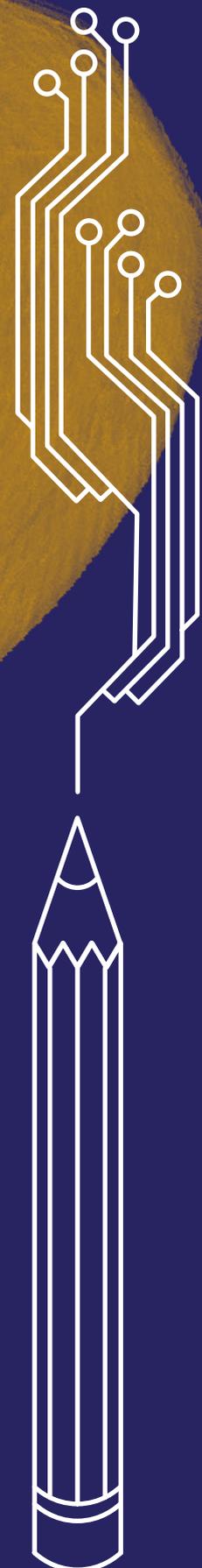Bournemouth University

Shared-Posthuman Imagination: Human-AI Collaboration in Media Creation

White Paper

Szilvia Ruszev, Liam Rogers, Selin Gurgun, Melanie Stockton-Brown, Maxine Gee, Tom Davis, Xiaosong Yang, Catherine Griffiths, Kejun Zhang, Boyuan Cheng, James Slaymaker and Stephanie Prajitna
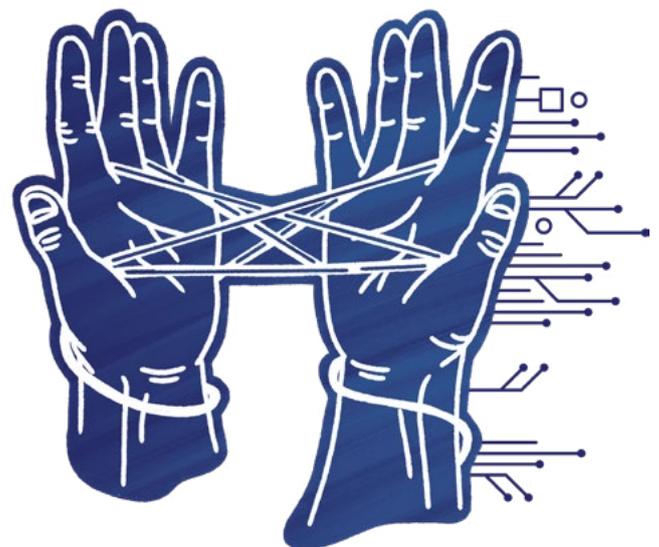
October 2024

Bournemouth University

BRAID
Bridging Responsible AI Divides

UKRI
Arts and Humanities Research Council

Contents

List of Tables and Figures

Executive Summary

The UK creative industries contribute over £124 billion to the UK economy annually and account for more than 5.7% of the total UK Gross Value Added (GVA).[1] Yet, there are feelings of uncertainty regarding what impact generative AI tools like ChatGPT, Midjourney and Suno are going to have on the creative industries. Undoubtedly, generative AI is revolutionising media production, changing the parameters of what constitutes creativity, authorship and ownership in unprecedented ways. Now more than ever, users of generative AI can produce stories, scripts, images, music and even entire films simply by prompting widely available, often free-to-use AI models that have been trained on large datasets through machine learning processes. While the capabilities of generative AI tools to perform these tasks to a human-level quality are, in some instances, still limited, they nonetheless bring with them a series of moral, ethical and legal challenges that need to be addressed as a matter of urgency. At present, there currently exists little in the way of guidance, regulation and recommendations for best practice regarding how to integrate generative AI, specifically, into UK media production in a responsible and ethical way. Hence, this research emerges as a timely intervention into the creative industries by assessing the limitations and opportunities of using generative AI in media creation.

---

[1] Creative Industries Council, 'UK Creative Industries Add £124bn to UK Economy', 18 June 2024 <https://www.thecreativeindustries. co.uk/site -content/uk-creative-industries-add-ps124bn-to-uk- economy> [Accessed 15 August 2024].

The aim of this report is to document the findings of 'Shared Post-Human Imagination: Human-AI Collaboration in Media Creation' as a research project. The project set out to interrogate and understand the impact that generative AI tools are having on the concepts of creativity, bias and collaboration within media production, where questions of control, agency, skill, labour, exploitation and representation are particularly pertinent to the hopes and fears of the creative industries. It did so by bringing together stakeholders from different parts of the generative AI media landscape, seeking to foster relationships amongst developers, filmmakers, policymakers and end users. By facilitating discussions across AI divides, this project revealed a complex landscape in which the perceptions, impact and applications of AI tools vary across different industrial contexts. The purpose of this report is to scope this complex landscape, considering what responsible media production with AI might look like and how different stakeholders can play a part in making this a reality. It does so by providing a set of practical guidelines,

practices and policy recommendations designed to promote the responsible use of generative AI in media creation.

As a research team comprised of an international and interdisciplinary group of academics and filmmakers working at Bournemouth University, Zheijang University (China) and Columbia University (US) across the disciplines of Media Production, Computer Animation, Data Mining and Machine Learning, Music and Audio Technology, and Law, we adopted a mixed-methodology approach to conduct this research.

Through a series of four two-day workshops, we invited academics, media professionals and end users to participate in practical-reflective exercises designed to both facilitate participant experimentation with generative AI tools and ascertain the prevailing hopes and concerns about generative AI's impact on screen-writing, image creation, editing, and sound and music through systems thinking. After each workshop, these critical reflections on the topics of creativity, bias and collaboration were analysed through a combination of quantitative and qualitative methods, affording our research team valuable insights into the scope of generative AI media production as an issue.

Following these workshops, we convened an Expert Bridging Group of selected filmmakers, producers, academics, lawyers, policymakers and union representatives, designed to workshop solutions regarding best working practice and potential policy recommendations for responsible AI in media production.

Concurrently, our research team combined this engagement of diverse stakeholders with creative practice, through which we are creating a short film in collaboration with various generative AI tools. In doing so,

we have sought to follow the responsible AI principles and guidance suggested to us across the four workshops and the Expert Bridging Group in order to reflect upon the practicalities of what responsible media creation with generative AI means for independent filmmakers and media practitioners.

We present the detailed findings of our research in this report. Chapters One and Two will offer readers a survey of the literature and our methodologies, respectively, while Chapters Three, Four and Five will chart through the industry desires and concerns as revealed from our research. Then, in Chapter Six, we present a list of necessary interventions that can help foster responsible AI use in media production. A preliminary summary of these findings follows below.

Within the media landscape, there exists both concern and optimism regarding how generative AI tools will change and are currently changing the role of creativity. Many creatives critique generative AI's inability to reproduce 'authentic' human emotion, leading to a potential degradation of artistic value as algorithmic computation is unable to capture subtext and ambiguity. This leads to concern regarding AI outputs being unoriginal and 'generic', a consequence of how generative AI models are trained to produce the most likely outcome to a human prompt based on synthesising pre-existing data. While some see this as a negative impact on creative skills, for others, generating media with AI constitutes not a loss but a shift in what we mean when we talk about 'skill'. Most of the end users who took part in our research recognised that successfully integrating AI into one's workflow requires a high degree of technical skill and capability, indicating that working with AI is more complex than people might assume. Participants also praised generative AI's potential to spark ideation through

experimentation and the production of things like loglines, storyboards, pitch decks and mood boards, as well as freeing up human creativity by undertaking more automated work of organisation and management, tasks well-suited to algorithmic computation. Yet, this last point is inherently tied up with assumptions that certain parts of the film production process (e.g. screen-writing and music production) are more creative than others (e.g. editing), an assumption that this report seeks to challenge.

Bias in AI systems often stems from biases in the training data and biased training procedures by technicians and engineers. This can lead not only to a lack of diversity in the generated outputs of AI media, but also potentially harmful productions that perpetuate stereotypes and disproportionately disadvantage marginalised communities. Since generative AI tools are trained to produce the most likely desired outcome, and training data is based on a pre-existing canon of public domain images and narratives of the past, then outcomes will inevitably reproduce the same ideological conditions that the original works were produced under. As such, creatives expressed concern about media produced with generative AI being more likely to be misogynistic, racist, homophobic, transphobic and ableist, potentially shaping perceptions about what certain identity groups look like for future generations.

While this emerged as near universal concern amongst participants and experts, our research found that there lacked a clear direction about how to address bias as an issue beyond the need for intervention. For some, this means integrating an awareness of how power and prejudice function intersectionally when working with AI models, as well as increasing the diversity of those involved in the training and development

phases of AI production. For others, more radical interventions are needed, such as mandatory education on ethics and bias for those at board level in the tech industry, and the establishment of risk assessments for generative AI models to determine and correct harmful biases before a model is released for public use. While generative AI tools have the potential to be democratising for those historically under-represented in the media landscape, imbalances of power and commercial interests often stand in the way of redressing the issue of bias, prompting the need for policy intervention.

Questions of collaboration were largely concerned with where to place the AI's role in media production and the subsequent ethical and legal ramifications that arise from this. For some, generative AI serves as a tool that facilitates ideation, aiding curiosity through educative experimentation and enabling amateurs to produce outputs of high quality. However, the possibility of AI collaboration to diminish human-human collaboration remained a consistent theme across our research, such that positions of authorship and ownership are rendered unstable with the incorporation of generative AI tools. There exists a feeling of insecurity within the creative industries that the widespread use of generative AI in media production will lead to a circumnavigation of human labour and widespread loss of employment if outputs can be produced faster and cheaper by big studios.

Similarly, the fact that the data being used to train generative AI models is being sampled often without the original creative's consent has led to feelings of distrust amongst the creative industries. This poses significant copyright concerns about whether this amounts to unlawful copyright infringement of the works in the dataset. As well as clear disclosure mechanisms for both data and

AI outputs themselves, creatives want to see fair compensation and remuneration of any work used to train Large Language Models (LLMs). The current ambiguity around copyright ownership of generative AI outputs, and of whether the training of these LLM models on the datasets amounts to copyright infringement, has resulted in a some hesitation to use generative AI for media production until these issues are clarified for fear of both illegal activity and of undertaking unethical working practices.

In order to tackle these problems pertaining to creativity, bias and collaboration, this report sets out a series of solutions designed in consultation with our workshop participants and members of our convened Expert Bridging Group. These solutions can be separated into two categories: recommendations for best working practice (designed with independent filmmakers and practitioners working with generative AI models) and proposals for governance and policy interventions (designed to promote legislative and regulatory change within the UK media landscape).

Based on our research both within the field of responsible AI and gathering the thoughts of different stakeholders within the UK creative industries, we propose the below recommendations for best working practice.

When working with generative AI tools for media production, we recommend following these core principles:

- **Accountability:** Keeping clear data records tracking the timeline and responsibilities of key decisions to prevent retrospective 'ethics washing' practices.

- **Transparency:** Make these data records easily accessible, including disclosing how generative AI has been used in any production process at the point of media release.

- **Redress bias:** Use of generative AI tools with an active awareness that generative AI models can produce biased, stereotyped and sometimes harmful outputs, as they are replicating underlying bias within the datasets on which they are trained. Redress this by embedding intersectionality and cultural specificity into your prompts as much as possible.

- **Collaboration:** Recognise the need for both human and machine labour in production processes, including the material financial consequences for human creatives.

- **Interdisciplinarity:** Ensure diversity of stakeholders in the production process, both in terms of identity and technical background, to mitigate potential bias and prevent homogenised outputs.

- **Informed participation:** Ensure that the use of any stakeholder data, including images and voices that can be manipulated digitally with AI, falls clearly within the original terms as set out in any contract or agreement. Consent should be informed, affirmative and opt-in, rather than opt-out. Ensure legal compliance with data protection, privacy, copyright and

Intellectual Property laws when handling this data.

- **Open datasets (where applicable):** For those with the technical ability, training and access to do so, develop and use your own localised Large Language Model, trained on data you already have the copyright for, that adheres to the principle of informed consent (and fair remuneration where appropriate) for artists whose work you may be using in the training process. Make this dataset open and transparent to the public upon release of any finished artefacts. For users who do not have the means to do this, further training and education are required to help people transition toward the development of their own models.

As well as following these principles for best working practice, media practitioners seeking to embed responsible AI practices into their workflows can also ask themselves critical-creative reflective questions on the purpose of their project, the necessity of using particular generative AI tools to achieve certain tasks, and whether these same tasks can be achieved without the use of AI. These critical-creative reflective questions will be unpacked in greater detail in Chapter Six.
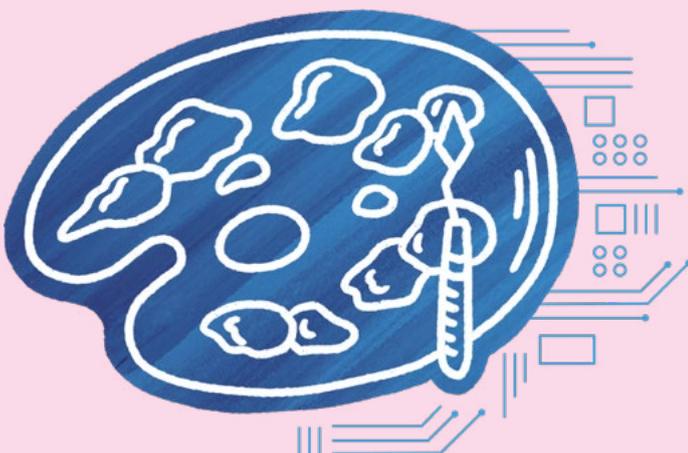
While our research findings suggest that these recommendations for best practice will help foster a responsible AI climate within the UK creative industries, they need to be accompanied by the development of a regulatory framework for generative AI if we want to bring about lasting societal, cultural and economic change. As such, our research proposes the following **12 policy interventions** that we call on the UK government to investigate and implement as a matter of urgency.

1. **Amendments to current UK copyright law**, aligning to the relevant provisions of the 2024 EU AI Act's position on AI transparency and data labelling (Article 50), which includes the disclosure of any copyrighted training materials used by generative AI developers in the training of new and existing AI models.

2. **New UK legislation** to protect actors and voice actors against the storage and/or use of their data beyond the agreed scope of a given media project.

3. **Set up a UK Generative AI Regulator for the Creative Industries**, who would establish a sector-wide Code of Practice based on a collective licensing model to assist in the monitoring, regulation and governance of the UK media industry's use of generative AI.

4. **Establish a UK Generative AI Expert Council**, formed of interdisciplinary experts in the fields of computer science, education, philosophy, law and media, who can be used for consultation by large technology companies seeking to integrate issues of AI ethics into their development of new and existing AI models.

5. **The diversification of generative AI developers**, particularly during hiring processes for trainers and technicians. Investigate the possibility of incentivising access schemes in education to achieve this.

6. **Call on UK generative AI developers to consult with leading experts in disability and accessibility** when developing, modelling and training new and existing generative AI models.

7. **Call on the UK government to champion UK Higher Education and further education institutes** to become leaders in educating filmmakers on how best to use generative AI tools in their craft, which would be aided by embedding experimentation with AI and responsible AI education into the national school curriculum.

8. **Develop a scalable, sector-wide ethical standard on the use of generative AI in media production** that imposes higher standards on large filmmaking and media projects with higher budgets.

9. **Build public confidence and trust in generative AI filmmaking** through thought leadership and media campaigns focused on positive narratives about generative AI.

10. **Develop a sector-wide 'Bias and Risk Matrix'** of different generative AI tools that can be updated continuously to enable UK filmmakers and media practitioners to make informed decisions about their own ethical use of generative AI.

11. **Crowd source data labelling for Large Language Models** to improve the quality of meta-data.

12. **Establish an Ethical AI Accreditation Scheme**, in line with BAFTA Albert, to assess the ethical and sustainable use of generative AI in UK-based film and media productions.

While pursuing and implementing these strategies will not eradicate the issues that generative AI tools are having on UK-based media production, given the dynamic nature of AI as an ever-changing technological landscape, we believe that they will go some way to redressing the current knowledge, access and power imbalances that exist within the creative industries and start the process of facilitating a truly equitable and responsible use of generative AI in media.

The main objective of this work is to scope the limitations and opportunities of using generative AI tools in media creation, determining the needs of different stakeholders so that we can gain insight into what interventions are required to foster a responsible AI media landscape. If these needs are not addressed, we risk further destabilising the creative industries and their social, cultural and economic contributions to the UK more broadly. The chapters that follow in this report translate these findings into a practical framework designed to shape guidelines, policy and best working practice. By doing so, we hope to foster important discussion regarding the principles of responsible AI in the development implementation and regulation of AI technology in media production.

Research Team and Acknowledgements

This research was primarily conducted at Bournemouth University, a leading public university in the UK whose Media Production Department and National Centre for Computer Animation, both of which host esteemed research communities with national and international recognition. As part of this research, we assembled an interdisciplinary group of scholars that bridges the fields of computer science, media production, law, music, film and television studies and animation. Doing so afforded us the opportunity to approach responsible AI in media production from a range of perspectives, thus diversifying our research and strengthening its impact. This holistic approach to research is reflected in how this White Paper is written, such that, as the product of a collaborative project undertaken across discipline and sector divides, this report highlights our team's multiple perspectives and voices in ways that cannot be tied down to a single semantic style. We sincerely thank the advisory board members, Prof. Holly Willis, USC, Prof. Yueting Zhuang, Zhejiang University and Prof. Lijun Sun, Beijing Film Academy for their invaluable guidance and feedback throughout the project.

Principal Investigators

Dr Szilvia Ruszev is a film editor, media artist and practice-based researcher who completed her PhD at USC in the Interdivisional Media Arts and Practice Program. She was previously part of the Google-funded research project 'Snowflake', a collaborative design environment for transmedia workflows combining machine learning and AI for search, image analysis and scene comparison.

Dr Melanie Stockton-Brown is a Lecturer in Intellectual Property Law at the University of Reading. She has extensive research experience concerning copyright's influence within film archiving and filmmaking.

Dr Maxine Gee is a screenwriter and researcher. As well as being a Principal Academic in Screenwriting, she developed a short screenplay titled, A Meeting, in collaboration with an AI generator as part of the "AI as Author: Bridging the Gap Between Computational Creativity Techniques and Literary Theory" project.

Dr Tom Davis is an Associate Professor in Music and Audio Technology whose practice-based work utilises machine listening and machine learning as part of his creative process. He is particularly interested in exploring notions of the performer-instrument relationship as a co-constitutive coupling between humans and technologies within a post-phenomenological framework..

Professor Xiaosong Yang leads the MSc Artificial Intelligence for Media at Bournemouth University. With over 30 years of experience in research, education and professional practice, he has developed new techniques using AI in the animation production pipeline.

International Co-Investigators

Dr Catherine Griffiths is a media artist, designer and researcher at the Columbia University, USA, where she explores critical code and algorithmic aesthetics in the context of machine learning ethics.

Professor Kejun Zhang is a leading expert in AI and emotional computing in music from Zhejiang University, China, where he focuses on developing the intelligent generation of audio-visual digital content.

Research Team

Dr James Slaymaker is a Teaching Fellow in Film Studies at Trinity College Dublin. He received his PhD from University of Southampton in 2022, where his research focused on essay cinema in the digital age. He is currently in the process of developing it into a monograph.

Dr Selin Gurgun is a scholar in Human-Computer Interaction who received her PhD from Bournemouth University in 2024. She focused on her research on designing solutions that can encourage users to confront misinformation on social media.

Dr Liam Rogers is a scholar of film and television studies who received his PhD from the University of Warwick in 2024, where he explored the interconnected relationship between onscreen depictions of technology and critical posthumanism as a philosophy.

Boyuan Cheng is a second-year PhD student at the National Centre for Computer Animation, Bournemouth University, focusing on the use of AI deep learning techniques to generate animations from natural language. His research covers multimodal data fusion and innovative applications of generative models, with potential applications in fields such as film, gaming, and education content creation.

Stephanie Prajitna is a postgraduate student at Bournemouth University, where she conducts research in media production, animation, music, digital marketing and intellectual property, drawing on her background in animation. Additionally, her studies at Bournemouth have led her to explore entrepreneurship and innovation.

Illustration and Design

Ellie Shipman designed this report. Ellie is an artist and illustrator interested in sustainability, community and womanhood. Ellie often collaborates with researchers on design, social practice and public engagement projects.

Acknowledgements

How to Cite this Report

Ruszev, S., Rogers, L., Gurgun, S., Stockton-Brown, M., Gee, M., Davis, T., Yang, X., Griffiths, C., Zhang, K., Cheng, B., Slaymaker, J. and Prajitna, S., 2024. *Shared-posthuman imagination: Human-AI collaboration in media creation*. Poole: Bournemouth University. http://dx.doi.org/10.18746/epxn-da67

CHAPTER ONE

Responsible AI and Media Production

Amidst the well-publicised 2023 SAG-AFTRA (Screen Actors Guild – American Federation of Television and Radio Artists) and WGA (Writers Guild of America) strikes that protested, amongst other things, the use of AI in script production and synthetic performances, concerns of labour, creativity and ethics have been at the forefront of recent AI debates within mainstream media. In March and April 2024, two prominent international film releases garnered controversy concerning their implementation of generative AI tools for production and/or marketing purposes: Late Night with the Devil (Cameron Cairnes and Colin Cairnes, 2024) and Civil War (Alex Garland, 2024). Those who critiqued the decision to turn to AI argued that the use of AI-generated images circumnavigated human creative labour, removing paid work for visual artists who could have manually produced these creative elements.[2] The following month, the Netflix documentary What Jennifer Did (Jenny Popplewell, 2024) came under fire for its undisclosed use of AI to manipulate archival footage. This act was largely interpreted as an attempt on the filmmakers' part to intentionally deceive audiences by tampering with photographic evidence, thereby undermining 'truth claims' integral to true

crime documentary as a genre[3]. Later that same month, prominent actress Scarlett Johansson began legal proceedings against OpenAI, the company behind ChatGPT, for its unlicensed use of her voice in its new voice-based assistant. The voice purposefully echoes Johansson's performance as a virtual assistant in the now canonical science-fiction film, Her (Spike Jonze, 2013), but as critics pointed out, the fact that OpenAI did so without and actively against Johansson's consent exemplifies the potential for generative AI to not only impose upon pre-existing image and intellectual property rights, but also to perpetuate harmful stereotypes about gender and identity. [4]

---

[2] James Hibberd, 'A24's New AI-Generated Civil War Ads Generate Controversy', The Hollywood Reporter, 17 April 2024 <https://www.hollywoodreporter.com/movies/movie-news/a24-civil-war-posters-controversy-1235876340/> [Accessed 15 August 2024].

[3] Ellie Muir, 'Netflix embroiled in alleged AI scandal with What Jennifer Did', The Independent, 19 April 2024 <https://www.independent.co.uk/arts-entertainment/tv/news/netflix-what-jennifer-did-ai-image-true-story-b2531363.html> [Accessed 15 August 2024].

[4] Kate Devlin, 'OpenAI's Scarlett Johansson Update Wasn't About Bridging the Gap Between Tech and Creatives – It Was Just Sexist', Byline Times, 25 May 2024 <https://bylinetimes.com/2024/05/25/openais-scarlett-johansson-update-wasnt-about-bridging-the-gap-between-tech-and-creatives-it-was-just-sexist/> [Accessed 15 August 2024].

These examples speak to the prescient need to address generative AI's impact on media production. The accessibility and widespread proliferation of generative AI tools inspire large communities of creative users to experiment with media production in new and exciting ways. However, they also come with important concerns manifesting within an industry seeking to keep up with the rapidly evolving and exponentially developing phenomenon that is generative AI. This chapter will begin interrogating some of these hopes and concerns by way of establishing the terminological and contextual parameters of 'Shared Post-Human Imagination' as a research project. It begins by introducing this report's terms of engagement (1.1) before outlining the project's research questions (1.2), tracing through the key stakeholders within the generative AI media landscape (1.3), and offering a four-part review of literature focused on AI in media production (1.4), ethics (1.5), the posthuman (1.6), and responsible AI (1.7).

1.1: Terms of Engagement

In order to talk about generative AI and media production, some notes on terminology are needed. The first part of this section will provide an overview of the technical vocabulary used throughout this report, including definitions of key industry terms and explanations of how generative AI models work. The second part of this section aims to provide readers with a baseline understanding of more fluid, open and constable concepts that have been used to guide this research.

1.1.1: Technical Vocabulary

Artificial Intelligence (AI) is a field of science concerned with the development of machines that can reason, learn and act in ways that replicate human-level intelligence. Over time and through rigorous training processes, machines can be taught to recognise patterns

and subsequently make decisions based on what they have 'learned', enabling them to distinguish between different sets of data on the basis of predicting expected outcomes. For example, if one wanted to teach a computer to distinguish between a cat and a dog, this might be achieved by showing it a substantial quantity of images containing cats and dogs, and, over time, the computer will begin distinguishing patterns and features, such as the shape of the animal's ears or the length of their tails. This will result in the machine being able to identify the subject of an image. As a subset of AI, generative AI uses these training processes of understanding patterns and relationships within large datasets to generate new outputs based on what it has been inputted. Generative AI tools are powered by Neural Networks, a set of virtual synapses that help the AI learn and make decisions based on connections to pre-existing knowledge; the more data these networks are fed, the better they become at generating content. ChatGPT and Poe are prominent examples of generative AI tools (GAIT) known as Large Language Models (LLMs). These are models trained on a substantial amount of text, such as novels, textbooks and websites, to enable the AI to learn the structural and semantic relationships between words and concepts in order to generate a response to a user's prompt. While GAIT have the potential to produce accurate and relevant outputs based on the user's prompt, they can also produce incorrect information and present it as fact, or make things up that were not part of the original prompt's instructions. This phenomenon is known as 'hallucinating'. For more information on technical vocabulary within the computer sciences, see Appendix 1.1.

Where generative AI tools enable the production of new text, images, music or even entire videos based on pre-existing datasets combined with a user's prompt, they can be

used to help with different parts of the media production process. For example, text-to-text based generation can be used to aid the screen-writing process, in which creatives produce scripts that detail the narrative, settings, characters, dialogue, actions and transitions that comprise individual scenes and entire films. Generative AI tools may also be used to develop images by simulating aspects of the visual production process such as composition, lighting, location, camera angle and aesthetic of a given shot or frame. These images may then be edited together using generative AI, which can handle the process of cutting, arranging and assembling them into a coherent sequential order to tell a narrative and communicate meaning. Editing with generative AI may mean the use of AI to label certain shot types or produce techniques such as cuts, fades, dissolves and wipes from shot to shot or scene to scene. While these techniques may result in a moving-image film, media practitioners will often build on the initial AI output by adding elements such as dialogue, synchronised sound effects, voiceover, original score or soundtrack music. While the use of one single tool to produce an entire feature-length film from start to finish is not yet a reality, media practitioners can follow these sequential steps in the filmmaking process and, by doing so, generate a film in collaboration with AI. For more technical vocabulary on the use of AI within film production, see Appendix 1.2.

While these technological advancements appear to offer practitioners a set of new tools with which to produce media, the use of generative AI in the film production process poses a challenge to pre-existing judicial concepts and legal frameworks. For instance, in the context of generative AI, the author may be any one of: the human user(s) who prompted the tool or refined its output in accordance with their original concept; the AI model itself, such that the generation of said

piece of work could not have taken place were the model not to have intervened; and/or the people(s) whose work comprise the datasets used to train AI models, such that any outputs produced by AI are inherently derivative of original ideas, stories and images produced by human authors.[5] There exists no singular definition of an author that can be applied to all circumstances of AI media production across all jurisdictions. The question of ownership, in the sense of the person(s) or entity that holds the rights to a piece of work, is also a contentious issue in relation to generative AI production. Ownership rights over a media text may be held the author, but this is not always the case. Authors may hand over the copyright (that is, the right to print, publish, perform, film, record or reproduce original materials) to a given piece of work before, during or after production via formal agreements or contracts. In the context of generative AI media production, the question of who owns the final output will depend upon the contractual agreement that the user

---

[5] Michael Murray, 'Generative AI Art: Copyright Infringement and Fair Use', SMU Science and Technology Law Review, 26: 2 (2023), 2023, 259-315 (p. 298).

enters with the developer at the point of sign up or use. For more information on technical vocabulary pertaining to AI within UK law, see Appendix 1.3.

As well as unresolved legal tensions, the use of generative AI to produce media also brings to the fore a series of questions about what responsible or ethical use should look like. For many, responsible use of generative AI tools requires that they be accessible to a diverse audience, holding the potential to break down barriers of access for those who have previously not had the means to produce media outputs. Designing, practicing and producing AI outputs with accessibility in mind can lead to a more just and fair media landscape, whereby everyone is given equal access to the opportunities afforded by generative AI.[6] Equally important to understanding what responsible AI practices might entail is a consideration of the concepts of accountability and transparency. In the context of generative AI, this means being open and clear in communications regarding the ways in which important decisions were made in collaboration with AI, building trust by way of allowing different stakeholders to inspect the mechanisms through which AI outputs come to be produced and holding those responsible accountable for key decisions.[7] Practices of accountability and transparency may entail reliability, such that end users have a clear understanding of what to expect from successive collaborations with AI tools. They also enable AI systems to be produced with safety and sustainability in mind, whereby AI tools are developed in ways that mitigate against potential risks to mental and physical health worldwide. This may be achieved on a local scale through safeguards designed to protect those more vulnerable in society from potentially harmful AI outputs and practices, as well as on a more global scale by redressing the high output of carbon emissions that come with storing, processing and training LLMs.[8] For more information on technical vocabulary within responsible AI, see Appendix 1.4.

1.1.2: Key Concepts

This report categorises the array of pressing concerns and opportunities that generative AI tools bring to media production according to three overarching concepts. Each of these concepts will be introduced below. Whereas the technical vocabulary outlined above in 1.1.1 offers a set of empirical, sometimes objective, definitions, here, definitions of key concepts are more open, contestable and subject to debate, highlighting the need for individual chapters dedicated to unpacking their nuances and complexities.

First, generative AI models are changing our understanding of **creativity** within media production. Chapter Two will unpack what we mean when we talk about creativity in more detail, but for now, it will suffice to recognise that creativity pertains to set of elusive, hard to define qualities regarding the ability to generate novel and useful (which may also mean use as art or performance) ideas, possibilities or outcomes.[9] Within the context of media production, creativity entails associations with innovation, skill and artistry, which intersect with concerns regarding generative AI's (in)ability to reproduce such qualities. Like creativity, there exists

---

[6] Luciano Floridi, The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities (Oxford: Oxford University Press, 2023), p. 62.

[7] Virginia Dignum, Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way (Cham: Springer, 2019), pp. 53-54.

[8] Eshan Navabi, Katherine Daniell, Elizabeth Williams and Caitlin Bentley, 'AI for Sustainability: A Changing Landscape', in Artificial Intelligence: For Better or Worse, Future Leaders, 157-176 (2019) (p. 170).

[9] Giorgio Franceschelli and Mirco Musolesi, 'Creativity and Machine Learning: A Survey', ACM Computing Surveys, 56: 11 (2024), 1-41 (p. 2).

no singular definition of art, meaning that adjacent questions regarding what constitutes 'AI art' are equally open to interpretation.[10] Across the media landscape, there are those for whom generative AI serves to bolster human creativity, facilitating ideation and generating artistic outputs that are innovative, creative and impactful, reminiscent of what human artists are capable of. Yet, for others, AI art is an oxymoron; since generative AI lacks the necessary skill, craft and cultural embeddedness needed to create work that is novel and useful (that is, socially meaningful), it cannot produce art. Along these lines, there exists a prevailing view that generative AI has the potential to not only augment and challenge human creativity but also displace human artists from creative roles, such that, if stakeholders in positions of power view AI-generated outputs as art, it might lead to less human artists being involved in the creative industries. This, in turn, has the potential to devalue societal conceptions of art and creativity. Chapter Three builds upon and interrogates these preliminary concerns, offering unique insights into how practitioners working within the creative industries have responded to the changing concept of creativity in the age of generative AI.

Secondly, generative AI models are reproducing and perpetuating **bias**. Running counter to values of fairness, equality and accessibility, bias constitutes the tendency to favour a particular group of things or people over others. Bias in generative AI tools entails the perpetuation of prejudicial outcomes concerning how certain identity groups are systematically excluded from fair representation in the design of media outputs. Such biases urgently need to be recognised and redressed. Bias in AI can therefore produce outcomes that are not only distortions of reality, through the perpetuation of stereotypes, but that also fuel societal problems of misogyny, racism, homophobia, transphobia, ableism, classism and ageism, amongst other things. Part of the problem regarding how best to redress AI bias comes from the fact that generative AI models are trained on human-created data of the past and by human engineers, for whom biases may influence their training processes. As well as biased data sets and developers, there also exists the potential for media production to heighten biased outcomes of labour, such that the most culturally under-represented people within society will be those at greater risk of job precarity.[11] The task is, then, to find ways to address bias in generative AI tools, ensuring equality and fairness for all, without halting the progression of technological innovation.[12] At present, there exists a lack of clarity, transparency and understanding about how best to undertake this task. Hence, Chapter Four will address these issues by way of scoping what different stakeholders see as the best way to tackle AI bias in media production.

Thirdly, generative AI models are forcing us to rethink what it means to 'use' AI. Traditionally, the use of AI entails a top-down, hierarchical process in which human agents do something with a particular model, interface or piece of software in order to produce an outcome or complete a task. Under this model, the 'use' of AI retains a sense of human control and agency, where the AI's role is to be a tool to help an individual agent carry out a particular function. In the context of generative AI, these roles become more fluid and harder to distinguish, leading to human-AI **collaboration**. Like creativity, collaboration in the context of generative AI is difficult to define, as it means different things to different people. For some, collaboration necessitates

[10] Marcel Danesi, AI-Generated Popular Culture: A Semiotic Perspective (Cham: Palgrave Macmillan, 2024), p. 10.

[11] Atoosa Kasirzadeh, Charlotte Bird and Eddie Ungless, 'Policy Report on Generative Artificial Intelligence', BRAID (July 2024), pp. 7-8.

[12] Christoph Trattner et al., 'Responsible Media Technology and AI: Challenges and Research Directions', AI and Ethics, 2 (2022), 585-594 (p. 589).

a symbiotic flow of exchange between two or more entities, such that to collaborate with AI is not merely to use or exploit AI for one's own purposes but rather to work in tandem with AI, producing an end product that is the shared output of human and machine intelligence, effort or creativity.[13] For others, collaboration with AI tools brings with it a potential loss of control for human creatives that may inevitably lead to the exploitation of individuals' work or intellectual property rights for financial gain or even the displacement of human labour from the creative industries. As this brief survey suggests, there are still important questions that need to be answered about where to place human rights and responsibilities in collaborative processes with AI. Chapter Five will unpack these questions and point towards potential solutions that enable the implementation of equitable and just forms of collaboration with generative AI.

1.2: Research Questions

The presence of AI in news media, notions of deepfakes and algorithmic echo chambers, and political manipulations of digital misinformation are, by now, widely known and critically discussed. The presence of AI in fictional media, however, is less well known and therefore brings with it a series of unanswered questions regarding opportunity, impact and what constitutes best working practice. As such, this research aims to map out the technological challenges, design considerations, legal implications and ethical concerns in the use of generative AI in film and media. It seeks to do so by providing answers to the following research questions:

- **How is the notion of media creativity being re-evaluated within a context of responsible AI, and how can we ensure that augmentations to human creativity happen in ways that protect against extractive database practices, intellectual property infringements and displacements of human labour?**

- **How does the use of generative AI in media production perpetuate social biases, and how can we ensure that there is justice, transparency and safety regarding the training of the large language models on which AI tools are built?**

- **What are the implications of human-AI collaboration, and how can we make sure that collaborative work including AI tools is accountable, just and accessible to all?**

By thinking about these driving questions, the aim of this research is to gain insights into possible interventions that could foster a responsible AI media landscape. In turn, this paper translates these findings into a practical framework designed to help developers, media practitioners and policymakers undertake responsible AI practices.

1.3: Stakeholders

The vast landscape of generative AI in media production mainly cuts across two disciplines: generative AI development in computer science, and its application in film and media production as a creative industry. Yet, as new generative AI models are increasingly designed with film and media production in mind, and as media practitioners look to integrate AI into their workflow practices, these two disciplines become one, and many overlapping stakeholders emerge.

---

[13] Florent Vinchon et al., 'Artificial Intelligence & Creativity: A Manifesto for Collaboration', Journal of Creative Behaviour, 57: 1 (2024), 472-484 (p. 479).

As part of our research, we hosted four, two-day workshops designed to both facilitate participant experimentation with different generative AI tools and ascertain the prevailing hopes and concerns about generative AI's impact on screen-writing, image creation, editing, and music through systems thinking methods (more on this in Chapter Two). We invited a diverse array of stakeholders, from academics and filmmakers to media professionals and end users, to reflect the diverse landscape of AI media production.

In Workshop 1 (Screen-writing) and Workshop 2 (Image Creation), we asked participants to reflect upon who they thought were the key stakeholders involved in AI media production and, subsequently, where wider dynamics of communication, collaboration and power imbalances could be mapped onto these different stakeholders. Participants of both workshops recognised a myriad of different stakeholders, including, but not limited to:

- **Filmmakers, including screenwriters, cinematographers, editors, animators, directors and sound technicians**
- **Actors and performers**
- **Software developers and technical engineers**
- **Guilds, advocacy groups and unions**
- **Studios, producers, publishers and commissioners**
- **Policymakers, lobbyists, government officials and local authorities**
- **Distributors, exhibitors and festivals**
- **Lawyers, legal departments and insurance companies**
- **Teachers and academics**
- **Students**
- **Philosophers and ethicists**
- **Critics, journalists and bloggers**
- **Audiences and end users of GAI**

This list shows that, overall, participants were astutely aware of the different stakeholders involved in media production with AI. They sought to emphasise the importance of accessibility and redressing power imbalances amongst these stakeholders, placing a particular emphasis on utilising generative AI to break down the gender, race, class, age and disability barriers faced by many members in the media industry.[14] Along these lines, they also sought clarification regarding questions of ownership, as the notion of which stakeholders own individual creative outputs remains both legally and discursively uncertain.

The similarities in responses across Workshops 1 and 2 instigated a pivot in our research approach for Workshops 3 and 4, where we instead presented a map of these stakeholders (Appendix 6) to the participants and asked them to comment upon the efficacy of its understanding of power, alignment and influence. For the final version of the stakeholder map see Appendix 6.

Participants in Workshop 3 (Editing) reflected upon the important role that education plays in generating relationships between different stakeholders, as well as the need to recognise that any map of potential stakeholders will inherently be culturally and nationally specific (for example, if we were to ask participants who are residents within China to map out the power and influence relations between different stakeholders, it would be significantly different from a UK context).

Participants in Workshop 4 (Music Production) stressed the importance of industry-specific stakeholder maps, such that one dedicated

---

14 Anna Ozimek, 'Equality, Diversity and Inclusion in the Screen Industries'. Research Report, The University of York (2020) <https://screen-network.org.uk/wp-content/uploads/2021/02/Equality-Diversity-and-Inclusion-in-the-Screen-Industries.pdf> [Accessed 16 August 2024].

specifically to film and television music production would need to account for the particular roles of music producers, composers, open-source developers and collective agencies, such as PRS for Music.

