DETAILED DESCRIPTION

[in English]

AI: Authorship and Interpretation

AIAI is an interdisciplinary project addressing fundamental questions concerning the status of authorship and interpretation of works (primarily texts and images) created by or with the aid of generative artificial intelligence such as ChatGPT, DallE or Midjourney. It combines theoretical work rooted especially in the poststructuralist tradition with interviews and empirical work, allowing a multifaceted understanding of the analyzed phenomena.

1. scientific goal of the project (description of the problem to be solved, research questions and hypotheses)

a. description of the problem to be solved

The emergence of generative AI – large language models such as GPT or text-to-image models, among others – has raised a number of pertinent questions in the field of cultural production. One of these questions concerns authorship – who or what is the author of a given text or image? Is it the person who put in the prompt (the user), the company which produced the app or tool in which the work was generated, the authors of the texts or images on which the app or tool was trained, or someone/something else altogether?

Answering such questions – and even posing them in the right way – is important for a number of domains. The current discussions focus on legal issues of authorship, especially concerning copyright (see “state of the art” below) – so they are, in fact, focussed on the question of who is the owner and who has the rights to the product, and especially the profit that is gained from it.

The proposed research asks the question of authorship on a more fundamental level, connecting it to interpretation and the status of the produced texts and images itself. As such, it finds inspiration not only in the current debates concerning AI explicitly, but also and especially in the debates concerning authorship and interpretation which have been the object of the humanities since the 19th century.

In traditional hermeneutics, the premise was that we are trying to infer "what the author meant to say", regardless of the author being human or divine. In this discourse, the notion of the author was considered as a classic, modern subject – conscious and in control of his or her intentions for the written text. While this is still a big part of how we understand authorship societally (as evidenced, for example, by the understanding of copyright issues, discussed below), theorists of authorship and interpretation complicated this problematic on a number of levels.

Crucially for the themes analyzed in this proposal, the problematization of the modern conception of the human subject lead to the undermining of the problematic of authorship, and therefore showed that – to put it very generally – that “something else” than the human was, in fact, the author, and needed to be reached through the work of interpretation. A key example of this tendency is what Paul Ricoeur famously called “hermeneutics of suspicion” (Ricoeur 1970), represented by Friedrich Nietzsche, Karl Marx and Sigmund Freud. Nietzsche’s genealogy of morality claimed that our moral judgments are not objective but can be traced to a historical process of which our culture is an effect (Nietzsche 2007); Marxist theory of ideology suggested that what “speaks" in the author depends not on their free will and individuality but rather the

economic realities of the society in which the “author” is situated (Marx and Engels 1998); Freudian psychoanalysis sees unconscious processes as the true *locus* of authorship (e.g., Freud 1900). Interpretation, thus, is not determining what the author wants to say, but what speaks through the author.

This tendency in understanding authorship and the role of interpretation has been amplified in the 20th century thanks to, among others, the advent of structuralism and post-structuralism. The structuralist belief in the privileging of general structure over individual content (e.g., Levi-Strauss 1963) has led directly to the total rejection of the notion of the author, most famously proclaimed as the “death of the author” (Barthes 1977). Building on that foundation, post-structuralist theorists have posited the goal of interpretation as deconstructing binary oppositions on which texts rely (e.g., Derrida 2016), or opting for a pragmatic reading, privileging the potential effects of the text, especially in the political realm – for example finding its “revolutionary force” (Deleuze and Guattari 1983, 106).

Throughout the 20th century, the notion of authorship has also been put into question by writers and artists, from the Dada movement and its practice of the cut-up poem (Tzara 1918) and their continuators such as William S. Burroughs (e.g., Burroughs 1989) – relying on chance rather than deliberate authorial intent, these works are perhaps closest to authorless works of art, and yet can be and have been successfully interpreted. Other artistic strategies that incorporate chance and aleatory effects are, for example, John Cage’s musical works, in which the composer uses “randomly tuned radios,” or a “random selection of electronic sounds” produced through “a series of programmed transparent cards" (Britannica 2023).

