debates about machine creativity, particularly regarding its comparison to and potential surpassing of human creative abilities in the autonomous production of artistic work. Yet, the very notion of creativity is ambiguous and to attribute creativity to a machine could be misleading. Conventional approaches often conceptualize creativity as an isolated phenomena, traceable in the internal functioning of an individual system—whether human or artificial—while disregarding its embeddedness within material and social contexts. By treating human or technological creativity as self-contained processes, such analysis overlooks the dynamic entanglement of artifacts, practices, and socio-economic arrangements that give rise to artistic products [1].

To advance our argument, we will focus on audiovisual production—specifically, Filmmaking—as a key case study. Following Vlad Glăveanu's perspective on creativity [2], audio-visual work is not an individual act, but a collaborative and distributed process that involves aspects both technical (lenses, tripods, sensors, lighting) and human (the experience and sensitivity of cinematographers, focus pullers or camera operators). In line with a Science and Technology Studies (STS) approach, creativity here arises from the dynamic interaction between the materiality of technology, subjective interpretations, and broader social and economic environments. In this framework, AI tools represent a technological shift that can reshape existing creative practices—both affording and constraining aesthetic and working possibilities.

With this theoretical framework in mind, we will provide an overview of some AI techniques and show how they might be useful to overcome common problems met in film shootings and editing with an impact on film aesthetics. Ultimately, we argue that AI's impacts on creative processes demand examination beyond technical functioning, requiring instead a situated analysis of how these technologies may become part of everyday creative practices.

References

[1] C. Celis Bueno, P.-S. Chow, and A. Popowicz. Not "what", but "where is creativity?": towards a relational-materialist approach to generative AI. *AI & SOCIETY*, Mar. 2024.

[2] V. P. Glăveanu. *Distributed Creativity: Thinking Outside the Box of the Creative Individual*. Springer, Cham, 2014 edition, Apr. 2014.

T10

Understanding Geometric Compression in Neural Network Dynamics

Federico Milanesio[1], **Marco Gherardi**[2], **and Matteo Osella**[1]

[1]*(Presenting author underlined) Department of Physics, University of Turin & INFN, Turin*
[2]*Department of Physics, University of Milan*

Neural networks are the most used algorithm in modern machine learning and have achieved incredible performance on different tasks. However, their development relies not on a deep theoretical understanding but on trial and error. One of the main difficulties of understanding NNs lies in the problematic nature of their training dynamics, which seeks optima in a very high-dimensional rough landscape. At the same time, the models avoid becoming trapped in suboptimal minima and converge to points with good generalization, thus escaping the so-called curse of dimensionality.

Intuitively, in regression tasks, the optimal representation of the data would be a low-dimensional manifold in which the points align, allowing for easy linear regression. We introduce a PCA-based measure to capture geometric compression and investigate this prediction. Our observations during training reveal a non-monotonic behavior, namely a first compression phase, followed by a subsequent decompression phase where our measure increases again. Such a result is remarkable and unobserved in regression NNs but aligns with previous results on classification tasks [**?**]. We prove that this behavior is a property of feature learning and is quite general for changes in hyperparameters and different datasets. We show that the epoch of inversion always happens after the network has learned to predict the best linear regression possible, and the subsequent decompression phase may indicate a phase of generalization, in which the model becomes more flexible and needs representations decompressed to accommodate more complex functions. This behavior aligns with current literature suggesting that neural networks learn the data distribution's moments sequentially [2]. We theorize that inversion happens when the network has learned the first two moments and starts learning from moments of higher order.

[1] S. Ciceri, L. Cassani, M. Osella, et al., Nature Machine Intelligence, **6(1)**, 40–47 (2024).
[2] M. Refinetti, A. Ingrosso, S. Goldt, PMLR, **202**, 28843 (2023).

The Myth of the Creativity Dial: Temperature in LLMs

Federico Milanesio[1], **Filippo Valle**[1], **Matteo Osella**[1]

[1]*(Presenting author underlined) Department of Physics, University of Turin & INFN, Turin*

Large Language Models (LLMs) have emerged as the frontier tool in natural language processing. Even compared to other promising fields in AI, the performance of LMMs in tasks such as language generation, translation, and sentiment analysis remains a forefront example of machine learning possibilities to revolutionize fields both within academia and in the job market.

LLMs generate text by predicting the next token (word or subword unit) by assigning a raw score to each potential token in the model vocabulary based on prior text. These scores are then converted into probabilities using the softmax function, which is regulated by a parameter called *temperature*. This temperature controls the randomness of the output, effectively influencing the degree to which the generated text is predictable. A lower temperature makes the model more deterministic, favoring the most likely next word based on its training, resulting in more conservative responses. In contrast, a higher temperature increases randomness by allowing less probable words to be selected more frequently.