Across all four workshops, participants recognised that the very notion of a stakeholder is itself fluid and open to change within an AI media landscape. Stakeholders may occupy different and even competing positions within the landscape of media production depending on certain contexts, such that the filmmaker may also be an academic, the actor may also be part of distribution and exhibition, and policymakers may themselves be audiences and end users of AI-produced media. This means that the dynamics of influence, alignment and power are also shifting at any given time. As such, participants often expressed a shared desire to seek common ground and facilitate critical dialogue between different stakeholders, many of whom occupy more than one cultural, economic or social position, since generative AI impacts everyone in the media landscape. As AI ethicist and computer scientist Virginia Dignum notes,

> AI will affect everybody. This demands that the development of AI systems ensures inclusion and diversity, that is, truly considers all humankind when determining the purpose of the systems. Therefore, Responsible AI also requires informed participation of all stakeholders.[15]

1.4: AI Development and Media Production

In recent years, AI development has made significant advancements across various fields, including text, image, video and music generation. Breakthroughs in these areas are transforming the way we interact with and experience the technology.

Over the past decade, the development of Large Language Models (LLMs) has undergone remarkable advancements. Early models, such as Recurrent Neural Networks (RNNs)[16] and Long Short-Term Memory (LSTM) networks,[17] laid the groundwork by allowing machines to remember information for a short time, helping them understand the flow of language better. As we moved forward, Transformer architecture emerged, which allowed models to process much larger sections of text at once and grasp complex relationships between words. This is how models like GPT (Generative Pretrained Transformers) came to be, taking generative AI to a whole new level.

The introduction of word embeddings, such as Word2Vec and GloVe, further refined this capability by capturing semantic relationships between words.[18] A pivotal moment came in 2018 with the release of BERT, which employed a bidirectional approach to context,[19] setting the stage for even more powerful pre-trained models like T5, RoBERTa, XLNet and ALBERT.[20] By 2023, the impact of generative AI was fully realised with the

---

[15] Dignum, Responsible Artificial Intelligence, pp. 47-48.

[16] Jeffrey Elman, 'Finding Structure in Time', Cognitive Science, 14: 2 (1990), 179-211; Michael Jordan, 'Serial Order: A Parallel Distributed Processing Approach', Approaches in Psychology, 121 (1997), 471-495.

[17] Sepp Hochreiter and Jürgen Schmidhuber, 'Long Short-Term Memory', Neral Computation, 9: 8 (1997), 1735-1780.

[18] Tomas Mikolov, Kai Chen, Greg Corrado and Jeffrey Dean, 'Efficient Estimation of Word Representations in Vector Space' (2013); Jeffrey Pennington, Richard Socher and Christopher Manning, 'GloVe: Global Vectors for Word Representation', Proceedings of the 2014

Conference on Empirical Methods in Natural Language Processing (EMNLP), Qatar (2014) 1532-1543.

[19] Jacob Devlin, Ming-Wei Chang, Kenton Lee and Kristina Toutanova, 'BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding' (2018).

[20] Yinhan Liu et al., 'A Robustly Optimised BERT Pretraining Approach' (2019); Zhilin Yang et al., 'XLNet: Generalised Autoregressive Pretraining for Language Understanding', 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Canada (2019); Zhenzhong Lan et al., 'ALBERT: A Lite BERT for Self-Supervised Learning of Language Representations', ICLR 2020 Conference (2019).

widespread use of ChatGPT, which demonstrated the ability of LLMs to generate human-like text and handle complex tasks.[21] As LLMs continue to evolve, they are expected to become even more intelligent, personalised, and multimodal, further enhancing their role in AI-driven applications.

The evolution of AI image generation has also witnessed significant milestones in recent years, starting in 2012 when Google researchers Andrew Ng and Jeff Dean trained a deep learning model to generate the first blurry images of cats marking the first substantial step in AI's ability to create visual content.[22] This was followed by a breakthrough in 2014 with Ian Goodfellow's introduction of Generative Adversarial Networks (GANs), which enabled AI to generate more realistic images through a process of adversarial training between a generator and a discriminator.[23] In 2015, Google's Deep Dream further explored AI-generated art by producing surreal and dream-like images, enhancing patterns found in existing pictures. The field reached a new height in 2021 with the launch of OpenAI's DALL-E, which revolutionised AI images

by allowing the generation of detailed and contextually relevant images from simple text descriptions.[24] This progression has significantly expanded the creative possibilities of AI in visual content generation.

In terms of AI-generated video technology, before 2016, developments mainly focused on basic video frame generation using deep learning techniques. Between 2016 and 2019, the introduction of GANs and Variational Autoencoders (VAEs) brought substantial improvements in video quality and realism.[25] For example, during this period, tools like Google's DeepMind demonstrated early video generation capabilities.[26] Post-2019, the field advanced with diffusion models leading to even more sophisticated outputs. Notably, in 2022, Google introduced Imagen Video and Meta launched Make-A-Video, showcasing the ability to create high-quality videos from textual descriptions.[27] By 2023, tools like Runway's Gen-1 Runway AI and Gen-2 Runway AI had further pushed the boundaries, enabling the generation of more complex and realistic videos.[28] Looking ahead to 2024, AI video generation is expected to

---

[21] Tom Brown et al., 'Language Models are Few-Shot Learners' (2020).

[22] Quoc Le et al., 'Building High-Level Features Using Large Scale Unsupervised Learning', Proceedings of the 29th International Conference on Machine Learning, Scotland (2012).
[23] Ian J. Goodfellow et al., 'General Adversarial Nets' (2014).

[24] Jonathan Ho, Ajay Jain and Pieter Abbeel, 'Denoising Diffusion Probabilistic Models', Advances in Neural Information Processing Systems, 33 (2020), 6840-6851.

[25] Diederik Kingma and Max Welling, 'Auto-Encoding Variational Bayes' (2013).

[26] Yuval Tassa et al., 'DeepMind Control Site' (2018).

[27] Jonathan Ho et al., 'Imagen Video: High Definition Video Generation with Diffusion Models' (2022); Uriel Singer et al., 'Make-a-Video: Text-to-Video Generation Without Text-Video Data' (2022).

[28] Anastasis Germanidis, 'Gen-1: The Next Step Forward for Generative AI', Runway (February 2023) <https://runwayml.com/research/gen-1> [Accessed 2 September 2024]; Anastasis Germanidis, 'Gen-2: Generate Novel Videos with Text, Images or Video Clips', Runway (February 2023) <https://runwayml.com/research/gen-2> [Accessed 2 September 2024].

incorporate multi-modal input capabilities and higher resolution outputs, as evidenced by ongoing developments in tools like OpenAI's SORA and Kuaishou's KLING AI.

The journey of AI-generated music can be traced back to 1938, when Bell Labs developed the Voder, the first electronic speech synthesiser, which laid the foundation for future electronic voice synthesis technology. Over the years, companies like Google, Meta, and Stability AI have advanced AI music generation with tools like Google DeepMind's Lyria, Dream Track and Meta's AudioGen, which create music or sound effects from text prompts.[29] The most significant breakthrough came with Suno, which can generate complete songs, including lyrics, vocals, instruments, and harmony, surpassing previous expectations and indicating that AI has unlocked new possibilities in music creation.

In the midst of these technological advancements and breakthroughs in research, the rapid development of AI use in media has brought about a scholarly impulse to interrogate what it might mean to work with AI in media production. Generative AI tools operate by predicting and producing the most likely outcome, designed to satisfy the needs of the end user. Alongside the ethical and moral ramifications of using AI in media production, this functionality has led to scholarly debates about whether generative AI should be used to generate art. Responses to this question vary widely across academic literature and depend upon contextual definitions of what constitutes art and labour, concepts that will be explored in greater detail across Chapters Three and Five, respectively. For instance, proponents of generative AI champion the technology's ability to elicit feelings of awe, challenging our anthropocentric notions that artistic creativity is unique to humanity by generating 'novel' rearrangements of existing pieces of art in collaboration with human input.[30] Of particular note in this regard is Florent Vinchon et al.'s manifesto for human-machine 'co-cre-AI-tion', which looks ahead toward an optimal vision of working with and alongside GAI that avoids the extremes of either rejecting or exploiting AI in media production.[31] Similar responses emerge within literature on music and screen-writing, as generative AI can be used to experiment with popular music production and help writers through inspiration and world building, respectively.[32] However, for some, generative AI tools 'diminish the complexity of human creativity' by merely parroting human artists whose work has been

---

[29] Amara Angelica, 'Transforming the Future of Music Creation', Mindplex (November 2023) <https://magazine.mindplex.ai/mp_news/transforming-the-future-of-music-creation/?> [Accessed 2 September 2024]; Shyam Nandan Upadhyay, 'Google DeepMind Launches Lyria, Transforming the Future of Music with AI', Analytics India (November 2023) <https://analyticsindiamag.com/ai-news-updates/google-deepmind-launches-lyria-transforming-the-future-of-music-with-ai/> [Accessed 2 September 2024]; Meta, 'Musicgen: Advanced AI Music Generation' (2023) <https://musicgen.com/> [Accessed 2 September 2024].

[30] Kobe Millet, Florian Buehler, Guanzhong Du and Michael Kokkoris, 'Defending humankind: Anthropocentric bias in the appreciation of AI art', Computers in Human Behaviour, 143 (2023); Xiaolei Zhao, and Zhao, Xin, 'Application of Generative Artificial Intelligence in Film Image Production', Computer-Aided Design and Applications, 21 (2024), 15-28; Mark Coeckelbergh, 'The Work of Art in the Age of AI Image Creation', Journal of Human-Technology Relations, 1: 1 (2023), 1-13; Ignas Kalpokas, 'Work of art in the Age of its AI Reproduction', Philosophy and Social Criticism (2023), 1-19 (p. 4).

[31] Vinchon et al., 'Artificial Intelligence & Creativity: A Manifesto for Collaboration', (p. 476).

[32] Emmanuel Deruty, Marteen Grachten, Stefan Lattner, Javier Nistal and Cyran Aouameur, 'On the Development and Practice of AI Technology for Contemporary Popular Music Production', Transactions of the International Society for Music Information, 5: 1 (2022), 35-49; Piotr Mirowski, Kory Mathewson, Jaylen Pittman, and Richard Evans, 'Co-Writing Screenplays and Theatre Scripts with Language Models: Evaluation by Industry Professionals', Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, April 23-28, 2023, Hamburg, Germany (2023).

used to train LLMs without their consent.[33] Along this line of argument, generative AI's ability to produce imitations and copies of pre-existing artistic material constitutes an act of plagiarism, copyright infringement, or at least the unethical appropriation of an artist's intellectual property.

Questions such as these have also dominated scholarship on AI and the law in recent years. Literature has emerged, largely from the United States, in an effort to understand what integrating AI into media design is doing to copyright law, particularly in light of high-profile lawsuits filed against OpenAI, Suno and Udio, as well as ongoing concerns about AI avatars eroding pre-existing protections against imitation for performers and actors.[34] Yet, such scholarship has seldom focused on the UK legal system, where concepts such as fair use have no legal protection. Likewise, the 1988 Copyright, Designs and Patents Act suggests that outputs generated by a computer without intervention of a human author might be protected under copyright, but the applicability of this Act to generative AI depends on whether human intervention via prompting is deemed to be an arrangement that requires enough skill or judgement needed to classify the human as an author.[35] As a result, there exists a general lack of clarity amongst both media practitioners and legal professionals when it comes to understanding where AI-generated media sits within current UK copyright law. As we shall see in Chapter Six, UK law will need to change and adapt to keep up with the urgent needs of media practitioners and stakeholders in this regard. Yet, this is not merely a question of judiciary and regulation; it is also a question of what constitutes ethical working practice.

1.5: AI and Ethics

AI ethics is a broad and wide-ranging field. Most commonly, AI ethics refers to ethical concerns regarding the research on, design of, and use and misuse of autonomous machines, as well as attempts to imbue ethical decision making within autonomous machines themselves.[36] From this perspective, much has already been written on ethical and unethical use of AI in society, with the vast majority of this research coming since 2018.[37] 2023 saw the UK hold the world's first global summit on AI safety, which saw international leaders and AI developers gather in Bletchley Park, Milton Keynes, to discuss how to mitigate the risk of unregulated AI. Following this, The Bletchley Declaration was signed by all participating nations as a call for fairness, accountability, regulation and safety in the future design of AI systems.[38] Alongside this, academic scholarship has reflected upon the application of pre-existing ethical frameworks to AI decision making capabilities, questions of responsibility and liability when AI decisions go 'wrong', and the potential positive and

---

[33] Harry Jiang et al., 'AI Art and its Impact on Artists'. AIES '23 August 08-10 (2023). 363-374 (p. 363).

[34] Jon Garon, 'A Practical Introduction to Generative AI, Synthetic Media, and the Messages Found in the Latest Medium', SSRN (2023); Murray, 'Generative AI Art', 259-315; Richard Arnold, 'Performers' Rights and Artificial Intelligence', in Research Handbook on Intellectual Property and Artificial Intelligence, edited by Ryan Abbott and David Geffen, 218-224 (Cheltenham: Edward Elgar Publishing, 2022).

[35] Kristofer Erickson, 'Copyright Protection in AI-Generated Works', Creative Industries Policy and Evidence Centre (January 2024) <https://pec.ac.uk/blog_entries/copyright-protection-in-ai-generated-works/> [Accessed 29 August 2024].

[36] Vincent Bonnemains, Claire Saurel, and Catherine Tessier, 'Embedded ethics: some technical and ethical challenges', Ethics and Information Technology, 20: 1 (2018), 41–58 (p. 41).

[37] Mona Ashok, Rohit Madan, Anton Joha and Uthayasankar Sivarajah, 'Ethical Framework for Artificial Intelligence and Digital Technologies', International Journal of Information Management, 62 (2022) 1-17 (p. 4).

[38] 'The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023', 1 November 2023 <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023> [Accessed 14 August 2024].

negative impacts, or what Fabio Tollon calls 'promises or perils',[39] that advancements in technology and AI might have on society as a whole.[40]

Most notably, philosopher Luciano Floridi offers a meta-theoretical account of the ethical issues concerning AI in his book, The Ethics of Artificial Intelligence (2023). **From his analysis of six different published reports on ethical AI between 2017 and 2018, Floridi sets out what he deems to be the five core principles of ethical AI:**
- **beneficence (promoting well-being, preserving dignity and sustaining the planet)**
- **nonmaleficence (privacy, security and capability caution)**
- **autonomy (retaining the human power to decide)**
- **justice (avoiding unfairness)**
- **and explicability (rendering AI processes intelligible and accountable to society).[41]**

Out of this emerges what Floridi calls his framework for building a 'Good AI Society', or AI for Social Good; a list of twenty policy recommendations designed around enhancing human agency and cultivating social cohesion through AI without devaluing human ability and removing human control.[42] While some of these recommendations accord with the approach media production taken within this report – such as introducing regulatory frameworks and bodies, developing auditing mechanisms for locating

AI bias, establishing a board dedicated to ethics consultation and accreditation, and promoting further education and public awareness about AI risks – Floridi's framework is one that speaks to AI ethics more broadly, and as such, its application to the specific context that is generative AI in media production is limited.[43]

When it comes to assessing the ethics of generative AI tools, prevailing literature has tended to focus on the topic of labour. Catherine Flick and Kyle Worrall, for instance, assess generative AI's impact on creative communities in relation to the digital replication of artistic styles or intellectual property, locating a key concern within the industry to be the demise of artistic labour and the subsequent rise in dangerous content that might be created without positive moral agents.[44] Such an assessment resonates with broader concerns about both current and future industry labour practices, in which dataset labelling is being outsourced to the Global South for work in unfair and often traumatising conditions[45] while fears of an automated workforce replacing human labour continue to persist within both academic and popular discourse.[46]

Generative AI tools are also at risk of spreading misinformation. Popular models have been known to 'hallucinate' incorrect information, a practice that Michael Townsen Hicks, Jake Humphries and Joe Slater have dubbed 'bullshitting' based on the fact that

---

[39] Fabio Tollon, 'Technology and the Situationist Challenge to Virtue Ethics', Science and Engineering Ethics, 30: 10 (2024).

[40] Bonnemains, Saurel and Tessier, 'Embedded ethics: some technical and ethical challenges'; Joanna Bryson, 'Patiency is not a virtue: the design of intelligent systems and systems of ethics', Ethics and Information Technology, 20: 1 (2018), 15–26; Patrick Lin, Ryan Jenkins and Keith Abney, eds., Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence (Oxford: Oxford University Press, 2017).

[41] Floridi, The Ethics of Artificial Intelligence, pp. 56-63.

[42] Ibid., pp. 170-173.

[43] Ibid., pp. 174-179.

[44] Catherine Flick and Kyle Worrall, 'The Ethics of Creative AI', in The Language of Creative AI: Practices, Aesthetics and Structures, ed. by Craig Vear and Fabrizio Poltronieri, 73-94 (Cham: Springer, 2022).

[45] Floridi, The Ethics of Artificial Intelligence, p. 47.

[46] Florian Butollo and Sabine Nuss, eds., Marx and the Robots: Networked Production, AI, and Human Labour, trans. by Jan-Peter Herrmann and Nivene Raafat (London: Pluto Press, 2022).

they are designed to produce convincing, but not necessarily truthful, outputs and information.[47] While this poses a particular problem for the use of generative AI in education,[48] its impact on media production can be seen in the practice of documentary filmmaking, journalism and algorithmic decision making, possessing the potential to erode public trust in media sources dedicated to factual reporting.[49] Similarly, generative AI tools risk perpetuating unethical biases and stereotypes about certain identity groups based on mischaracterising outputs. The limitations and opportunities for change that emerge from this will be unpacked in greater detail in Chapter Four.

As Luciano Floridi writes, ethics must 'not be a mere add-on, an afterthought, a latecomer', by which time potentially harmful and practices are already taking place; nor should it be 'a mere exercise in questioning'.[50] Instead, AI ethics needs to work towards finding and potentially implementing 'shareable solutions' that different stakeholders can turn to as ways of establishing best working practices.[51] While this review of literature has thus far located many questions circulating within discourse on AI and media production, it will now turn to consider two important frameworks that shaped this project's approach in formulating the 'shareable solutions' laid out in the forthcoming chapters: theories of the posthuman and practices of 'responsible AI'.

1.6: AI and the Posthuman

While the term 'posthuman' has been subject to a range of different debates and definitions across both academic literature and popular culture, it can broadly be described as a theoretical framework that seeks to comprehend how cultural, biological and technological developments are challenging what we previously thought it meant to be human. Scholars in the fields of critical theory and new materialism have turned to the term as a way of making sense of recent shifts away from distinctly humanist, anthropocentric and dualistic understandings of the human.[52] Within this broad assemblage, there are those who conceptualise the posthuman as an ontological succession to the human species, a next step in human evolution;[53] an ethical way of being in the world that we ought to strive toward by rethinking the exclusory nature of what it means to be human;[54] and a category of being that we have always been, such that being human has always comprised a collaboration with technologies, tools and 'more-than-human elements' since the dawn of the *Homo sapiens* species.[55]

The posthuman forces us to consider what a union of the human and the machine ought to look like, what it might do to our humanist understanding of concepts like creativity and essence, and how we might utilise the

---

[47] Michael Townsen Hicks, James Humphries and Joe Slater, 'ChatGPT is bullshit', Ethics and Information Technology, 26: 38 (2024), 1-10 (p. 3).

[48] Tama Leaver and Suzanne Srdarov, 'ChatGPT Isn't Magic: The Hype and Hypocrisy of Generative Artificial Intelligence Rhetoric', M/C Journal, 26: 5 (2023).

[49] Dominic Lees, 'Deepfakes in Documentary Film Production: Images of Deception in the Representation of the Real', Studies in Documentary Film, 18: 2 (2024), 108-129; Anandana Kapur and Nagma Ansari, 'Coding Reality: Implications of AI for Documentary Media', Studies in Documentary Film, 16: 2 (2022), 174-185; Bronwyn Jones, Rhianne Jones and Ewa Luger, 'Generative AI and Journalism: A Rapid Risk-Based Review' (June 2023) <https://www.research.ed.ac.uk/en/publications/generative-ai-amp-journalism-a-rapid-risk-based-review> [Accessed 29 August 2024].

[50] Floridi, The Ethics of Artificial Intelligence, p. 90.

[51] Ibid., p. 91.

[52] Rosi Braidotti, The Posthuman (Cambridge: Polity Press, 2013); Francesca Ferrando, Philosophical Posthumanism (London: Bloomsbury, 2019).

[53] Newton Lee, ed., The Transhumanism Handbook (Cham: Springer, 2019).

[54] Patricia MacCormack, Posthuman Ethics: Embodiment and Cultural Theory (Farnham: Ashgate, 2012).

[55] Alan Smart and Josephine Smart, Posthumanism: Anthropological Insights (Toronto: University of Toronto Press, 2017), p. 66.

potentials that emerge from within this union for good. As N. Katherine Hayles writes,

My dream version of the posthuman embraces the possibilities of information technologies without being seduced by fantasies of unlimited power and disembodied immorality, that recognizes and celebrates finitude as a condition of human being, and that understands human life is embedded in a material world of great complexity, one on which we depend for our continued survival.[56]

Like Donna Haraway's infamous manifesto on the cyborg, Hayles' posthuman contains a doubled vision which contains the potentials for both perpetuating exploitative practices and liberation from said practices. It offers these potentials not as a straightforward technophobic/technophilic dichotomy but rather as a state of simultaneity, such that, to quote Haraway, seeing 'both perspectives at once […] reveals dominations and possibilities unimaginable from the other vantage point'.[57]

As one of the defining philosophical movements of the past three decades, posthumanism and the posthuman offer some insights into how we might ethically integrate GAI as a technology into working practices. Namely, they enable us to conceptualise tool use not as a hierarchical state of manipulation and configuration but instead as a collaborative process between human and machine, in which different qualities and skill sets become entangled to produce an end product. Framing GAIT use in this way enables us to see the generative possibilities

that come with re-evaluating concepts like human creativity, artistry and skill, without losing sight of the very real material impacts that GAIT are having on labour and copyright within the creative industries. It is from this perspective that this paper seeks to set out what responsible AI practice might look like for a range of different stakeholders within the context of creative media production.
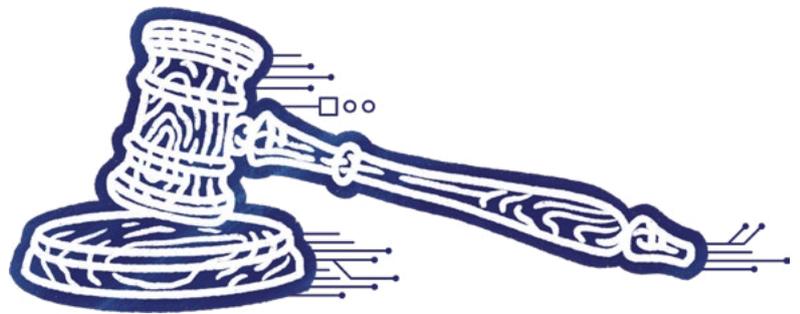
1.7: Responsible AI

In her book, Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way (2019), Virginia Dignum recognises that 'responsible AI means different things to different people', with the term serving as 'an overall container for many diverse opinions and topics'.[58] As such, there is no one way in which to approach or implement responsible AI practices. There are those for whom a focus on explainability, or 'XAI', enables the building of trust and fairness between users and developers by making the design processes of AI models understandable, comprehensible and transparent.[59] There are also those for whom responsible AI is inherently tied up with sustainability. Given the fact that a single 'training run of GPT-3' is said to have produced the same amount of

---

[56] N. Katherine Hayles, How We Became Posthuman (Chicago: University of Chicago Press, 1999), p. 5.

[57] Donna Haraway, Simians, Cyborgs, and Women: The Reinvention of Nature (New York: Routledge, 1991), p. 154.

[58] Dignum, Responsible Artificial Intelligence, p. 93.

[59] Alejandro Barredo Arrieta et al., 'Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI', Information Fusion, 58 (2020), 82-115.

carbon emissions as almost 50 cars across a single year,[60] and that use costs are far more carbon-intensive than training costs,[61] it is imperative that we document and redress the negative environmental impact of developing and using GAIT in relation to energy consumption.[62]

Responsible AI practices can therefore take place in the development, manufacture, selling, use and regulation of the entire spectrum of AI systems, such that its focus will depend upon what engineers and practitioners define as 'responsible'.[63] For Dignum, this necessitates that a 'code of behaviour' be established in advance; that is, a consensus on what we want to see as best working practice for AI, how it can work in favour of our agreed upon societal values, and what system requirements or technical functions might we need to interrogated, rethought or potentially regulated in order to make responsible AI a reality.[64] Hence, we can think about responsible AI as a form of ethical praxis, since responsible AI 'implies the need for mechanisms that enable AI systems to act according to ethics and human values'.[65]

While Dignum suggests that responsible AI requires interdisciplinary commitment from all stakeholders in its conceptualization and implementation, she also recognises that, first and foremost, responsible AI is an issue of governance. She states, 'a new and more ambitious form of governance is one of the most pressing needs in order to determine that inevitable AI advances will be accessible to all'.[66] Alongside pre-existing AI initiatives and governance bodies, such as IEEE's Ethics of Autonomous and Intelligent Systems initiative, The European Commission's High-Level Expert Group on AI, or the British House of Lords Select Committee on AI, Dignum calls for clearer regulation and guidance on issues of liability, transparency and accessibility, proposing a certification of responsible AI systems alongside education of said certification so that, at the very least, users 'would then have the choice of what system' to use based on their own ethical criteria.[67]

Since Dignum's formative book, academic interest in responsible AI governance has tended to focus on responsibility as an ongoing, relational process rather than a fixed set of rules reserved for individual agents.[68] Where specific strands of AI have been the subject of working papers, these have tended to focus on the role of responsible AI in education.[69] A rising number of policy papers, position papers and consultation responses have also been published in the last two years that seek to propose a responsible AI rubric, such as those published by Partnership on AI,[70] based in the US, and the Trustworthy Autonomous Systems Hub in the UK.[71]

---

[60] Floridi, The Ethics of Artificial Intelligence, p. 187.

[61] Alexandra Luccioni, Yacine Jernite and Emma Strubell, 'Power Hungry Processing: Watts Driving the Cost of AI Deployment?', ACM FAccT '24, Brazil (2024), 1-15 (p. 9).

[62] Navabi, Daniell, Williams and Bentley, 'AI for Sustainability'. Navabi et al. do recognise that the issue of AI sustainability also needs to consider the positive impact that AI can have on forging an environmentally sustainable future, such as its use in detection and monitoring processes.

[63] Dignum, Responsible Artificial Intelligence, p. 57.

[64] Ibid., pp. 48, 62.

[65] Ibid., p. 104.

[66] Ibid.

[67] Dignum, Responsible Artificial Intelligence, p. 98.

[68] Shannon Vallor, 'Edinburgh Declaration on Responsibility for Responsible AI', Medium, 14 July 2023 <https://medium.com/@svallor_10030/edinburgh-declaration-on-responsibility-for-responsible-ai-1a98ed2e328b> [Accessed 9 August 2024].

[69] Caitlin Bentley et al., 'A Framework for Responsible AI Education: A Working Paper', SSRN Working Paper Series (2023).

[70] Partnership on AI, 'PAI's Responsible Practices for Synthetic Media: A Framework for Collective Action' (2023).

[71] Trustworthy Autonomous Systems Hub, 'Response to: AI regulation: a pro-innovation approach – policy proposals', submitted in Open Consultation from Department for Science, Innovation and Technology and Office for Artificial Intelligence (2023).

Responsible AI, a UK-based funding body, have recently started funding multidisciplinary research across the Arts, Social Sciences, Computer Sciences and Engineering, taking a Responsible Research and Innovation (RRI) approach to consider the consequences of incorporating AI within different industries, sectors and national public services.[72] Similarly, as part of their 'Artificial Intelligence (Safe and Ethical)' centre of excellence, the Alan Turing Institute has been undertaking research in line with the UK Government's industrial strategy to examine what a responsible approach to AI defence, national security and health might look like.[73] Such institutional endeavours serve as important and timely interventions into the AI landscape, offering insight into the need for disclosure mechanisms and meta-data, outreach programmes designed to educate businesses and developers on responsible AI, the integration of technical staff into responsible AI regulators, and the anchorage of AI governance within UN Charter and International Human Rights law. However, in adopting a broad approach that attempts to account for all types and modalities of AI, these groups also overlook some of the particularities that emerge within a media production context.

Building upon these advancements in responsible AI, in 2022, the UK's Arts and Humanities Research Council set up BRAID (Bridging Responsible AI Divides). Working in collaboration with both the Ada Lovelace Institute and the BBC, BRAID seek to integrate the Arts and Humanities into the responsible AI ecosystem by bridging divides between divergent stakeholders and industries. Through a combination of demonstrator projects, fellowships and scoping research, of which this paper and project are part, BRAID investigate responsible AI across different sectors, such as education, transportation, policing, surveillance, curation, and both public and private workplaces.[74] In their work with the Ada Lovelace Institute, they have called on the UK government to create incentives for AI risk assessments[75] and infrastructure to establish a centralised AI monitoring programme, designed to monitor both the inputs and outputs of AI systems in line with a responsible AI approach.[76] Such interventions are indicative of BRAID's policy approach to AI regulation and governance more broadly, highlighting the necessity for government intervention to 'fill the responsibility gap' created by large tech companies so that 'everyone can benefit from the economic development and social value brought about by these new advances'.[77] While BRAID's remit goes far beyond the topic of generative AI, the programme's interest in the creative arts means that many of its recent projects and research outputs have sought to investigate what responsible AI use might look like within newsrooms, museums, creative

---

[72] Responsible AI UK, 'Responsible AI Governance: A Response to UN Interim Report on Governing AI for Humanity', submitted in Open Consultation from the UN AI Advisory Body' (2024). See also https://rai.ac.uk/ for information on their guiding principles, impact accelerator projects and international partnership projects.

[73] The Alan Turing Institute, 'Artificial Intelligence (Safe and Ethical)'. <https://www.turing.ac.uk/research/research-programmes/artificial-intelligence-ai/safe-and-ethical> [Accessed 29 August 2024].

[74] For examples of BRAID's research on responsible AI, see John Zerilli, Iñaki Goñi and Matilde Masetti Placci, 'Automation Bias and Procedural Fairness: A Short Guide for the Public Sector', BRAID (July 2024); Karen Gregory and Cailean Gallagher, 'Mitigating Harms in On-Demand Delivery Platforms: AI Regulations, Data Protection, and Workers' Tools', BRAID (July 2024).

[75] Ada Lovelace Institute, 'AI Assurance? Assessing and Mitigating Risks Across the AI Lifecycle' (July 2023) <https://www.adalovelaceinstitute.org/report/risks-ai-systems/> [Accessed 29 August 2024].

[76] Ada Lovelace Institute, 'Keeping an Eye on AI: Approaches to Government Monitoring of the AI Landscape' (July 2023) <https://www.adalovelaceinstitute.org/report/keeping-an-eye-on-ai/> [Accessed 29 August 2024].

[77] Bhargavi Ganesh, 'Policy Approaches for Building a Responsible Ecosystem: Contextualising AI Governance Challenges Within Other Regulatory/Governance Sectors and Histories', BRAID (July 2024).

writing, sound design and public service media.[78] In particular, their July 2024 'Policy Report on Generative Artificial Intelligence' established the need for the UK government to adapt its pro-innovation framework on AI regulation in order to mitigate and offset the potential risks posed by text-to-image AI tools when it comes to producing discriminatory content, misinformation and infringements on privacy and copyright laws.[79] Their recommendations to the UK government integrated pre-existing responsible AI concepts, such as XAI, audit mechanisms and digital literacy education programmes, into the specific technological innovation that is generative AI. In doing so, they emphasised the importance of taking a holistic approach to responsible AI development and regulation, one which informs this report's understanding of the need for both examining generative AI media production under existing responsible AI rubrics and formulating new frameworks that respond to the specific needs of the media industry in particular.

Aside from BRAID, there are some other groups working both internationally and nationally for whom AI in media production is of the utmost concern. For example, UK Music, the collective voice for the UK's music industry, have published policy recommendations that they would like to see, including protections for artist copyright and personality rights in UK law, accessible record keeping of what music has been used to train AI LLMs and clear labelling of all music generated using AI.[80] Meanwhile, the Content Authenticity Initiative was set up in 2019 as a cross-industry coalition of over 2000 media companies who

have sought to develop content credentials that accurately label AI-generated images and content as such.[81] Most notably, the MediaFutures centre in Bergen, Norway, works to develop responsible media technology practices in the age of AI. Through an interdisciplinary approach that cuts across different aspects of the AI media environment, they seek to develop organizational structures and initiatives that embed 'hybrid human-AI workflows that reflect domain values'.[82] Yet, like many prior publications on responsible AI, the work of MediaFutures is predominantly focused on AI-based personalisation and algorithmic decision making. The result is that, despite growing public and academic interest, the issue of what governance, regulation and best practice ought to look like specifically for the responsible use of generative AI tools in media production still remains unclear.

---

[78] See https://braiduk.org/ for information on the programme's existing fellowship and scoping research.

[79] Kasirzadeh, Bird and Ungless, 'Policy Report on Generative Artificial Intelligence', p. 3.

[80] UK Music, 'UK Music Policy Position on Artificial Intelligence' (July 2023).

[81] Content Authenticity Initiative, <https://contentauthenticity.org/> [Accessed 19 August 2024].

[82] Christoph Trattner et al., 'Responsible Media Technology and AI'.

1.8: Conclusion

From this review of literature, it is clear that both academic and popular narratives concerning AI ethics, posthuman collaboration and responsible AI are not necessarily in short supply. What is missing, however, is a consideration for these concepts from the perspective of generative AI specifically. It is at this juncture that this research operates, integrating questions about what generative AI media does to creative labour, artistry and UK legal frameworks into responsible AI frameworks in order to offer a set of guidelines and policy recommendations intended start the creative industries on a journey toward responsible AI practice. By scoping out the key issues concerning different stakeholders within the media landscape and potential solutions to their concerns, this paper seeks to understand what generative AI tools are doing to our understanding of concepts like creativity, bias and collaboration. This will be the task of Chapters Three, Four and Five, which will interrogate each of these concepts in turn by reflecting upon the contributions of our workshop participants, Expert Bridging Group and our own creative practice in the process. Each chapter will provide a critical survey through the key concerns and opportunities for change that exist across the media landscape in order to establish what needs to be changed we want to facilitate responsible generative AI use in media production. Yet, before we do this, we need to consider the methodologies we used to undertake this research project, which will be the subject of Chapter Two.

CHAPTER TWO

Methodologies

In order to understand what different stakeholders within the UK creative industries see as the possibilities and limitations posed by generative AI tools to their production practices, it is necessary to examine the methodologies that we used to collect and collate their views. As such, this chapter will outline the main methods that were used for data collection and analysis during the research period of 'Shared Post-Human Imagination: Human-AI Collaboration in Media Creation'.

The chapter starts by tracing through the methods we used to collect data from the participants who agreed to take part in the project across our four main workshops (2.1) as well as a brief survey of our participant numbers and demographics (2.2). Following this, it examines the design process of our workshops, including structure, tool use, and the pre- and post-surveys that we utilised to gather our participants' thoughts on AI in media production (2.3). After unpacking the methods and tools we used to analyse our wide range of data sources (2.4), it offers some preliminary observations regarding the analysis of our quantitative data sources for the workshops (2.5). It then turns to consider how we convened an Expert Bridging Group as a way of offering possible solutions to the problems thrown up over the course of our workshops (2.6), before ending by reflecting on these methodologies in light of our own creative practice, outlining why we decided to undertake a creative project in order to interrogate what responsible AI collaboration would mean for independent film and media producers (2.7).
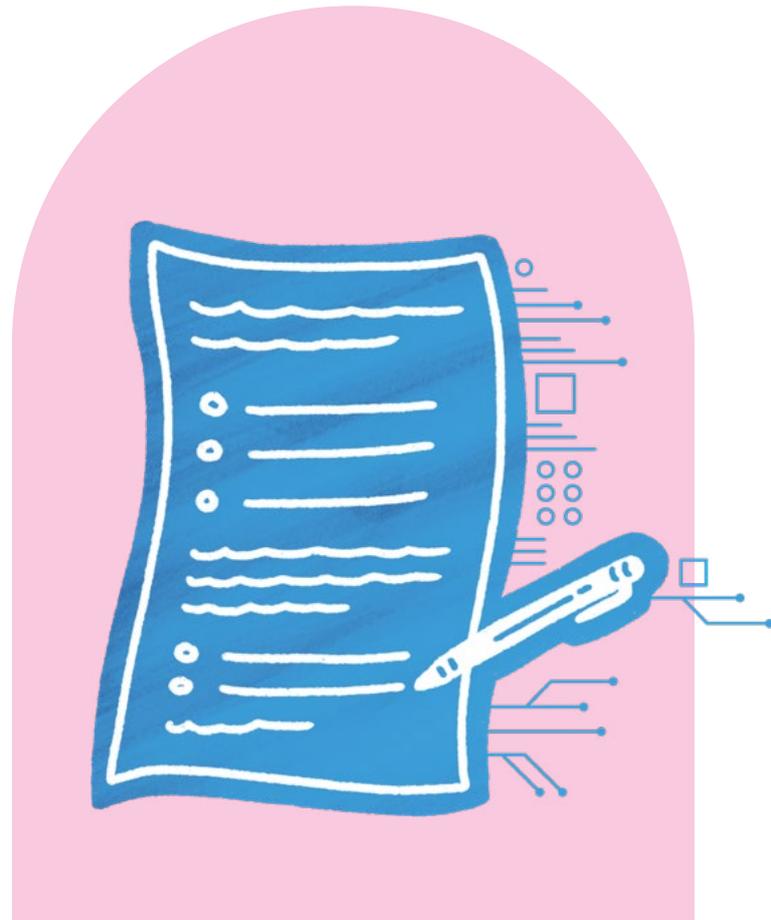
2.1: Workshop Methods

In this study, we employed a mixed methods approach to data collection, combining quantitative and qualitative data sources. Our research design included four workshops and one Expert Bridging Group. Quantitative data were gathered through structured questions in pre- and post-workshop surveys, providing measurable insights into participants' knowledge, attitudes, and perceptions before and after the workshops. Qualitative data were collected from multiple sources: open-ended questions in the surveys, collaborative Miro boards used during the workshops, transcriptions of workshop discussions, and the transcriptions of Expert Bridging Group session. The quantitative survey data offered statistical insights into participants' demographics, their familiarity with AI, and their perceptions towards using AI in media production, while the qualitative data from Miro boards, open-ended survey responses, and transcriptions provided contextual information and captured the participants' opinions, concerns and experiences.

2.2: Participants

In line with the research project's goal of bridging responsible AI divides, the participants across our workshops ranged from academics, media producers, filmmakers, industry professionals and end users. A total of 192 people registered for the workshops, while we received a total of 110 responses (age range 18-81, 58 male, 47 female and 1 'agender'[83]) were gathered from our surveys. These responses were distributed as follows: 45 for Workshop 1, 18 for Workshop 2, 27 for Workshop 3, and 20 for Workshop 4.

Bournemouth University Research Ethics Committee approved the ethics (Ethics ID 54026). Data collection took place between April 26th 2024 – August 2nd 2024. Participant registration took place via Eventbrite hosting pages, which were circulated to potential participants through established mailing lists within academia, media and the creative industries. Upon registration for each workshop, participants were provided with a Participant Agreement Form (Appendix 4), which outlined the funding, purpose and structure of the research project, as well as how their data will be recorded and managed in line with GDPR regulations and Bournemouth University's Research Participant Privacy Notice. Additionally, they were informed of their rights for confidentiality and anonymity, their right to withdraw from the study, and their right to access the findings of the study.