In fact, earlier AI text “generators” were more akin to these techniques than actually generating text – programs such as ELIZA (see e.g., ELIZA, nd) produced replies through a simple process of using templates to which (parts of) user prompts were copied. More complicated programs such as Racter, credited as the author of the book *The Policeman’s Beard is Half-Constructed* (Racter 1984) were pre-programmed with ready-made syntactic formulas into which they input random words, creating syntactically correct, yet mostly meaningless prose.

All this means that the current generation of generative AI has not landed in an empty landscape when it comes to theories of interpretation and theories and practices of authorship – many of already existing theories have tried in their own way to understand those terms in ways that do not presuppose a typically human author, or at least an author understood as a Modern, autonomous, conscious subject. However, all of these theories have been created with regards to works that have been produced by human authors; moreover, the aforementioned theories have overwhelmingly involved historical or social factors as the *loci* of authorship, therefore remaining anthropocentric.

The question, therefore, becomes, (1) does the advent of AI-generated works create an entirely new situation, or is it just a continuation or fulfillment of the landscape of authorship and interpretation that has already been described by existing theories; and conversely (2) can the theories that sought to move away from the model of the Modern vision of the autonomous, conscious author provide an adequate understanding of authorship and interpretation in the era of generative AI, or do they need to be changed or at least supplemented because of this new situation?

b. research questions and hypotheses

The research questions in the project will be grouped into three work packages (WP), each corresponding to a cluster of interconnected problems.

WP1. Determining the status of the author of AI-generated texts and images.

Research questions: Can we understand the process of creating texts and images with the aid of AI in terms of authorship? Are theories of authorship developed in the humanities – whether hermeneutic, psychoanalytic, structuralist or post-modern – suited to understanding the problem of authorship of AI-generated texts and images? Can ideas developed within the field of copyright law help us to understand the ontological status of the author of AI-generated material? How do (human) artists who create works with the aid of AI tools understand the problem of authorship in the process?

WP2. Understanding the status of interpretation of AI-generated texts and images.

Research questions: Does it make sense to apply methods of interpretation devised for human-made texts and images to AI-generated material? What is the status of interpretation of AI-generated material both as a practice and as a product? Can this kind of interpretation be likened to the interpretation of other non-human generated material, such as animal art? How do experts at interpretation (e.g., literary scholars or art critics) understand the process of interpretation of AI-generated work?

WP3. Developing a new framework for understanding authorship and interpretation of AI-generated works.

Research questions: What theories of authorship and interpretation should look like to encompass the changes brought about by the advent of AI-generated content? Should theories of authorship concentrate more on the process of creation or the beliefs of the person engaged in the process? Should theories of interpretation focus on the work itself, the author (whoever or whatever it is), the person interpreting, or the act of interpretation itself?

Hypotheses:

We can expect that the 20th-century schools of interpretation, such as those mentioned above, will not be directly applicable to the creations made using generative AI. They are promising, however, as sources of inspiration, focused on what *kind* of information is needed to form a relevant theory of interpretation that does not concentrate on a classically understood subject. For example, there seem to be promising parallels between the psychoanalytic idea of an unconscious structure of the mind, of which a work of art can be considered a product, and the multifaceted structure of production of text in generative AI. In more culturally-oriented theories, instead of an unconscious structure, one can see an interlinked web of ideology and cultural beliefs which “speak through” the work and are responsible for its axiological aspect; in the case of generative AI there seems to be a complicated link between the materials used for training the models (bearing the imprint of dominant culture), cultural norms embedded in how output is moderated (i.e., what can and cannot be “said” by the model) and the norms of the person prompting the model. We expect that such

analogies, treated with caution and awareness of the uniqueness of the current situation, will provide a good preliminary understanding of how authorship and interpretation of AI-generated content can be understood.