Temperature is often referred to as a *creativity parameter*. Although recent scientific literature has begun to challenge this characterization [1], it remains the de facto understanding of practitioners. This aligns with the guidance given by model developers, who recommend specific temperature settings for different use cases. Temperature is also linked to some poorly understood behaviors of LLMs. Among these are the apparent phase transition-like behavior in the models' output as the temperature changes and the tendency of models to generate looping and repetitive outputs for very low temperatures. Although the latest language models solved these problems through architectural and post-processing techniques [2], the fundamental causes remain poorly understood.

In this talk, we will examine the current understanding of the temperature parameter, discussing its practical role in controlling output diversity and its contested status as a measure of creativity. Finally, we will present new results exploring the phase transition-like behaviors observed in LLM outputs, proposing simple models to shed some light on these problems.

[1] M. Peeperkorn, T. Kouwenhoven, D. Brown, A. Jordanous, in Proceedings of the 15th International Conference on Computational Creativity, 226-235 (2024).

[2] A. Holtman, J. Buys, L. Du, et al., in International Conference on Learning Representations (2020).

T11

Can Machines Philosophize?

Michele Pizzochero [1,2] **& Giorgia Dellaferrera** [3]

[1] *Department of Physics, University of Bath, United Kingdom*
[2] *School of Engineering and Applied Sciences, Harvard University, United States*
[3] *McKinsey & Co, United Kingdom*

Inspired by the Turing test, we present a novel methodological framework to assess the extent to which a population of machines mirrors the philosophical views of a population of humans. The framework consists of three steps: (i) instructing machines to impersonate each human in the population, reflecting their backgrounds and beliefs, (ii) administering a questionnaire covering various philosophical positions to both humans and machines, and (iii) statistically analyzing the resulting responses. We apply this methodology to the debate on scientific realism, a long-standing philosophical inquiry exploring the relationship between science and reality. By considering the outcome of a survey of over 500 human participants, including both physicists and philosophers of science, we generate their machine personas using an artificial intelligence engine based on a large-language generative model. We reveal that the philosophical views of a population of machines are, on average, similar to those endorsed by a population of humans, irrespective of whether they are physicists or philosophers of science. As compared to humans, however, machines exhibit a weaker inclination toward scientific realism and a stronger coherence in their philosophical positions. Given the observed similarities between the populations of humans and machines, this methodological framework may offer unprecedented opportunities for advancing research in experimental philosophy by replacing human participants with their machine-impersonated counterparts, possibly mitigating the efficiency and reproducibility issues that affect survey-based empirical studies.

T12

Generative AI: collaborator or tool? The experience of interaction in the process of artistic creation

Iulia-Dana Pușcașu[1]

[1]*Babeș-Bolyai University, Cluj-Napoca (Romania)*

This study begins by analysing the ambiguous use of the word 'collaboration' in relation to generative AI in artistic creation. Identifying some of the main contexts in which the term is used reveals the concept of interaction to be of key interest, since in all cases, at a minimum, collaboration implies a form of interaction between a human and an AI. While I do not take a position on whether it is appropriate to designate any human-AI interaction as collaboration, I aim to show that a sense of collaboration emerges from the experience of interaction. Broadly defined as "mutual or reciprocal action or influence," the concept of interaction we are interested in here covers human–non-human exchange processes. Following an intentionalist approach to art [1], my claim is that, in the context of production and creation—and in light of current technological developments—the interaction between an artist and an AI can fundamentally be considered the same as the interaction between an artist and their regular tools, but that it is the experience of interacting with AI systems that differs. The first part of my claim is based on the fact that generative AI systems appear in the continuation of a long line of technologies that have enabled automatic image-making techniques [2], and which, in retrospect, can be said to constitute an art historical tradition. Furthermore, while AI systems produce formal features that could be considered artistically relevant, they do not exercise knowledge-how or evaluative abilities in producing these features as such, and therefore do not qualify as artistically creative agents [3]. The second part of my claim is explained, first, by the ongoing efforts to simulate interpersonal communication in the modelling of interfaces for human–machine interaction [4]. Second—and this is my focus—is the tendency to infer decision-making capabilities and intention from the system's manner of responsiveness through misdirected interpretation, i.e., as if a generative AI were operating on a semantic level. This issue also concerns skill, and the difficulty of scaling both the functioning and the possibilities that sophisticated tools like generative AI systems offer.

[1] A. Danto, Jaac, **33**, 139 (1974).
[2] K. Wojtkiewicz, Jaac, **81**, 454 (2023).
[3] C. Anscomb, Odradek, **8**, 13 (2022).
[4] S. Schleidgen, O. Friedrich, S. Gerlek, et al., Humanit. Soc. Sci. Commun. **10**, 551 (2023).