2.3: Workshop Design

2.3.1: Structure and Materials (see Appendix 7)

We conducted a series of four workshops that began on April 26th and concluded on June 29th. Each workshop took place over two days: the first day lasted roughly 6.5 hours with a 1.5 hour break, while the second day lasted approximately 3 hours with a 30 minute break. Each workshop covered a different sequential aspect of a typical film production: Screen-writing, Image Creation, Editing, and Sound and Music with AI. The goal of each workshop was to scope specific generative AI tools by using and misusing them, reflecting upon the issues that arise in the process and producing prototype artefacts through experimentation. In doing so, we sought to foster dialogue and feedback loops among various stakeholders to develop best working practices for responsible AI use in media production. We achieved this by combining invited guest speakers with practical experimentation and collaborative discussions regarding the possibilities and limitations of each tool being experimented with. Guest speakers were invited based on our research team's pre-existing network of experts across various fields within the AI media landscape, including renowned scriptwriters, artists, editors and musicians, as well as computer science PhD students, union representatives, technicians, lawyers and academics. In doing so, we sought to strike a balance between considering the technical efficacy of different generative AI tools for different purposes and reflecting upon their cultural, social and legal ramifications.

---

[83] This is the self-reported term that this individual participant used to describe their gender

The first day of each workshop was dedicated to presentations and discussions, while the second day was set aside for in-depth discussions by taking a systems thinking approach to AI. **By simulating media production processes with AI tools, these workshops sought to understand how AI can be used responsibly, addressing concerns around copyright, Intellectual Property, bias, and the impact on human creativity and labour. Where our workshops were designed to provide participants with a space to critically reflect upon their own use of generative AI tools in media production, they also served as a research methodology for our team. Where this was the case, the level of participant involvement ranged from mere consultation to full collaboration and control. Therefore, it was important for researchers to strike a right balance, facilitating discussions collaboratively while ensuring participants have a greater voice and role.[84] This is why, in the workshops, participants first discussed the topics in small breakout rooms as groups with our researchers, before sharing the insights gathered from these group discussions with a larger audience.**

At the start of each new workshop, participants were introduced to the online collaboration platform, Miro, which facilitates collaborative and interactive discussions.[85] It enables asynchronous collaboration, allowing participants to engage with the content on their own time, and supports a variety of activities such as brainstorming and visual organisation of ideas.[86] This aligns well with our workshops that involve complex discussions and the exploration of multifaceted concepts. By using our curated and specifically designed Miro boards,

participants were able to summarise the key points and outcomes from their discussions. They also enabled participation in our workshops in ways beyond the confines of oral discussion, as the Miro boards remained open after the end of each workshop to allow participants to add any further comments or queries.

After our expert presentations and practical workshops designed to interrogate each respective AI tool, the first day of each workshop ended with a series of guiding questions aimed at gaining insights regarding the advantages and disadvantages of working with AI tools, current and potential applications of AI in media creation, and ethical considerations arising from the use of these technologies. These initial prompts served to frame the subsequent discussions and activities for day two.

The second day of each workshop consisted of in-depth explorations of five primary themes: stakeholders, collaboration, creativity, bias and representation, and future directions. We utilised a systems thinking approach to our discussions in order to encourage a holistic understanding of AI as not just an isolated set of algorithms or tools but as part of a larger socio-technical system to consider how AI interacts with people, processes and data across its entire lifecycle of development, implementation and regulation.

By working through the key stakeholders involved in AI media production and, subsequently, where dynamics of communication, collaboration and power imbalances could be mapped onto these stakeholders, participants helped us

---

84 Rikke Ørngreen and Karin Levinson, 'Workshops as a Research Methodology', The Electronic Journal of e-Learning, 15: 1 (2017), 70-81.

85 www.miro.com

86 Thomas Chan, Jason Ho and Michael Tom, 'Miro: Promoting Collaboration Through Online Whiteboard Interaction', RELC Journal (2023).

produce a stakeholder map (Appendix 6) that captured the complex ecosystem of AI and media creation. In examining collaboration, participants discussed the impact of AI on authorship, copyright, and intellectual property. They also considered the broader implications for human labour, skills development, and the evolving job market in creative industries. The theme of creativity prompted participants to reconsider their understanding of these concepts in light of AI capabilities. This included debates on how to evaluate the quality of AI-produced media and how their own creative processes and outputs might be influenced by AI tools. When addressing bias and representation, the workshop explored AI's potential to either perpetuate or challenge existing biases and stereotypes in media creation. Ethical considerations were at the forefront of this discussion, with participants expressing concerns about the nature of training data and the implications of using systems trained on others' creative works. The future directions component encouraged participants to envision how they might shape the future of AI in media production.

The data collected from across all four workshops, in line with Bournemouth University's ethics approval and informed consent from each participant (Appendix 5), consisted of recorded group discussions, Miro board reflections, and a set of surveys, comprised of both quantitative and qualitative elements.

2.3.2: Pre-Survey

Prior to conducting the workshops, a pre-workshop survey was circulated to everyone who registered via our Eventbrite pages. This survey, which included both open-ended and closed-ended responses, aimed to
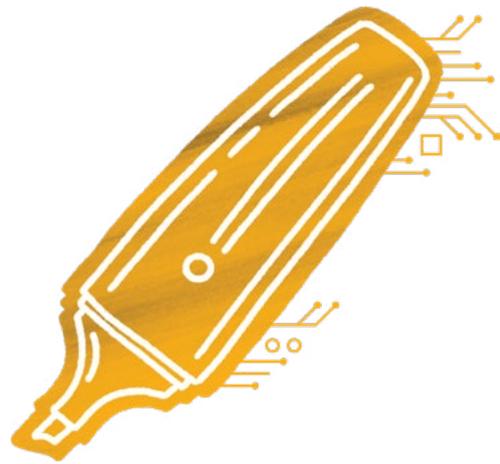
gather preliminary information about each participants' attitudes, understandings, experiences with generative AI tools and viewpoints on responsible AI. See Appendix 3.1 for our pre-survey questionnaire.

2.3.3: Post-Survey

After conducting the workshops, we also circulated a post-workshop survey to everyone who attended the workshop. This survey also included both open-ended and closed-ended responses, and aimed to gather follow up information regarding whether our participants' initial assessments of attitudes, understanding, experiences and viewpoints had changed in light of the activities undertaken throughout the course of the workshop. See Appendix 3.2 for our post-survey questionnaire.

2.4: Data Analysis

Qualitative data, including open-ended survey responses, Miro board content, and transcriptions from workshops and the expert focus group, were analysed by a team of researchers using thematic analysis using NVivo version 20, a qualitative data analysis software. Virginia Braun and Victoria Clarke's six step framework was followed which included data familiarisation, coding, theme identification, review, and definition, concluding with report generation.[87] We conducted an inductive analysis of the participants' responses to uncover recurring themes and patterns that emerged across

---

[87] Virginia Braun and Victoria Clarke, 'Using Thematic Analysis in Psychology', Qualitative Research in Psychology, 3: 2 (2006), 77-101.

Table 1

the qualitative data.[88] To ensure reliability, the primary author conducted the initial analysis, followed by collaborative discussions among all co-authors to develop a comprehensive codebook. This codebook guided the subsequent in-depth thematic analysis, which was then reviewed and validated by co-authors to ensure consistency and thoroughness in the interpretation of the data.

Quantitative data from pre- and post-workshop surveys were analysed using SPSS version 28, with a focus on descriptive statistics to provide an overview regarding participants' demographics, their familiarity of AI and their perceptions towards AI in media production.

2.5: Quantitative Results

2.5.1: Participant Demographics

In total, 110 participants completed the online survey. 4 people did not prefer writing their age. Table 1 summarises demographics.

Participant demographics	
N	110
Age: M(SD)	47 (13.3)
Age: Range	18 – 81
Gender: Males (%)	58 (52.7)
Females (%)	47 (42.7)
Race	
White (%)	88 (80)
Black or African American (%)	4 (3.6)
Asian (%)	9 (8.2)
Other (%)	9 (8.2)
Profession	
Academic (%)	41 (37.3)
Academic, Media production (%)	23 (20.9)
Media production (%)	28 (25.5)
Tech developer (%)	1 (0.9)
Other (%)	17 (15.5)
Frequency of using AI (%)	
Daily	14.5
2-3 times a week	13.6
Once a week	12.7
Once a month	10
Less than once a month	49.1

---

88 Lisa Given, The Sage Encyclopedia of Qualitative Research Methods (London: SAGE, 2008).1.

2.5.2: Familiarity with Generative AI

The majority of our workshop participants (45.5%) reported being **moderately familiar** with generative AI. 13.6% reported being **very familiar**, and 1 participant (0.9%) reported being **extremely familiar**. However, a notable proportion of participants (10.9%) reported being **not familiar at all** with generative AI, suggesting that there is still a need for education and training in this area. Overall, the mean familiarity score was 2.65, indicating a moderate level of familiarity among our participants on a five-point scale.

Figure 1

Familiarity with generative AI
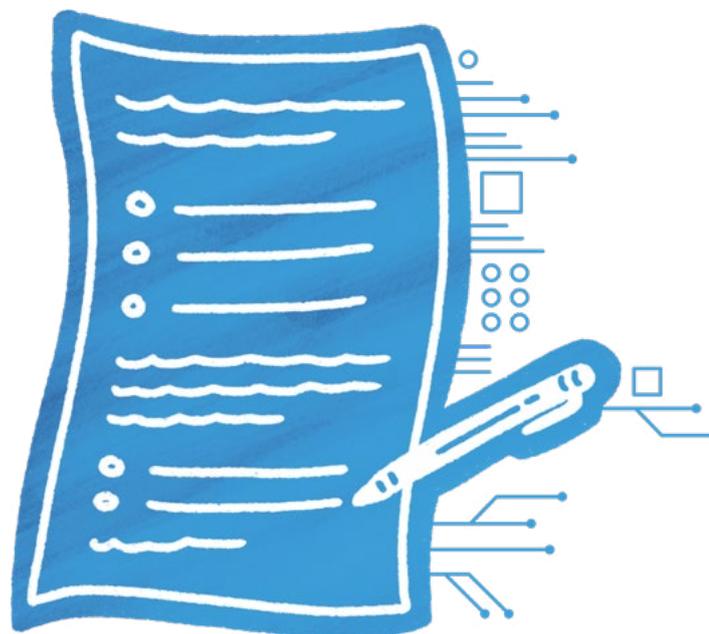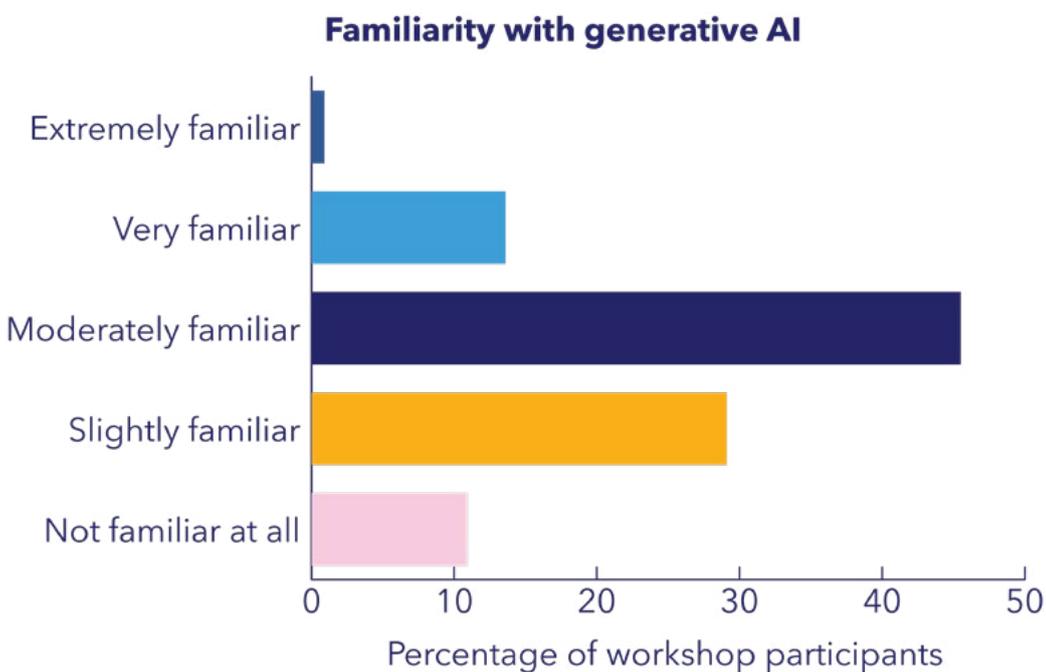
2.5.3: Attitudes Towards Generative AI

When asked to rate their attitudes towards the use of generative AI in media production, on a scale from 0 being the most extreme negative view to 100 being the most extreme positive view, our participants expressed a generally positive attitude, with a mean score of 61.94. This suggests that the majority of our participants see the potential benefits of using generative AI in media production.

Figure 2

Distribution of Attitudes with Mean Highlighted
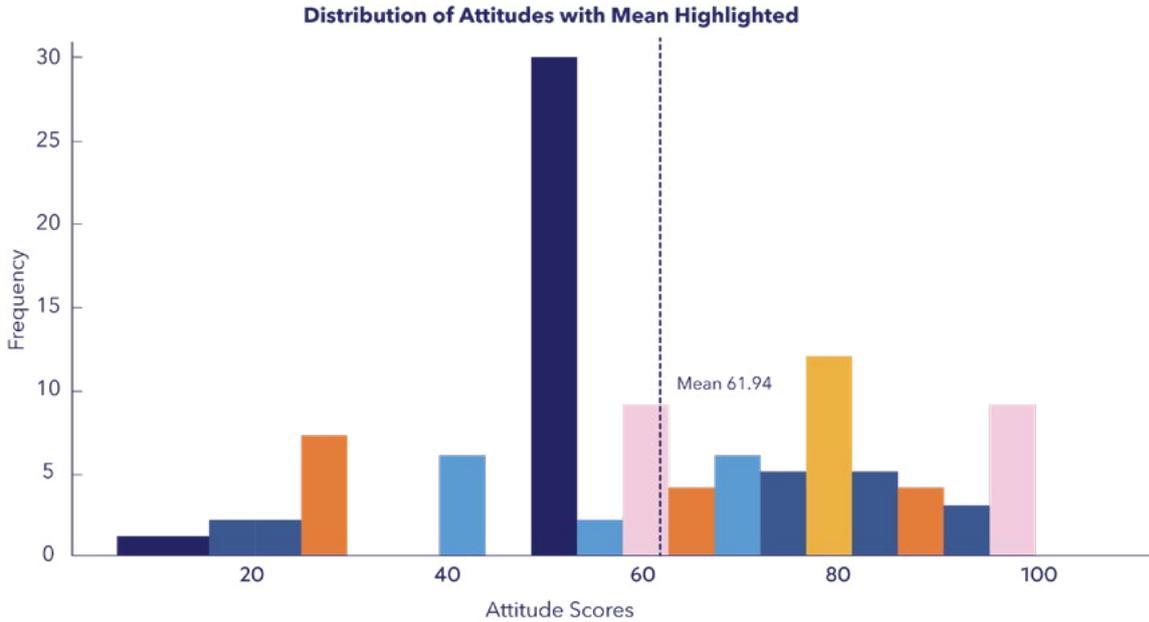
2.5.4: Generative AI Tool Preferences

When it comes to the specific tools they are using, ChatGPT was the most popular, with 31.25% of participants reporting its use. Other popular tools included Midjourney, DALL-E, and Claude. Additionally, a notable proportion of participants (20%) reported using a variety of other tools, including but not limited to Leonardo, AI Horde, Ollama. This suggests that there is a diverse range of tools being used in the field, and that our participants are actively exploring different options to find the ones that work best for them.

Figure 3

Percentage Distribution of Tools
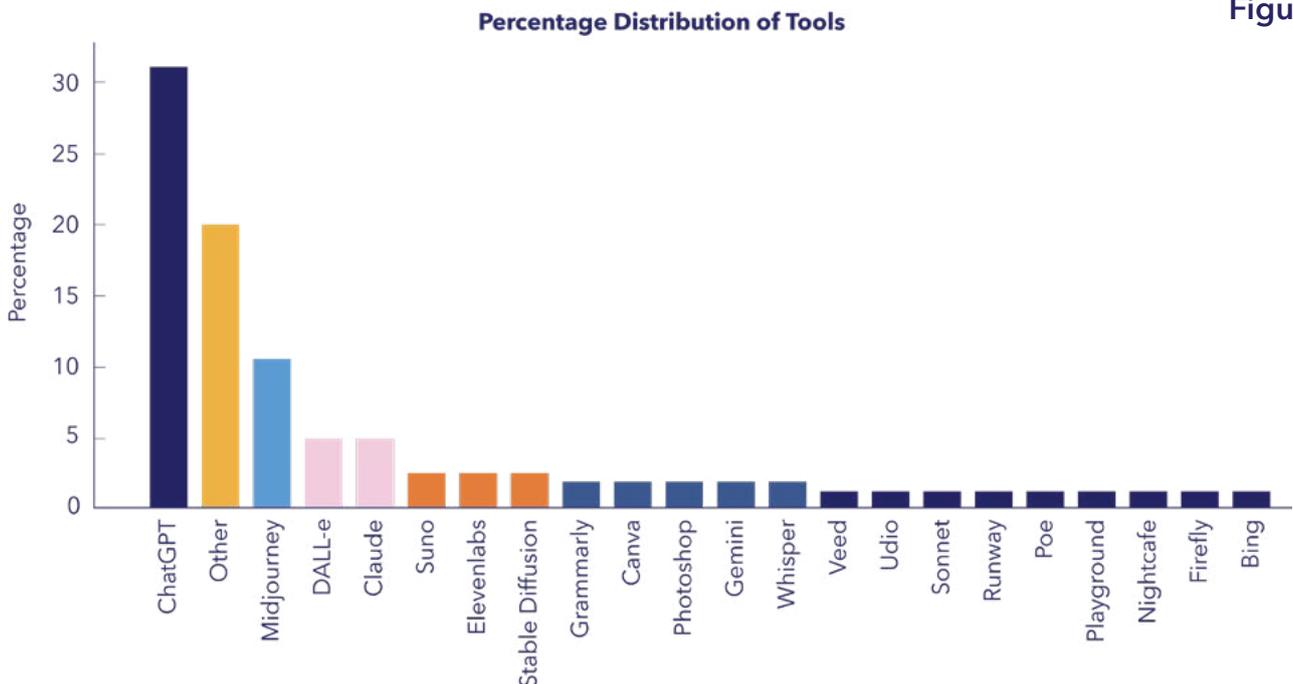
2.5.5: GAIT Competency

The results of the surveys reveal a striking lack of confidence among participants in their ability to use generative AI. 81.8% of participants reported feeling **very incompetent**, **incompetent**, or only **somewhat competent**, indicating that they predominantly do not see themselves as technically proficient in using generative AI tools. This suggests that the majority of participants are struggling to effectively use generative AI and may be in need of additional training or support to build their skills and confidence. Only 18.2% of participants reported feeling competent or very competent.

Figure 4

Perceived Competence in Using GAIT
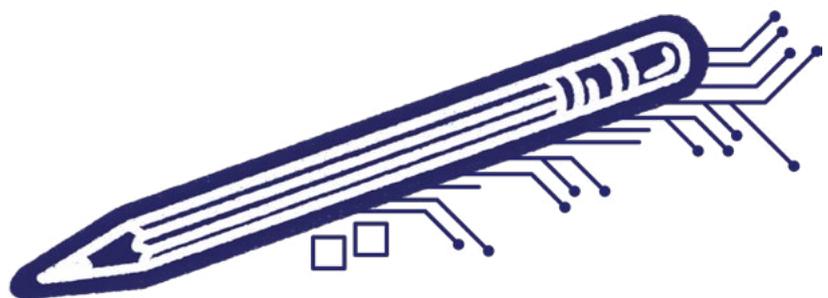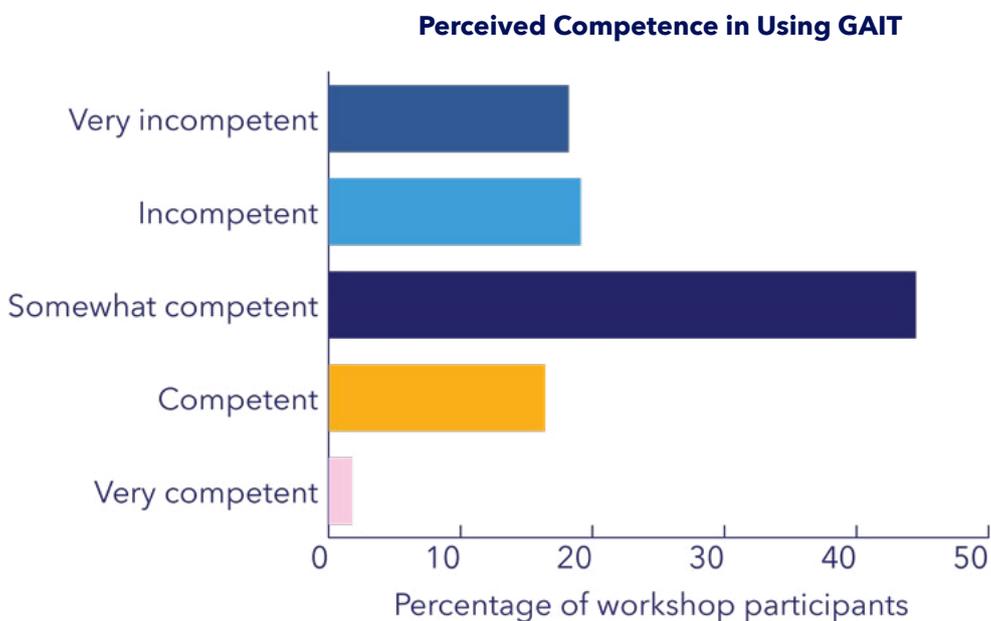
2.5.6: Significance of Key Topics

The survey explored the significance of key topics including **collaboration, creativity, copyright, productivity, labour and bias** in creators' roles. Participants were asked to rate the importance of these issues on a 5-point scale (1 = not at all important, 5= extremely important). The results suggest that all these topics are seen as important, as the mean scores for each item were above the midpoint of 3. Creativity emerged as the most significant topic with a mean score of 4.54, indicating that creators prioritise the creativity aspect of AI when working AI tools. Collaboration was also rated highly with a mean score of 4.21, highlighting the importance of collaboration and shared efforts in AI-driven media production. Copyright, scoring a mean of 4.17, is also a key concern, likely because generative AI tools introduce challenges around IP and ownership. While productivity (M=4.07) and labour (M=3.98) are rated slightly lower, they remain critical, indicating that efficiency and the impact of AI on the workforce are still important. Bias with the mean score 3.79, has the lowest score, suggesting that further education on the potential harms of generative AI tool use is needed. However, the fact that it still scored above the midpoint of 3 indicates that there is a growing awareness of ethical considerations and fairness in AI.

Figure 5

Significance of Key Topics from the Survey Results
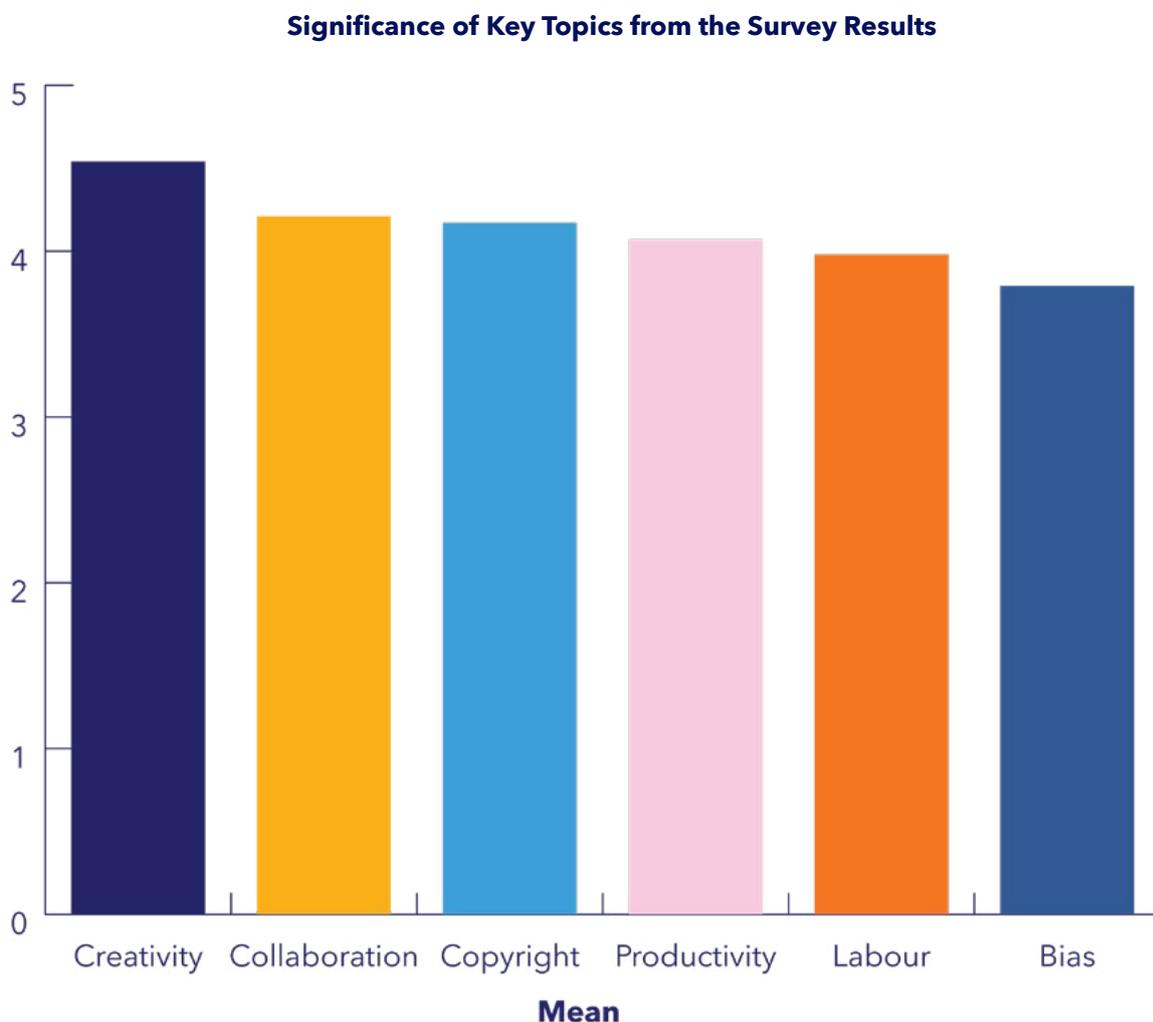
2.5.7: Responsible AI

In our survey, we also asked participants about their understanding of the term 'responsible AI' through an open-ended question asking them to define the concept in their own words. The responses ranged from expressions of unfamiliarity with the term to attempts to provide a definition, focusing on ethical considerations, fairness and the impact of AI on creativity and intellectual property. A significant proportion of responses indicated some level of unfamiliarity with the concept of responsible AI. Many participants either admitted to having no knowledge of the term or they provided speculative answers, associating it with legal or ethical responsibilities. This highlights a gap in awareness and suggests the need for further education. Those who provided more defined answers often conceptualised their understanding of 'responsible AI' in relation to broad ethical principles, as they used associative terms such as fairness, transparency and accountability. For example, participants noted the need for AI to be 'traceable and devoid of bias', to ensure 'transparency and fairness' and to avoid harm or misuse.

A common theme from the participants' answers to this question was the protection of intellectual property and the importance of financial remuneration for human artists. Many participants mentioned that responsible AI should respect copyright and not infringe upon the rights of content creators. The risk of AI replacing human labour, especially in creative fields, was another major concern. Participants suggested that responsible AI should serve as a collaborative tool rather than a replacement for human labour.

Ethical AI use in terms of accurate and truthful outputs also emerged as a key component of our participants' initial understanding

of responsibility. Several participants highlighted the need for AI to avoid spreading misinformation and they pointed out the importance of fact-checking AI outputs. In contrast, a few participants framed responsible AI from a more sceptical perspective. Some considered the term 'useless', while others pointed out that responsibility might lie more with the user than with the AI itself. These statements highlight the need for a holistic, systems-thinking approach to responsible AI as a collective, ongoing and relational process involving all stakeholders.

2.6: Expert Bridging Group

After conducting our four key workshops, we convened an Expert Bridging Group in order to scope out tangible solutions and interventions that could be realistically implemented, both within best working practice and UK policy, to promote responsible AI use in media production and mitigate some of our participants' concerns.

This took the format of an in-person focus group, which took place on July 22nd 2024 at Bournemouth University, Talbot Campus. The discussion consisted of 12 invited members, who were compensated for their preparation for, and participation in, the Expert Bridging Group. The group brought together academics, filmmakers, industry professionals and end users together to discuss the key issues that emerged from our four workshops. In doing so, it built upon the diversity of the workshops by bridging across areas of technical, ethical, philosophical, legal and industrial expertise. As part of the formation of our Expert Bridging Group, we made the conscious decision as a research team to re-invite a chosen selection of our workshop participants and previous guest speakers with the goal of generating conversations between people who were already familiar with our research and those who were entirely new to

the project. As part of the agreement upon joining the Expert Bridging Group, members have been afforded anonymity in this report to protect the disclosure of any sensitive or confidential information.

While the Expert Bridging Group was largely successful in bringing together disparate groups and having different stakeholders reflect upon each other's unique positionalities when it comes to the AI media landscape, one limitation of the group was our inability to attract generative AI developers to the conversation (only one member of our Expert Bridging Group worked for an AI company). We reached out to a number of technology companies who have developed generative AI tools for media production, but whereas smaller companies declined citing a lack of time, larger companies did not respond at all. This latter point regarding well-known and established AI developers was particularly disappointing given the enthusiasm with which many other stakeholders responded to our invitation to discuss these urgent issues. Our workshop participants acknowledged that while software developers and technology companies hold the influence of power over producers, distributors and audiences of contemporary media, they simultaneously noted the possibility for influence that researchers, creators and policymakers can have on AI development if given the chance (see Appendix 6 for more details). This indicates that part of the 'bridging' of AI divides that our participants deem as the most important is involving AI developers in the discussions and actions regarding responsible AI. Their decision to largely not take part in our research also means that their voices, needs and concerns are not necessarily reflected in our recommendations for best practice and policy intervention, prompting the need for further research.

Across three discussion sessions chaired and facilitated by different members of our research team, our Expert Bridging Group convened to answer semi-structured interview questions that both ensured consistency in preparation and allowed for flexibility and the chance to respond to other members' comments. These interview questions were provided to the Expert Bridging Group members in advance of the session and centred around our three main topics for the research project: creativity, bias and collaboration.

Data collection for the Expert Bridging Group discussion replicated the qualitative data collection methods undertaken across the four workshops. These methods included a digital Miro board, accompanied by in-person sticky notes; audio recordings that were later transcribed and coded by a team of researchers; written notes from a member of our research team present on the day; and any preparatory materials that our members used as part of their discussions that they also shared with us after the event.

While the results of the Expert Bridging Group will be discussed in more detail across Chapters Three, Four and Five, a summary of the key proposed interventions and policy recommendations that came out of the focus group can be found in Appendix 4. Many of these have proceeded to influence our recommendations for governance and policy intervention, which can be located in Chapter Six.

2.7: Creative Practice

The final part of our mixed methodology for this research project was to approach the main research questions with a practice-based research methodology. By critically reflecting on the contributions of our workshop participants, Expert Bridging Group and our

own reflections as media practitioners and academics, we have drafted a list of best working practice principles and guiding questions, which will be examined in further detail in Chapter Six. In an iterative process, we applied and updated the list of best working practices, while working collectively as a team to create a short film in collaboration with different generative AI tools.

The process began with an ideation session without the use of AI, in which we recognised a shared desire within our research team to depart from the traditional, auteurist, oftentimes hierarchical format of film and media production in which a singular director's vision is followed from start to finish. Instead, we agreed to utilise generative AI tools not as a replacement for human-human collaboration but as a supplement to it, such that all members of the research team would engage creatively in all parts of the film's production process. During our ideation session, we agreed to critically explore contemporary themes of credit and ownership in a meta-reflection that draws inspiration from the life and works of Mary Shelley, infamous both for the renowned literary classic, Frankenstein: Or, The Modern Prometheus (1818), and for the Shelley family's historic ties to the town of Bournemouth. This spawned a thematic connection the technoscientific, cautionary-tale narrative of Frankenstein with modern developments in generative AI, out of which formed a series of conceptual imagery and ideas with which to tease out these associative connections.

This initial ideation session laid the groundwork from which members of our research team collaborated with text-, image- and music-generating tools to produce a series of artefacts. The production processes of these artefacts were subsequently documented, including reflection on responsible AI practices, and our main themes of collaboration, creativity and bias.

We shared the experience that integrating generative AI tools in our already existing creative practices was not necessarily intuitive nor easy. Members of the team pointed out the initial lack of control of the final output combined with the necessity of a steep learning curve and considerable time investment to personalise the AI tools. We have discussed the fact that prompting, the prevalent interaction form with AI tools, expects a language-based, nuanced and explicit formulation of ideas which in some creative practices minimises the possibility of a more intuitive, serendipitous process. Team members also reflected the fact that in some instances, the generated output is set in a way that hinders further collaboration or development of the artefact. For instance, some members of the team noted the generic quality of the image outputs, in particular, experimenting with images of Frankenstein's monster that closely resembled his depiction within popular media. These examples highlighted how deeply AI models are trained on pre-existing cultural artefacts without actually citing their sources, producing a further ethical dilemma regarding the appropriation of copyrighted material without consent or recognition. Likewise, our team highlighted that the lack of diversity within AI-generated outputs, unless explicitly stated, places the onus on the creator to actively counteract inherent bias.

In the second round of this process, we responded to the AI-generated artefacts created in the first round through text-based reflections and further image and music creation that built upon some of the abstract visual metaphors established in our earlier ideation session. We followed an approach akin to the 'exquisite corpse' method of art production made popular by the Surrealist movement.[89] This formed the basis from which

89 Tate UK, 'Cadavre Exquis (Exquisite Corpse)' <https://www.tate.org.uk/art/art-terms/c/cadavre-exquis-exquisite-corpse> [Accessed 10 September 2024].

to produce a screenplay and images with the help of generative AI, the latter of which were subsequently edited together and scored/soundtracked using generative AI tools. While this process necessitated that we understand what our participants and Expert Bridging Group members deemed to be 'responsible' use of generative AI, the process of producing a film based on responsible AI principles and guidelines also allowed us to test out their practicalities and subsequently update them as necessary. This mixed methodology approach enabled us not only to see the pressing needs and concerns of those working within the film and media industry in the UK, but also to consider what practical steps, guidance and recommendations might help people in embedding further transparency and combating bias when creating media with generative AI.

2.8: Conclusion

This chapter sought to provide readers with an overview of our mixed methodologies approach to our research. It prefaced the following three chapters by contextualising our key findings within both quantitative and qualitative analysis taken from a range of sources, such as surveys, Miro boards and collaborative discussions amidst workshop participants. It also established some of the preliminary findings of our data analysis, demonstrating the need to capitalise on peoples' general desire to use generative AI tools while providing further clarity on what constitutes best practice. By combining the scoping of key concerns and possibilities within the creative industries with both the formation of an Expert Bridging Group to workshop possible interventions and our creative practice experimentation, we forged a multidisciplinary framework from which to approach responsible human-AI collaboration. The following three chapters will utilise framework to enrich this scope of generative AI in media creation and interrogate our main research questions in further detail, paving the way for our recommendations for best practice and policy intervention, which we present in Chapter Six.

CHAPTER THREE

Creativity

As AI developers design new tools to generate stories, make music and produce images, and users and media practitioners begin to integrate these same tools into their workflows, the notion that creativity is an inherently human trait is increasingly placed into doubt. While a series of opportunities emerge at the precipice of AI-generated creative outputs, there exists a wave of uncertainty within the UK media industry regarding what these new tools will do to the role of the human creative that has, up until now, laid at the heart of media production. Fundamental to these debates are the concepts of artistic value, skill, originality and authenticity. It will be the task of this chapter to survey concerns and hopes within the UK creative industries relating to the rise of GAIT, so as to better understand how the notion of media creativity is being re-evaluated within a context of responsible AI, and, therefore, to explore how we can ensure that augmentations to human creativity are happening in ways that still protect against extractive database practices, intellectual property infringement and the displacement of human labour.

This chapter is divided into two sections. The first section, '3.1: Workshop Findings', details the results and discoveries from across our four workshops (see Chapter Two for more details on the methodology employed in these workshops). After unpacking academic definitions of creativity within different contexts (3.1.1), the chapter charts a trajectory through our participants' understanding of the changing nature of creativity (3.1.2), the role of prompting in new creative processes (3.1.3),
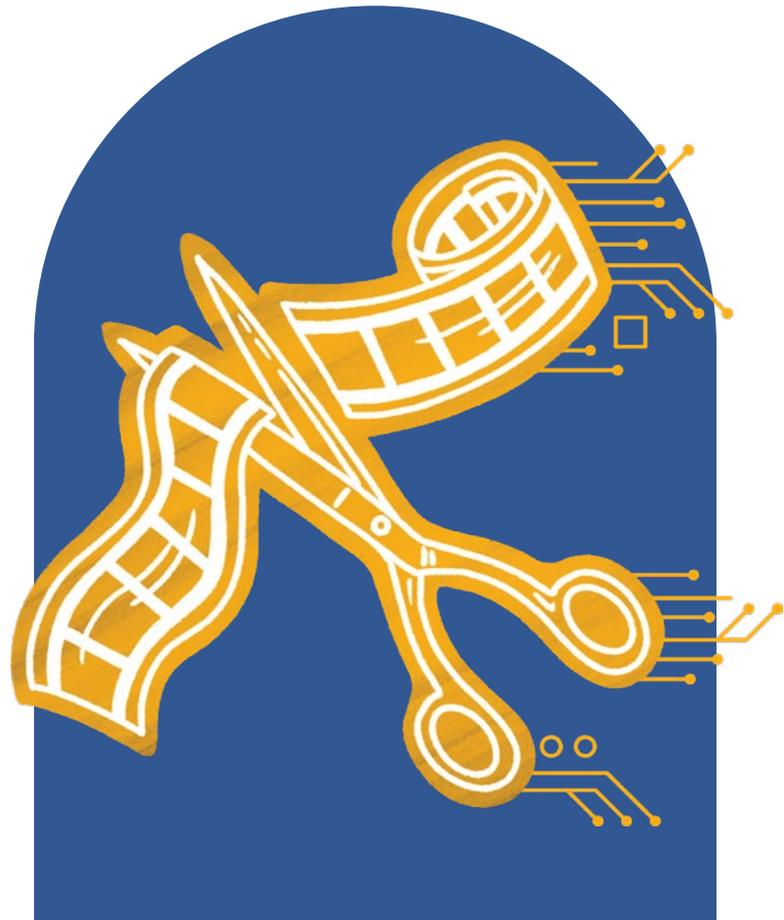
and their differing perspectives on creativity as a fundamental human trait (3.1.4). Following this, the section pivots to consider potential use cases for AI creativity, such as managing automated routines that leave more time for human creative works (3.1.5), enhancing human creativity (3.1.6), and the potential democratisation of creative production (3.1.7). Yet, amidst these opportunities, a prevailing concern persists; how does AI-generated creative outputs devalue art? The section unpacks this question further, considering generative AI from the perspectives of generic outputs and CGI (3.1.8), before examining our participants' reflections on authenticity, originality and accuracy (3.1.9). Section two of this chapter, '3.2: EBG and Creativity', builds upon and contributes to the debates that circulated within our workshops by examining some potential solutions to these concerns that were offered by members of our Expert Bridging Group. It outlines the questions we circulated to the EBG in advance of the 'Creativity' discussion session (3.2.1), before considering the various ways in which members of the EBG have sought to

integrate, or restrict, the use of generative AI for creative purposes in their own work (3.2.2). Throughout this chapter, we offer reflections on the thoughts of our research participants from our own creative practice experiences, unpacking some of the underlying assumptions and potentials for change that emerge when utilising AI for creative tasks.