In our empirical work, we expect to find a more pronounced attachment of our research subjects to a more traditional understanding of authorship. Thus, we expect some factors that will be controlled in the experimental procedure to influence the perception of authorship. Some of these factors concern the type of work that is being done with the tool, e.g.: (1) The amount of work, including the amount of post-processing that was done by the human and by the AI tool – the more work, the more one may perceive oneself to be the author; (2) The type of interaction with the tool, especially when it comes to setting the goal of the work; (3) Amount of control over the generation process (one-shot generation from the input in one go; generation in multiple steps with an option of further control; an option to regenerate the output; a choice from multiple suggested/generated options; an option to change/refine the input; an option to request further modifications); (4) Understanding how the tool works – the more the tool is perceived as a “black box” (or its work as “magic”), the less the subject may perceive themself as the author.

Other factors concern the design of the tool itself. For example, if the tool is designed to work as if it has an opinion/preference (calling some of its own products better or more beautiful than others) or a “personality,” the subject may consider themself less of an author. If the tool is designed as “deterministic” – i.e., gives a very similar output every time – the subject may consider themself less of an author. Finally, if the tool involves user-generated data in the final product, the user may consider themself more of an author.

2. significance of the project (state of the art, justification for tackling a specific scientific problem, justification for the pioneering nature of the project, the impact of the project results on the development of the research field and scientific discipline);

a. state of the art

Literature discussing authorship in LLMs is often implicitly or explicitly (e.g. Soos et al., 2023) using the premise that the LLM itself cannot in principle have any communicative intent and understanding of language, acting only as a “stochastic parrot” (Bender et al., 2021). As intent is often considered an important part of authorship, these authors dismiss the possibility of the LLM being itself an author. However, the premise of no communicative intent and no understanding in LLMs has been contested in various ways.

Andreas (2022) shows that LLMs can be regarded as models of (populations of) agents, modeling their desires, beliefs and intentions. The representations of beliefs and desires that the model uses for generating the text can be influenced by manipulating the inner state of the model, or giving various instructions in the model’s generation prompt. Depending on the inner state and/or the prompt, the model may display different “personas" in the generated content. Søgaard (2022) shows that while LLM’s ability to learn referential semantics is limited, it is not impossible; also, learning inferential semantics is possible based on textual inputs only. We intend to investigate the possibility of regarding LLMs as authors in various frameworks based on the views that oppose the “stochastic parrot” view.

The discussion concerning AI and authorship has a long history in the field of copyright law, which focuses on questions of ownership of the given work or the material that was used to generate it. The issue is currently (as of 2023) a main focal point on the agenda of the US Copyright Office (US Copyright Office 2023). Aside from general inquiries into the matter (e.g., D'Agostino 2023, Thorne 2020) and proposals for legal frameworks (eg. Kasap 2019, Ciani 2019, Lu 2021), researchers ask specific questions, for example, what kind of recognition should be given if AI is used in the creation of an artistic work (Deltorn and Macrez 2019, Ajani 2020, Lee 2023,), or a scientific paper (e.g., Solomon et al. 2023, Polonsky et al. 2023, Balat and Bahşi 2023, Pourhoseingholi et al 2023); or to what extent the use of AI in different kinds of works constitutes plagiarism (e.g., Anson 2022). These questions, however, do not problematise the question of the human author, remaining firmly embedded in the Modern paradigm of authorship; they usually (with few exceptions e.g., Murphy 2022/23) do not acknowledge any of the theories of the author that were exemplified in the first part of this proposal.

The discussion of interpretation is limited to an experiential perspective of the receiver of a given work of art (e.g., Di Dio et al, 2023) or a journalist text (Henestrosa et al, 2023); the experience of reading AI texts was also the topic of numerous pieces of opinion journalism (notably Chomsky 2023).

b. justification for tackling a specific scientific problem

The popularization of LLM-based AI tools and image generators has created a dire need for researchers to understand the impact of those tools on a number of activities and professions. The proposed research is based on the premise that the problems of authorship and interpretation are of key importance especially for such disciplines in the humanities as cultural studies, literary studies, philosophy and art history. The researchers practicing these disciplines rely on robust theories of interpretation which serve as foundations for their work – a given interpretation is only valid in light of knowledge of what interpretation is. Given the plausible expectation that in the near future more and more of texts, images and other works we encounter will be produced by or with the aid of AI tools, the understanding of what it means to interpret such work will be paramount.

c. justification for the pioneering nature of the project

As clear from the "state of the art" above, the project takes a more fundamental approach than usual discussions of the problems of authorship and interpretation; the latter problem is rarely, if ever, analyzed on its own – it seems safe to say that this will be the first attempt to fundamentally analyze the status of interpretation of AI-generated works.