3.1: Workshop Findings

The workshops revealed multiple tensions relating to the perceived impact of generative AI on the UK creative industries. What emerged as a benefit for some proved to be a drawback for others. This tension was a recurring theme that manifested in various contexts throughout all of the workshops. For instance, the democratisation of creative tools was seen as a positive development by some, but it also raised concerns about the devaluation of professional expertise by others. Such tensions underscore the need for a thorough and nuanced investigation into the usage of generative AI in media production, exploring the complex issues and views surrounding AI-driven creativity so that we can ensure that generative AI tools are carefully used and designed to address the diverse needs of different users and stakeholders.

3.1.1: Creativity as a Concept
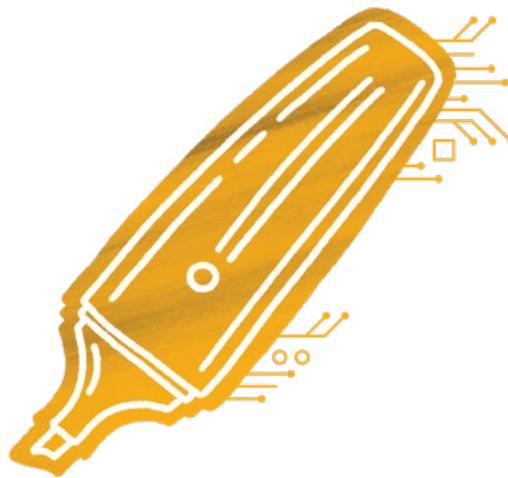
Creativity is a complex concept, and although there exists decades of research and countless attempts at definition, a singular, universally accepted understanding of the term remains elusive. Despite the existence of numerous competing and often contradictory definitions, with hundreds documented to date, a comprehensive understanding of creativity continues to be a subject of ongoing research and inquiry across the Arts, Humanities, Social Sciences and STEM.[90] The Merriam-Webster dictionary defines creativity as 'the ability to create', a circular definition that does not go far into understanding what makes creativity unique as a concept.[91] The Cambridge dictionary takes this one step further and instead describes it as 'the ability to produce or use original and unusual ideas', out of which one can extrapolate associations of creativity with imagination and innovation.[92] One of the most widely accepted definitions within cognitive science is summarised by Margaret Boden, who writes, 'creativity is the ability to come up with ideas or artefacts that are new, surprising, and valuable'.[93] Although formative in understanding the complex web of associations and connections that make up creativity as an umbrella term, these definitions are nevertheless subject to intense debate depending on context and discipline, as will be demonstrated when we consider the views of our workshop participants.

Yet, for decades, two determinants have remained consistent in characterising

90 Andrei Alienikov, Sharon Kackmeister and Ron Koenig, Creating Creativity: 101 Definitions (what Webster Never Told You) (Midland: Alden B. Dow Creativity Centre, 2000); Giorgio Franceschelli and Mirco Musolesi, 'Creativity and Machine Learning: A Survey', ACM Computing Surveys, 56: 11 (2024), 1-41.

91 Merriam-Webster, 'Creativity' (2024) <https://www.merriam-webster.com/dictionary/creativity> [Accessed 23 September 2024].

92 Cambridge Dictionary, 'Creativity' <https://dictionary.cambridge.org/dictionary/english/creativity> [Accessed 23 September 2024].

93 Margaret Boden, The Creative Mind: Myths and Mechanisms 2nd edn. (London: Routledge, 2004).

creativity within psychology: novelty and appropriateness.[94] Keith Sawyer and Danah Henriksen, prominent psychologists in the field of creativity and innovation, describe creativity as 'part of what makes us human',[95] building upon Mihalyi Csikszentmihalyi's formative understanding of creativity as a uniquely human ability that separates humanity from other species.[96] Within this understanding, Sawyer and Henrikson categorise creativity research into two main approaches: the individualist and the sociocultural.[97] The sociocultural approach to creativity highlights the importance of novelty and appropriateness, such that what is considered 'creative' is contextually dependent on what has come before it and its situated position within a broader sociocultural environment. Sawyer and Henrikson note that, under the sociocultural position, 'new' creations always remain tied to convention, either through adherence of departure, indicating an emphasis on the maintenance of order and intention within the creative output. In contrast, the individualist or 'Romantic' approach to creativity is marked by a combination of existing thoughts and concepts into an idea in ways that have not yet been previously combined by an individual. The reality, as Paul Thompson and Phillip McIntyre recognise, is often a 'dynamic interaction' of the individual with the socio-cultural.[98] As we shall see both later in this chapter and when we come to consider collaboration in Chapter Five, the interface nature of generative AI tools updates both the individualist and socio-cultural view of

creativity through its synthesising function and requirement of a human input.

While this lack of a single shared definition opens up avenues for debate and critique across different disciplines, it nevertheless hinders the establishment of a common rubric among researchers, making it difficult to compare quantitative findings across studies. This ambiguity is particularly evident in the field of Human-Computer Interaction (HCI), where the focus on collaborative creativity and the development of supportive tools has led to a renewed interest in understanding the complexity of creative processes.[99] In their comprehensive survey of 998 creativity-related publications, Frich et al. found that when the HCI community engages with the topic of creativity, only few (28.37%) contributions directly define or clarify the concept of creativity.[100]

Given the challenges in defining and understanding human creativity, it is not surprising that the rise of AI-generated creative content has sparked intense debate about the nature of creativity itself. If creativity is understood as the process of combining existing thoughts and concepts in novel ways, as Sawyer and Henrikson noted regarding the individualist approach, then it raises the possibility that AI could also be considered creative on the basis of synthesising pre-existing material to generate new outputs. While some would dismiss this synthesising or sampling function as a lack of creative potential, due to the nature of AI models

[94] Frank Barron, 'The Disposition Toward Originality', The Journal of Abnormal and Social Psychology, 51: 3 (1955), 478-485 (p. 478); Mark Runco and Garrett Jaeger, 'The Standard Definition of Creativity', Creativity Research Journal, 24: 1 (2012), 92-96; James Kaufman, Creativity 101 2nd edn. (New York: Springer, 2016).

[95] R. Keith Sawyer and Danah Henriskon, Explaining Creativity: The Science of Human Innovation 3rd edn. (Oxford: Oxford University Press, 2024).

[96] Mihaly Csikszentmihalyi, Creativity: Flow and the Psychology of Discovery and Invention (New York: HarperCollins, 1997), 1-16.

[97] Sawyer and Henriskon, Explaining Creativity.

[98] Paul Thompson and Phillip McIntyre, 'Rethinking Creative Practice in Record Production and Studio Recording Education: Addressing the Field', Journal on the Art of Record Production, 8 (2013).

[99] Jonas Frich, Michael Biskjaer and Peter Dalsgaard, 'Twenty Years of Creativity Research in Human-Computer Interaction: Current State and Future Directions', Proceedings of the 2018 Designing Interactive Systems Conference (2018), 1235-1257.

[100] Ibid.

being trained on existing creative works, it nevertheless has the ability to offer outputs that appear new and original to some people on the basis that what they are seeing, reading or hearing may be unique based on their lived experience in the world. Yet, if we follow the perspective that creativity is an intrinsically human trait, characterised by the *Homo sapiens'* unique ability to generate ideas either from sociocultural embeddedness or from the originality of their own imagination, then the idea of AI being creative becomes more controversial to anthropocentric viewpoints.

According to a 2018 study, artists from across the world have been using AI tools as assistants for tasks that were deemed to not require a high level of creativity, such as image searching and, more controversially, editing, for years.[101] This finding suggests that the artistic community recognises AI's potential to augment, rather than supplant, human creativity. Yet, developments in AI have come a long way since 2018, and with the advent of AIs capable of producing (albeit imperfect) scripts, images, edits and soundtracks following guidance from a human end user, these claims need to be rethought both within a context of UK-specific media production and specifically generative AI as a model. In the following sections, this paper traces through the attitudes, behaviours and practices of participants across our four workshops, providing unique insight into how UK media practitioners engage with and use generative AI in their own creative processes.

3.1.2: The Changing Nature of Creativity

Across all four workshops, participants noted a shift in the creative process, highlighting how AI tools have transformed the traditional

creativity process. The increasing proliferation of AI-generated content has led to a blurring of lines between creators and prompters, raising essential questions about the nature of creativity, effort, and value. AI's role as an 'assistant' or 'tool' was mentioned frequently, with participants acknowledging its potential to enhance creative productivity as an add-on or compliment to pre-existing creative processes, rather than replacing these creative roles. Within the group discussions, many of our participants debated whether AI-generated content should even be considered truly creative given that it is based on pre-existing creative material, suggesting an inherent belief in creativity as a concept defined by how much one individual can ideate or produce based on their own original, unique ideas. Yet, as an important corollary to these discussions, one participant noted that 'in the shadow of AI, the role of creativity should be redefined'. Such comments suggest a shift in not only the role of creativity but also in how creativity is perceived whenever AI tools and models are involved.
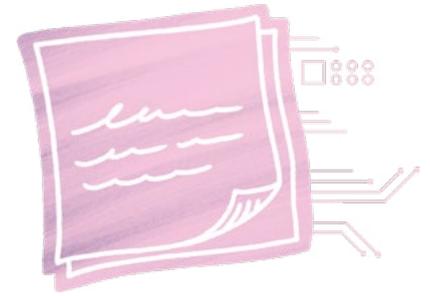
3.1.3: The New Creativity: Prompting

Within traditional models of creativity, the creative process has been associated with thinking outside the box, generating novel ideas, and bringing them to life through various forms of artistic expression.[102] However, with the rise of AI tools, the creative process has become increasingly dependent on prompting (see Appendix 1 for a definition of prompting). As one professional movie editor noted in the workshop, 'my everyday job has shifted from editing to prompting to an AI, verbalising, describing what I

---

[101] Andreas Pfeiffer, Creativity and Technology in the Age of AI (2018).

[102] Kaufman, Creativity 101.

imagine and how to edit'. This shift highlights the emergence of a new creative role, the prompter, but questions still remain about whether prompting can and should be seen as a creative process in and of itself, and from this, whether prompting can be considered a form of artistic expression.

Within our workshops, but particularly across Workshop 1 (Screen-writing), there existed a split amongst participants regarding their views on prompting and creativity. On the one hand, prompting an AI successfully requires a deep understanding of the individual tools' capabilities and the ability to communicate complex ideas and emotions through text. Many of our participants commented that they were surprised to see just how much time, effort and skill goes into constructing a successful prompt, particularly after our practical demonstration by an internationally recognised screenwriter. This process can be seen as a form of creativity, where the prompter must find innovative ways to guide the AI away from a generic output and towards the desired outcome. On the other hand, some participants appeared to express the view that prompting of an AI tool is less of a creative skill than a technical one, a view that was held particularly prominently in Workshop 3 (Editing). Under this understanding, the AI is the agent doing the actual creating, while the prompter is simply providing the input with which to make that creation happen. This split then led our discussion to ask the following question: who is the author and creator of this process? Is it the prompter, who came up with the idea and guided the AI towards a desired outcome? Is it the AI, which actually produced the final product? Or is it a combination of humans and AI, made all the more complex by the fact that AI models are trained to synthesise and augment pre-existing intellectual property?

3.1.4: Creativity as a Fundamentally Human Trait

Adjacent to concerns over authorship is the question of whether something other than human could ever be considered capable of creating. Despite the impressive human-like and photo-realistic capabilities of some generative AI tools, many of our participants still believed that creativity is profoundly human. Along these lines, while AI can generate output, it is the human input, human judgment, and human touch that are essential elements in the creative process. Participants who held this view noted that AI is simply a tool, and that its value lies in how it is used by the creative person. As one participant put it, 'it is a tool like crayon or pencil or paints. How do you, as a creative, use this tool as an extension of your storytelling?'. This framing speaks to the view that generative AI models, as tools, serve human goals as prosthesis, indicated by the use of the term 'extension', in which the human retains control and agency over the final output.

According to our participants, it is the human user who provides the context, nuance, and emotional depth that makes the output meaningful. Moreover, participants emphasised that AI cannot be creative if the prompt is not creative, returning us back to the notion that the quality of the output is proportional to the quality of the input, and if the prompt is dull or unoriginal, the output will likely be the same.

These ideas are underpinned by a view that creativity is about the generation of new and novel meaning. Based on our own creative practice experience, this suggests that creativity as a value is inherently human-centred, in which the very notion of creative meaning is endowed onto, and decoded from, works that resonate with our individual or socio-cultural understanding of novelty

or quality. This may change from person to person and culture to culture, such that creativity as a concept is constituted by the act of noticing value. The parameters of value may also be expressed differently between what a human would deem to be a creative output and what a generative AI tool would label as such. Within this context, the human user may be better thought of as a curator, rather than a creator, on the basis that they are assigning creative value to the works produced in collaboration with AI through 'acts of noticing'.[103]

A 46-year-old male academic shared: "The two workshops presented issues of creativity very differently. That was eye opening as I had not thought about issues of creativity too much before." Participants expressed that while AI may not possess creativity inherently, it can occasionally facilitate serendipity and novelty, leading to unexpected creative insights in the creative process. One participant shared their experience with AI in image creation (Workshop 2), noting that, although they often spend hours working towards a specific goal, they sometimes encounter what they called 'lucky moments' where the AI unexpectedly produces a desirable result. They highlighted that while creativity remains a distinctly human trait, AI, lacking inherent creativity, can only occasionally provide the desired unexpected outcome after hours of effort. This speaks to a view of human-AI collaboration that returns to humanist and anthropocentric conceptions of creativity as uniquely human, suggesting that the best way to approach generative AI in media production may not be through the lens of posthumanism, or the 'post-human', after all, but rather a retention of human control.

Amidst the possibilities that generative AI tools brought to creative processes within media production, there were also some concerns expressed by participants when it comes to creativity. One of the central concerns that arose was the belief that GAITs lack the ability to understand subtext and context, which was perceived to be a crucial element in creative decision making. Participants noted that this is particularly problematic in creative fields like scriptwriting or music production, in which emotional connection and authenticity are essential and may bound up in concepts of ambiguity, open to interpretation. A striking example given to us by one participant was the example of editing for comedy, in which the timing and pacing of a cut can make or break the humour of a given scene, suggesting a departure from assumptions of editing as a solely technical and uncreative task in media production.

3.1.5: AIs Managing the Routine, Freeing Up Time for Human Creative Work

Many of our participants expressed that they were using AI tools for what they perceived to be 'mundane' tasks, such as data management, organisation and time management. Those who did often stated a desire to maintain control over the creative aspects of their work, highlighting a dynamic where the AI serves as a tool to enhance productivity without overshadowing human creativity.

One of the main advantages of using AI, as perceived by our workshop participants, was its ability to handle technical and routine tasks, allowing creators to focus on the more imaginative, creative parts of their work. As one statement suggests, the 'main advantage [is] to use AI as a technician so that as a composer I can focus on the creative side of the music production'. From this viewpoint, AI allows human creators to focus their energy

103 Tyler Cowen, 'Rick Rubin on Listening, Taste, and the Act of Noticing (Ep. 169)', Conversations with Tyler (18 January 2023) <https://conversationswithtyler.com/episodes/rick-rubin/#:~:text=RUBIN%3A%20I%20think%20they're,isn't%20making%20the%20thing> [Accessed 1 October 2024].

on creativity and innovation, rather than being burdened by technical details. One of our participants describes this as the 'opportunity for faster technical process to have more time to bring out the creativity'. Moreover, the use of AI can lead to significant improvements in efficiency and productivity. By automating repetitive tasks, AI was seen as being capable of speeding-up work processes and providing creators with more free time for their personal time. As one statement notes, AI can 'help human beings create faster and have more free time for their families', indicating AI's potential to enhance, rather than threaten, the material lifestyles of human creatives.

By taking on these mundane yet necessary tasks, AI is reshaping workflows in media production, allowing creators to dedicate more time and energy to the aspects of their work that require creativity. While it is tempting to believe that removing certain tasks perceived as repetitive, such as background noise cleaning or image retouching, allows creators to focus solely on more creative' work, research also shows that engaging with these tasks can stimulate creative thinking in subtle but significant ways.[104] Additionally, certain forms of creative inspiration often come from the hands-on interaction with tools and materials. In filmmaking, directors and editors who involve themselves in tasks like trimming footage or correcting colour may encounter serendipitous discoveries. Therefore, AI doing the automated tasks could create a disconnect between the creator and the material, leading to a loss of creative insight.

3.1.6: Creativity is Human, AI Makes it Stronger

While many participants believe that creativity is fundamentally a human trait, they also acknowledged the potential of AI when it comes to enhancing and expanding creativity through ideation and fresh perspective. According to participants, AI plays a supportive role in the creative process as it 'fuels creativity' and offers the opportunity to 'reach new heights in art and artistic expression'. While AI may offer various choices, one participant noted that 'it doesn't reduce the time you spend; it just gives you a lot of choices to consider', indicating that AI does not replace the creative effort but enhances the range of options available.

Participants also highlighted that AI can open new creative possibilities that might not have been considered otherwise. One participant reflected that AI allows creators to 'go where we could not imagine creatively', suggesting that AI enables artists to explore and realise previously unimaginable possibilities. Another participant saw opportunities in 'generating images not otherwise possible for creative ends', reflecting the view AI builds upon human creativity. In line with this, many participants highlighted that AI enhances the creative productivity. As one participant enthusiastically stated, 'it's going to tell wonderful stories. It's going to save time and lift us, if we use it right'. Yet, precisely what sorts of stories it is going to tell, and how they are going to tell them, is bound up to questions of bias and collaboration, which we will address in Chapters Four and Five, respectively.

---

[104] Steven Smith and Thomas Ward, 'Cognition and the Creation of Ideas', in The Oxford Handbook of Thinking and Reasoning, ed. by Keith Holyoak and Robert Morrison, 456-474 (Oxford: Oxford University Press, 2012).

3.1.7: Democratisation of Creativity

The way we discuss and engage with creativity is transforming in the twenty-first century. The traditional view of creativity, which celebrates the individual 'geniuses' of figures like Albert Einstein and Pablo Picasso, is giving way to a more collaborative and inclusive understanding of creativity.[105] This new perspective highlights the creative potential of everyday interactions and activities, recognising that creativity can be found in everyday contexts and is not limited to a select few individuals. Now, with the advancement of generative AI, creative tools and technologies have the potential to become more accessible to a broader range of people regardless of their background, education, or professional experience, enabling them to engage in creative activities and artistic expression that were previously exclusive to experts.

What some participants termed this 'democratisation of creativity' means that creativity is no longer limited to a few 'creative' people but belongs to everyone. This increased accessibility can lead to new opportunities for people to express themselves creatively. As one participant stated, 'GAI gives creative voice to people who otherwise could not express their creativity'. As another participant noted, in the creative sector, it is important to provide opportunities to people without prior long-term training, allowing them to engage in artistic creations and the creative experience. The idea that 'anybody can create' fosters a more inclusive and diverse cultural landscape, particularly given the historical barriers that people from diverse backgrounds have faced when it

comes to entering a particularly insular UK media industry.
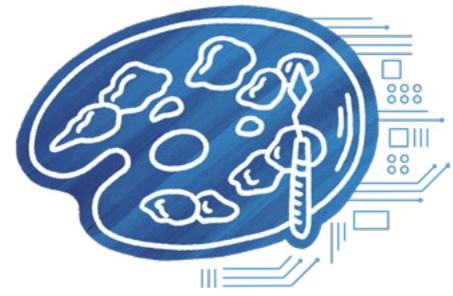
However, with this democratisation also comes questions about the role of professional creators in creating art. Historically, the value of art has often been linked to technical mastery and originality, qualities that are typically gained over years of practice and training.[106] When the barriers to entry are lowered and anyone can be said to produce art, the distinction between amateur and professional blurs. This leads to concerns about how professional artists, who have spent years advancing their craft, can differentiate themselves and make a living from their work.

> **How do professional creators earn a living when the market is flooded with content from millions of new creators? In an environment where 'anybody can create', where do we place the role of those who have gone through intensive training, effort and experience in order to create art?**

This debate regarding the artistic value of democratisation is not a new phenomenon. Previous discussions that pre-dated generative AI tools debated the extent to which democratisation of creativity can lead to a 'cult of the amateur', in which artists simply become 'content providers' and quality could diminish, particularly in what many see as 'postmodern' or 'postnormal' times characterised by complexity, chaos, and contradiction.[107] The rapid advancement and availability of GAIT has only intensified this concern. As creativity is transformed to a new model that embraces

---

[105] Alfonso Montouri, 'Beyond Postnormal Times: The Future of Creativity and the Creativity of the Future', Futures, 43: 2 (2011), 221-227.

[106] Aaron Kozbelt, 'Originality and Technical Skill as Components of Artistic Quality', Empirical Studies of the Arts, 22: 2 (2004), 157-170.

[107] Ziauddin Sardar, 'Welcome to Postnormal Times', Futures, 42: 5 (2010), 435-444; Montouri, 'Beyond Postnormal Times'.

everyday interactions by everyday people using accessible tools, the traditional markers of art may be challenged and new paradigms of art could emerge.

3.1.8: AI-Art: Devaluation or a Continuation of Digital Imaging Practices?

A key concern raised by our participants was the potential devaluation of art in an age of AI-generated creativity. Participants worried that the value of human-created art will decrease. This concern is rooted in the idea that the value of art is tied to the effort, skill and time that goes into its creation, and research across social psychology and economics enforces this idea. This is known as the 'effort heuristic'.[108] When people perceive that a piece of art has been created with ease, they tend to devalue it. One participant summarised this concern by stating that 'a broader public that cannot differentiate and/or does not care about the difference between AI-generated music and human-created music will further devalue the arts'. This is reflective of broader trends in our workshop participants, in which people argued that, despite the amount of effort needed to produce AI-generated art, if people believe that AI-generated art requires less effort and skill than human-generated art, both will inevitably be devalued as a result. This raises important questions such as if AI-generated art were to become indistinguishable from human-created art, where would we place the economic value of art created by a human?

Amidst these concerns over labour are also concerns of quality. A common perception amongst our participants was that AI risks increasing the output of generic creations, given that it is trained to produce a most likely outcome based on pre-existing creative material. Similarly, participants expressed concern over a rise in potential 'laziness' amongst creatives who use generative AI tools and become reliant upon them to produce art. As one participant noted, 'it's so easy to ask AI for more ideas, that we might become lazy and let it do the work'. This idea is rooted in the belief that creativity requires effort and dedication, and when creators rely too heavily on AI, there will inevitably be a drop in originality, artistic depth and a decrease of overall creativity. Another participant echoed this sentiment, suggesting that 'creatives might tend to be more 'lazy' in their approaches because the system does so much of the work'. Along this line, when tasks are perceived as being too 'easy', they run the risk of perpetuating boredom and demotivating creative practitioners, a belief supported by research in psychology.[109]

The debates had between our workshop participants reflect growing trends in the fields of art history and digital media studies. Digital media theorist Lev Manovich and philosopher Emanuele Arielli have documented the rise of GAI aesthetics in real-time with their book, Artificial Aesthetics: Generative AI, Art and Visual Media, published continuously between 2021 and 2024. While recognising the often stereotypical or idealised images produced by models like Midjourney and DALL-E, Manovich and Arielli examine how, far from constituting an end to aesthetic practice, GAI brings about a resurgence in 20th century

---

108 Justin Kruger, Derrick Wirtz and Leaf Van Boven, 'The Effort Heuristic', Journal of Experimental Social Psychology, 40: 1 (2004), 91-98; Daniel Kaheman and Shane Frederick, 'A Model of Heuristic Judgement' in The Cambridge Handbook of Thinking and Reasoning, ed. by Keigh Holyoak and Robert Morrison, 267-293 (Oxford: Oxford University Press, 2012).

109 Julia Haager, Christof Kuhbandner and Reinhard Pekrun, 'To Be Bored or Not To Be Bored – How Tas-Related Boredom Influences Creative Performance', The Journal of Creative Behaviour, 52: 4 (2018), 297-304.

modernist practices, such as database art or composite photography, by accumulating and remixing pre-existing artefacts.[110] In contrast, Marcel Danesi argues that generative AI's ability to synthesise popular culture on which they have been trained means that they resonate aesthetically with what Jean Baudrillard called the 'hyperreal', in which there exists only copies for which there are no longer any originals.[111] Although Danesi interrogates the aesthetics of AI-generated media more broadly, his chapter on 'AI-Generated Cinema' argues that, at least stylistically, AI-generated films ought to be thought of as akin to animation in the sense that, like animation, CGI and other digital imaging practices, films made in collaboration with generative AI were never 'filmed' at all.[112] That is, even where the strive for an aesthetics of perceptual realism, AI-images and films are entirely artificial, altering the ontological relationship between photography and the real world that photography captures or records.[113] Yet, Manovich and Arielli argue that a clear difference emerges between GAI and animation in that, where animation and CGI produced a transition from photographic recording to digital simulation, GAI constitutes a shift away from simulation and toward prediction; that is, it's generative function works by attempting to predict what its end user might want to see next.[114]

3.1.9: Authenticity, Originality and Accuracy

Within these debates on AI-generated art, we located three key themes emerging from our participant discussions: authenticity, originality, and the accuracy of outputs.

In our discussions, authenticity was primarily framed as a distinguishing factor between AI- and human-generated content. Participants expressed concern that AI-produced outcomes might not be seen as authentic because, in their view, creative media should reflect personal characteristics and individual expression. One participant noted that 'the nuanced and emotive qualities inherent in human expression may not be fully replicated by AI systems', highlighting the concern that AI might not capture the essence of human creativity. Furthermore, participants felt that if AI could produce art indistinguishable from human creations, it could potentially undermine the very foundations of what we consider to be an authentic experience in and of the world, as it erodes the human connection in artistic production.

Originality, on the other hand, was seen differently by participants. While originality could be considered part of authenticity, participants associated the term more with the uniqueness of the produced outcomes themselves, regardless of whether they were created by humans or AI. Their understanding of originality was not tied to the human-AI distinction but rather to the uniqueness and distinctiveness of each artistic creation. One participant argued that AI-generated content 'lacks originality, trading creativity for generic, mash-up images, stories, and ideas'. This critique points to the limitations of generative AI systems, which rely on pre-existing data and models to create new content. In the context of media creation, this leads to homogenised works that fail to introduce novel ideas or innovative storytelling techniques. The

---

[110] Lev Manovich and Emanuele Arielli, Artificial Aesthetics: Generative AI, Art and Visual Media (2021-2024) <https://manovich.net/index.php/projects/artificial-aesthetics> [Accessed 16 August 2024].

[111] Marcel Danesi, AI-Generated Popular Culture: A Semiotic Perspective (Cham: Palgrave Macmillan, 2024), p. 10.

[112] Ibid., p. 52.

[113] For debates about CGI's impact on cinema's relationship to realism, see: Lev Manovich, The Language of New Media (London: MIT Press, 2001); Stephen Prince, Digital Visual Effects in Cinema: The Seduction of Reality (New Brunswick: Rutgers University Press, 2012); Lisa Purse, Digital Imaging in Popular Cinema (Edinburgh: Edinburgh University Press, 2013); Deborah Levitt, The Animatic Apparatus: Animation, Vitality, and the Futures of the Image (Alresford: Zero Books, 2017).

[114] Manovich and Arielli, Artificial Aesthetics, pp. 11-12.

question, 'can AI create truly original work based on lived experience?', encapsulates the debate over whether AI can ever genuinely replicate the human creative process, with most participants feeling that it falls short of delivering original content on the basis that it is unable to generate new or unique ideas.

From the perspective of computer science, the question of whether generative AI can produce original work is contingent upon viewpoint. While it is easy to dismiss AI-generated work as original or creative on the basis that it is produced based on data within its limited knowledge domain, the perception of this work as original or creative will depend upon what we individually or socially perceive to be original or new from our inherently subjective, limited perspectives of the world. An analogy might be to think of all of the data held by an AI model as a beach, where each individual human artist's work may be imagined as a grain of sand. Where a generative AI tool produces a piece of work based multiple grains of sand, while it is technically not new or original in the sense that it did not originate from outside of its knowledge domain (or beach), it may nevertheless be perceived as original on the basis that thousands or millions of people may have never experienced these grains of sand before.

Accuracy also emerged as a frequent concern among participants. One participant observed that AI 'brings out lots of garbage', noting the challenge in managing the accuracy of content generated by generative AI tools. This aligns with current research, which suggests that while AI systems can produce information quickly, they often do so at the expense of accuracy.[115] From our own creative practice, we found maintaining object permanence or temporal continuity to be significant challenges for generative AI tools, which were more adept at generating symbolic or metaphorical creative outputs that could deviate from these expectations of realism. Yet, this is further complicated by the sometimes convincing nature of individual AI outputs, especially as they develop in technical proficiency, as they may appear accurate but require diligent fact-checking. Participants emphasised the importance of individuals being educated in distinguishing misinformation. Without proper human oversight, AI-generated content risks disseminating incorrect information on a large scale. This suggests that enhancing users' literacy in this area is crucial.

In the midst of all of these debates, the future of creativity and art will have to find ways of navigating through such concerns, empowering those who seek to integrate AI into their workflow practices while still promoting the value of human labour, skill and effort. A responsible approach to AI creativity is one that recognises the opportunities that generative AI offer creatives in terms of time, freedom and democratisation, while instilling human decision-making at the heart of the creative process. It is only by doing so that we can build a vision of generative AI use that enhances, rather than replaces, human creative skill.

3.2: EBG and Creativity

3.2.1: Creativity Interview Questions

In advance of the Expert Bridging Group meeting, we circulated the following

---

[115] Gary Marcus, 'AI Platforms Like ChatGPT Are Easy to Use but Also Potentially Dangerous', Scientific American (19 December 2022) <https://www.scientificamerican.com/article/ai-platforms-like-chatgpt-are-easy-to-use-but-also-potentially-dangerous/> [Accessed 27 September 2024].

preliminary information from our workshop findings, followed by a set of guiding questions that we asked each member of the Expert Bridging Group to prepare answers to.

What we discussed:

There are concerns regarding AI-generated content quality and its limited ability to innovate beyond its training data. AI tools also face difficulties in adapting to diverse artistic styles and genres, which can result in technically proficient but creatively uninspired content. AI systems struggle with originality, as they tend to 'copy' and reproduce elements from their training data, leading to output that can seem like a patchwork of existing works, referred to colloquially as 'generic shit' in our workshop. The changing definition of professional quality content is another concern, as AI might not create space for human creatives by taking over creative tasks. This raises the potential devaluation of human creativity and questions whether the lack of human effort in creating content diminishes its perceived creativity.

In terms of authenticity and originality, AI systems are expected to generate novel ideas rather than merely replicating existing works. Despite this expectation, AI often lacks creativity, leading to the homogenisation of works. The presence or absence of human authorship significantly influences the perceived authenticity of media produced, at least in part, using artificial intelligence. AI-generated content may lack the human touch, including the emotional and contextual understanding that human creators bring. There is also potential scepticism from audiences about the originality of AI-generated content and the risk of AI-generated misinformation.

Questions

- **To enhance the creative quality of generated content, what tools and processes can be developed?**

- **How can we ensure that AI-generated content maintains a high level of creativity?**

- **What do creators need to enhance their creativity when using AI as a tool, while allowing AI to handle mundane tasks?**

- **What specific criteria should be established to define and measure 'professional quality content' in an era where AI-generated and AI-assisted content is becoming prevalent?**

3.2.2: Discussion on Creativity

As the discussion topic that opened the Expert Bridging Group, 'Creativity' began with introductions from each participating member, including an initial scoping exercise designed to ascertain how different stakeholders make use of generative AI to enhance or augment their professional creativity.

A leading industry professional working with generative AI tools in the editing process discussed how their company are unable to use AI for many post-production processes because they are technically difficult to automate. Tasks such as rotoscoping, match-moving and rigging, although deemed in the industry as 'entry level' post-production roles, are currently unable to be successfully performed by generative AI tools, they explained, while some of the more 'creative' roles are easier to automate through tools such as Midjourney. This runs counter to the prevailing view amongst our workshop participants, who sought to harness AI for more tedious and time-consuming tasks

in order to assist human creativity. Other professionals stated that they used tools such as Avid Ada or Premiere Pro Adobe to help with editing tasks, but reflected that there needs to be further differentiation between AI 'editing assistants' and digital 'editing tools' that have been around for decades and are being subsumed under the label of AI. Such reflections speak to the desire to utilise generative AI tools as a means toward efficiency, even where these possibilities do not yet quite exist.

Some members of the Expert Bridging Group discussed their own use of generative AI tools through calls to other media. A UK-based photographer and director of photography likened the process of using generative AI tools in image creation to photography, such that, anyone can technically pick up a camera and produce an automated image of the world, but it still requires skill, training and groundwork to develop sensory acuity and produce a valuable and artistic creation at the end of it. Similarly, an academic working within film exhibition described the use of generative AI for creative means as 'curation', in which the artefacts being curated may not have been produced directly by a human hand but it still requires human intervention and decision making in the form of prompting, selecting and editing in order to generate something of value. These views reflect the tensions that emerged from our workshop regarding the creative quality of AI-generated content.

Some members of the Expert Bridging Group recognised that while generative AI tools have the potential to increase creativity on an individual level, they risk decreasing creativity on a societal or collaborative level. While this report will unpack these comments in greater detail in Chapter Five, it is important to consider that an overreliance on generative AI models as ideation tools, soundboards or virtual assistants has the capacity to displace human-to-human connection and supplant the creativity that emerges through social interaction. This may also lead to further exploitative practices of media production, in which the individual(s) in charge of the production are given freedom to use generative AI tools to enhance their vision, while those whose duties are deemed 'uncreative', such as editors, actors or voiceover artists, are seen as merely data extractors or data to be extracted. On this basis, members of the Expert Bridging Group wanted to see responsible AI practices that kept in-tact the socio-cultural element of creative endeavours.

In response to the question, 'how can we ensure that AI-generated content maintains a high level of creativity?', some members of the Expert Bridging Group noted that there exists a societal distrust regarding AI-generated content and media amidst feelings of being 'cheated' when the use of AI in a media production has not been disclosed upfront. A leading UK union representative for creative artists pointed to the fall out that The Prince Charles Cinema, London, received for scheduling a screening of The Last Screenwriter (Peter Luisi, 2024), purporting to be the first film based on a screenplay entirely produced through ChatGPT.[116] The cinema

---

[116] James Kelly, 'London Cinema Drops AI-Written Film After Backlash', The BBC, 19 June 2024 <https://www.bbc.co.uk/news/articles/cjll3w15j0yo> [Accessed 12 September 2024].

received a high volume of complaints about this decision, particularly regarding concerns relating to the lack of compensation for artists whose work was used to train the LLMs used during the production of the film. The scale of the backlash was so high that the theatre decided to cancel the planned screening entirely. This suggests that although media practitioners are keen to integrate AI into their workflow practices where appropriate, there still exists the hurdle of overcoming AI as a 'pejorative', as one workshop participant noted, to describe works that people do not like or whose stylistic tendencies are found to be reductive.

Part of this, as numerous members of the Expert Bridging Group explained to us, is because there is a lack of clarity regarding where copyright and intellectual property sit within the landscape of AI-produced creative outputs. This leads to feelings of uncertainty both within and outside of the creative industries regarding whether the use of AI to create content is ethical, not to mention legal, given that use of original artists' whose to train LLMs is happening without their consent or fair remuneration. A union representative within the group proposed that if AI-produced content can legally be considered 'source material', then 'format fees' may be taken away from artists and transferred over to producers or companies leading to the further perpetuation of existing inequalities of labour. Similarly, prompts that ask generative AI tools to produce works 'in the style of' a particular author or artist could potentially lead to successful imitations that allow producers to bypass paying the original artist altogether.

Such concerns highlighted the urgent need for legal and governance protections to human creative labour. While members of the Expert Bridging Group varied on how they viewed generative AI and its various limitations and possibilities when it comes

to creativity, one thing that all members of the group agreed upon was the need to protect creativity as a human value. For this to happen, propositions need to be put in place, incentives need to be created, and UK legal precedents need to be established. Chapter Six will examine what these interventions might look like by way of regulatory and governance change, as well as what individual users can do to empower human creativity through the responsible use of AI.

3.3: Conclusion

This chapter has considered the changing role of creativity in light of generative AI tools being used within the UK media industry. It explored how many media practitioners conceptualise creativity not as something that is lost when using generative AI tools but, rather as something that is recalibrated, as producing prompts and re-editing AI produced materials constitute considerable creative processes in themselves. Yet, such possibilities are problematised by the potential degradation of artistic value that comes with AI-generated content producing generic and likely outcomes. This possible loss of novelty and originality needs to be considered in line with material consequences for human creative labour, as

stakeholders from different aspects of the creative industries consistently warned us of the pressing need for tightly regulated legal protections for human creativity. Given our own creative practice in the form of generating a film using different AI tools and models, we echo these calls for protections of human creativity, so that human creatives may feel empowered to AI to enhance their workflows in responsible ways. Chapter Five returns to some of the concerns addressed here by considering the role of collaboration in human-AI media production, but before this, Chapter Four examines responsibility over the harmful social biases that are being perpetuated as a result of generative AI development, training and use processes.

CHAPTER FOUR

Bias

In the last chapter, we considered what impact generative AI tools are having on industrial conceptions of creativity within the UK media landscape. It examined how best to protect the value of human creative productions while empowering users to capitalise on generative AI's accessible possibilities, concluding that, in order to avoid overly 'generic' outputs, human creative skill, labour and experience are still needed. In this sense, questions of creativity also overlap with questions of bias, as generative AI tools often produce stereotyped outcomes that risk perpetuating social biases. Yet, in a landscape where AI developers are reluctant to disclose training datasets and processes, it can be difficult to know precisely how to redress the issue of AI bias. The task of this chapter is to consider how different stakeholders within film and media production conceptualise generative AI bias as both a problem and an opportunity for the future. It explores potential solutions to the issue of bias in AI models with the goal of implementing a more just, transparent and safer AI media landscape.

As with the previous chapter, this chapter is divided into two sections. The first section, '4.1: Workshop Findings', examines what our workshop participants had to say about the issue of bias in relation to GAIT. After contextualising the section with prevailing academic literature on stereotypes and bias (4.1.1), the chapter unpacks precisely why our workshop participants deemed bias to be a harmful prospect in AI-generated content (4.1.2), before considering some of
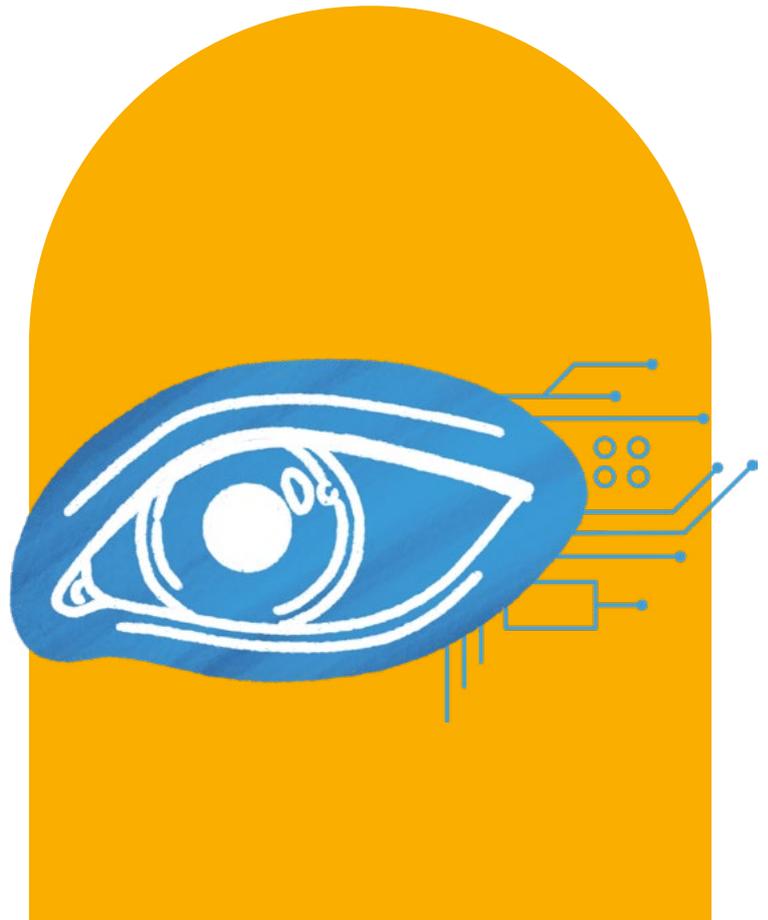
the opportunities for change that emerge from GAIT in terms of increasing diversity within the media ecosystem (4.1.3). Following this, the chapter considers how our Expert Bridging Group thought through issues of representation, accessibility, inclusivity and training processes. Section two, '4.2: EBG and Bias', begins by outlining the questions we circulated to the EBG in advance of the 'Bias' discussion session (4.2.1), after which the section looks at strategies of implementation and best working practices that can be developed to help mitigate bias and prevent the reinforcement of stereotypes (4.2.2). As with the previous chapter, this chapter interweaves our reflections from our creative process as a research team amidst observations and analysis from our research participants.

4.1: Workshop Findings

4.1.1: Stereotypes and Biases

Stereotypes and biases influence how we perceive and interact with others. Stereotypes

refer to widely held, overgeneralised beliefs and expectations about the attributes, behaviours or characteristics of individuals who belong to a particular social group.[117] These groups are often defined by shared physical or sociocultural traits such as ethnicity, race, gender, social class, nationality, or other common characteristics. While they may serve as cognitive shortcuts, allowing people to quickly categorise and process information about others, they have the potential to be harmful in the sense that they might unfairly mischaracterise entire groups of people and oversimplify complex patterns or behaviours.[118]

While biases and stereotypes often overlap, biases refer to more personally-held systematic deviations from objectivity that occur during the processing of information.[119] Biases shape how individuals interpret and assign meaning to events, influenced by both internal factors (e.g., physiological sensations, emotional states) and external factors (e.g., social context, environmental cues).[120] Biases can also influence how people process information and make decisions about others, often in ways that are unfair or prejudiced. Biases can both be rooted in, and perpetuate practices of, racism, misogyny, ableism, classism, homophobia, transphobia and ageism, amongst other discriminatory behaviours.

Building upon this definition, AI bias, also

referred to as machine learning bias or algorithmic bias, can be defined as the production of distorted or prejudiced results that emerge from AI systems, models and tools. The result is that certain identity groups become systematically misrepresented or excluded from fair representation in the design or outputs of AI media. AI biases stem from the wider societal biases and prejudices that are embedded in both the human-created datasets used to train AI models and the design of the AI algorithm itself. As AI systems start to play a greater role in computational processes, including decision-making and content creation, the issues of stereotype and bias take on increased importance. Human-Computer Interaction (HCI) research has primarily focused on usability, task efficiency, and the functional aspects of user interfaces, rather than on the broader social implications like stereotypes and bias. These concerns emerged when algorithms began to reflect and perpetuate pre-existing societal biases due to the data they were trained on and the design choices of the systems.

Recent studies have demonstrated how algorithmic biases perpetuate existing inequalities, with AI systems often amplifying biases related to race, gender, and socioeconomic status.[121] These biases can arise from training data, model design, or even the interpretations and usage of AI-generated content by media creators. For instance, much has already been written on the fact that voice-

[117] Charles Stangor and Mark Schaller, 'Stereotypes as Individual and Collective Representations', in Stereotypes and Prejudice: Essential Readings, ed. by Charles Stangor, 64-82 (Hove: Psychology Press, 2000); Perry Hinton, Stereotypes, Cognition and Culture (Hove: Psychology Press, 2013).

[118] David Hamilton, 'Stereotyping and Intergroup Behaviour: Some Thoughts on the Cognitive Approach', in Cognitive Processes in Stereotyping and Intergroup Behaviour, ed. by David Hamilton, 333-353 (Hove: Psychology Press, 2015).

[119] Chiara Acciarini, Federica Brunetta and Paolo Boccardelli, 'Cognitive Biases and Decision-Making Strategies in Times of Change: A Systematic Literature Review', Management Decision, 59: 3 (2021), 638-652.

[120] Łukasz Gawęda and Martyna Krężołek, 'Cognitive Mechanisms for Alexithymia in Schizophrenia: Investigating the Role of Basic Neurocognitive Functioning and Cognitive Biases', Psychiatry Research, 271 (2019), 573-580.

[121] Safiya Noble, Algorithms of Oppression: How Search Engines Reinforce Racism (New York: New York University Press, 2018); Ruha Benjamin, Race After Technology: Abolitionist Tools for the New Jim Code (Cambridge: Polity Press, 2019).

based generative AI assistants that produce deepfake content, of which 96% is believed to be pornographic,[122] overwhelmingly impacts women and girls.[123] This presents the risk of re-entrenching a culture of sexual violence and violent gender biases.[124] Yet, even those collaborating with AI in good faith may still find themselves unknowingly producing biased content on the basis that generative AI tools are being trained on biased datasets and by engineers for whom unconscious bias filters into their training practices.

It has also been widely reported that generative AI tools produce stereotyped images and narratives around race. For instance, as far back as 2018, data training processes designed around labelling content for Google Photos incorrectly labelled images of Black people as gorillas, an incident that sparked debate about how design and training methods can result in the perpetuation of racist stereotypes.[125] Similarly, in The Black Technical Object: On Machine Learning and the Aspiration of Black Being, Roman Amaro outlined how principles of machine learning for surveillance purposes are predicated upon institutional racism and racial characterisation.[126] While not pertaining to generative AI specifically, Amaro's contention, that AI tools are disproportionately trained to recognise and label Black people as objects of suspicion, speaks to the prevailing issue that is racial bias within the development of AI systems as a whole.
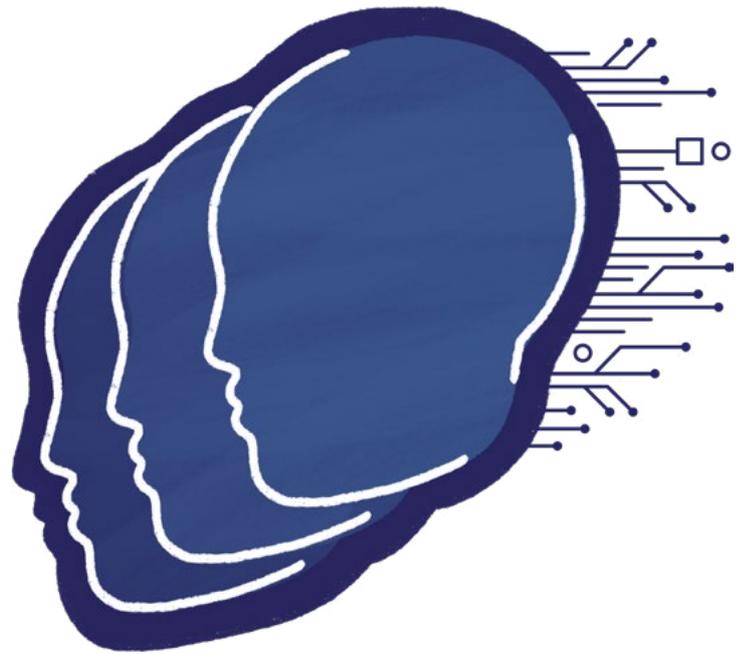
The same can be said for categories of sexuality, gender identity, disability, class and age; where generative AI tools are designed to produce a most likely desired outcome, and where said design is trained on historical texts in which these identity groups rarely received fair representation, their outputs, if not corrected, will inevitably be rooted in stereotype and bias. While generative AI tools have the potential to democratise media production by offering accessible access to a broader range of users, as explored in Chapter Three, their tendency to produce biased and stereotyped outputs undermines this potential. While there exists no singular way to combat this, proponents of ethical generative AI practices emphasise the importance of regulations surrounding the disclosure of training data and the production of AI-image watermarks to signal detectability.[127] This chapter will examine some of these solutions in section 4.2. Before

---

[122] Of these 96% pornographic deepfakes, 100% were found to be images and videos of women and girls. See Dominic Lees, 'Deepfakes in Documentary Film Production: Images of Deception in the Representation of the Real', Studies in Documentary Film, 18: 2 (2024), 108-129 (p. 121).

[123] As Atoosa Kasirzadeh, Charlotte Bird and Eddie Ungless note, 'searching for 'girl' on promptbase.com, a marketplace for prompts, returns a much larger number of sexualised images with childlike features' when compared with the prompt 'boy'. See Atoosa Kasirzadeh, Charlotte Bird and Eddie Ungless, 'Policy Report on Generative Artificial Intelligence', BRAID (July 2024), p. 12.

[124] Kevin Rawlinson, 'Digital Assistants Like Siri and Alexa Entrench Gender Biases, Says UN', The Guardian, 22 May 2019 <https://www.theguardian.com/technology/2019/may/22/digital-voice-assistants-siri-alexa-gender-biases-unesco-says> [Accessed 19 August 2024]; Amy Schiller and John McMahon, 'Alexa, Alert Me When the

Revolution Comes: Gender, Affect, and Labour in the Age of Home-Based Artificial Intelligence', New Political Science, 41: 2 (2019), 173-191; Kate Devlin, 'Power in AI: Inequality Within and Without the Algorithm', in The Handbook of Gender, Communication, and Women's Human Rights, ed. by Margaret Gallagher and Aimée Vega Montiel, 123-139 (London: Wiley Blackwell, 2023).

[125] Alex Hern, 'Google's Solution to Accidental Algorithmic Racism: Ban Gorillas', The Guardian, 12 January 2018 < https://www.theguardian.com/technology/2018/jan/12/google-racism-ban-gorilla-black-people> [Accessed 19 August 2024].

[126] Roman Amaro, The Black Technical Object: On Machine Learning and the Aspiration of Black Being (London: Sternberg Press, 2023), pp. 20-21.

[127] Harry Jiang et al., 'AI Art and its Impact on Artists'. AIES '23 August 08-10 (2023). 363-374 (pp. 371-372).

then, it is necessary to consider how AI bias is viewed by our workshop participants as not just a technical challenge but a socio-cultural imperative that requires an interdisciplinary perspective.

4.1.2: Bias in AI: Limitations

Many of our participants were aware that AI tools can reinforce stereotypes through the prediction of desired outcomes, as several people raised concerns about the echo-chamber effect of AI-generated media. One participant explained that 'it definitely is doing a very similar processing […] giving us what it thinks we want'.

Comments such as these suggest that there exists a recognition within the creative industries that generative AI tools work by anticipating likely intent during prompting. While this can threaten creativity through the production of generic outputs, as we saw in Chapter Three, the risk here is that what constitutes generic is predicated upon cultural and ideological hegemony.

This comment also highlights the risk of AI tools reinforcing users' preferences or biases over time, as instead of offering exposure to diverse viewpoints, the production of a most likely desired outcome simply creates a self-perpetuating loop of bias. This issue is compounded by the fact that, where generative AI tools use prompts and outputs as part of their ongoing training processes, biased content is often fed into algorithms repeatedly. One participant commented on this phenomenon, stating that 'there is a vicious cycle of biases […] the ones in our thoughts, the ones in our language, and the ones in images'. This statement suggests that biases in AI are not simply a product of the original datasets on which they were trained, but are also exacerbated by users' interactions with the tool themselves.

Of particular concern amongst our participants was the racial and gender biases in AI-generated images. One participant explained that 'if you type into Midjourney […] everything it brings out to you will be a white person unless you type in Black specifically'. This underscores AI's default preference for white representations unless otherwise specified, speaking to wider systematic assumptions about whiteness as a universal non-race and the need to specify if you want divergences from this position. Where AI models are trained on a historical canon of data, which has typically tended to be English-language, Western, modern texts written and imagined from a white male perspective, then AI tools will inevitably reproduce the types of stories and images populating the historical public domain without correction.

Another participant shared their own example of how a woman used generative AI tools to create an image of a 57-year-old Black character but ended up producing an image that looked significantly older, explaining that 'she's [the person who prompted] no Black; she cannot pick that up […] that's where AI fails us'. Examples such as these demonstrate how racial and age biases already prevalent amongst training datasets may intersect and lead to misrepresentations that then reinforce societal biases, rather than challenge them. This participant's particular comments also speak to a belief that where end users do not have the lived, conscious experiences of those who they are representing through generative AI tools, biases and mischaracterisations may slip through the cracks.

These examples suggest the importance of greater awareness and education, both about AI's role in the perpetuation of bias and about the wider impact of societal biases more broadly. As one of our participant noted, 'creators need to know about the set of learning images' for bias awareness. The

implication here is that, if developers are made aware about the specific contents and demographic makeup of their training data, they would be in a better position to redress these issues and potentially mitigate biased outputs. While this might be true, from a technical perspective, computer scientists are keen to emphasise that bias and stereotypes are arising from problems in the training data itself, in which samples from certain categories or groups may be more or less representative than others. Another participant reflected that 'we still have to have consultants look at that work', highlighting the view that external regulation, auditing and the developing of bias matrix could help both developers and practitioners when it comes to ensuring that AI-generated content is both accurate and respectful of different identity groups. All of these proposed interventions, regardless of their practicalities or efficacies, demonstrate a clear view that it is not enough to simply develop an AI tool and leave it to work out issues of bias and lack of diversity on its own; in order to enact effective change in diversity and representation, human oversight is fundamental.

4.1.3: Bias in AI: Opportunities

Surprisingly, some participants raised the possibility that AI bias might actually be useful to media production within certain contexts. One participant provided an example from the advertising industry, noting that stereotypes can actually help advertisers better appeal to their intended target audience. This suggests that bias, when understood and controlled, might serve a purpose in creating consistency or maintaining specific stylistic choices. However, it was also proposed that creators need to be extremely careful and responsible with this understanding of bias. As one participant stated, 'people need a better understanding of exactly how these terms are used within AI and technology'.

The suggestion here, then, is that stereotypes might theoretically be helpful in a very narrow set of circumstances, but only when they are recognised, managed and intentionally applied within a context-appropriate manner.

Another participant discussed the possibility of using biased outputs to their 'advantage' such that the overtly biased nature of many generative AI tools allows them to know what not to produce in their own work. That is, if AI tools produce stereotypes or an average person of a certain target group, then users attuned to different biases can actively go against this production and know what sorts of images and narratives to avoid. Probing the AI for its biases can therefore be a potentially fruitful creative approach, an exercise undertaken by one of our Guest Speakers in Workshop 2 (Image Creation), whose meta-reflective film about AI's inherent biases held up a mirror and exposed the deep rooted issues of bias that already existed within the system and more widely within society. Along these lines, generative AI tools may end up establishing a system of creative resistance, in which creators expose and highlight their issues. Confronted with both societal and their own biases, creators may be prompted to undertake critical reflection during the prompting stage, training them away from their ingrained biases and furthering the creativity of the product in the process.

While the general consensus amongst our participants was that bias is an issue, there remained an open question about how to both determine and correct bias within AI-generated media. Some participants raised the possibility of training LLMs on local, diverse datasets, while others highlighted the need for increased bias awareness, using educational programs to help users choose their inputs more carefully and with a commitment to diversity when prompting. Many of our participants emphasised the need

for more diverse teams in AI development, coupled with industry regulation and further education about the biases that AI tools may perpetuate, as a solution. The following section examines what this might look like in further detail.

4.1.4: Workforce Diversity and Biases

The lack of diversity in the AI industry has become a critical issue, contributing to the perpetuation of biases in AI systems. As of 2023, women made up only 22% of AI professionals globally, leaving the field overwhelmingly male-dominated at 78%, according to the World Economic Forum.[128] The disparity extends to racial and ethnic representation as well. McKinsey's 2022 report highlights that just 25% of AI developers are racial or ethnic minorities, with 29% of organisations reporting no minority employees working on their AI solutions.[129] Moreover, a systematic review by Rifat Shams, Didar Zowghi and Muneera Bano for AI and Ethics revealed that the integration of AI with diversity and inclusion is an under-researched area in the academic arena as well.[130]

These industry and academic trends were reflected in our workshop findings. Across all our workshops, participants were unified in their call for more diverse teams involved in the development and training of generative AI models. In particular, they wanted to see greater diversity regarding who trains the generative AI tools on what constitutes the most productive strategy to mitigate against unfair, inaccurate or harmful outputs regarding

identity groups and protected characteristics. Yet, one participant reflected upon this possibility by asking: 'how do you ensure that even if diversity is included in the workforce, that they are actually listened to, and their voices are valued?'. This question takes us back to our discussion of stakeholders in Chapter One and the map of stakeholder power imbalances as outlined in Appendix 6. It underscores the need not simply for surface-level diversity, implemented purely for optics or to fulfil any proposed regulation, but rather for the meaningful inclusion and integration of diverse perspectives through the entire AI development process.

However, from a computer science perspective, this is not the only way to resolve bias. Where biased outputs are the result of biased data, training processes that privilege unsupervised learning strategies can help ensure parity amongst different clusters of training samples and data automatically. The focus, from this viewpoint, should be more on accurate data checking and a diversity of data itself. Yet, even here, issues emerge with the role of human oversight, such that, where humans are overseeing tasks like clustering according to categories, labelling, and deciding what constitutes a 'successful' attempt by an AI output, the potential for unconscious biases to unknowingly enter this process remains.

Another participant emphasised the need for state engagement on the issue of lack of diversity in AI development, stating that the 'government needs to be active to

---

[128] Gabriela Ramos, 'Why We Must Act Now to Close the Gender Gap in AI', World Economic Forum (22 August 2022) <https://www.weforum.org/agenda/2022/08/why-we-must-act-now-to-close-the-gender-gap-in-ai/> [Accessed 26 September 2024].

[129] McKinsey, The State of AI in 2022 – and a Half Decade in Review (6 December 2022) <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review> [Accessed 26 September 2024].

[130] Rifat Shams, Didar Zowghi and Muneera Bano, 'AI and the Quest for Diversity and Inclusion: A Systematic Literature Review', AI and Ethics (2023), 1-28.

legislate to have more diversity within the AI workforce'. Another echoed these sentiments, highlighting the 'need for government to be proactive – need to force a diverse workforce'. While the specifics of what this legislation ought to look like were not commented upon, the fact that our participants wanted to see drastic change highlights the industry-wide desire for regulatory intervention to ensure diverse representation within generative AI.

Where this is not yet in place at a regulatory level, participants proposed taking a proactive approach amongst themselves to embed more diversity into film and media productions that use generative AI tools. One participant rather poignantly reflected that, as an industry, we need to 'create projects which recruit under-represented voices, so that we can shift the status quo before it is established'. This call for early intervention as a way of shaping a more inclusive film and media industry is indicative of the responsibility gap that currently exists amongst AI developers and legal structures. The task now is to consider precisely how to redress this responsibility gap by building upon some of the insights and solutions offered here through our Expert Bridging Group discussion.

4.2: EBG and Bias

4.2.1: Bias Interview Questions

In advance of the Expert Bridging Group meeting, we circulated the following preliminary information from our workshop findings, followed by a set of guiding questions that we asked each member of the Expert Bridging Group to prepare answers to.

What we discussed:

Bias in AI systems often stems from biases in the training data, leading to a lack of diversity in the generated content, which,

in turn, reinforces harmful stereotypes. When AI is trained on data that does not adequately represent diverse perspectives, the outcome can become homogenised, stifling creativity and limiting the range of artistic expressions. This lack of diversity not only perpetuates stereotypes but also results in biased outcomes that disproportionately disadvantage specific population groups. Moreover, these biases exacerbate accessibility challenges, making it difficult for marginalised communities to benefit from AI advancements. Compounding these issues is a general lack of transparency about the potential risks and harms associated with AI-generated content, as well as limited traceability of creative outcomes, which makes it challenging to identify and rectify biases. Additionally, there is often insufficient information and education about the proper use of AI tools, along with a lack of clear guidelines and ethical frameworks to guide their development and deployment. Addressing these issues requires comprehensive efforts to ensure diverse and representative training data, transparent and traceable AI processes, and robust educational resources and ethical guidelines to promote fairness and inclusivity in AI applications.

Questions

- **How can we develop techniques for identifying and mitigating biases in training data, ensuring that AI systems are trained on diverse and representative datasets?**

- **What are the best ways to provide detailed information about the data used in their training, enabling users to understand the potential biases and limitations of the generated content?**

- **What are the best practices for developing AI systems that are sensitive to the potential for bias and stereotype reinforcement, and how can we integrate these practices into the design and development process?**

- **How can we design AI systems that are accessible to diverse users, regardless of their technical expertise or cultural background, promoting inclusivity and equitable access to AI advancements?**

4.2.2: Discussion on Bias

Within the Expert Bridging Group discussion on bias, members debated the merits and legal possibilities of changing training and labelling processes. In terms of labelling, an AI policy advisor in the UK highlighted concern that AI developers are relying on cheap, outsourced labour, particularly from countries in the Global South, to label and moderate often biased, harmful or traumatising content. This resonates with wider reports of companies like OpenAI exploiting workers in Kenya by paying them less than $2 an hour to label graphic and sexual content, leaving workers 'disturbed' with little recourse to counselling.[131] Such global supply chain issues are not necessarily new, established during earlier digital movements like the birth of social media. Yet, they speak to a tension within attempts to redress AI bias through labelling, such that keeping generative AI outputs 'safe' and 'equitable' is, in its current form, built off the backs of exploited workers. Amidst broader legislative change needed to tackle AI developers from procuring outsourced labour under exploitative rates and conditions, one potential solution offered

was the pre-processing and cleaning of data by crowdsourcing, adding an additional layer of accuracy to AI labelling and meta-data. Chapter Six will examine in further detail how this could work as a practice of informed citizen science.

In terms of training, our participants expressed a desire to see further education and media literacy for the developers and workers involved in creating generative AI tools, which they believed could change the practices of existing developers when it comes to catching and redressing bias. Along these lines, a representative for a leading UK creative union argued for the need to embed a right to human review in all training processes, such that human labourers educated on bias may be the ones to redress the situation, rather than relying on algorithmic decision making. They also discussed the possibility of retraining existing models in order to address the biases already prevalent within datasets of prominent generative AI tools. Questions surrounding the efficacy and ethics of removing data from existing AI models, as well as putting pressure on AI developers to retrain models for our own gain, were raised, with a contention emerging that such ideas may be difficult due to the lack of consensus regarding what constitutes an 'objective' dataset free from bias.

Building upon this, one member of the Expert Bridging Group with a background in responsible AI development explained the difficulty in determining bias in the first place. Once bias is detected, it is technically possible to change the model, but the problem is that we are only able to pick up on biases to which we are already attune. This, they argued, leads to an inevitable information asymmetry, such that there are times when we may not even be aware of our own biases. One potential solution to this problem was the proposal of AI bias interpreters or translators who would

---

131 Billy Perrigo, 'OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic', TIME (18 January 2023) <https://time.com/6247678/openai-chatgpt-kenya-workers/> [Accessed 25 September 2024].

consult with AI developers in order to help spot and redress any identity-based biases to which the developers are unaware. Another member took this proposal one step further by considering what that might look like on a structural level, in which unions, organisations, a regulator or an expert council would come together to determine where different biases sit and the rubrics on how to challenge them.

On an individual level, members reflected some of the proposed ideas put forth by our workshop participants when they called for users to prompt with a greater awareness of the dangers of biased AI-generated media. A leading UK academic on AI in education described how representations within generative AI media are rooted in extremes, such that descriptive terms such as 'young' or 'old' tend to generate exaggerated embodiments of those terms. This caricaturing function can lead to dehumanised outputs when it comes to issues of race, disability or sexuality, along with age and weight. Members of the Expert Bridging Group placed an emphasis on recognising the intersectionality of biases and prejudices during the prompting stage, including being attune to the fact that different members of an identity group may experience different forms of oppression depending on their other intersecting and overlapping social identities. As with our workshops, one member raised concerns over overcorrections of bias leading to inaccurate or perhaps even offensive representations, citing the example of Google's Gemini producing images of Black Nazis based on the company's attempts to produce more diverse representations.[132] The task for end users, as our Expert Bridging Group saw it, is to prompt in ways that challenge biased portrayals of identity groups,

but the task for AI developers is to ensure that users can mitigate bias in ways that are culturally sensitive and respectful.

This leads us to the question of how best to embed and incorporate diversity in AI development. Echoing sentiments expressed across our workshops, members noted that the lack of diversity and perpetuation of bias are two interconnected symptoms of a wider societal problem within the tech industry. Where AI developers are homogenous both in their expertise and lived experience, the result is that people already denied equal access to employment opportunities within the sector are further marginalised. Where supervised training processes do not account for diversity and inclusivity, the result might be that images and narratives are more likely to reflect the views, including any biases, of the creators, developers and data. Culturally specific models, such as those developed in consultation with and designed for certain demographics including children or populations of specific countries, were proposed as one way to ensure fairness in the production of AI-generated content. Yet, on a broader scale, what our Expert Bridging Group recognised was needed was diversity embedded at the source; of media production, technology development and AI use.

A UK-based filmmaker and producer who uses generative AI tools in their media production shared their personal story about how media representation influences popular perception of communities. This includes not only who has the resources to tell their story and how they are telling them, but also what stereotypes they are perpetuating or challenging in the production of that story. Their concluding remark, that 'diverse users will create diverse stories', highlighted that, while attempts to redress bias and inequalities require regulatory change at the level of

---

[132] Chris Gilliard, 'The Deeper Problem with Google's Racially Diverse Nazis', The Atlantic (26 February 2024) <https://www.theatlantic.com/technology/archive/2024/02/google-gemini-diverse-nazis/677575/> [Accessed 25 September 2024].

training and development, the question of diversity comes back to issues of access and practice during the creative process.

4.3: Conclusion

This chapter examined how and why generative AI tools come to be imbued with stereotypes and biases. It considered how different stakeholders within the media industry conceptualised these biases as both limitations that need to be addressed and opportunities for change when creating AI-generated content. The solutions offered by our workshop participants regarding how to build a responsible AI media landscape were expanded upon by members of our Expert Bridging, who proposed that, in order to redress systematic bias in the outputs of generative AI models, we need to rethink practices of labelling, training, prompting and accessibility. In our own creative practice, we trialled some of these potential solutions to examine their efficacy for independent practitioners. The result of these mixed

methodologies was that bias and stereotype are issues that those working across AI divides are actively invested in redressing, but there currently lacks a clear trajectory for how to go about redressing them. Chapter Six will outline some principles, practices and policy recommendations that are designed to start this process, but before these can be explored, we need to consider the cultural, ethical and legal implications of working with and alongside generative AI tools in media production. As such, the next chapter will interrogate the last of our research questions, exploring ways that we can make collaborative work with AI tools accountable, just and accessible to all.

CHAPTER FIVE

Collaboration

In order for creatives to feel both protected and empowered using generative AI tools, and for issues of bias and stereotype to be addressed, we need to consider the implications of human-AI collaboration on concepts like productivity, labour, copyright and accessibility. This means grappling with the specific role that generative AI tools play. in UK-based media production in an effort to understand the subsequent legal and ethical ramifications that arise from human-AI collaboration. While GAITs may increase creative productivity and open up new opportunities for non-professionals, there exists a general feeling of unease and distrust regarding their infringement upon intellectual property. The result is that, where many UK creatives express a desire to collaborate with generative AI, they are held back by a lack of clarity regarding the potential ethical implications and legal ramifications of doing so. As such, this chapter delves deeper into the views and possible solutions offered by those working within the UK creative industries when it comes to issues of collaboration, as it examines the need for both best practice and regulatory intervention to ensure that collaborative work including AI tools remains accountable, just and accessible to all.

As with Chapters Three and Four, this chapter is divided into two sections. The first section, '5.1: Workshop Findings', details the results and discoveries from across our four workshops (see Chapter Two for more details). Following a theoretical engagement with collaboration as a concept (5.1.1), this chapter
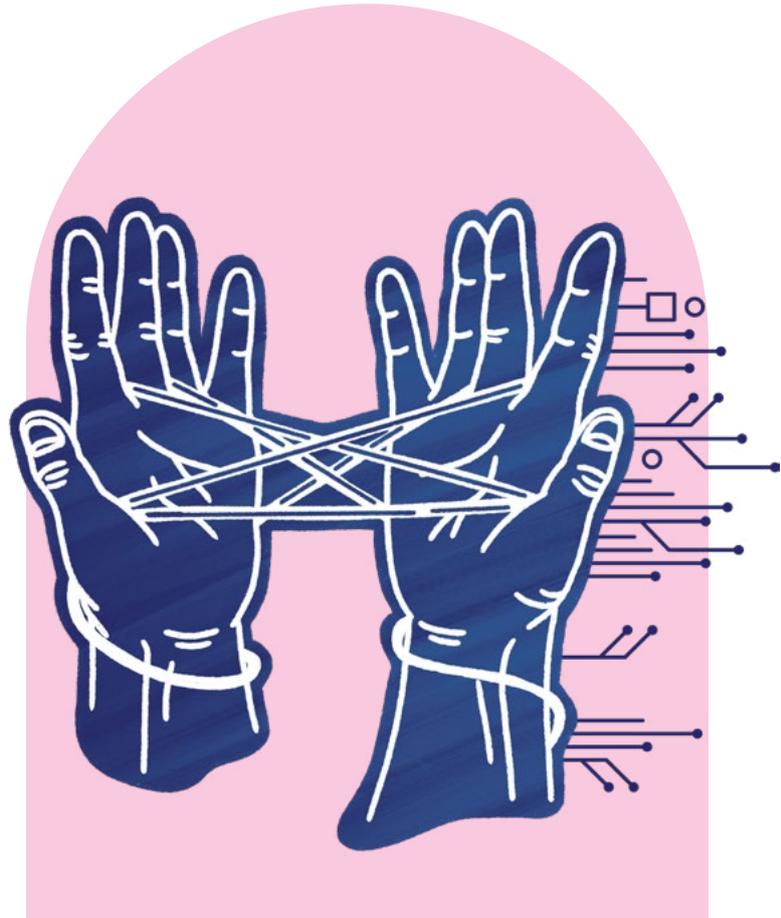
proceeds to consider how our workshop participants viewed collaborating with generative AI tools. It outlines the pervasive feeling that generative AI tools enhance productivity and efficiency by undertaking practical management work (5.1.2), before unpacking the complex issues of copyright and intellectual property that arise during processes of collaboration (5.1.3). Following this, the chapter examines what participants saw as the future of human labour (5.1.4), accessibility (5.1.5) and education (5.1.6) under the lens of human-AI collaboration. Section two of this chapter, '5.2: EBG and Collaboration', pivots to consider some of the practical and tangible outcomes that members of our Expert Bridging Group want to see to tackle the issues raised by our workshop participants. It does so by providing readers with an overview of the questions we circulated to the Expert Bridging Group in advance of the 'Collaboration' discussion session (5.2.1), after which it considers possible solutions and answers to these questions by way of different topics of conversation (5.2.2). As with the last two

chapters, our own consideration for these issues in light of our practical creative work will be embedded where appropriate to reflect upon the practicalities of responsible AI for independent media practitioners.

5.1: Workshop Findings

5.1.1: Collaboration as a Concept

Like creativity, the term 'collaboration' holds multiple connotations regarding its definition. It denotes a process of co-operation between more than one party or agent, working together to complement each other's skill and experience. In this sense, collaboration entails a symbiotic entanglement of both human and more-than-human elements, as espoused by philosophies of posthumanism and theories of the posthuman, which were explored in detail in Chapter One. Accompanying this context, philosophies of new materialism and post-phenomenology examine the agency of technology as part of a creative act, problematising the boundaries we draw between ourselves and the technologies depending on whether we consider them extensions of ourselves, agents we control, or agents that are outside of our control.[133] Questions of agency also bring to the fore attributions of intention, inscribing anthropomorphic qualities onto pieces of technology, like LLMs, through language and culture.[134]

Extracted from this academic context, on a more colloquial level, collaboration concerns the practice of 'using' AI tools. Yet, even here, the connotations of 'use' differ depending on context, such that, to collaborate with generative AI tools may mean to enter into an equal flow of exchange to create an end product that is the shared output of both human and machine intelligence. Equally, for those with particular concerns about a loss of human agency, control, ownership and intellectual property rights, collaboration may be a particularly daunting possibility as it implies the endowment of certain creative rights to a machine that, to return to Chapters Three and Four, may be incapable of creative output and may produce biased outcomes.

Amidst practical user guides[135] and publications concerning AI's impact on legal precedents,[136] understandings of collaboration run across a broad range of scholarship on generative AI and media and intersect with questions of what exactly constitutes a creative act.[137] A notable reflection of this trend is Florent Vinchon et al.'s aforementioned 'A Manifesto for Collaboration'.[138] Here, Vinchon et al. question what it might mean for an AI to interact with human creatives, as they hypothesise four potential collaboration scenarios: (1) the celebration of 'organic' creativity as a totalising rejection of generative AI tools; (2) an overreliance on generative AI as a replacement for human skill, or what they call 'Plagiarism 3.0'; (3) creative 'shut down' in which people feel disillusioned to even try and be creative in the face of generative AI; or (4) human-AI 'co-cre-AI-tion', a symbiotic

---

[133] Tom Davis, 'The Feral Cello: A Philosophically Informed Approach to an Actuated Instrument', NIME'17: Proceedings of the International Conference on New Interfaces for Musical Expression (2017), 279-282.

[134] Daniel Dennett, 'Intentional Systems', The Journal of Philosophy, 68: 4 (1971), 87-106.

[135] Eric Sarrion, Exploring the Power of ChatGPT: Applications, Techniques and Implications (Apress: New York, 2023).

[136] Michael Murray, 'Generative AI Art: Copyright Infringement and

Fair Use', SMU Science and Technology Law Review, 26: 2 (2023), 2023, 259-315.

[137] John McCormack et al., 'Design Considerations for Real-Time Collaboration with Creative Artificial Intelligence', Organised Sound, 25: 1 (2020), 41-52.

[138] Florent Vinchon et al., 'Artificial Intelligence & Creativity: A Manifesto for Collaboration', Journal of Creative Behaviour, 57: 1 (2024), 472-484.

collaboration between humans and machines and what they see as the optimal future.[139] While inherently speculative in nature, these possible futures offer a way into understanding how our workshop participants came to view human-AI collaboration in the context of media production.

5.1.2: Enhancing Productivity and Efficiency

The results of the workshops reveal that participants viewed generative AI as a valuable tool for enhancing productivity and efficiency in creative workflows. Many participants emphasised how generative AI has the potential to streamline workflows, save time, and allow them to focus on high-level creative tasks, a concept that we unpacked in greater detail in Chapter Three. For example, several participants noted that AI could 'speed up previous mundane tasks', 'speed up workflows' and 'make the process more efficient'. At the same time, generative AI can assist creators in various aspects of the creative process, such as subtitling, data-heavy procedures, brainstorming and generating ideas. By taking on these tasks, AI can help people to work more efficiently and effectively, leading to increased productivity and potentially better outcomes.

Generative AI's ability to handle data-heavy procedures and manage complex media tasks was frequently highlighted as one of its main benefits across our workshops. Within the specific context of video editing, for example, one participant noted that the ability for an editor to sort through hundreds of hours of footage has become a standard industry expectation, and AI could provide significant assistance in this area. Along these lines, GAIT and platforms like Hugging Face offer editors and content creators assistance in managing large amounts of data through labelling, categorising and classifying based on shot type or composition. However, where

many of our workshop participants recognised the need to have a high degree of technical proficiency and skill in order to use these tools effectively, and therefore GAITs may also be perceived as hinderances to the productivity of the creative process. That is, as with any new technology, integrating AI tools into one's workflow can just as easily be frustrating and slow things down as it can enhance productivity and efficiency, depending on the tool being used and the user's familiarity with that tool.

While AI may be an effective, though also difficult to master, collaborator by automating routine tasks that require consistency and logical processing, it still currently encounters challenges in more nuanced or ambiguous scenarios where contextual awareness is needed to handle unexpected situations.[140] As we saw in Chapter Three, this limitation is aligned with how participants think about using generative AI in their creative work.

While participants recognised AI as a tool for enhancing productivity, there remained a pervasive sense of caution about AI's role as a collaborator in creative decisions.

As one participant noted in Workshop 3 (Editing), 'editing is, a lot of the times, deciding when not to cut as it is when to cut and what decision to make'. This suggests that while AI can handle technical tasks, human judgment remains crucial for the artistic aspects of creative work. This indicates that human-AI collaboration constitutes what Vinchon *et al.* termed a 'co-cre-AI-tion' process of each agent filling in the others' shortcomings.

---

[139] Ibid., pp. 475-476.

[140] Thomas Davenport, Abhijit Guha, Dhruv Grewal and Timna Bressgott, 'How Artificial Intelligence Will Change the Future of Marketing', Journal of the Academy of Marketing Science, 48 (2020), 24-42; Nisreen Ameen, Gagan Sharma, Shlomo Tarba, Amar Rao and Ritika Chopra, 'Toward Advancing Theory on Creativity in Marketing and Artificial Intelligence', Psychology & Marketing, 39: 9 (2022), 1802-1825.

5.1.3: Copyright and IP Issues

When discussing collaboration with GAITs in the creative process, many participants expressed concerns relating to the issues of copyright and intellectual property, particularly relating to the identification of ownership. Several themes connected to this issue emerged, such as the ambiguity of authorship, the necessity for transparency in the training process of new and existing generative AI models, and the development of legal frameworks to protect both human creators and content generated in collaboration with AI.

One of the most pressing concerns expressed by participants was the question of who constitutes the creator when users collaborate with generative AI models to produce an artistic output. Almost all participants expressed uncertainty regarding how this was currently regulated. One participant exemplified this by asking, in relation to AI-generated music, 'ethically AI brings the copyright issue to mind: who is the creator? AI or the composer using AI?'. Statements such as these highlight a central dilemma regarding the perception of AI-generated content, highlighting how the boundaries of authorship and ownership are becoming increasingly blurred. They also indicate a pervasive desire to have these issues clarified as a matter of ethical and legal urgency, as participants indicated that they would be more willing to collaborate with generative AI tools if they could be sure that such practices would not be infringing upon anyone else's copyright, whose work may have been used to train the LLM in question.

Yet, in human-human collaborative projects, an understanding of precisely who owns what at the end is not necessarily always clear either. Within a framework of collaboration in which agents are working collectively toward a common goal, perhaps part of the collaborative process is giving away an individual sense of ownership or Intellectual Property in favour of a shared one. From our own creative practices, we recognise that our perception of ownership over a collaborative output may shift from case to case, depending on how much individual skill, effort, time and labour has gone into its production. If a piece of work is relatively 'easy' to create, does this detract from the value of ownership as perceived by an artist? While the question of who constitutes the owner and creator of AI-generated outputs remains open to such philosophical debate, and where legal ownership will shift according to the specific tool being used and the legal jurisdiction in place, our workshop participants recognised the need to credit at least some part of the AI in the process so as to avoid a scenario like Vinchon et al.'s 'Plagiarism 3.0'.