The interdisciplinary nature of the research is also a novelty, especially the combination of the psychoanalytic, structuralist and poststructuralist theories of interpretation with social-scientific methods of inquiry, since the two types of work are usually performed by scholars with a very critical attitude to the approach that they are not using – poststructuralism especially is shunned as "unscientific" by more classically minded social scientists, and classic social science is considered "uncritical" by poststructuralists. We strongly believe that only through a critical combination of the two attempts can we rise to the challenge of generative AI.

d. the impact of the project results on the development of the research field and scientific discipline

The project, interdisciplinary in its nature, will impact a number of fields and disciplines. From the perspective of philosophy, as well as literary and art theory, it will (1) provide an assessment of currently existing theories of interpretation with regards to their utility to AI-generated content; (2) determine what a theory of interpretation should consist of to provide an understanding of the stakes of interpreting AI-generated content; (3) provide a better understanding of the concept of authorship of works that have been created with the aid of AI tools. From the perspective of theory of education, it will help to understand the usefulness of AI tools in teaching of interpretation. Indirectly, the proposed research will influence the discussion of authorship in other domains, e.g., in the field of copyright law.

3. concept and work plan (general work plan, specific research goals, results of preliminary research, risk analysis);

a. general work plan

The first two work packages will be divided into three phases – (1) theoretical research; (2) expert interviews; (3) empirical research. The first year of the project will focus on the first two phases of WP1 and WP2 – the theoretical research (explained in 4b below) will help to design a frame for discussing the topic with external experts. These interviews, in turn, will lead to the construction of the empirical part of the research, which will involve testing the awareness of authors in a controlled, experimental environment.

The empirical research will be planned, prepared and performed in year two of the project during two meetings conducted by the whole project team (i.e., both Czech and Polish teams) in Warsaw and Prague. This will allow two iterations of the same research in two different millieus and languages (Czech and Polish) to ensure a better understanding of both general and country/culture-specific elements in the results.

The third year of the project will be entirely devoted to WP3, which will include a summary of the results of the two previous phases and the establishment of a new theoretical framework for interpretation and authorship of works created by or with the aid of AI tools.

Each year, we will organize two physical project meetings, one in Prague and one in Warsaw. These meetings will serve for a close collaboration of the research teams, but will also involve workshops with external participants: (1) In the first year, we will invite experts with practical experience in the field and jointly conduct some of the expert interviews. (2) In the second year, we will jointly perform some of the participant studies. (3) In the third year, we will invite relevant members of the scientific community to present and discuss our findings.

b. specific research goals

Year 1: WP1 and WP2 theoretical research will yield an in-depth theoretical understanding of the problem of authorship and interpretation of AI-generated works. This will allow us to design questions for external experts and to formulate hypotheses informing the empirical part of the research. In this phase, we will write no less than two theoretical scientific papers concerning the usefulness of current theories of interpretation and authorship for tackling problems related to AI-

generated and -assisted works. We will also rework parts of the expert interviews into at least one article meant for a non-scientific outlet.

Year 2: WP1 and WP2 empirical research will allow us to comprehend these theoretical problems in the practical sense, i.e., to understand the perspective of the authors of works and of interpretations. The results will be published in at least two scientific papers.

Year 3: WP3 will result in developing a new understanding of authorship and interpretation with reference to AI-generated works. This framework will help to pose new questions in the area and thus continue empirical research. The results will be published in at least two scientific papers. We will also seek to promote our results in non-scientific outlets, from print media to popular podcasts.

c. results of preliminary research

The members of the Czech team have been previously involved in the THEaiTRE project, which focused on generating theater play scripts using a newly devised AI-based tool operated by experienced playwrights and theater dramaturges. As the goal of the project was to produce a theater play “written by an artificial intelligence system”, the question of authorship naturally implicitly arised within the project (Schmidtová et al, 2022, chapter 5), even though research on authorship and interpretation was not the main goal of the project and thus was only discussed to a limited extent.