The existing literature regarding the IP and ownership is notably scarce within a UK context, primarily because of the uncertainty and novelty of the AI generated art and music.[141] This gap in the legal arena highlights the necessity for new and adaptive regulation that can address the current challenges. Moreover, existing research highlights that the legislation surrounding AI-generated works is highly region-dependent, with no clear consensus in many jurisdictions.[142] For instance, in the UK, computer-generated works which are defined as 'generated by computer in circumstances such that there is no human author of the work' are, at least theoretically, protected under the Copyright, Designs and Patents Act 1988. Yet, as outlined in Chapter One, the applicability of this Act to generative AI depends on whether

---

[141] Roser Batlle-Roca, Emila Gómez, Wei-Hsang Liao, Xavier Serra and Yuki Mitsufuji, 'Transparency in Music-Generative AI: A Systematic Literature Review' (2023).

[142] Ibid.

human intervention, such as via prompting, is deemed to be an arrangement that requires enough skill or judgement needed to classify the human as an author, returning us to the debates circulating around prompting as the new creativity that we explored in Chapter Three. Hence, as Pamela Samuelson notes, where generative AI models continue to evolve, they raise significant challenges to the legality of using and producing outputs from copyrighted material.[143] This raises the need for transparency in disclosing training and development processes.

Transparency is a concept that carries diverse meanings across different fields (see Appendix 1.4 for more details). In the realm of Government, Policy, and Social Sciences, the IEEE glossary offers a specific definition:

a process whereby information is requested and then disclosed completely within the limits of public law, without distortion, and with respect to the computational and cognitive capacities of the information recipient in order to enable those recipients to interpret the information so that they are able to make rational and informed decisions.[144]

In placing emphasis on transparency as a process with the end goal of enabling users to 'make rational and informed decisions', the IEEE's definition resonates with many of the desires of our workshop participants when it comes to ethical standards.

Our participants highlighted three key aspects of transparency that they feel that are essential for fostering trust in generative AI tools. One aspect is transparency in authorship, which is important not only for the individual ability to make ethical decisions based on collaboration but also for maintaining trust in the creative industries. Participants emphasised the significance of disclosing any involvement of generative AI in creative works, with one participant stating, 'it should be that someone has used AI, and if they don't declare it, there should be a massive scandal'. Statements like these highlight the growing desire to distinguish between human-created content and works produced in collaboration with generative AI. The 'scandal', as this participant put it, comes from a failure to be open and explicit to stakeholders about how certain decisions were made in light of the controversy surrounding generative AI training processes and the displacement of human labour. To address this challenge, participants recommended implementing industry-wide 'watermarking' for AI-assisted works, which would allow users to identify non-AI-generated content and help preserve end user choice as to whether or not they consume work that has been made in collaboration with AI.

Some companies and institutions are already working on watermarking features. For example, Meta introduced a feature on Instagram allowing users to label AI-generated content. The platform enables automatic labelling when AI indicators are present and allows users to manually label content they share, thus promoting transparency.[145] Similarly, as explored in Chapter One, the Content Authenticity Initiative was set up in 2019 as a cross-industry coalition of over 2000 media companies who have sought to develop content credentials that accurately label AI-generated images and content as such.[146] Members who have signed up to the

---

[143] Pamela Samuelson, 'Legal Challenges to Generative AI, Part I', Communications of the ACM, 66: 7 (2023), 20-23.

[144] Institute of Electrical and Electronics Engineers, 'A Glossary for Discussion of Ethics of Autonomous and Intelligent Systems, Version 1' (2017).

[145] Meta, 'Label AI-Generated Content on Instagram', Instagram Help Centre (2024) <https://help.instagram.com/761121959519495> [Accessed 23 September 2024].

[146] Content Authenticity Initiative, <https://contentauthenticity.org/> [Accessed 19 August 2024].

Content Authenticity Initiative include Adobe, the BBC, Getty Images, Microsoft, Shutterstock and The New York Times, amongst others. While such practices are undoubtedly positive steps in the fulfilment of explainability, allowing stakeholders to make informed decisions about their own use of generative AI, the fact that they were not already known by our participants highlights the need for more media coverage and education on AI disclosure and watermarking mechanisms.

Another aspect highlighted by participants was the transparency surrounding the training data used in machine learning processes (for definitions of these terms, see Appendix 1.1). As AI algorithms are typically trained on large datasets that may include copyrighted materials, there is a legitimate concern amongst those within the UK creative industries about potential copyright infringement in AI-generated outputs. This leads to fears that AI tools are 'stealing' an artist's unique style or likeness, as one participant reflected upon their own use of AI-generated music that closely resembled the style of popular vocalists. Another participant expressed these worries in the following: 'I still need strict rules about the copyright and remuneration for authors whom [sic] works are used to train the AI and improve data sets'. Such statements resonate with our own creative practice with generative AI tools, as the lack of transparency in regards to which original materials were used to train the datasets gave rise to concerns about unintentional plagiarism and conflicted feelings about claims of authorship. Workshop participants highlighted the urgent need for new regulatory frameworks that can address these ambiguities. Chapter Six will

outline some potential ways of mitigating these concerns of human-AI collaboration when it comes to training transparency.

For the third aspect, participants expressed a desire for greater transparency from AI developers regarding the ethical use of their technologies. They seek clear guidance on how to use AI responsibly, which includes comprehensive guidelines for ethical practices, information on the potential implications of AI-generated content, and best practices for integrating AI into creative workflows. As well as regulatory and governance interventions, Chapter Six will also address these desires by way of offering our readers a series of recommendations of best practice when it comes to collaborating responsibly with generative AI for media production.

So drawing from participants' statements, we can define transparency in media creation using generative AI as:

The open and honest disclosure of information related to the use of AI in the creative process, including the clear and explicit disclosure of AI's involvement, the sources and methods used to train AI algorithms, and the provision of guidelines for responsible and ethical use.

An AI-media landscape that accords with the understanding of transparency above would, then, facilitate the implementation of substantial legal protections for human creative labour.

5.1.4: The Future of Labour

Of the concerns expressed amongst our participants, the potential displacement of jobs due to AI and automation appeared to cut across all aspects of the creative industries. Participants expressed fear that a rapid rise in media practitioners using generative AI, or being expected to use generative AI, would lead to a significant reduction in employment opportunities, particularly across areas like post-production. For instance, one participant declared that generative AI would 'decimate cutting rooms. You're not going to have editing assistants; you're not going to have trainees', while another participant noted that they were 'apprehensive about AI and how it will take jobs away from the media production industry'. As a recurring concern, some participants saw collaborating with AI as a potential erasure of the traditional career path, in which one would start out as an assistant and work their way up. This trajectory was seen to be at risk with the widespread use of AI assistants, in which these jobs 'can be easily optimised, automated', as one of our Workshop 3 (Editing) participants reflected. Such perspectives align with the widespread concern with mass unemployment being the result of the widespread proliferation of GAITs, which refers to a situation in which our capacity to find ways to reduce the need for human labour is greater than our ability to create new jobs or find new works for human labour.[147]

A potential loss of human labour would inevitably have wide-ranging economic implications, both for the individual whose role would be at risk to replacement with creative generative AI and for society as a whole. Our participants expressed a desire to see fair compensation be protected in collaborations with generative AI, which they viewed as a pathway to avoid potential devaluation of the skills required to undertake creative tasks. The risk that large tech companies, concerned primarily with profit under capitalism, would seek to exploit the development of generative AI to prevent paying a human creative was seen as a very real possibility amongst many of our participants, as reflected by one participant, who stated that 'human skill is expensive so generative AI will be used instead, meaning jobs currently regarded as 'creative' will decline'. Other participant highlights the devaluation of skilled workers by noting 'there is a concern that perhaps the perception is that you don't need skilled workers when actually there is still a level of skill that's required, an eye for design and detail and a need for touching up'.

Like our participants' concerns over transparency and intellectual property, participants highlighted the need for clear guidelines and updated legislation to 'protect those whose work is possibly exploited', as one of our participants put it. Media practitioners, in particular, wanted to see legal frameworks implemented to safeguard creators from potential misuse of their work and avoid collaborative 'shut down'. Their articulation of the need for 'clear legislation and rules of use that keep up with the development of the technology' reflects a broader concern that current intellectual property laws are not adequate enough to address this complex way of collaborating with AI to generate content.

5.1.5: Accessibility and Inclusivity

While no single conception of accessibility within AI has been universally accepted, and thorough research on the topic is currently limited, we can draw insights from related

---

[147] Yuri Lima, Carlos Babosa, Herbet dos Santos and Jano Moreira de Souza, 'Understanding Technological Unemployment: A Review of Causes, Consequences, and Solutions', Societies, 11: 2 (2021).

fields. Prevailing scholarly literature offers different interpretations for AI accessibility, particularly concerning 'web accessibility', which may provide a starting point. Julio Abascal, Myriam Arrue, Nestor Garay-Victoria and Jorge Tomás define web accessibility as 'the possibility of accessing any web content by anyone, regardless of circumstances such as impairments, platforms, devices, browsers etc.'.[148] Building upon this, Helen Petrie, Andreas Savva and Christopher Power gathered over 50 definitions of web accessibility and synthesised them into a single understanding that:

'[A]ll people, particularly disabled and older people, can use websites in a range of contexts of use, including mainstream and assistive technologies; to achieve this, websites need to be designed and developed to support usability across these contexts'.[149]

While pertaining to website design and use, these preliminary understandings of web accessibility can help us unpack what our participants saw as the key issues concerning AI accessibility.

From our analysis of the workshops, we found that our participants located three different perspectives of accessibility when collaborating with generative AI tools: accessibility for users with different abilities; accessibility regardless of socioeconomic factors, such as income, experience or age; and geographic accessibility. Notably, questions of disability, considering how AI may be more accessible to people with visual, auditory, motor or cognitive disabilities, was not a central concern within these discussions, as this topic was only mentioned explicitly

a few times. There might be several reasons why our participants did not focus much on this topic. It could be due to a societal lack of awareness for disability, such that understandings of accessibility may centre around a generalised ease of use that stems from what developers see as an 'average' user, regardless of ability.[150] Alternatively, participants might have already assumed that because of the rapid pace that AI is developing as a cluster of technologies, generative AI tools may already be accessible to all users. Whatever the reason, disability-specific considerations of accessibility remained less prominent in our discussions of accessibility than other factors, such as socio-economic diversity. This reflects a gap in awareness about the challenges faced by disabled users when collaborating with generative AI tools, highlighting the need for more extensive discussions on these topics, as well as further education about what inclusive design practices of AI models that take into account the needs of all users, including those with disabilities, ought to look like.

Our participants' focus on accessibility mainly centred around socio-economic factors, such as whether an AI tool is made available to individuals with varying experiences and skill levels, as well as ages and incomes. One of the most frequently mentioned opportunities that generative AI tools give media creators was the chance for non-professionals or those lacking the necessary artistic skills to enter the industry. As one participant stated, collaborating with generative AI 'lowers the bar for entry level which helps getting more people involved who might have felt intimidated by this'. This draws upon discussions of AI equity or

[148] Julio Abascal, Myriam Arrue, Nestor Garay and Jorge Tomás, 'A Web Service for Automatic Accessibility Analysis of Web Pages Based on the Use of XML Structures' (2003), 925-929.

[149] Helen Petrie, Andreas Savva and Christopher Power, 'Towards a Unified Definition of Web Accessibility', W4A '15: Proceedings of the

12th International Web for All Conference (2015), 1-13.

[150] Sheryl Burgstahler, 'Universal Design: Process, Principles, and Applications', DO-IT (Disabilities, Opportunities, Internetworking and Technology) (2009).

the democratisation of AI art, as explored in Chapter Three, in which we explored counter-claims to these sentiments that sought to protect the necessary skill, effort and talent required to produce creative outputs in the face of generative AI. Another participant similarly explained that generative AI 'presents a shortcut between people who would have had to have spent 20 years learning an instrument or notation for music […] to being completely democratised to somebody'. Where such comments reflect a prevailing view amongst media practitioners who saw the positive potentials of generative AI when it comes to accessibility, they nevertheless return us to concerns about job replacement and precarious labour, which already disproportionately disadvantages those who have previously been denied access to the UK creative industries.

These discussions also coincided with debates about generative AI tools being free-to-use vs. requiring a paid subscription. Many of our participants praised the fact that some of the most popular generative AI tools, such as ChatGPT, Poe and Midjourney, are free to use, while AI tools with costs involved presented a barrier to those on lower incomes. The problem, as our participants reflected, is that 'different AIs have different qualities that you might want to have, like, a mixture of them, to achieve your proposal', and that, 'at £20 each month […] you can really end up paying a mortgage to have the assistance, so it is

the same cost of having, almost, a human assistant'. Add to this the fact that generative AI tools often come with different tiers that increase technical proficiency in exchange for paid subscription.

Finally, participants reflected upon generative AI's relationship to geographic accessibility. This included considerations about whether AI tools are truly available to all members of society across the world, or only accessible to certain privileged groups. While recognising the persistence of problems like bias and stereotype, as explored in Chapter Four, our participants largely viewed generative AI as possessing a global reach. Yet, they also expressed concern about the current state of inclusivity that generative AI media entails. Some of our participants expressed concern over the current lack of representation, particularly for Black people, in the collaboration with, and end product of, generative AI media. They recognised that fostering inclusivity could 'enrich the diversity of voices and perspectives represented within media landscapes', ultimately leading to a more representative and equitable media production industry. Combining the trends outlined here with the exploration of bias that took place in Chapter Four, the next chapter will outline some potential ways to mitigate these concerns and embed accessibility and inclusivity into the AI media landscape.

5.1.6: Education

As a collective of different stakeholders, many of whom were end users of generative AI tools seeking to both learn about responsible AI use and improve their own technical proficiency, our participants recognised AI as a 'complex' technology that requires a substantial amount of time and skill to use effectively. One participant highlighted this as a knowledge gap, noting that 'full understanding [is] only in the hands of the developers'. Comments

like these reflect the belief that redressing the knowledge gap is about shifting the dynamics of power and understanding, a task that requires active participation from all stakeholders involved. This same participant recommended training and education to teach people about how to ethically use generative AI, while another participant emphasised the importance of incorporating AI training into the formal education system. They argued that generative AI should be 'incorporated into educational practices to ensure that the next generation are prepared for the world of work', indicating a belief that, as generative AI tools change what constitutes creativity and collaboration within the UK creative industries, those new to the fields need to be adequately equipped to deal with these changes.

When discussing AI education, we also heard the perspectives of educators and teachers. Some educators perceived generative AI as a useful way to help students overcome learning and developmental barriers in creative processes, such as visualising ideas quickly or getting words onto a page that can later be edited. Yet, at the same time, our participants raised concerns about the impact of generative AI on assessment methods, with suggestions that traditional essay-based evaluations may need to be reconsidered and plagiarism conditions re-evaluated. Such concerns are backed-up by the plethora of scholarship that is seeking to address this issue.[151] While these fall outside of the remit of this research project, they nonetheless speak to a prevailing concern that users may become reliant upon collaborating with generative AI tools in ways that undermine traditional workflow practices.

5.1.7: Sustainability as an Emerging Issue

While the environmental impact of AI in media production was not the primary focus of our research, as we engaged with participants, several related issues organically emerged, highlighting the need to address sustainability in AI media practices.

Concerns such as the excessive energy consumption required for training AI models, the increasing computational demands, and the significant data storage needs were frequently mentioned. One participant pointed out the 'unspoken environmental impact of generating hundreds of images', highlighting an awareness for how the rich output created by AI can lead to intensified data storage requirements and, consequently, far greater energy consumption. Participants' worries are supported by studies highlighting the substantial energy consumption and carbon emissions associated with training large AI models. Research has shown that training a single large AI model like a language model can consume thousands of megawatt hours of electricity and emit hundreds of tons of carbon.[152] These concerns become particularly relevant when considering the role of AI in media production. As AI tools become integrated into the media creation process, the environmental footprint expands significantly.

Initiatives such as WeAreAlbert and Small File Media Festival make valuable contributions to sustainability in media production by targeting different aspects of the media lifecycle.[153] While WeAreAlbert focuses on the production phase, encouraging greener practices behind the scenes of

---

[151] Caitlin Bentley et al., 'A Framework for Responsible AI Education: A Working Paper', SSRN Working Paper Series (2023); Tama Leaver and Suzanne Srdarov, 'ChatGPT Isn't Magic: The Hype and Hypocrisy of Generative Artificial Intelligence Rhetoric', M/C Journal, 26: 5 (2023).

[152] Carole-Jean Wu et al., 'Sustainable AI: Environmental Implications, Challenges and Opportunities', Proceedings of Machine Learning and Systems, 4 (2022), 795-813.

[153] https://wearealbert.org/; https://smallfile.ca/.

content creation, Small File Media Festival addresses the consumption phase, pushing for reduced data use and environmentally friendly streaming habits. Where our insights and research has shown that sustainability, in terms of environmental responsibility and resource efficiency, must be considered a key factor in AI media production, these kinds of initiatives are incredibly valuable. However, there is still a strong need to educate users about the environmental impact of their media consumption and the role AI plays in this process. Moreover, more initiatives like these must emerge, bringing greater attention to sustainability in AI media production and ensuring that this critical topic receives the focus it deserves.

5.2: EBG on Collaboration

In advance of the Expert Bridging Group meeting, we circulated the following preliminary information from our workshop findings, followed by a set of guiding questions that we asked each member of the Expert Bridging Group to prepare answers to.

5.2.1: Collaboration Interview Questions

What we discussed

A significant issue here is the potential loss of control by human collaborators, which can diminish the sense of true partnership. The perception of AI tools also varies; some may see them as sophisticated plugins, others as assistants, or simply as tools, affecting the dynamics of collaboration. Ensuring that AI systems are seen as augmentative rather than replacements can enhance the collaborative experience. Transparency in AI systems is crucial for fostering trust and understanding. AI systems must be open and transparent about their processes and outcomes, promoting traceability, explainability, and effective user communication. A lack of

information about the potential risks and harms of AI-created content is a major concern. Limited traceability of creative outcomes and ambiguity regarding the ownership of these outcomes further complicate matters. Users often lack adequate information and education about the tools, guidelines, and ethical frameworks governing AI use. Moreover, there is often no consent mechanism for individuals depicted in AI-generated media, leading to ethical dilemmas. The potential for exploiting individuals' work for financial gain and the lack of transparent data provenance are also significant issues, as is the inaccessibility of information about the computational background of AI tools. Copyright and intellectual property (IP) issues are critical when it comes to AI systems.

These systems must respect and adhere to copyright and IP laws, ensuring that the ownership and rights of both AI-generated and human-created content are clearly defined and protected. However, the ownership of AI-generated content often remains unclear, leading to potential copyright infringement and plagiarism. Protecting the rights of human creators and identifying copyrighted works used in AI training are challenging tasks. The impact of AI on labour and job displacement in the creative industries cannot be overlooked. AI systems should be developed with consideration for their effects on employment, addressing changes in the job market, skill requirements, and economic impacts on human workers. The potential loss of jobs in traditional creative roles is a major issue, as is the shift in required skills for media creation jobs. AI can cause economic disruption in creative industries, raising questions about whether media production with AI is easier or harder compared to traditional methods, and what technical skills are required to achieve high-quality results with AI. Environmental well-being and sustainability are also important

considerations for AI systems. These systems should promote sustainability and ecological responsibility, avoiding any harm. The high energy usage and non-environmentally friendly practices associated with AI are significant concerns. There is often a lack of information about the environmental footprint of AI technologies in terms of resource usage, creating ambiguity about the balance between AI benefits and environmental costs.

Questions

- What are the most effective ways to integrate AI systems into existing creative workflows, enabling seamless collaboration and allowing users to leverage AI's capabilities without losing control?

- How can we establish clear legal frameworks for the ownership and licensing of AI-generated content, ensuring that both human creators and AI systems are treated fairly?

- How can we design AI systems that encourage shared ownership of creative outcomes, blurring the lines between human and AI contributions and fostering a sense of shared authorship?

- What measures can be taken to protect the intellectual property rights of human creators whose works are used in training AI systems?

- Who should hold the copyright for content generated by AI systems: the developer, the user, or the AI itself?

- How can we establish legal requirements for AI systems to maintain transparent data provenance, providing clear and accessible information about the sources and origins of the data used in their training?

- How can we develop policies and programs that mitigate the potential for job displacement in creative industries due to the rise of AI technologies?

- How can we update existing copyright and intellectual property laws to address the unique challenges posed by AI-generated content, ensuring that both human creators and AI systems are treated fairly?

- What tools and frameworks can be developed to help users to understand the footprint of their AI projects?

5.2.2: Collaboration Discussion

As the discussion topic that closed the Expert Bridging Group, 'Collaboration' began with each participant member offering their key proposal for change that they wanted to see implemented as a result of generative AI's impact on the UK creative industries. This meant that many of the questions above were addressed from the perspective of practical solutions, rather than further elaborations on the problems caused by human-AI collaboration.

Discussion began with expressions of interest regarding the future development of generative AI as a technology, as members offered their own visions of future human-AI collaboration or 'co-cre-AI-tion'. For one UK-based editor, collaboration with generative AI tools could be improved by the development of voice-activated editing, opening the possibility of a natural flow of exchange that replicates human-human collaboration. Similarly, an independent filmmaker and producer of documentary cinema wanted to see AI assisted plug-ins for pre-existing digital filmmaking tools, allowing the user to set contextually specific parameters to achieve creative outputs. These desires reflect a trend within practitioners of the UK

creative industries to make generative AI tools more advanced in their media production capabilities.

Aside from technical improvements, a topic that dominated our discussion on the subject of 'Collaboration' was the possibiltiy of mitigating risk when integrating AI systems into creative workflows, particularly in relation to obscured training processes. Practitioners, academics, researchers and union representatives all wanted to see greater transparency regarding the use of pre-existing creative works for the training of Large Language Models by AI developers. In order to build audience trust and prevent societal 'shut down' when engaging with generative AI outputs, a leading academic on AI education proposed the implementation of meta-data as a disclosure mechanism for all digital creative outputs, ensuring that, where a product is the result of human-AI collaboration, it is clearly marked as such. Coupled with the suggestion that all AI companies should reveal the sources upon which they have trained their LLMs, this would enable traceability of creative lineage, potentially mitigating the personal risk that comes with using AI-generated content that has been trained on works that do not align with your own ethical or moral values.

Going one step further, many of our members hypothesised about the possibility of not only revealing training data but paying for it, suggesting that fair remuneration for artists whose works have been used to train LLMs, would facilitate a greater sense of fairness and parity in the model development process. Yet precisely how much this would be and how this would work remained a point of contention. One avenue of exploration is the establishment of common rates of remuneration in which AI developers negotiate with Collective Management Organisations to grant access on behalf of

rights holders for the use of their work in AI training. Where it to be feasible to trace the source data for each and every AI-generated output, a different mechanism could involve the end user accepting copyright responsibility of the final output, including the price to pay to the artist(s) whose work was used or contributed to the end product. This would involve a technical challenge regarding how to trace each and every data source for a generated output, but it would nonetheless ensure that users are only paying for materials that went into their productions.

In response to the question, 'how can we establish legal frameworks for the ownership and licensing of AI-generated content, ensuring that both human creators and AI systems are treated fairly', our EBG members all recognised the urgent need for change. Members wanted to see the passing of laws mandating clearer copyright stakes, as this is currently holding creatives back from fully engaging with GAITs in media production. As one academic and film distributor noted, clearer copyright law could allow the AI industry to potentially avoid a Napster-style distribution of data that infringes upon creative copyright. It was agreed that terminology surrounding this update to UK copyright law needs to be specific, so as to prevent 'future-proofing' agreements that allow AI companies to circumvent paying creative artists by anticipating any future changes to the AI legal landscape, but not so restrictive that it limits the use of GAITs altogether. Under this framework, copyright could potentially be assigned to creative prompts. A UK union representative presented their position as implementing legal regulation so that explicit permission must be sought from individual creatives before their work is used by AI developers. This would form the basis of a voluntary, opt-in licensing framework, giving human creatives agency over their own data that can

be sold to developers at a negotiable price for training as a secondary income stream. Yet, other members emphasised the need not for the generation of new laws but rather on the practical implementation of pre-existing intellectual property laws and updates to regulation that emphasise product safety and consumer and creative rights as a priority.

Amidst these discussions, a key issue raised was what human-AI collaboration means for the human actor or voice actor, a stakeholder position that our workshop design failed to adequately account for. A leading academic on digital performance rights raised the point that any contractual rights about copyright and intellectual property need to allow for pathways in actor and performer co-design in the media production process. The issue of remuneration in this regard is difficult within a UK legal context, which has no statute-based protection of image rights or personality rights. Yet, when voicemod and video-based AI technologies can make actors 'do' and 'say' things they would not have originally performed, as we saw in Chapter Four, this can lead not to the spread of harmful misinformation and violent outputs. It also perpetuates imbalances of power between producers and performance artists, in which an actors' likeness can be used beyond the agreed upon terms and conditions as stated in their original contract. An example given was that of a voiceover artist who signed a contract for research purposes but later found out that their voice was sold on to be used and cloned by an AI developer and appropriated for commercial purposes. Other examples include the manipulation of line readings or performances by generative AI for use in pornography. While there existed a general agreement amongst our members that we need greater collective bargaining and the possibility for actors and voice actors to withdraw consent, producers expressed a desire to retain the right to be flexible in how

they use an actors' image rights, suggesting a tension between different stakeholders within the UK creative industries.

In this sense, human-AI collaboration necessitates that the rights of actors and rights of producers be placed into conversation with one another when tackling issues of copyright and intellectual property. Amidst a responsibility gap, which needs to be filled as a matter of urgency by legal intervention, members of our Expert Bridging Group expressed their commitment to self-regulation based on what they deemed to be ethical human-AI collaboration practices. Doing so would help prevent the wider societal effects 'shut down'; that is, of pushing people away from generative AI in light of the complex legal landscape that does not adequately address the technology's impact on media practitioners. In this space, members discussed the need for clearer communications for artists in how they should or should not be using generative AI tools in their own work; increased education opportunities to emphasise responsible human-AI collaboration; and further opportunities for divergent stakeholders to have difficult conversations and learn from one another across the AI divide.

5.3: Conclusion

This chapter examined how the concept of 'collaboration' may be discussed in relation to media production processes that use generative AI tools. It investigated what media practitioners and end users thought about the potential promises and drawbacks of collaborating with GAITs, exploring the tensions that emerge within the UK creative industries between positioning generative AI as a facilitator of productivity and accessibility, on the one hand, and seeking to find solutions to issues of copyright infringement and labour extraction, on the other. It subsequently placed these tensions under the lens of potential solutions from our Expert Bridging Group. What emerged was a consensus amongst stakeholders regarding the need for greater guidance and communication

on how best to collaborate with AI in a responsible way. Yet, as our participants and Expert Bridging Group members recognised, this needs to be combined with legal and regulatory clarification from the UK government designed to mitigate exploitation, enforce copyright laws, and embed protections for creative artists and media practitioners. While the necessity of these developments is readily apparent, the specifics about what they ought to look like need to be clarified as a final point. Hence, Chapter Six concludes this White Paper by way of addressing the current academic, legal and professional vacuum that exists within the UK creative industries, as it imagines the future possibilities for developing a responsible AI media climate.
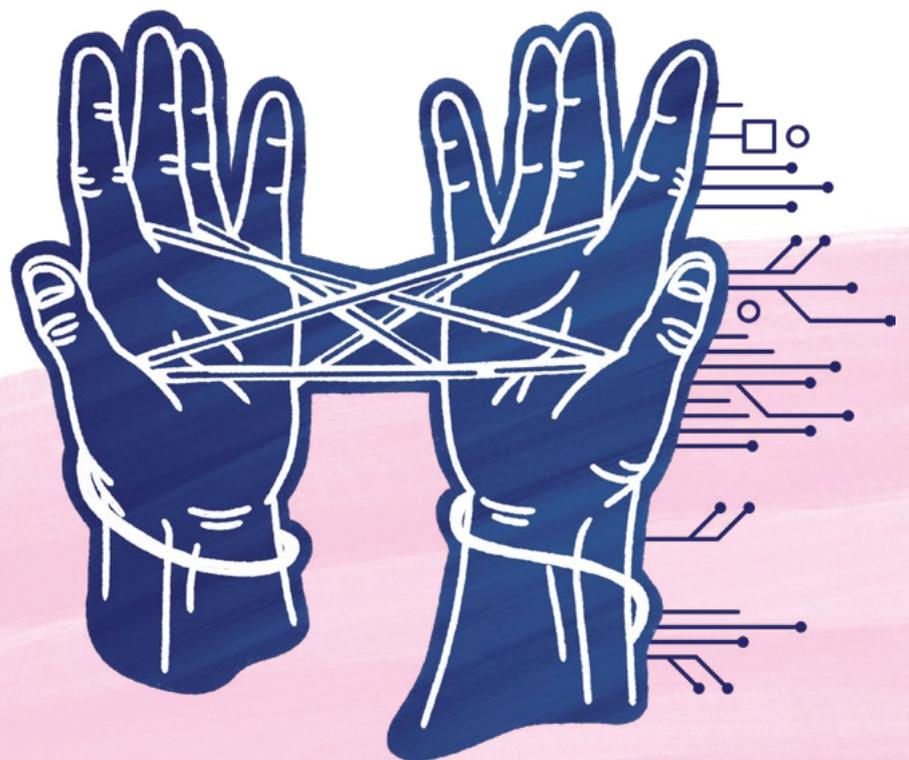
CHAPTER SIX

Possible Futures

6.1: Principles and Recommendations for Responsible AI Practice

Based on our extensive research investigating the needs and concerns of the UK film and media industry, we have devised a series of principles and recommendations for responsible AI practices. These principles are by no means exhaustive; their applicability will be context dependent, such that different stakeholders working across different stages of the media production process (screen-writing, image generation, editing, music creation) will need to tailor these recommendations accordingly. Nor are these recommendations designed to definitively address industrial or academic discussions about what 'responsible' AI use looks like within the creative industries, a task that will be an ongoing, relational and collaborative process so long as advancements in generative AI keep developing at their current rate. Rather, we offer them as a way of synthesising the views expressed across our workshop, the Expert Bridging Group, and our experience during creative practice.

6.1.1: Accountability

Comprising what Virginia Dignum calls 'the first condition for responsible AI', accountability concerns the ability to explain and justify key decisions to stakeholders for whom these decisions may impact.[154] This allows involved agents and parties to participate in the process of making the use of generative AI responsible. One way to achieve this is by keeping clear logs and data records when using generative AI tools. Before, during and after your usage, record who is doing what, as well as how key decisions are made (such as tool use, prompting and revisions), so that those involved in the media production can later explain these decisions to relevant stakeholders, such as audiences and production crew. Doing this can help what AI ethicist Luciano Floridi calls 'ethics shopping' or 'ethics washing', the practice in which developers and producers adopt a vague and open 'ethical' standard to retrofit pre-existing behaviours to prevent any need for change or critical reflection.[155]

6.1.2: Transparency

While accountability via clear data records helps hold those in positions of power answerable to involved stakeholders, this

---

[154] Virginia Dignum, Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way (Cham: Springer, 2019), p. 54.

[155] Luciano Floridi, The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. Oxford: Oxford University Press, 2023), pp. 69-70.

principle needs to be combined with specific measures taken to ensure transparency about said data records in order to build trust. As such, media practitioners using generative AI tools should exercise a principle of transparency through clear and open communication about how important decisions have been made. This means making data records available in an accessible format, which might involve publishing these alongside the release of your finished production or the ongoing documentation of your process via blogs, websites or social media. Transparently publishing your data records, including tools and prompts used and the sequential process from start to finish, will provide your stakeholders with a traceable overview of your creative process without infringing upon any commercially sensitive information. Most importantly, we urge media practitioners to always disclose when a specific generative AI tool has been used to produce a creative output. Following this principle will not only prevent potential audience distrust in your output; it will also go some way to fostering a climate of generative AI use that is open to development and debate as a whole.

6.1.3: Redressing AI Bias

As Chapter Four's key findings demonstrated, redressing the bias inherent within popular generative AI models is important so that AI-generated media does not perpetuate harmful stereotypes in its portrayal of already marginalised societal groups. If uncorrected, generative AI outputs are more likely to produce misogynistic, racist, homophobic, transphobic and ableist outputs, based as they are on biased datasets and training processes. While correcting this requires regulatory intervention on training processes (which will be addressed in 6.3), it also necessitates active intervention from all stakeholders within the AI media landscape. As such, we urge users

of generative AI to do so with a conscious awareness that generative AI tools can and most likely will produce biased outcomes.

One way to do this is to prompt with the active intention of incorporating diversity into your outputs; doing so with respect to an awareness that prejudice functions intersectionally across axes of gender, race, sexuality, class and ability can help mitigate biased and potentially harmful outputs. For example, users looking to generate an image of a historically white- and male-dominated profession may prompt an image-generator with the clarification of producing, say, a 'Native American woman' in that role. While this might go some way to redressing the output of predominantly white men, depending on the specific AI tool used, it may also risk perpetuating further biases, such as reinforcing patriarchal and ableist beauty standards by producing only slim, young and able-bodied women or stereotyping Native American cultures through generic apparel. A participant from Workshop 4 (Music) gave a different example, highlighting how, while a prompt asking for the production of 'African music' may initially seem like an endeavour towards diversity actually does a disservice to the diversity of cultures that exist across the African continent. Prompts should therefore be more attuned to cultural specificity and, as much as possible, avoid generic, stereotyped descriptions or labels as these can homogenise large groups of diverse people. Upon the production of an output, a right to human review should always be implemented, facilitating the process in which an end user can edit or tweak AI-generated content that appears biased.

6.1.4: Collaboration

As Chapter Five's key findings demonstrated, media practitioners want to see a path forward for collaboration that simultaneously

recognises the potentials of integrating generative AI tools into one's workflow and values the human workers at the heart of creative labour. In this sense, we encourage creatives to adhere to the principle of collaboration, using generative AI tools to enhance, rather than replace, human creativity in media production. Rather than putting an end to human-human collaboration, generative AI tools should be used to help us collaborate more fully with other human creatives, such as the connectivity capabilities that AI translation or closed captions afford people across the world. Those in the position of hiring or managing human staff should not see generative AI tools as a replacement for human labour but rather its capacity to augment the creative process, championing AI as a helpful technology that supplements the creative, authentic and original experience and skill of the human. In terms of tangible actions, this would mean ensuring that human creatives are still being hired, and paid, for their role within the media production at every stage of the process, so as to not lose sight of the material financial consequences of generative AI, particularly for those in positions of job precarity. Where generative AI tools are used to assist with tasks such as ideation, subtitling, logging, colour correction or the production of a demo score, they need to be integrated into the production in ways that do not appropriate or compromise the human role in decision-making.

6.1.5: Interdisciplinarity

In line with leading research on responsible AI and the findings of our workshops that encouraged further dialogue across different facets of the AI media landscape, we call on practitioners to exercise interdisciplinarity when collaborating on media projects. Engaging critically with different stakeholders, such as filmmakers, performers, software developers and technical engineers, unions,

studios, academics, ethicists and end users, will enable you to incorporate a diverse range of views into a final output and prevent generating an echo chamber. One way to achieve this would be to develop a team with wide diversity, both in terms of positionality and specialisms. Bridging AI divides across media, policy and development of generative AI tools also means that your project would be making use of the vast array of expertise and specialist knowledge that sit within different fields and disciplines.

6.1.6: Informed Participation

In line with the principles of accountability and transparency, media practitioners looking to facilitate a responsible use of generative AI in their productions should provide clear communication with all involved parties in the project, ascertaining their expressed consent to be involved, and ensuring that the use of any of their data falls clearly within the original terms as set out in a legal contract or agreement. This extends from consent to be part of a project that uses generative AI in its production process, as well as consent from a stakeholder to have their image, voice or likeness manipulated by a specific generative AI tool. Consent must be fully informed and affirmative, meaning people actively opt-in, as opposed to it being a default position that someone has to opt-out of. Obtaining informed participation from actors, staff, crew or other creatives prior to their involvement in the project will enable you to adhere to intellectual property and privacy laws. It will also set an ethical standard in which any use of personal data would meet the expectations of all participants, especially actors and voice actors, for whom the risk of exploitation and data extraction is particularly high.

6.1.7: Open Datasets (Where Applicable)

For those with the technical ability, training

and access to do so, we would encourage media practitioners to develop and use your own localised LLM, trained on data that you own the copyright in or have acquired the necessary permissions from the copyright owner, so as to mitigate against any potential copyright infringement that comes with using publicly available generative AI models. An example would be using an image dataset of photographs that you have taken. Another example may be using images or works that are in the legal 'public domain', meaning that they are no longer in copyright duration and everyone is free to use them.

We recognise that this will not be an opportunity open to everyone within the AI media landscape, and the functionality of this principle will be contextually specific, but the use of personally trained LLMs and the development of one's own model for internal purposes reduces the ethical and legal implications that come with popular models which are trained on creative works that AI companies do not own the copyright for. We therefore advocate for further training and education designed to help people transition toward the development of their own, locally-trained LLMs.

6.2: Critical-Creative Reflective Questions

As well as following the above principles and practices, we recommend that users contemplate the purpose and necessity of collaborating with a specific generative AI tool. In doing so, users can try to mitigate potential harms and risks by asking themselves critical-creative reflective questions that work to ensure self-regulation of ethical and responsible AI use. To begin with, we would urge users to consider the following questions:

Why is the use of generative AI necessary? Could a non-AI tool be used for the same purpose?

Doing so allows the user to stop and think about the purpose of their project in a critical light. In ascertaining the necessity or lack thereof of using a specific generative AI tool, users might want to reflect upon the overall project's aims and intentions, and whether or not these align with the implications of using generative AI, as well as the functionality, transparency and ethical considerations for the specific AI tool that is being used. Appendix 2 (AI Tools Matrix) details some preliminary findings on some of the most popular generative AI tools in this regard, but we would urge users to do their own research into their tool of choice. Asking yourself this central question will not only lead you to develop a greater sense of your own project; it will also enable you to consider what you want to get out of generative AI if you decide that its use in your project is necessary. For example, you may decide that, after some critical reflection, the use of a specific generative AI tool is necessary because it saves resources by speeding up otherwise tedious processes; it enhances your creative potential through ideation, research or review processes; it offers unique aesthetic capabilities that other technologies do not; or it enables you to access media creation processes that were not previously available to you.

Following this, we would encourage you to ask yourselves the following questions at every stage in the production process. We divide these questions (as well as example scenarios and solutions) into three distinct stages.

6.2.1: Before Using AI

How transparent are the developers behind the AI tool about the dataset(s) they have used for training?

This information is not always easy to identify, and many of the most popular generative AI tools are not open about this to end users. Understanding what datasets have been used to train the tool you are using can help you determine any potential copyright and intellectual property risks; as well as any ethical concerns you have about inherent bias.

To be clear, currently in the UK there is not a legal requirement that organisations make this data publicly available. Some organisations may choose to make this information available.

You could consider whether the company behind the AI tool you are using are signed up to the Content Authenticity Initiative, a cross-industry coalition of over 2000 media companies dedicated to transparent labelling of AI-generated images.[156] Adobe Firefly (image generation) is just one example of an AI tool that publicly highlights the diversity of its dataset, but even here, the specific details about its data sources are not disclosed in depth.

How does the tool use the end user's input?

As outlined above, many generative AI tools train their Large Language Models in real-time based on the end user's data. Since generative AI tools work by synthesising and amalgamating large amounts of data to produce an aggregated output based on pre-existing material, users who are cautious about their own creative work being used to

help develop AI-generated outputs may want to consider how the model proposes to use their prompts or data.

Reading the terms of use on the website/ license agreement should clearly state what uses will or will not be made with the outputs you generate using the tool. There may also be information in these terms about how to exclude your outputs being used in further training for the LLM models.

What information can you find about the company's commitment to sustainability and other ethical issues? Are there any mechanisms in place for you to provide feedback?