Nevertheless, we empirically arrived at a clear distinction of the human versus machine authorship as perceived by the operators of the tool, depending on the design of the interface and the options that the user had at their disposal. We worked with two distinct designs of the tools, varying the amount of control over the generation process. In the first variant, the play script was generated line by line, with the operator selecting one of five generated variants for each line to be used at each step. In the second variant, 10 lines were always generated at once, with the operator having the choice of either using them, or discarding them and having a new set of lines generated. Within the project, all the operators unanimously agreed that with the first variant of the tool, they considered themselves to be the main author of the script (with the tool rather providing suggestions than acting as an author), while in the second variant, they found the AI tool to be the main author (with the operator being rather a beta-reader merely guiding the artificial author), even though the underlying text-generating language model was identical in both variants.
We believe that this shows the grave importance of the nature of the interaction of the user and the machine on the attribution of authorship.

Starting in 2024, the members of the Czech team will be involved in a follow-up “EduPo” project, focused on neural modeling and generation of poetry. Again, the project is rather applied, with little focus on theoretical research questions, but it does involve literature experts as well as artists. Therefore, we believe that the EduPo project will serve as a further source of valuable practical experience with text generation, which it will provide through the experts involved in the proposed project. The proposed project might also use tools developed within the EduPo project for further empirical experiments.

d. risk analysis

We envisage four main types of risk: (1) Theoretical risk involved in the premise of the research; (2) Practical risk stemming from the quickly advancing knowledge in the field of AI; (3) Practical risk related to the character of the empirical research; (4) Possible bias stemming from the relative homogeneity of the possible research subjects and research teams.

(1) Theoretical risk revolves around the pertinence of the research questions and possible answers that the research can provide; it relates predominantly to the problem of interpretation (since the pertinence of the question of authorship is well established in literature), in which case it might turn out that no meaningful or useful understanding of interpretation of AI-generated works can be found – i.e., interpretation of such works is meaningless. However, even in this improbable scenario, such a nihilist view of interpretation would be of scientific value.

(2) While studying relevant literature will be the basis for our work, the current development in the field is so fast that the findings published in literature are considerably lagging behind current developments. We will update the research questions and methodology based on the current development in the field and we envision close contact with theoreticians and practitioners in the field so that we stay maximally up to date.

(3) The empirical research will face two challenges in particular which both stem from the quickly changing legal and ethical rules and regulations surrounding AI:

a) The exact effects of the EU AI Act and other legal regulations cannot be exactly envisaged at this point, but they may mean that some parts of our methodology might become legally complicated or impossible to carry out. We will need to adapt these by changing the methodology, or changing the selection of participants and participant consents. In the unlikely case that there will be no meaningful way to carry out the envisioned experiments due to regulatory changes, we will need to omit experiments and instead focus on extended interviews, or our own autoethnographies of work with AI.

b) Similarly to law regulations, the views on AI and AI tools from the ethical point of view are quickly changing. In itself it is an immensely pertinent object of study and our research team will stay on top of current developments. However, it also constitutes a threat to our research plans, as we will make sure to abide by the ethical principles valid at the time and place of the research. Therefore, we will need to watch closely the development of the ethics dimension of our planned research, and adjust our plans accordingly. We also intend to voluntarily undergo ethics committees reviews when the project is approved, and whenever we feel unsure about specific ethical aspects of our planned research.

(4) Risk stemming from the relative homogeneity of the research subjects and researchers groups:

(a) Wherever possible, we will make our participant groups as balanced and representative as possible, but given the group sizes and the qualitative instead of quantitative research methodology, a perfect balance cannot and will not be ensured. The proposed methodology inherently introduces biases towards inhabitants of Central Europe, with higher education, with interest in artificial intelligence and modern technologies. We also expect our

participant groups to be biased towards young and middle age. On the other hand, we believe to be able to ensure a good gender balance in the groups, including minorities.

(b) All of the currently planned team members are white males from Central Europe. The Polish team will seek to balance this bias by the inclusion of a relevant third member. Even so, the relative homogeneity of the group may limit our views and skew our research proposals and findings. To compensate for this, we will make efforts when researching existing literature and interviewing experts external to the project to explicitly also include views of non-white and non-male researchers and experts. However, we expect not to be able to get rid of the Central European bias, and thus we will underscore this limitation in discussing our findings.

4. research methodology (underlying scientific methodology, methods, techniques and research tools, methods of results analysis, equipment and devices to be used in research);

a. underlying scientific methodology

The research methodology – reflected by the composition of the teams – is based on the premise that only a multifaceted and interdisciplinary approach can be successful in shedding light on the problem of authorship and interpretation of AI-generated content. This approach allows to understand the problem both on an external/ontological level (i.e., asking who *is* or *can be considered* an author and what constitutes valid interpretation) and an internal/anthropological level (i.e., asking who and under what circumstances *considers themself* the author or what experts consider themselves doing when interpreting AI-generated content).

The proposed research will be informed by in-depth knowledge of both the workings of LLM systems and contemporary theories of interpretation to ascertain the external/ontological understanding and include empirical research (interviews and experimental procedures described below) to provide the internal/anthropological understanding.

b. methods, techniques and research tools

The project will proceed through the use of interdisciplinary methods, ensuring a multifaceted approach to the problem of authorship and interpretation. Theoretical research will begin with a thorough analysis of the potential of already existing theories of interpretation and authorship to help us understand those concepts in reference to AI-generated content. Focus will be placed on theories which complicate the notion of the author, i.e., point to processes or structures other than individual human subjectivity as the locus of authorship (see above for candidates for such theories).

Following the first phase of theoretical research, we will design the structure and content for a set of expert interviews, consisting of a combination of questionnaires and structured and semi-structured interviews. We will then interview experts with first-hand experience of authorship of works co-created by AI (i.e., creators, such as writers, playwrights and artists) and their interpretation (i.e., literary scholars and art critics). In this phase, we will not manipulate the creation process used by the creators in any way; the information we gain will thus pertain to a potentially highly varied range of human-AI co-creation setups and approaches, as used by the

creators in practice, providing us with an understanding of the usual approaches. The goal of these interviews is to confront our theoretical findings and initial hypotheses with the real-world experience of the experts, testing some of our assumptions as well as gaining further knowledge and opinions that will help us to refine our understanding of the problem and reformulate our hypotheses.

Based on our research of existing literature and discussions with experts, we will propose a set of features that may have effect on the ascription of authorship by the user of the AI system; a preliminary set of candidates for such features is listed within the Hypotheses section. We will then empirically investigate the effect of these features on the authorship ascription through surveys of non-expert participants who will be debriefed after periods of work as creators using AI tools in a controlled environment. For these experiments, we will adapt or design a set of AI tools and interfaces, as well as tasks to fulfill by the participants, differing in some of the features that we identify as potentially important for authorship. Contrary to the expert interviews, we will now manipulate the creation process so that we can reliably investigate the particular features that we are interested in; even though some features cannot be manipulated, such as the creator's background or the particular result of the creation process, and will need to be collected, observed and taken into account. We will collect data from the participants through a set of questionnaires composed of close-ended questions as well as open-ended questions, including both their immediate reflections on the creation process and the authorship and interpretation of the results, as well as various data about their background which we identify as potentially relevant (socioeconomic factors, experience with AI tools, creative experience, understanding of the AI technology, etc.). We plan to survey 80 participants in total in this way.

c. methods of results analysis

The results of the participant surveys will be processed and analyzed. Our primary interpretation of the survey results will be qualitative, as we will operate with a low number of participants and in a situation with a large number of factors influencing the observations. However, we will also try to induce small-scale quantitative findings if the collected results offer sufficient support for that, which will be determined by appropriate statistical methods such as significance testing and measuring correlation coefficients.