As with obtaining information on training, finding out about the environmental impact of using a specific generative AI tool can be difficult. At present, little is known about the true environmental impact of using generative AI tools, partly because there exists a large degree of variation in terms of energy consumption and carbon emissions depending on the models being used and the purpose of their deployment. However, new research is emerging to suggest that image generation uses a significantly greater amount of carbon to produce outputs than text generation or any classifying, modelling or extraction processes.[157]

A good place to start is by seeing if the AI model you are intending to use has any information published on their website about their commitment to sustainability and other ethical issues, or whether they are in affiliation with recognised organisations and institutions committed to global sustainability. If not, does the company invite feedback and allow you to ask questions about its commitment

156 Content Authenticity Initiative, <https://contentauthenticity.org/> [Accessed 19 August 2024].

157 Alexandra Luccioni, Yacine Jernite and Emma Strubell, 'Power Hungry Processing: Watts Driving the Cost of AI Deployment?', ACM FAccT '24, Brazil (2024), 1-15.

to sustainability? This may be a sign that the company in question is actively involved in understanding the energy and emissions involved in developing, training and using its tool.

How might you mitigate potential labour issues?

This question enables you to consider the material consequences that your use of generative AI may have on other stakeholders involved in the creative process. In conjunction with 6.1.4 and the principle of 'Collaboration', ask yourself how you might combine the possibilities that generative AI offers in terms of costs and productivity with a commitment to fair remuneration of human creatives for their time, effort and skill. As an example, when utilising voiceover in your production, you could consider a blended method of both hiring a human voiceover artist and experimenting with generative AI tools and the creative possibilities they enable. However, any manipulation or experimentation with the voiceover artists' voice needs to be communicated clearly and openly to the artist beforehand, and must not violate the agreed terms of the contract, as set out in the principle of Informed Participation (6.1.6).

6.2.2: While Using AI

How is the tool reinforcing stereotypes and biased representations? How can you mitigate those?

As outlined above in 6.1.3, it is important to be aware of how the generative AI tool you are using may potentially reinforce stereotypes and biases. While using the tool, be mindful of the sorts of outputs you are receiving. For instance, if you are prompting a text-to-text or text-to-image AI tool to produce a description or an image of a doctor without using any clarifying terms about the doctor's identity, is

the tool only producing one type of doctor, assumed to be white, male, able-bodied etc.? On the other hand, ask yourself: is the introduction of identity-based descriptive terms generating outputs that can be said to be stereotyped?

These questions are important to consider during the use process so that you can try to mitigate these through consideration of intersectional prejudice and cultural specificity where appropriate. Through using different generative AI tools for the same goal, you may also find tools that are more or less likely to produce biased outputs, which can then inform your own future usage.

How can you mitigate against copyright issues caused by undisclosed datasets?

Asking yourself this critical-creative reflective question can help mitigate against copyright infringement when using popular generative AI tools. While this will not always be possible, users can think about potential solutions to this problem that involve only referencing styles or aesthetics that sit within the legal public domain during the prompting stage.

The legal public domain means that there is no copyright on a work or it is out of copyright duration, such as the text of The Great Gatsby (F. Scott Fitzgerald, 1925). The legal public domain differs from the social concept of the 'public domain' that people often refer to, which means the shared works, ideas and conversations of popular culture. Something may be in the social 'public domain', such as a song by The Beatles that millions of people enjoy, but it can still be in copyright duration, and therefore not in the legal 'public domain'.

How might you offset any sustainability issues caused by your use of generative AI?

As previously mentioned, the carbon footprint of generative AI tools is difficult to measure at present due to limited research. Yet, depending on which metrics are used, experts estimate that a single text-based query of ChatGPT uses around 4.32g of CO_2 emissions while averaging '50 million unique visits per day'.[158] As Vinnie Wong notes, 'if each visit resulted in 10 queries on average, you'd have 15 trillion queries each month'; for comparison, 139 queries is the equivalent to one full load of laundry in a washing machine.[159]

Equally, generating 1000 images using an image-generator such as Stable Diffusion is equivalent to driving roughly 4.1 miles in an average gas-powered car and uses the same amount of energy as charging your smartphone by 16%.[160]

Amidst the need for further research into and legal regulation of the sustainability of generative AI tools, users should think about what steps they can make, as media producers, to reduce the environmental impact of their AI use. This might include carbon offsetting schemes, based on the above estimations, or limiting the number of prompts by thinking carefully about your prompt in detail before you submit it, considering that every new prompt results in a new computation and, therefore, more energy and carbon is used.

6.2.3: After Using AI

How can you make your use of generative AI transparent to your audience?

This question pertains to the principles of accountability and transparency as set out in section 6.1 of this chapter. Our research suggests that obscuring the use of generative AI in media production can lead to feelings of being cheated by some contemporary audiences. Being transparent about the use of GAITs, for example by watermarking AI-generated content, will help media practitioners build trust with audiences.

How can you ensure a fair and responsible use of your outputs?

While you may have adhered to the principles of responsible AI as set out in this report, you may not always be able to guarantee the responsible future use of your AI-generated outputs. Thinking about how you can ensure a fair and responsible use of your outputs will help media practitioners respect both the copyright of any used materials and the copyright of your own end output.

If you are wanting to encourage a sequential use of your own AI-generated outputs by other creatives collaborating with generative AI tools, you might want to consider publishing a statement on how other creatives can do this responsibly. This keeps the cycle of responsibility alive within a landscape of media production designed around responding to other creatives' materials.

What might you do differently next time you use generative AI for media production?

Consider how you might change your processes for future generative AI use.

---

[158] Vinnie Wong, 'Gen AI's Environmental Ledger: A Closer Look at the Carbon Footprint of ChatGPT', Piktochart, 3 May 2024 <https://piktochart.com/blog/carbon-footprint-of-chatgpt/> [Accessed 18 September 2024].

[159] Ibid.

[160] Luccioni, Jernite and Strubell, 'Power Hungry Processing', p. 5.

[161] Shannon Vallor, 'Edinburgh Declaration on Responsibility for Responsible AI', Medium, 14 July 2023 <https://medium.com/@svallor_10030/edinburgh-declaration-on-responsibility-for-responsible-ai-1a98ed2e328b> [Accessed 9 August 2024].

Responsibility, as Shannon Vallor reminds us, is not a fixed, stable, one-time action to be achieved as a tick-box exercise; rather, responsible AI use constitutes a successive, ongoing set of practices and processes based upon subjective criteria for what constitutes 'responsible' use of AI.[161] This means that the work of the creative is never fully complete, as generative AI use necessitates that we consistently reflect upon our own practices and the practices of others in order to accept responsibility as an evolving duty of care.

6.3: Policy Recommendations for a Responsible AI Media Climate

While our research findings suggest that these recommendations for best practice will help foster a responsible AI climate within the UK creative industries, they need to be accompanied by the development of a regulatory framework for generative AI if we want to bring about lasting societal, cultural and economic change. As such, our research proposes the following 12 policy interventions that we call on the UK government to investigate and implement as a matter of urgency. Amidst these 12 interventions are a combination of legislative change, policy change, industry regulation, media thought leadership, and education and media literacy. These 12 interventions are also designed with a holistic, systems-thinking approach to generative AI in mind, such that each one complements and builds upon the last to chart a regulatory course through what responsible AI use might look like. While many of these policy interventions derived from the conversations we had with our Expert Bridging Group, they have been

further shaped and refined by our holistic findings from our workshop participants. For a list of interventions as proposed by our EBG, see Appendix 4. For a shortened version of our policy recommendations, see Figure 6 (overleaf).

1. Amendments to current UK copyright law, aligning to the relevant provisions of the 2024 EU AI Act's position on AI transparency and labelling (Article 50), which includes the disclosure of any copyrighted training materials used by generative AI developers in the training of new and existing AI models.

Current UK copyright law is ill equipped to deal with the wide-ranging legal implications of generative AI. As such, we propose that UK copyright law be amended in line with the 2024 EU AI Act. In particular, Article 50 of the EU AI Act, concerning 'Transparency Obligations for Providers and Deployers of Certain AI Systems', which will come into effect on 2 August 2026, should be used as a guiding model with which to adapt UK copyright law.

Sub-section 1 of Article 50, Art. 50 (1), legislates that generative AI must be developed in a transparent way unless 'authorised by law to detect, prevent, investigate or prosecute criminal offences'.[162] This means that outputs of generative AI tools must be marked as 'artificially generated or manipulations', although this does not apply to assistive functions of 'standard editing'.[163]

Alongside this, Recital 107 of the EU AI Act states that AI developers must publish 'a

---

[162] The European Parliament and the Council of the European Union, 'Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), Official Journal of the European

Union (June 2024), p. 82. See also <https://artificialintelligenceact.eu/ai-act-explorer/>

[163] Ibid.

Figure 6

Human-Gen AI Media Collaboration Interventions

1 **New UK AI copyright law**

Aligned to EU AI Act on transparency and labelling; and Gen AI tech companies to disclose copyright-protected training materials.

2 **UK legislation for actors and voice actors**

Legal prohibition on storing and using voice actors' voices beyond the scope and length of the agreed project.

3 **Gen AI regulator for the creative and entertainment industries**

Development of sector wide Gen AI Code of Practice; liaise with policymakers on a voluntary licensing model; and respond to new AI.

4 **Council of experts for Gen AI in media**

An expert body that Gen AI tech developers can consult with on usage, accessibility and removing bias.

5 **Breaking down barriers: diversification**

Ensuring diversity of the Gen AI tech developers and academics in hiring decisions, and through access schemes.

6 **Breaking down barriers: ensuring accessibility**

Ensuring disability and accessibility consultants engaged with by Gen AI tech developers.

7 **Education and embedding AI in curriculum**

Universities as leaders in educating filmmakers using these tools; and embedding responsible Gen AI tools education in national school curriculum.

8 **Sector wide ethical standards**

Sector wide ethical standards, distinguished by size of the project. Development of ethical Gen AI that impose higher standards on large film-making / media projects.

9 **Build public confidence and trust in Gen AI film-making**

The Gen AI Regulator to work with the media to encourage positive responses to Gen AI film-making, particularly for widening access and accessibility capabilities.

10 **Sector wide bias and risk matrix of specific Gen AI tools**

Development and continual updating of a risk matrix relating to Gen AI tools for filmmakers. Gen AI developers can choose to opt in, as a form of ethical accreditation.

11 **Crowd sourcing data labelling**

Crowd sourcing of data labelling for LLMs in Gen AI tools, to improve the quality of the meta-data labelling.

12 **Ethical environmental AI use accreditation scheme**

Development of a film/ creative media wide accreditation scheme for ethically and environmentally aware use of Gen AI tools.

sufficiently detailed summary of the content used for training the general-purpose AI model'.[164] It continues to state that, while 'taking into due account the need to protect trade secrets and confidential business information', these summaries must be:

> generally comprehensive in […] scope instead of technically detailed to facilitate parties with legitimate interests, including copyright holders, to exercise and enforce their rights under Union law, for example by listing the main data collections or sets that went into training the model, such as large private or public databases or data archives, and by providing a narrative explanation about other data sources used'.[165]

We call for the adoption of similar amendments to UK copyright so as to protect the rights of human creatives, whose works are being used to train Large Language Models without consent or fair remuneration. These combined amendments would necessitate that content generated or modified using generative AI tools needs to be clearly and transparently labelled as such, and we believe that digital watermarking, in line with current practices established by the Content Authenticity Initiative, would successfully achieve this. Where the European Commission's AI Office will provide a template for AI developers to adhere to these regulations, we want to see a similar guidance and assistance for UK-based generative AI companies to do the same, as we propose that this is a task that could be performed by a UK Generative AI Regulator for the Creative Industries (proposal 3).

2. New UK legislation to protect actors and voice actors against the storage and/or use of their data beyond the agreed scope of a given media project.

Where actors and voice actors are at high risk of having their voices, images and likeness manipulated by generative AI tools against their will, we believe that new UK legislation will mitigate against these practices. To support our principle of 'Informed Participation', we want it to be made unlawful to use an actor or voice actors' data beyond the contractually agreed scope of a given media project, thereby preventing the appropriation of their likeness for further commercial purposes.

Media practitioners and our research team expressed unease at generating outputs that closely resemble a real person's likeness, such that, without knowing the origins of the creation, they could not be sure if the AI had based it on a composite of real people/actors or another source entirely. This legislation would not only prevent misuse of peoples' likeness but also empower end users to be certain that the images and voices they are generating fall within the remit of explicit consent.

Where a media producer wants to utilise generative AI tools to manipulate an actor's image, voice or likeness, we recommend that they need to obtain the explicit expressed permission to do so. We believe that this will help redress the imbalances in power and resources that currently befall UK-based actors and voice actors within film and media production.

3. Set up a UK Generative AI Regulator for the Creative Industries, who would establish a sector-wide Code of Practice based on a collective licensing model to assist in the monitoring, regulation and governance of the UK media industry's use of generative AI.

---

[164] Ibid., p. 28.

[165] Ibid.

In order to build upon the recommendations for best practice outlined in 6.1, we propose the establishment of a UK Generative AI regulatory body, designed specifically for the creative industries, who would ensure the implementation of legislation pertaining to generative AI use within film and media production. This regulatory body would also oversee how AI developers are complying with updated laws and regulations, as well as providing ways for said developers to report on their compliance with laws and regulations.

Part of the task for this regulatory body would then be to establish a sector-wide Code of Practice, based on a collective licensing model, in consultation with AI developers, users and UK-based filmmakers and media practitioners. In the absence of sufficient engagement from certain stakeholders across the generative AI divide, the regulatory body, along with the Code of Practice, would function to strengthen standards within the creative industries and ensure that all parties' voices are taken into consideration when assessing the impact of new legislation. We envision this Code of Practice as a continuation of the work begun within this research project, updated to account for new legislation implemented based on proposals 1 and 2.

4. Establish a UK Generative AI Expert Council formed of interdisciplinary experts in the fields of computer science, education, philosophy, law and media, who can be used for consultation by large technology companies seeking to integrate issues of AI ethics into their development of new and existing AI models.

In order to assist AI developers and large technology companies in integrating issues of ethics into their development practices, we propose setting up a Council of Experts formed across the fields of computer science, education, philosophy and media. We see this as a continuation of our work with our Expert Bridging Group, who were instrumental in both establishing many of the outputs of this report and indicative of the sorts of cross-disciplinary conversations that we feel will lead to lasting, impactful change. In line with prevailing literature on responsible AI, the interdisciplinary focus of this Expert Council is paramount if we are to facilitate open dialogue across AI divides. This Expert Council would meet regularly to discuss and appraise any new innovations and developments that pertain to generative AI use in the UK, and could be used for consultation by developers committed to responsible AI development and practice.

5. The diversification of generative AI developers, particularly during hiring processes for trainers and technicians. Investigate the possibility of incentivising access schemes in education to achieve this.

Alongside biased datasets, a lack of diversity when it comes to training and development processes can lead to biased outcomes, such that labels may be based on the unconscious bias of those doing the labelling or stereotyped outputs may not be corrected. As outlined in Chapter Four, the statistics from the World Economic Forum in 2023 showed that women made up only 22% of AI professionals globally,[166] while McKinsey's 2022 report noted that 29% of organisations reported having zero minority ethnic employees working on AI.[167] Although up-to-

---

[166] Gabriela Ramos, 'Why We Must Act Now to Close the Gender Gap in AI', World Economic Forum (22 August 2022) <https://www.weforum.org/agenda/2022/08/why-we-must-act-now-to-close-the-gender-gap-in-ai/> [Accessed 26 September 2024].

[167] McKinsey, The State of AI in 2022 – and a Half Decade in Review (6 December 2022) <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review> [Accessed 26 September 2024].

date UK-based statistics on AI demographics are currently unknown, according to the US-based careers website Zippia's advice on Artificial Intelligence, only 9% of AI specialists in America are women, while 67% are white.[168] As our research suggests, in order to ensure that culturally accurate, fair and just representations of already marginalised communities exist within a responsible AI landscape, further support is needed to increase diversity at the development and training stages.

As a way to tackle a lack of diversity in both the development and conventional outputs of generative AI tools, we propose investigating the possibility of education access schemes. This might involve the Generative AI Regulator working in collaboration with UK-based AI companies in the running of free summer schools and vacation schemes for people in school, college and university, with a specific focus on locating those from wider communities with an interest in working in tech and encouraging those individuals to apply for roles in related generative AI fields. This may also extend to the establishment of mentoring schemes, in which those involved within diversity, equality and inclusion initiatives work with schools to establish student outreach and engagement of those under-represented within AI development.

6. Call on UK generative AI developers to consult with leading experts in disability and accessibility when developing, modelling and training new and existing generative AI models.

As our workshop participants recognised, through its use in transcription, keyboard

navigation, speech generation and alternative text, generative AI tools have the potential to help people with visual, auditory, motor or cognitive disabilities enter a media industry that has previously denied them equal access and opportunities. Yet, at the same time, our participants noted the inaccessibility of certain generative AI tools, particularly those involved in image generation and editing, which runs counter to their potential to support people with disabilities. The tension gets further muddled by the fact that disabled people are frequently erased from the generic outputs of AI-generated stories, images or videos.

All of this highlights the need for AI developers to consult with leading experts in disability when developing, modeling and training new and existing AI models in order to capitalise fully on their accessible possibilities. We propose the design of further training and education pathways for AI developers in issues of disability and accessibility, as well as integrating a consideration for how generative AI tools may help and hinder people with disabilities into their development practices at every stage in the process. These tasks could be integrated into the remit of the Expert Council (proposal 4), who would then work with UK-based AI developers to help them create AI tools that empower, rather than further limit, people with disabilities.

7. Call on the UK government to champion Higher Education institutes to become leaders in educating filmmakers on how best to use generative AI tools in their craft, which would be aided by embedding experimentation with AI and responsible AI into the national school curriculums.

As institutions comprised of world leading AI researchers working across a range of different disciplines and perspectives, universities and colleges have the opportunity

[168] Zippia, 'Artificial Intelligence Specialist Demographics and Statistics in the US', 24 June 2024 <https://www.zippia.com/artificial-intelligence-specialist-jobs/demographics/> [Accessed 18 September 2024].

to drastically shape how future generations engage with generative AI tools for the benefit of society. In order for this to happen, we need to redress the trend of funding priorities which has seen a structural divestment in resources away from segregated Arts and Humanities courses in favour of STEM, since the cultivation of a responsible AI climate is contingent upon interdisciplinarity and the bridging of AI divides.[169]

As such, we call on the UK government to champion Higher Education institutions to become leaders in educating the filmmakers of tomorrow on how best to use generative AI tools in their own craft. Doing so would prevent future generations of creatives from both rejecting generative AI entirely and using AI tools in ways that perpetuate its problematic impact on creativity, bias and collaboration. Ways that the UK government can help UK universities and colleges to achieve this is through further support and funding to integrate advanced generative AI education into pre-existing media production courses, designed to focus on issues of ethics and responsible use. This would also entail the development of new courses dedicated specifically to media production with generative AI, such as Bournemouth's own MSc in Artificial Intelligence for Media.

In addition, we call on the UK government to embed responsible AI use into the national curriculum at Key Stage 3 and Key Stage 4 of secondary education. Our participants recognised the upmost importance of early intervention when it comes to educating young people about what responsible use of generative AI tools looks like. We believe that this could be achieved through the adoption of responsible AI frameworks,

including teaching young people about the potential harms of generative AI use and the principles of best practice as laid out in this White Paper. This could be undertaken by training educators already involved in Media, Music, Computing (or Information and Communication Technology), and Arts and Design subjects. Bringing responsible AI use into the classroom would prepare future generations and empower young people to make ethical decisions when collaborating with generative AI tools.

8. Develop a scalable, sector-wide ethical standard on the use of generative AI in media production that imposes higher standards on large filmmaking and media projects with higher budgets.

As part of the Generative AI Regulator for the Creative Industries (proposal 3), we recommend the development of a sector-wide standard on the ethical and responsible use of generative AI, including a set of minimum requirements that productions have to meet. Importantly, as our workshop participants and members of our Expert Bridging Group highlighted, this standard would have to be scalable on a case-by-case basis that would impose higher ethical standards on larger filmmaking and media projects with higher budgets to meet said standards. This would need to be designed to not only protect employment rights and opportunities for those working under large media productions, but also enable those involved in smaller, independent productions to collaborate with generative AI tools. Were this standard not scalable, the Regulator would risk perpetuating a hierarchy in which only productions with sufficient financial backing would be able to use generative AI tools, restricting accessibility in the process.

---

[169] While the current Labour government commitments on Higher Education funding are currently unknown at the time of writing this White Paper, these divestments date back as far as 2021. See Heidi Ashton, David Brownlee, Jack Gamble and Melanie Stavrou, 'The State of the Arts', Campaign for the Arts (2024), pp. 90-91.

9. Build public confidence and trust in generative AI filmmaking through thought leadership and media campaigns focused on positive narratives about generative AI.

To combat societal skepticism and distrust regarding the impact that generative AI tools are having in film and media production, we propose to develop thought leadership work to build public confidence and trust in responsible generative AI use. This is not to suggest that this distrust is misplaced; as this White Paper has demonstrated, there are significant limitations and drawbacks to using generative AI in media production that need to be addressed as a matter of urgency. However, narratives that centre around technophobia do not solve these issues but merely perpetuate them. What is needed, we suggest, is the development of nuanced media narratives that frame responsible AI use as a matter of consciously assessing the opportunities and limitations that these new tools offer the creative industries. Media narratives focused on issues of accessibility, technical literacy and case studies of those directly involved in, and affected by, the key issues, will go some way to achieving this. This task could also be spearheaded by the Generative AI Regulator for the Creative Industries as a means of collaborating with industry professionals and public service media to ensure fair and just reporting on generative AI developments.

10. Develop a sector-wide 'Bias and Risk Matrix' of different generative AI tools that can be updated continuously to enable UK filmmakers and media practitioners to make informed decisions about their own ethical use of generative AI.

In order to empower users to make informed decisions about the right tool for them, we propose the development of a sector-wide 'Bias and Risk Matrix' that details the

functionalities, transparency and ethical concerns for new and upcoming generative AI tools. We envision this as an extension of the research undertaken in part by this project, as Appendix 2 details a prototype of this matrix designed to help mitigate potentially harmful uses of generative AI. Where we propose an Expert Council of interdisciplinary academics and practitioners to audit the UK AI landscape, this matrix could be an extension of their remit and a way to document their discussions in an accessible and transparent way.

11. Crowd source data labelling for Large Language Models to improve the quality of meta-data.

Where generative AI models provide biased and incorrect information, this is largely a result of the training processes in which these models learn from large datasets. As members of our Expert Bridging Group highlighted, one way this can be tackled is by increasing the accuracy of data labelling for Large Language Models through crowd sourcing labelling. In order to improve the quality of meta-data, we propose the development of a 'citizen science' project as a way of developing accurate datasets through the engagement of the general public. Importantly, any large crowd sourcing or citizen science project needs to adhere to the principle of informed participation and consent, with participants being made explicitly aware of what they are doing and why they are doing it, in order to avoid the unknowing, unpaid contribution of technoscientific labour to large technology companies.[170] This would also need to be in conjunction with tackling the present unethical practices of labelling used large technology companies, in which the labelling of harmful,

---

[170] James Riley and Will Mason-Wilkes, 'Dark Citizen Science', Public Understanding of Science, 33: 2 (2023), 142-157 (p. 142); Daphne Ezer, 'Crowdsourced and Citizen Science', The Alan Turing Institute (2018) <https://www.turing.ac.uk/research/research-projects/crowdsourced-and-citizen-science> [Accessed 20 September 2024].

often traumatic material is being outsourced to the Global South for work in unfair and underpaid conditions.

12. Establish an Ethical AI Accreditation Scheme, in line with BAFTA Albert, to assess the ethical and sustainable use of generative AI in UK-based film and media productions.

Following the lead of BAFTA Albert, who support the UK film and television industry to reduce the environmental impacts of production and certify productions according to a 1-3 star system of sustainability, we propose the establishment of an Ethical AI Accreditation Scheme. This scheme would independently evaluate film and media productions based on their use of generative AI tools. It would work in tandem with the sector-wide standards, as set out in proposal 8, but would work to accredit those productions who go above and beyond the minimum standards and reward outstanding ethical AI use. Productions who are labelled 'Ethical AI Accredited' could then use this on media and marketing materials as an indication of their commitment to ethical and responsible generative AI use. The Ethical AI Accreditation Scheme would also publish a toolkit to help productions file for accreditation in line with established principles of ethical AI.

6.4: Conclusion

The main objective of this work has been to address the limitations and opportunities of using generative AI tools in media creation, determining the needs of different stakeholders so that we can gain insight into what interventions are required to foster a responsible AI media landscape. Our findings suggest that generative AI tools offer exciting possibilities to the creative industries, but concerns about creativity, bias and collaboration are threatening to destabilise UK media production. Creating a responsible AI landscape requires the informed participation of everyone involved. This means fostering interdisciplinarity and forging working relationships with different stakeholders across AI divides. End users can work towards this societal goal by adhering to, and building upon, the recommendations for best practice principles and critical-creative reflective questions outlined in this chapter. Yet, to complement responsible AI working practice, clear guidance and regulation is required from the UK government designed in ways to empower end users and facilitate their ethical, legal use of generative AI.

If these needs are not addressed, we risk further destabilising the creative industries and their social, cultural and economic contributions to the UK more broadly. Although implementing the recommended strategies outlined in this report will not eradicate the issues that generative AI tools are having on UK-based media production overnight, given the dynamic nature of AI as an ever-changing technological landscape, it is the contention of this research team that they will go some way to redressing the current knowledge, access and power imbalances that exist within the creative industries and start the process of facilitating a truly equitable and responsible use of generative AI in media.

References

Abascal, Julio, Myriam Arrue, Nestor Garay and Jorge Tomás, 'A Web Service for Automatic Accessibility Analysis of Web Pages Based on the Use of XML Structures' (2003), 925-929.

Acciarini, Chiara, Federica Brunetta and Paolo Boccardelli, 'Cognitive Biases and Decision-Making Strategies in Times of Change: A Systematic Literature Review', Management Decision, 59: 3 (2021), 638-652.

Ada Lovelace Institute, 'AI Assurance? Assessing and Mitigating Risks Across the AI Lifecycle' (July 2023) <https://www.adalovelaceinstitute.org/report/risks-ai-systems/> [Accessed 29 August 2024].

Ada Lovelace Institute, 'Keeping an Eye on AI: Approaches to Government Monitoring of the AI Landscape' (July 2023) <https://www.adalovelaceinstitute.org/report/keeping-an-eye-on-ai/> [Accessed 29 August 2024].

Adiyansjah, Alexander Gunawan and Derwin Suhartono, 'Music Recommender System Based on Genre Using Convolutional Recurrent Neural Networks', Procedia Computer Science, 157 (2019), 99-109.

Adobe, 'Text-Based Editing in Premiere Pro' <https://helpx.adobe.com/premiere-pro/using/text-based-editing.html> [Accessed 9 September 2024].

Adobe, 'The Power of AI Image Style Transfer' <https://www.adobe.com/products/firefly/features/generative-match.html> [Accessed 9 September 2024].

Agora, 'What Does Noise Reduction Do?' <https://www.agora.io/en/blog/what-does-noise-reduction-do/> [Accessed 9 September 2024].

Ahrefs, 'Free AI Grammar Checker' <https://ahrefs.com/writing-tools/grammar-checker> [Accessed 9 September 2024].

AI Maverick, 'A Comprehensive Guide to Latent Space', Medium (December 2023) <https://samanemami.medium.com/a-comprehensive-guide-to-latent-space-9ae7f72bdb2f> [Accessed 9 September 2024].

Alienikov, Andrei, Sharon Kackmeister and Ron Koenig, Creating Creativity: 101 Definitions (what Webster Never Told You) (Midland: Alden B. Dow Creativity Centre, 2000).

Amaro, Roman, The Black Technical Object: On Machine Learning and the Aspiration of Black Being (London: Sternberg Press, 2023).

Amazon Web Services, 'What is a Recurrent Neural Network?' <https://aws.amazon.com/what-is/recurrent-neural-network/> [Accessed 9 September 2024].

Ameen, Nisreen, Gagan Sharma, Shlomo Tarba, Amar Rao and Ritika Chopra, 'Toward Advancing Theory on Creativity in Marketing and Artificial Intelligence', Psychology & Marketing, 39: 9 (2022), 1802-1825.

Angelica, Amara, 'Transforming the Future of Music Creation', Mindplex (November 2023) <https://magazine.mindplex.ai/mp_news/transforming-the-future-of-music-creation/?> [Accessed 2 September 2024].

Anishnama, 'Understanding Gated Recurrent Unit (GRU) in Deep Learning', Medium (May 2023) <https://medium.com/@anishnama20/understanding-gated-recurrent-unit-gru-in-deep-learning-2e54923f3e2> [Accessed 9 September 2024].

Arnold, Richard, 'Performers' Rights and Artificial Intelligence', in Research Handbook on Intellectual

Property and Artificial Intelligence, edited by Ryan Abbott and David Geffen, 218-224 (Cheltenham: Edward Elgar Publishing, 2022).

Ashok, Mona, Rohit Madan, Anton Joha and Uthayasankar Sivarajah, 'Ethical Framework for Artificial Intelligence and Digital Technologies', International Journal of Information Management, 62 (2022), 1-17.

Ashton, Heidi, David Brownlee, Jack Gamble and Melanie Stavrou, 'The State of the Arts', Campaign for the Arts (2024).

Barredo Arrieta, Alejandro, et al., 'Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI', Information Fusion, 58 (2020), 82-115.

Barron, Frank, 'The Disposition Toward Originality', The Journal of Abnormal and Social Psychology, 51: 3 (1955), 478-485.

Batlle-Roca, Roser, Emila Gómez, Wei-Hsang Liao, Xavier Serra and Yuki Mitsufuji, 'Transparency in Music-Generative AI: A Systematic Literature Review' (2023).

BBC Newsround, 'Deepfake Technology: What is it, how does it work, and what it be used for?' <https://www.bbc.co.uk/newsround/69009887> [Accessed 9 September 2024].

Benjamin, Ruha, Race After Technology: Abolitionist Tools for the New Jim Code (Cambridge: Polity Press, 2019).

Bentley, Caitlin, et al., 'A Framework for Responsible AI Education: A Working Paper', SSRN Working Paper Series (2023).

Boden, Margaret, The Creative Mind: Myths and Mechanisms 2nd edn. (London: Routledge, 2004).

Bonnemains, Vincent, Claire Saurel, and Catherine Tessier, 'Embedded ethics: some technical and ethical challenges', Ethics and Information Technology, 20: 1 (2018), 41–58 (p. 41).

Braidotti, Rosi, The Posthuman (Cambridge: Polity Press, 2013).

Braun, Virginia and Victoria Clarke, 'Using Thematic Analysis in Psychology', Qualitative Research in Psychology, 3: 2 (2006), 77-101.

Brown, Tom et al., 'Language Models are Few-Shot Learners' (2020).

Bryson, Joanna, 'Patiency is not a virtue: the design of intelligent systems and systems of ethics', Ethics and Information Technology, 20: 1 (2018), 15–26.

Burgstahler, Sheryl, 'Universal Design: Process, Principles, and Applications', DO-IT (Disabilities, Opportunities, Internetworking and Technology) (2009).

Butollo, Florian, and Sabine Nuss, eds., Marx and the Robots: Networked Production, AI, and Human Labour, trans. by Jan-Peter Herrmann and Nivene Raafat (London: Pluto Press, 2022).

Cambridge Dictionary, 'Creativity' <https://dictionary.cambridge.org/dictionary/english/creativity> [Accessed 23 September 2024].

Chan, Thomas, Jason Ho and Michael Tom, 'Miro: Promoting Collaboration Through Online Whiteboard Interaction', RELC Journal (2023).

Cloudflare, 'What is Low-Rank Adaptation (LoRA)?' <https://www.cloudflare.com/en-gb/learning/ai/what-is-lora/> [Accessed 7 October 2024].

Coeckelbergh, Mark, 'The Work of Art in the Age of AI Image Creation', Journal of Human-Technology Relations, 1: 1 (2023), 1-13.

Content Authenticity Initiative, <https://contentauthenticity.org/> [Accessed 19 August 2024].

Cowen, Tyler, 'Rick Rubin on Listening, Taste, and the Act of Noticing (Ep. 169)', Conversations with Tyler (18 January 2023) <https://conversationswithtyler.com/episodes/rick-rubin/#:~:text=RUBIN%3A%20I%20think%20they're,isn't%20making%20the%20thing> [Accessed 1 October 2024].

Creative Industries Council, 'UK Creative Industries Add £124bn to UK Economy', 18 June 2024 <https://www.thecreativeindustries.co.uk/site-content/uk-creative-industries-add-ps124bn-to-uk-economy> [Accessed 15 August 2024].

Csikszentmihalyi, Mihaly, Creativity: Flow and the Psychology of Discovery and Invention (New York: HarperCollins, 1997).

Danesi, Marcel, AI-Generated Popular Culture: A Semiotic Perspective (Cham: Palgrave Macmillan, 2024).

Davenport, Thomas, Abhijit Guha, Dhruv Grewal and Timna Bressgott, 'How Artificial Intelligence Will Change the Future of Marketing', Journal of the Academy of

Marketing Science, 48 (2020), 24-42.

Davis, Tom, 'The Feral Cello: A Philosophically Informed Approach to an Actuated Instrument', NIME'17: Proceedings of the International Conference on New Interfaces for Musical Expression (2017), 279-282.

Dennett, Daniel, 'Intentional Systems', The Journal of Philosophy, 68: 4 (1971), 87-106.

Deruty, Emmanuel, Marteen Grachten, Stefan Lattner, Javier Nistal and Cyran Aouameur, 'On the Development and Practice of AI Technology for Contemporary Popular Music Production', Transactions of the International Society for Music Information, 5: 1 (2022), 35-49.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee and Kristina Toutanova, 'BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding' (2018).

Devlin, Kate, 'Power in AI: Inequality Within and Without the Algorithm', in The Handbook of Gender, Communication, and Women's Human Rights, ed. by Margaret Gallagher and Aimée Vega Montiel, 123-139 (London: Wiley Blackwell, 2023).

Devlin, Kate, 'OpenAI's Scarlett Johansson Update Wasn't About Bridging the Gap Between Tech and Creatives – It Was Just Sexist', Byline Times, 25 May 2024 <https://bylinetimes.com/2024/05/25/openais-scarlett-johansson-update-wasnt-about-bridging-the-gap-between-tech-and-creatives-it-was-just-sexist/> [Accessed 15 August 2024].

Dignum, Virginia, Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way (Cham: Springer, 2019).

Dufour, Pascal, 'Complete LLM Glossary', Step Up AI (November 2023) <https://stepup.ai/complete-llm-glossary/> [Accessed 9 September 2024].

Dutta, Ishan, Enrique Elias Ubaldo and Oğuz Akif, 'Text-to-Image', Hugging Face <https://huggingface.co/tasks/text-to-image> [Accessed 9 September 2024].

Elman, Jeffrey, 'Finding Structure in Time', Cognitive Science, 14: 2 (1990), 179-211.

Encord, 'Datasets' <https://encord.com/glossary/datasets-definition/> [Accessed 9 September 2024].

Enticknap, Anthony, 'What is AI Translation and How to Use It?', Centus (December 2023) <https://centus.com/blog/what-is-ai-translation> [Accessed 9 September 2024].

Erickson, Kristofer, 'Copyright Protection in AI-Generated Works', Creative Industries Policy and Evidence Centre (January 2024) <https://pec.ac.uk/blog_entries/copyright-protection-in-ai-generated-works/> [Accessed 29 August 2024].

Ezer, Daphne, 'Crowdsourced and Citizen Science', The Alan Turing Institute (2018) <https://www.turing.ac.uk/research/research-projects/crowdsourced-and-citizen-science> [Accessed 20 September 2024].

Ferrando, Francesca, Philosophical Posthumanism (London: Bloomsbury, 2019).

Flick, Catherine, and Kyle Worrall, 'The Ethics of Creative AI', in The Language of Creative AI: Practices, Aesthetics and Structures, ed. by Craig Vear and Fabrizio Poltronieri, 73-94 (Cham: Springer, 2022).

Floridi, Luciano, The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities (Oxford: Oxford University Press, 2023).

Foster, Kelsey, 'What is Automatic Speech Recognition? A Comprehensive Overview of ASR Technology', Assembly AI (September 2023) <https://www.assemblyai.com/blog/what-is-asr/> [Accessed 9 September 2024].

Franceschelli, Giorgio and Mirco Musolesi, 'Creativity and Machine Learning: A Survey', ACM Computing Surveys, 56: 11 (2024), 1-41.

Frich, Jonas, Michael Biskjaer and Peter Dalsgaard, 'Twenty Years of Creativity Research in Human-Computer Interaction: Current State and Future Directions', Proceedings of the 2018 Designing Interactive Systems Conference (2018), 1235-1257.

Gafton, Paul, and Osman Alenbey, 'Image-to-Image', Hugging Face <https://huggingface.co/tasks/image-to-image> [Accessed 9 September 2024].

Ganesh, Bhargavi, 'Policy Approaches for Building a Responsible Ecosystem: Contextualising AI Governance Challenges Within Other Regulatory/Governance Sectors and Histories', BRAID (July 2024).

Garon, Jon, 'A Practical Introduction to Generative AI, Synthetic Media, and the Messages Found in the Latest Medium', SSRN (2023).

Gawęda, Łukasz and Martyna Krężołek, 'Cognitive Mechanisms for Alexithymia in Schizophrenia: Investigating the Role of Basic Neurocognitive

Functioning and Cognitive Biases', Psychiatry Research, 271 (2019), 573-580.

Germanidis, Anastasis, 'Gen-1: The Next Step Forward for Generative AI', Runway (February 2023) <https://runwayml.com/research/gen-1> [Accessed 2 September 2024].

Germanidis, Anastasis, 'Gen-2: Generate Novel Videos with Text, Images or Video Clips', Runway (February 2023) <https://runwayml.com/research/gen-2> [Accessed 2 September 2024].

Gilliard, Chris, 'The Deeper Problem with Google's Racially Diverse Nazis', The Atlantic (26 February 2024) <https://www.theatlantic.com/technology/archive/2024/02/google-gemini-diverse-nazis/677575/> [Accessed 25 September 2024].

Given, Lisa, The Sage Encyclopedia of Qualitative Research Methods (London: SAGE, 2008).

Goodfellow, Ian et al., 'General Adversarial Nets' (2014).

Google Cloud, 'What is Artificial Intelligence (AI)?' <https://cloud.google.com/learn/what-is-artificial-intelligence?hl=en> [Accessed 9 September 2024].

Google for Developers, 'Overview of GAN Structure' <https://developers.google.com/machine-learning/gan/gan_structure> [Accessed 9 September 2024].

Gregory, Karen and Cailean Gallagher, 'Mitigating Harms in On-Demand Delivery Platforms: AI Regulations, Data Protection, and Workers' Tools', BRAID (July 2024).

Gretel, 'Conditional Generation FAQ', <https://docs.gretel.ai/create-synthetic-data/models/synthetics/conditional-generation-faq> [Accessed 9 September 2024].

Haager, Julia, Christof Kuhbandner and Reinhard Pekrun, 'To Be Bored or Not To Be Bored – How Tas-Related Boredom Influences Creative Performance', The Journal of Creative Behaviour, 52: 4 (2018), 297-304.

Hamilton, David, 'Stereotyping and Intergroup Behaviour: Some Thoughts on the Cognitive Approach', in Cognitive Processes in Stereotyping and Intergroup Behaviour, ed. by David Hamilton, 333-353 (Hove: Psychology Press, 2015).

Haraway, Donna, Simians, Cyborgs, and Women: The Reinvention of Nature (New York: Routledge, 1991).

Hayles, N. Katherine, How We Became Posthuman

(Chicago: University of Chicago Press, 1999).