Based on the results and analyses of the surveys, we will formulate our final set of findings and hypotheses. We will then use these to build a comprehensive theoretical framework for understanding the authorship and interpretation of works co-created by using AI tools, which will be the main result of our project. The framework will thus be based on previous theoretical works, on our theoretical research, as well as on our empirical research. We will explain the framework in scientific publications as well as publications aimed at the general public, and present it and discuss it with experts in the field.

d. equipment and devices to be used in research

We primarily rely on resources provided by our institutions, including access to relevant literature as well as hardware and software equipment. We will only purchase resources which we find to be unavailable at the institutions while being crucial for the research objectives. Among other, we

envision the need to buy access to paid AI tools, as state-of-the-art tools used by creators in practice tend not to be freely available.

For AI tools that can be obtained freely, installed and operated locally, we will use the large computational cluster of the Czech partner, which has sufficient capacity of operating current medium-sized neural models.

5. project literature (a reference list for publications included in the project description, with full bibliographic data).

Ajani, G. (2020). *Contemporary artificial art and the law: searching for an author*. Boston: Brill, 2020.

Andreas, Jacob. "Language Models as Agent Models." Findings of the Association for Computational Linguistics: EMNLP 2022. 5769–5779.

Anson, C. M. (2020). AI-Based Text Generation and the Social Construction of "Fraudulent Authorship": A Revisitation. Composition Studies, 50(1), 37-46.

Balat, Ayşe & Bahşi, İlhan (2023). We Asked ChatGPT About the Co-Authorship of Artificial Intelligence in Scientific Papers. *European Journal of Therapeutics* 29 (3):e16-e19.

Barthes, R. (1977). "The Death of the Author," in *Image, Music, Text*, Transl. Stephen Heath, New York: Hill & Wang,, pp. 142-148.

Bender, E. M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 610–623. https://doi.org/10.1145/3442188.344592

Britannica, T. Editors of Encyclopaedia (2023, October 13). *John Cage*. Encyclopedia Britannica. https://www.britannica.com/biography/John-Cage. Accessed November 24, 2023.

Burroughs, W. S. (1989) The Job: Interviews with William Burroughs, with D. Odier. New York: Penguin.

Chomsky, N.,Roberts, I. and Watumull, J. (2023). Noam Chomsky: The False Promise of ChatGPT. The New York Times, 8 March. https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html. Accessed November 30, 2023.

Ciani, J. (2019). Learning from Monkeys: Authorship Issues Arising from AI Technology. In: Moura Oliveira, P., Novais, P., Reis, L. (eds) *Progress in Artificial Intelligence*. EPIA 2019. Lecture Notes in Computer Science(), vol 11804. Springer, Cham. https://doi.org/10.1007/978-3-030-30241-2_24.

D'Agostino, S. (2023). AI Raises Complicated Questions About Authorship. *Inside Higher Education*, August 22. https://www.insidehighered.com/news/tech-innovation/2023/08/22/ai-raises-complicated-questions-about-authorship. Accessed 30 November 2023.

Deleuze, Gilles, and Felix Guattari. 1983. *Anti-Oedipus. Capitalism and Schizophrenia*. Transl. Robert Hurley, Mark Seem and Helen R. Lane. Minneapolis: University of Minnesota Press.

Deltorn, Jean-Marc and Macrez, Franck. (2019). Authorship in the Age of Machine learning and Artificial Intelligence). In: Sean M. O'Connor (ed.), The Oxford Handbook of Music Law and Policy, Oxford University Press. Available at SSRN: https://ssrn.com/abstract=3261329 or http://dx.doi.org/10.2139/ssrn.3261329.

Derrida, J. (2016). *Of grammatology*, trans. G. C. Spivak. Baltimore: Johns Hopkins University Press.

Di Dio, C., Ardizzi, M., Schieppati, S. V., Massaro, D., Gilli, G., Gallese, V., & Marchetti, A. (2023). Art made by artificial intelligence: The effect of authorship on aesthetic judgments. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000602.

ELIZA (nd). ELIZA: a very basic Rogerian psychotherapist chatbot. https://web.njit.edu/~ronkowit/eliza.html. Accessed 1 December 2023.

Freud, S. (1900). "The interpretation of dreams," in *The Standard Edition of the Complete Works of Sigmund Freud*, Vol. 4, 5, London: Hogarth Press.

Henestrosa A., Greving H., Kimmerle, J. (2023). Automated journalism: The effects of AI authorship and evaluative information on the perception of a science journalism article. *Computers in Human Behavior*. Volume 138, January, 107445. https://doi.org/10.1016/j.chb.2022.107445.

Kasap, A. (2019) Copyright and Creative Artificial Intelligence (AI) Systems: A Twenty-First Century Approach to Authorship of AI-Generated Works in the United States. *19 Wake Forest J. Bus. & Intell. Prop. L.* 19(4), 337-360.

Lee, Edward, Prompting Progress: Authorship in the Age of AI (October 22, 2023). Florida Law Review, Vol. 76, 2024 Forthcoming, Available at SSRN: https://ssrn.com/abstract=4609687 or http://dx.doi.org/10.2139/ssrn.4609687.

Levi-Strauss, C. (1963). *Structural Anthropology*. Trans. by C. Jacobsen & B. Grundfest Schoepf. Allen Lane: The Penguin Press.

Lu, B. (2021). A theory of 'authorship transfer' and its application to the context of Artificial Intelligence creations. *Queen Mary Journal of Intellectual Property*, 11(1), 2-24. Retrieved Nov 30, 2023, from https://doi.org/10.4337/qmjip.2021.01.01.

Marx, K. and Engels, F. (1998). The *German Ideology, including Theses on Feuerbach*. Amherst: Prometheus Press.

Murphy, P. (2022/23). Writers and Writers of Writers: Creativity and Authorship in the First AI Novel. *Kritikos. an international and interdisciplinary journal of postmodern cultural sound, text and image*. Volume 19, Fall/Winter 2022/2023, ISSN 1552-5112.

Nietzsche, F. (2007). *On the Genealogy of Morality*. Translated by Carol Diethe. Cambridge: Cambridge University Press.

Polonsky, M. J., & Rotman, J. D. (2023). Should Artificial Intelligent Agents be Your Co-author? Arguments in Favour, Informed by ChatGPT. Australasian Marketing Journal, 31(2), 91-96. https://doi.org/10.1177/14413582231167882.

Pourhoseingholi, M. A., Hatamnejad, M. R., & Solhpour, A. (2023). Does chatGPT (or any other artificial intelligence language tool) deserve to be included in authorship list?. *Gastroenterology and hepatology from bed to bench*, 16(1), 435–437. https://doi.org/10.22037/ghfbb.v16i1.2747.

Racter (1984). *The Policeman's Beard is Half-Constructed*, New York: Warner.

Ricoeur, P. (1970). *Freud & Philosophy: An Essay on Interpretation*. Translated by Denis Savage. New Haven and London: Yale University Press, 1970.

Schmidtová, P., Rosa, R., Košťák, D. et al. (2022). THEaiTRE: Generating Theatre Play Scripts using Artificial Intelligence. Praha: Institute of Formal and Applied Linguistics.

Solomon, D. H., Allen, K. D., Katz, P., Sawalha, A. H., & Yelin, E. (2023). ChatGPT, et al … Artificial Intelligence, Authorship, and Medical Publishing. *ACR open rheumatology*, 5(6), 288–289. https://doi.org/10.1002/acr2.11538

Soos, Carlin & Haroutunian, Levon. (2023). On the Question of Authorship in Large Language Models (LLMs). NASKO. 9. 1-17. 10.7152/nasko.v9i1.16299. https://doi.org/10.7152/nasko.v9i1.16299

Søgaard, A. Understanding models understanding language. *Synthese* **200**, 443 (2022). https://doi.org/10.1007/s11229-022-03931-4

Thorne, S. (2020). Hey Siri, tell me a story: Digital storytelling and AI authorship. Convergence, 26(4), 808-823. https://doi.org/10.1177/1354856520913866.

Tzara, T. (1918). “Manifeste Dada,” *Dada* vol. 3, December.

US Copyright Office. (2023). Copyright and Artificial Intelligence. Retrieved from https://www.copyright.gov/ai/. Accessed November 30, 2023.