Haywood, Sloan, Genevieve Warren and Alex Wolf, 'Understand Tokens', Learn Microsoft (May 2024) <https://learn.microsoft.com/en-us/dotnet/ai/conceptual/understanding-tokens> [Accessed 9 September 2024].

Hern, Alex, 'Google's Solution to Accidental Algorithmic Racism: Ban Gorillas', The Guardian, 12 January 2018 < https://www.theguardian.com/technology/2018/jan/12/google-racism-ban-gorilla-black-people> [Accessed 19 August 2024].

Hibberd, James, 'A24's New AI-Generated Civil War Ads Generate Controversy', The Hollywood Reporter, 17 April 2024 <https://www.hollywoodreporter.com/movies/movie-news/a24-civil-war-posters-controversy-1235876340/> [Accessed 15 August 2024].

Hill, Joshua, 'Revolutionise Your Projects with AI Colour Grading', Hit Paw <https://www.hitpaw.com/video-tips/ai-color-grading.html> [Accessed 9 September 2024].

Hinton, Perry, Stereotypes, Cognition and Culture (Hove: Psychology Press, 2013).

Ho, Jonathan, Ajay Jain and Pieter Abbeel, 'Denoising Diffusion Probabilistic Models', Advances in Neural Information Processing Systems, 33 (2020), 6840-6851.

Ho, Jonathan et al., 'Imagen Video: High Definition Video Generation with Diffusion Models' (2022).

Hochreiter, Sepp, and Jürgen Schmidhuber, 'Long Short-Term Memory', Neral Computation, 9: 8 (1997), 1735-1780.

Hugging Face, 'What is Image Segmentation?' <https://huggingface.co/tasks/image-segmentation> [Accessed 9 September 2024].

Hugging Face, 'Inpainting' <https://huggingface.co/docs/diffusers/using-diffusers/inpaint> [Accessed 9 September 2024].

IBM, 'What is Machine Learning?' <https://www.ibm.com/topics/machine-learning> [Accessed 9 September 2024].

Institute of Electrical and Electronics Engineers, 'A Glossary for Discussion of Ethics of Autonomous and Intelligent Systems, Version 1' (2017).

Jam Canda, 'AI in Music: Composition, Production, and Recommendation', Medium (April 2024) <https://

medium.com/@jam.canda/ai-in-music-composition-production-and-recommendation-4b5bbde1b10b> [Accessed 9 September 2024].

Jiang, Harry, et al., 'AI Art and its Impact on Artists'. AIES '23 August 08-10 (2023). 363-374.

Jordan, Michael, 'Serial Order: A Parallel Distributed Processing Approach', Approaches in Psychology, 121 (1997), 471-495.

Jones, Bronwyn, Rhianne Jones and Ewa Luger, 'Generative AI and Journalism: A Rapid Risk-Based Review' (June 2023) <https://www.research.ed.ac.uk/en/publications/generative-ai-amp-journalism-a-rapid-risk-based-review> [Accessed 29 August 2024].

Kaheman, Daniel, and Shane Frederick, 'A Model of Heuristic Judgement' in The Cambridge Handbook of Thinking and Reasoning, ed. by Keigh Holyoak and Robert Morrison, 267-293 (Oxford: Oxford University Press, 2012).

Kalpokas, Ignas, 'Work of art in the Age of its AI Reproduction', Philosophy and Social Criticism (2023), 1-19.

Kapur, Anandana, and Nagma Ansari, 'Coding Reality: Implications of AI for Documentary Media', Studies in Documentary Film, 16: 2 (2022), 174-185.

Kasirzadeh, Atoosa, Charlotte Bird and Eddie Ungless, 'Policy Report on Generative Artificial Intelligence', BRAID (July 2024).

Kaufman, James, Creativity 101 2nd edn. (New York: Springer, 2016).

Kelly, James, 'London Cinema Drops AI-Written Film After Backlash', The BBC, 19 June 2024 <https://www.bbc.co.uk/news/articles/cjll3w15j0yo> [Accessed 12 September 2024].

Kim, JJ, 'What is Denoising', NVIDIA (November 2022) <https://blogs.nvidia.com/blog/what-is-denoising/> [Accessed 9 September 2024].

Kingma, Diederik and Max Welling, 'Auto-Encoding Variational Bayes' (2013).

Kozbelt, Aaron, 'Originality and Technical Skill as Components of Artistic Quality', Empirical Studies of the Arts, 22: 2 (2004), 157-170.

Kruger, Justin, Derrick Wirtz and Leaf Van Boven, 'The Effort Heuristic', Journal of Experimental Social Psychology, 40: 1 (2004), 91-98.

Lan, Zhenzhong et al., 'ALBERT: A Lite BERT for Self-Supervised Learning of Language Representations', ICLR 2020 Conference (2019).

Le, Quoc et al., 'Building High-Level Features Using Large Scale Unsupervised Learning', Proceedings of the 29th International Conference on Machine Learning, Scotland (2012).

Leaver, Tama, and Suzanne Srdarov, 'ChatGPT Isn't Magic: The Hype and Hypocrisy of Generative Artificial Intelligence Rhetoric', M/C Journal, 26: 5 (2023).

Lee, Newton, ed., The Transhumanism Handbook (Cham: Springer, 2019).

Lees, Dominic, 'Deepfakes in Documentary Film Production: Images of Deception in the Representation of the Real', Studies in Documentary Film, 18: 2 (2024), 108-129.

Levitt, Deborah, The Animatic Apparatus: Animation, Vitality, and the Futures of the Image (Alresford: Zero Books, 2017).

Lima, Yuri, Carlos Babosa, Herbet dos Santos and Jano Moreira de Souza, 'Understanding Technological Unemployment: A Review of Causes, Consequences, and Solutions', Societies, 11: 2 (2021).

Lin, Patrick, Ryan Jenkins and Keith Abney, eds., Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence (Oxford: Oxford University Press, 2017).

Liu, Yinhan et al., 'A Robustly Optimised BERT Pretraining Approach' (2019).

Luccioni, Alexandra, Yacine Jernite and Emma Strubell, 'Power Hungry Processing: Watts Driving the Cost of AI Deployment?', ACM FAccT '24, Brazil (2024), 1-15.

MacCormack, Patricia, Posthuman Ethics: Embodiment and Cultural Theory (Farnham: Ashgate, 2012).

Manovich, Lev, The Language of New Media (London: MIT Press, 2001).

Manovich, Lev, and Emanuele Arielli, Artificial Aesthetics: Generative AI, Art and Visual Media (2021-2024) <https://manovich.net/index.php/projects/artificial-aesthetics> [Accessed 16 August 2024].

Marcus, Gary, 'AI Platforms Like ChatGPT Are Easy to Use but Also Potentially Dangerous', Scientific American (19 December 2022) <https://www.scientificamerican.com/

article/ai-platforms-like-chatgpt-are-easy-to-use-but-also-potentially-dangerous/> [Accessed 27 September 2024].

Math Works, 'Time Series and Sequence Data Networks' <https://uk.mathworks.com/help/deep-learning-hdl/time-series-and-sequence-data-networks.html> [Accessed 9 September 2024].

Math Works, 'What is a Convolutional Neural Network?', <https://uk.mathworks.com/discovery/convolutional-neural-network.html> [Accessed 9 September 2024].

McKinsey, The State of AI in 2022 – and a Half Decade in Review (6 December 2022) <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review> [Accessed 26 September 2024].

McCormack, John et al., 'Design Considerations for Real-Time Collaboration with Creative Artificial Intelligence', Organised Sound, 25: 1 (2020), 41-52.

Merriam-Webster, 'Creativity' (2024) <https://www.merriam-webster.com/dictionary/creativity> [Accessed 23 September 2024].

Merritt, Rick, 'What is a Transformer Model?', NVIDIA <https://blogs.nvidia.com/blog/what-is-a-transformer-model/> [Accessed 9 September 2024].

Meta, 'Label AI-Generated Content on Instagram', Instagram Help Centre (2024) <https://help.instagram.com/761121959519495> [Accessed 23 September 2024].

Meta, 'Musicgen: Advanced AI Music Generation' (2023) <https://musicgen.com/> [Accessed 2 September 2024].

Mikolov, Tomas, Kai Chen, Greg Corrado and Jeffrey Dean, 'Efficient Estimation of Word Representations in Vector Space' (2013).

Millet, Kobe, Florian Buehler, Guanzhong Du and Michael Kokkoris, 'Defending humankind: Anthropocentric bias in the appreciation of AI art', Computers in Human Behaviour, 143 (2023).

Mirowski, Piotr, Kory Mathewson, Jaylen Pittman, and Richard Evans, 'Co-Writing Screenplays and Theatre Scripts with Language Models: Evaluation by Industry Professionals', Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, April 23-28, 2023, Hamburg, Germany (2023).

Montouri, Alfonso, 'Beyond Postnormal Times: The Future of Creativity and the Creativity of the Future', Futures, 43: 2 (2011), 221-227.

Muir, Ellie, 'Netflix embroiled in alleged AI scandal with What Jennifer Did', The Independent, 19 April 2024 <https://www.independent.co.uk/arts-entertainment/tv/news/netflix-what-jennifer-did-ai-image-true-story-b2531363.html> [Accessed 15 August 2024].

Murray, Michael, 'Generative AI Art: Copyright Infringement and Fair Use', SMU Science and Technology Law Review, 26: 2 (2023), 2023, 259-315.

Nandan Upadhyay, Shyam, 'Google DeepMind Launches Lyria, Transforming the Future of Music with AI', Analytics India (November 2023) <https://analyticsindiamag.com/ai-news-updates/google-deepmind-launches-lyria-transforming-the-future-of-music-with-ai/> [Accessed 2 September 2024].

Navabi, Eshan, Katherine Daniell, Elizabeth Williams and Caitlin Bentley, 'AI for Sustainability: A Changing Landscape', in Artificial Intelligence: For Better or Worse, Future Leaders, 157-176 (2019).

Noble, Safiya, Algorithms of Oppression: How Search Engines Reinforce Racism (New York: New York University Press, 2018).

NVIDIA, 'Text-to-Speech' <https://www.nvidia.com/en-us/glossary/text-to-speech/> [Accessed 9 September 2024].

Ørngreen, Rikke and Karin Levinson, 'Workshops as a Research Methodology', The Electronic Journal of e-Learning, 15: 1 (2017), 70-81.

Ozimek, Anna, 'Equality, Diversity and Inclusion in the Screen Industries'. Research Report, The University of York (2020) < https://screen-network.org.uk/wp-content/uploads/2021/02/Equality-Diversity-and-Inclusion-in-the-Screen-Industries.pdf> [Accessed 16 August 2024].

Papers With Code, 'Audio Source Separation' <https://paperswithcode.com/task/audio-source-separation> [Accessed 9 September 2024].

Papers With Code, 'Dialogue Generation' <https://paperswithcode.com/task/dialogue-generation> [Accessed 9 September 2024].

Papers With Code, 'Speech Emotion Recognition' <https://paperswithcode.com/task/speech-emotion-recognition> [Accessed 9 September 2024].

Partnership on AI, 'PAI's Responsible Practices for Synthetic Media: A Framework for Collective Action' (2023).

Pennington, Jeffrey, Richard Socher and Christopher Manning, 'GloVe: Global Vectors for Word Representation', Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Qatar (2014) 1532-1543.

Perrigo, Billy, 'OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic', TIME (18 January 2023) <https://time.com/6247678/openai-chatgpt-kenya-workers/> [Accessed 25 September 2024].

Petrie, Helen, Andreas Savva and Christopher Power, 'Towards a Unified Definition of Web Accessibility', W4A '15: Proceedings of the 12th International Web for All Conference (2015), 1-13.

Pfeiffer, Andreas, Creativity and Technology in the Age of AI Pfeiffer Report (2018).

Prince, Stephen, Digital Visual Effects in Cinema: The Seduction of Reality (New Brunswick: Rutgers University Press, 2012).

Purse, Lisa, Digital Imaging in Popular Cinema (Edinburgh: Edinburgh University Press, 2013).

Ramos, Gabriela, 'Why We Must Act Now to Close the Gender Gap in AI', World Economic Forum (22 August 2022) <https://www.weforum.org/agenda/2022/08/why-we-must-act-now-to-close-the-gender-gap-in-ai/> [Accessed 26 September 2024].

Rawlinson, Kevin, 'Digital Assistants Like Siri and Alexa Entrench Gender Biases, Says UN', The Guardian, 22 May 2019 < https://www.theguardian.com/technology/2019/may/22/digital-voice-assistants-siri-alexa-gender-biases-unesco-says> [Accessed 19 August 2024].

Responsible AI UK, 'Responsible AI Governance: A Response to UN Interim Report on Governing AI for Humanity', submitted in Open Consultation from the UN AI Advisory Body' (2024).

Riley, James, and Will Mason-Wilkes, 'Dark Citizen Science', Public Understanding of Science, 33: 2 (2023), 142-157.

Runco, Mark, and Garrett Jaeger, 'The Standard Definition of Creativity', Creativity Research Journal, 24: 1 (2012), 92-96.

Samuelson, Pamela, 'Legal Challenges to Generative AI, Part I', Communications of the ACM, 66: 7 (2023), 20-23.

Sardar, Ziauddin, 'Welcome to Postnormal Times', Futures, 42: 5 (2010), 435-444.

Sarrion, Eric, Exploring the Power of ChatGPT: Applications, Techniques and Implications (Apress: New York, 2023).

Sawyer, R. Keith, and Danah Henriskon, Explaining Creativity: The Science of Human Innovation 3rd edn. (Oxford: Oxford University Press, 2024).

Schiller, Amy, and John McMahon, 'Alexa, Alert Me When the Revolution Comes: Gender, Affect, and Labour in the Age of Home-Based Artificial Intelligence', New Political Science, 41: 2 (2019), 173-191.

Shams, Rifat, Didar Zowghi and Muneera Bano, 'AI and the Quest for Diversity and Inclusion: A Systematic Literature Review', AI and Ethics (2023), 1-28.

Sharma, Aditya, '11 Best Generative AI Tools and Platforms', Turing <https://www.turing.com/resources/generative-ai-tools> [Accessed 9 September 2024].

Singer, Uriel et al., 'Make-a-Video: Text-to-Video Generation Without Text-Video Data' (2022).

Smart, Alan, and Josephine Smart, Posthumanism: Anthropological Insights (Toronto: University of Toronto Press, 2017).

Smith, Steven, and Thomas Ward, 'Cognition and the Creation of Ideas', in The Oxford Handbook of Thinking and Reasoning, ed. by Keith Holyoak and Robert Morrison, 456-474 (Oxford: Oxford University Press, 2012).

Stangor, Charles and Mark Schaller, 'Stereotypes as Individual and Collective Representations', in Stereotypes and Prejudice: Essential Readings, ed. by Charles Stangor, 64-82 (Hove: Psychology Press, 2000).

Super Annotate, 'Introduction to Diffusion Models for Machine Learning' <https://www.superannotate.com/blog/diffusion-models> [Accessed 9 September 2024].

Synthesia, 'AI Script Generator' <https://www.synthesia.io/features/ai-script-generator> [Accessed 9 September 2024].

Tassa, Yuval, Yotam Doron, Alistair Muldal, Tom Erez, Yazhe Li, Diego de Las Casas, David Budden, Abbas

Abdolmaleki, Josh Merel, Andrew Lefrancq, Timothy P. Lillicrap and Martin A. Riedmiller. 'DeepMind Control Suite.' ArXiv abs/1801.00690 (2018), 1-24.

Tensor Flow, 'Intro to Autoencoders' <https://www.tensorflow.org/tutorials/generative/autoencoder> [Accessed 9 September 2024].

The Alan Turing Institute, 'Artificial Intelligence (Safe and Ethical)'. <https://www.turing.ac.uk/research/research-programmes/artificial-intelligence-ai/safe-and-ethical> [Accessed 29 August 2024].

'The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023', 1 November 2023 <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023> [Accessed 14 August 2024].

The European Parliament and the Council of the European Union, 'Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), Official Journal of the European Union (2024).

Thompson, Paul and McIntyre, Phillip, 'Rethinking Creative Practice in Record Production and Studio Recording Education: Addressing the Field', Journal on the Art of Record Production, 8 (2013).

Tollon, Fabio, 'Technology and the Situationist Challenge to Virtue Ethics', Science and Engineering Ethics, 30: 10 (2024).

Townsen Hicks, Michael, James Humphries and Joe Slater, 'ChatGPT is bullshit', Ethics and Information Technology, 26: 38 (2024), 1-10.

Trattner, Christoph, et al., 'Responsible Media Technology and AI: Challenges and Research Directions', AI and Ethics, 2 (2022), 585-594.

Trustworthy Autonomous Systems Hub, 'Response to: AI regulation: a pro-innovation approach – policy proposals', submitted in Open Consultation from Department for Science, Innovation and Technology and Office for Artificial Intelligence (2023).

UK Music, 'UK Music Policy Position on Artificial Intelligence' (July 2023).

UK Public General Acts, 'Copyright, Designs and Patents Act 1988' (1988) <https://www.legislation.gov.uk/ukpga/1988/48/contents> [Accessed 30 September 2024].

Vallor, Shannon, 'Edinburgh Declaration on Responsibility for Responsible AI', Medium, 14 July 2023 <https://medium.com/@svallor_10030/edinburgh-declaration-on-responsibility-for-responsible-ai-1a98ed2e328b> [Accessed 9 August 2024].

Vinchon, Florent et al., 'Artificial Intelligence & Creativity: A Manifesto for Collaboration', Journal of Creative Behaviour, 57: 1 (2024), 472-484.

Wang, Zihao, Teng Zou and Kai Sundmacher, 'A Novel Machine Learning-Based Optimisation Approach for the Molecular Design of Solvents', Computer Aided Chemical Engineering, 51 (2022), 1477-1482.

Wong, Vinnie, 'Gen AI's Environmental Ledger: A Closer Look at the Carbon Footprint of ChatGPT', Piktochart, 3 May 2024 <https://piktochart.com/blog/carbon-footprint-of-chatgpt/> [Accessed 18 September 2024].

Wu, Carole-Jean et al., 'Sustainable AI: Environmental Implications, Challenges and Opportunities', Proceedings of Machine Learning and Systems, 4 (2022), 795-813.

Yang, Zhilin et al., 'XLNet: Generalised Autoregressive Pretraining for Language Understanding', 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Canada (2019), 1-11.

Yasar, Kinza, 'Generative Adversarial Network', TechTarget <https://www.techtarget.com/searchenterpriseai/definition/generative-adversarial-network-GAN> [Accessed 9 September 2024].

Zerilli, John, Iñaki Goñi and Matilde Masetti Placci, 'Automation Bias and Procedural Fairness: A Short Guide for the Public Sector', BRAID (July 2024).

Zhao, Xiaolei, and Zhao, Xin, 'Application of Generative Artificial Intelligence in Film Image Production', Computer-Aided Design and Applications, 21 (2024), 15-28.

Zippia, 'Artificial Intelligence Specialist Demographics and Statistics in the US', 24 June 2024 <https://www.zippia.com/artificial-intelligence-specialist-jobs/demographics/> [Accessed 18 September 2024].

Appendix 1: Glossary of Technical Vocabulary

This glossary contains a catalogue of key terms and definitions that may be of use to those new to responsible AI and media production as a topic. It is divided into four sections: AI and Computer Science (1.1), AI in Media Production (1.2), Law and Copyright (1.3), and Responsible AI (1.4). In an effort to make the technical vocabulary of computer science and media production as accessible as possible, sections 1.1 and 1.2 take definitions of key terminology largely from popular sources, including blogs, articles and developer websites. In doing so, we recognise the prescient need for academic publications within the computer sciences to embed more accessible and interdisciplinary understandings of key terms and concepts within their scope of scholarship.

Appendix 1.1: AI and Computer Science

Basic Concepts

Artificial intelligence (AI): A field of science concerned with building computers and machines that can reason, learn, and act in such a way that would normally require human intelligence or that involves data whose scale exceeds what humans can analyse.[171]

Attention: The ability of a LLM to learn which words carry more importance in different contexts.

Conditional Generation: A technique where a generative model is asked to generate data according to some pre-specified conditioning, such as a topic, sentiment, or using one or more field values in a tabular, text, or image-based dataset.[172]

Dataset: A collection of data that is used to train and test algorithms and models. Datasets are crucial to the development and success of machine learning and AI systems, as they provide the necessary input and output data for the algorithms to learn from.[173]

Embeddings: A numerical matrix computed as a transformation of a text such that it's understandable by a LLM.[174]

Generative AI (GAI): A subset of AI that focuses on teaching machines to produce

---

[171] Google Cloud, 'What is Artificial Intelligence (AI)?' <https://cloud.google.com/learn/what-is-artificial-intelligence?hl=en> [Accessed 9 September 2024].

[172] Gretel, 'Conditional Generation FAQ', <https://docs.gretel.ai/create-synthetic-data/models/synthetics/conditional-generation-faq> [Accessed 9 September 2024].

[173] Encord, 'Datasets' <https://encord.com/glossary/datasets-definition/> [Accessed 9 September 2024].

[174] Pascal Dufour, 'Complete LLM Glossary', Step Up AI (November 2023) <https://stepup.ai/complete-llm-glossary/> [Accessed 9 September 2024].

original and creative content. Unlike traditional AI, which operates based on predetermined rules, generative AI has the ability to learn from data and generate content autonomously. This technology leverages complex algorithms and neural networks to understand patterns and produce outputs that mimic human-like creativity.[175]

Generative AI Tools (GAIT): Software applications that use generative AI models to create new content, such as text, images, audio, or video, based on patterns learned from existing data.

Hallucinations: Any nonsensical text or output generated by a LLM.[176]

Inference: The ability of a model to understand a text prompt.

Latent Space: A lower-dimensional space that captures the essential features of the input data. In simpler terms, it is a compressed representation of the original data where each dimension corresponds to a specific feature or characteristic.[177]

Large Language Model (LLM): A computational model capable of language generation or other natural language processing tasks. As language models, LLMs acquire these abilities by learning statistical relationships from vast amounts of text during a self-supervised and semi-supervised training process.

Low-Rank Adaptation (LoRA): Low-rank adaptation (LoRA) is a technique for quickly adapting machine learning models to new contexts. LoRA helps make complicated machine learning models more suited for specific uses. It works by adding lightweight pieces to the original model, as opposed to changing the entire model. LoRA helps developers quickly expand the use cases for the machine learning models they build.[178]

Machine learning (ML): Machine learning is a branch of AI and computer science that focuses on the using data and algorithms to enable AI to imitate the way that humans learn, gradually improving its accuracy.[179]

Deep Learning: A subset of machine learning that uses multi-layered neural networks, called deep neural networks, to simulate the complex decision-making power of the human brain.

Supervised Learning: A category of machine learning that uses labelled datasets to train algorithms to predict outcomes and recognise patterns.

Unsupervised Learning: A type of machine learning that learns from data without human supervision. Unlike supervised learning, unsupervised machine learning models are given unlabelled data and allowed to discover patterns and insights without any explicit guidance or instruction.

Neural Network: A method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain.

---

[175] Aditya Sharma, '11 Best Generative AI Tools and Platforms', Turing <https://www.turing.com/resources/generative-ai-tools> [Accessed 9 September 2024].

[176] Pascal Dufour, 'Complete LLM Glossary'.

[177] AI Maverick, 'A Comprehensive Guide to Latent Space', Medium (December 2023) <https://samanemami.medium.com/a-comprehensive-guide-to-latent-space-9ae7f72bdb2f> [Accessed 9 September 2024].

[178] Cloudflare, 'What is Low-Rank Adaptation (LoRA)?' <https://www.cloudflare.com/en-gb/learning/ai/what-is-lora/> [Accessed 7 October 2024].

[179] IBM, 'What is Machine Learning?' <https://www.ibm.com/topics/machine-learning> [Accessed 9 September 2024].

Prompt: The process of describing the task that you want an AI model to perform. This may be any form of text, question, information or coding that communicates your desired outcome.

Tokenisation: The process of slicing a text into token numbers in order to compute. A LLM is limited and usually judged by how many tokens (or a length of text) it can inference upon.[180]

Zero-Shot Learning: The ability of a LLM to learn a text-based task without any prior example.[181]

Models

Autoencoder: A type of neural network that is trained to copy its input to its output. For example, given an image of a handwritten digit, an autoencoder first encodes the image into a lower dimensional latent representation, then decodes the latent representation back to an image.[182]

Convolutional Neural Network (CNN): A network architecture for deep learning that learns directly from data. CNNs are particularly useful for finding patterns in images to recognise objects, classes, and categories. They can also be quite effective for classifying audio, time-series, and signal data.[183]

Diffusion Model: Diffusion models are prominent in generating high-quality images, video, sound, etc. They are named for their similarity to the natural diffusion process in physics, which describes how molecules move from high-concentration to low-concentration areas.[184]

Generative Adversarial Network (GAN): A machine learning model in which a generator and a discriminator compete with one another by using deep learning methods to become more accurate in their predictions.[185]

Generator: The generator learns to generate plausible data. The generated instances become negative training examples for the discriminator.[186]

Discriminator: The discriminator learns to distinguish the generator's fake data from real data. The discriminator penalises the generator for producing implausible results.[187]

Large Language Models

Generative Pretrained Transformer (GPT): A transformer model trained on unlabelled text to predict the next word.

Supervised Fine-Tuning (SFT): A GPT model trained on specific tasks with example texts for those specific tasks.

---

[180] Sloan Haywood, Genevieve Warren and Alex Wolf, 'Understand Tokens', Learn Microsoft (May 2024) <https://learn.microsoft.com/en-us/dotnet/ai/conceptual/understanding-tokens> [Accessed 9 September 2024].

[181] Pascal Dufour, 'Complete LLM Glossary'.

[182] Tensor Flow, 'Intro to Autoencoders' <https://www.tensorflow.org/tutorials/generative/autoencoder> [Accessed 9 September 2024]. 183 Math Works, 'What is a Convolutional Neural Network?', <https://uk.mathworks.com/discovery/convolutional-neural-network.html> [Accessed 9 September 2024].

[183] Math Works, 'What is a Convolutional Neural Network?', <https://uk.mathworks.com/discovery/convolutional-neural-network.html> [Accessed 9 September 2024].

[184] Super Annotate, 'Introduction to Diffusion Models for Machine Learning' <https://www.superannotate.com/blog/diffusion-models> [Accessed 9 September 2024].

[185] Kinza Yasar, 'Generative Adversarial Network', TechTarget <https://www.techtarget.com/searchenterpriseai/definition/generative-adversarial-network-GAN> [Accessed 9 September 2024].

[186] Google for Developers, 'Overview of GAN Structure' <https://developers.google.com/machine-learning/gan/gan_structure> [Accessed 9 September 2024].

[187] Ibid.

Reinforcement Learning from Human Feedback (RLHF): A GPT model further trained to adapt to human preferences using a less demanding approach such as a Likert scale.

Retrieval Augmented Generation (RAG): A GPT model with the ability to search databases to generate a based response.

Transformer: A neural network that learns context and thus meaning by tracking relationships in sequential data like the words in the sentence.[188]

Variational Autoencoder (VAE): A powerful generative machine learning method used to transfer a molecular structure into a continuous latent vector with an encoder and to convert the latent vector back to the molecule with a decoder.[189]

Recurrent Neural Network (RNN): A deep learning model that is trained to process and convert a sequential data input into a specific sequential data output.[190]

Long short-term memory (LSTM): A type of recurrent neural network that can learn long-term dependencies between time steps of sequence data.[191]

Gated Recurrent Unit (GRU): Like LSTM, GRU is designed to model sequential data by allowing information to be selectively remembered or forgotten over time. However, GRU has a simpler architecture than LSTM, with fewer parameters, which can make it easier to train and more computationally efficient.

Appendix 1.2: AI in Media Production

AI & Script writing

AI Correction: A process that identifies and rectifies grammatical errors such as misused tenses, subject-verb agreement issues, and faulty sentence structures.[193]

AI Translation: The translation performed by artificial intelligence without human involvement. Unlike human translation, AI translation is done instantly regardless of the complexity of the source text.[194]

Automated Dialogue Generation: The task of understanding natural language inputs within natural language processing in order to produce speech output.[195]

Script generator: A tool that uses artificial intelligence to generate scripts based on a few text prompts.[196]

[188] Rick Merritt, 'What is a Transformer Model?', NVIDIA <https://blogs.nvidia.com/blog/what-is-a-transformer-model/> [Accessed 9 September 2024].

[189] Zihao Wang, Teng Zou and Kai Sundmacher, 'A Novel Machine Learning-Based Optimisation Approach for the Molecular Design of Solvents', Computer Aided Chemical Engineering, 51 (2022), 1477-1482 (p. 1477).

[190] Amazon Web Services, 'What is a Recurrent Neural Network?' <https://aws.amazon.com/what-is/recurrent-neural-network/> [Accessed 9 September 2024].

[191] Math Works, 'Time Series and Sequence Data Networks' <https://uk.mathworks.com/help/deep-learning-hdl/time-series-and-sequence-data-networks.html> [Accessed 9 September 2024].
[192] Anishnama, 'Understanding Gated Recurrent Unit (GRU) in Deep Learning', Medium (May 2023) <https://medium.com/@anishnama20/understanding-gated-recurrent-unit-gru-in-deep-learning-2e54923f3e2> [Accessed 9 September].

[193] Ahrefs, 'Free AI Grammar Checker' <https://ahrefs.com/writing-tools/grammar-checker> [Accessed 9 September 2024].
194 Anthony Enticknap, 'What is AI Translation and How to Use It?', Centus (December 2023) <https://centus.com/blog/what-is-ai-translation> [Accessed 9 September 2024].

[195] Papers With Code, 'Dialogue Generation' <https://paperswithcode.com/task/dialogue-generation> [Accessed 9 September 2024].

[196] Synthesia, 'AI Script Generator' <https://www.synthesia.io/features/ai-script-generator> [Accessed 9 September 2024].

AI & Image

Deepfake: Videos, picture or audio clips made with artificial intelligence to look or sound real. While deepfakes can be used for the purpose of scientific research, their design, to impersonate real people, can be conducive to unethical manipulations of celebrity likeness and the deliberate spreading of often harmful misinformation.[197]

Denoising: An advanced technique used to decrease grainy spots and discoloration in images while minimising the loss of quality.[198]

Image Inpainting: A tool for replacing or editing specific areas of an image. This makes it a useful tool for image restoration like removing defects and artefacts, or even replacing an image area with something entirely new.[199]

Image Segmentation: A process that divides an image into segments where each pixel in the image is mapped to an object. This task has multiple variants such as instance segmentation, panoptic segmentation and semantic segmentation.[200]

Image Style Transfer: A tool that transfers use trained neural networks, which learn the style features of the reference image (including colours, textures, and brush strokes). The neural networks then apply those style features to the content of another image.[201]

Super-Resolution: Models that increase the resolution of an image, allowing for higher-quality viewing and printing.[202]

Text-to-Image: Models that generate images from input text. These models can be used to generate and modify images based on text prompts.[203]

AI & Editing

Automated Colour Grading: The process of using AI technology to enhance and manipulate the colours in videos. This technique leverages advanced algorithms to analyse and adjust colour tones, contrast, saturation, and other visual elements automatically.[204]

Text-based editing: The use of artificial intelligence to produce a transcript of your videos that provides a way to edit the video by selecting text.[205]

AI & Music/Sound

Automatic Composition: This technique utilises complex algorithms and machine learning to analyse vast datasets of music,

---

[197] BBC Newsround, 'Deepfake Technology: What is it, how does it work, and what it be used for?' <https://www.bbc.co.uk/newsround/69009887> [Accessed 9 September].

[198] JJ Kim, 'What is Denoising', NVIDIA (November 2022) <https://blogs.nvidia.com/blog/what-is-denoising/> [Accessed 9 September 2024].

[199] Hugging Face, 'Inpainting' <https://huggingface.co/docs/diffusers/using-diffusers/inpaint> [Accessed 9 September 2024].
200 Hugging Face, 'What is Image Segmentation?' <https://huggingface.co/tasks/image-segmentation> [Accessed 9 September 2024].

[201] Adobe, 'The Power of AI Image Style Transfer' <https://www.adobe.com/products/firefly/features/generative-match.html> [Accessed 9 September 2024].

[202] Paul Gafton and Osman Alenbey, 'Image-to-Image', Hugging Face <https://huggingface.co/tasks/image-to-image> [Accessed 9 September 2024].

[203] Ishan Dutta, Enrique Elias Ubaldo and Oğuz Akif, 'Text-to-Image', Hugging Face <https://huggingface.co/tasks/text-to-image> [Accessed 9 September 2024].

[204] Joshua Hill, 'Revolutionise Your Projects with AI Colour Grading', Hit Paw <https://www.hitpaw.com/video-tips/ai-color-grading.html> [Accessed 9 September 2024].

[205] Adobe, 'Text-Based Editing in Premiere Pro' <https://helpx.adobe.com/premiere-pro/using/text-based-editing.html> [Accessed 9 September 2024].

enabling them to identify patterns and styles. This knowledge is then used to compose pieces that push the boundaries of genre conventions, ranging from classical symphonies to modern pop beats.[206]

Audio Separation: The process of separating a mixture (e.g. a pop band recording) into isolated sounds from individual sources (e.g. just the lead vocals).[207]

Automatic Speech Recognition (ASR): The use of Machine Learning or AI technology to process human speech into readable text.[208]

Music Recommendation System: By using music recommender system, the music provider can predict and then offer the appropriate songs to their users based on the characteristics of the music that has been heard previously.[209]

Noise Reduction: The process that involves taking a signal and eliminating unwanted noise from it.[210]

Speech Emotion Recognition: A task of speech processing and computational paralinguistics that aims to recognise and categorise the emotions expressed in spoken language.[211]

Text to Speech (TTS): Models that generate high-quality, natural-sounding speech from text with low latency.[212]

Appendix 1.3: Law and Copyright

Authorship: The original creator of a 'work' in copyright is known as its author, e.g. the writer of a book, the artist of a drawing, the composer of a piece of music. For a film, the legal author is taken to be the producer and principal director; and for a sound recording the author is taken to be the producer, as set out in section 9 Copyright, Designs and Patents Act 1988.

Current UK copyright law does not expressly include or exclude Generative AI produced works. In s.9(3) of the CDPA, computer-generated works are included, stating: 'In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken'.[213]

Film and media creation tends to involve multiple copyright 'works' coming together, and often involves multiple original creators, and so often sees 'joint authorship' of film and media works. Around the world, there is not an agreed consensus within copyright law concerning whether an author must be human. How precisely authorship functions in

[206] Jam Canda, 'AI in Music: Composition, Production, and Recommendation', Medium (April 2024) <https://medium.com/@jam.canda/ai-in-music-composition-production-and-recommendation-4b5bbde1b10b> [Accessed 9 September 2024].

[207] Papers With Code, 'Audio Source Separation' <https://paperswithcode.com/task/audio-source-separation> [Accessed 9 September 2024].

[208] Kelsey Foster, 'What is Automatic Speech Recognition? A Comprehensive Overview of ASR Technology', Assembly AI (September 2023) <https://www.assemblyai.com/blog/what-is-asr/> [Accessed 9 September 2024].

[209] Adiyansjah, Alexander Gunawan and Derwin Suhartono, 'Music Recommender System Based on Genre Using Convolutional Recurrent Neural Networks', Procedia Computer Science, 157 (2019), 99-109.

[210] Agora, 'What Does Noise Reduction Do?' <https://www.agora.io/en/blog/what-does-noise-reduction-do/> [Accessed 9 September 2024].

[211] Papers With Code, 'Speech Emotion Recognition' <https://paperswithcode.com/task/speech-emotion-recognition> [Accessed 9 September 2024].

[212] NVIDIA, 'Text-to-Speech' <https://www.nvidia.com/en-us/glossary/text-to-speech/> [Accessed 9 September 2024].

[213] UK Public General Acts, 'Copyright, Designs and Patents Act 1988' (1988) <https://www.legislation.gov.uk/ukpga/1988/48/contents> [Accessed 30 September 2024].

copyright law and Generative AI interact is not currently clear.

Ownership: The person or company who owns the copyright of a particular work, such as a film or book, and has the legal ability to decide what others may or may not do with that work, such as deciding whether to allow other people to make copies of the work or broadcast it. The first owner of the copyright can choose to legally sell or transfer (assign) it to another person or company.

The author of a work is the first owner of any copyright in it, as set out in section 11(1) Copyright, Designs and Patents Act 1988. The author and first owner of copyright in a sound recording is the record producer.

Where a work, including a film, is made by an employee in the course of their employment, then it is the employer that is taken to be the first owner of any copyright in the work unless otherwise agreed in writing, as set out in s. 11(1) CDPA. Where work is commissioned by another person or organisation, the author and first legal owner of the work is taken to be the person who created the work, as opposed to the person or organisation who commissioned it, unless otherwise agreed in writing. For both employment and commissioning, it is therefore vital that there is a written contract or agreement that sets out the intention of who owns the copyright.

Appendix 1.4: Responsible AI and Ethics

Accessibility: The quality of being widely available or easily usable for a diverse audience. This might include being free at the point of access, hosting a clear layout or design, or having a user-friendly interface. Accessibility is about breaking down barriers of access, particularly for those who have previously been denied access due to gender, race, age, class or disability.

Accountability: The condition of being answerable, liable or culpable in relation to the design and use of AI, as well as its resulting wider social impacts.[214] Embedding accountability into AI usage means that those in positions of authority to make important decisions are held responsible for their actions. This might mean keeping clear data records of who is making what decision at any given time, so that those responsible can justify and explain key decisions to relevant stakeholders.

Fairness: The practice of ensuring the use and outputs of GAI are free from prejudice and discrimination. Fairness in AI entails the installation of accessibility, such that everyone in society is given equal access to GAI models. It may also overlap with questions of justice, whereby anyone who is a victim of harmful and unfair GAI practices and outputs may be able to seek reparations.[215]

Reliability: The quality of producing outcomes in line with the expected outcomes of the end user(s). In the context of GAI, this means being able to regularly give the same result based upon successive trials and similar prompts. While reliability may be desired in GAI media production for those desiring dependable outcomes and stylistic consistency, reliability may also hinder creativity in that its production of the most expected outcome may result in common, generic or stereotyped outputs.

---

214 Virginia Dignum, Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way (Cham: Springer, 2019), p. 53.

215 Luciano Floridi, The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. Oxford: Oxford University Press, 2023), p. 63.

Safety: The quality of being free from harm. In the context of AI more broadly, this may pertain to issues of cybersecurity, warfare and state and private surveillance.[216] In the context of GAI, specifically, safety concerns the production of inclusive and ethical systems, structures and outputs that protect against risks to mental and physical health. This may include safeguards designed to protect against children, marginalised communities and those more vulnerable in society from potentially harmful GAI outputs and practices, as well as the prevention of misuse of GAI to spread misinformation or bias.

Sustainability: The ability for a system's rate of change to be maintained and governed in order to support its existence. Sustainability in AI often refers to ecological sustainability and AI's impact on Earth's natural resources. This may be achieved by addressing the carbon footprint that comes with storing, processing and training large amounts of data for GAI models. It may also entail utilising AI to support global sustainability goals, such as monitoring and developing new ecological processes.[217] Sustainable AI overlaps with responsible AI in that it also encompasses the maintenance of social cohesion and existence.

Transparency: The practice of being clear and unobstructive to perceptions of information. In the context of GAI, transparency means being open and explicit to stakeholders about how certain decisions were made.[218] Transparency therefore entails explainability, and this may be achieved through the disclosure of engineering methods, training data or risk assessments.[219] Transparency in GAI can build an increased level of trust, as it allows all stakeholders to inspect the mechanisms through which AI outputs come to be produced.

---

[216] 'The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023', 1 November 2023 <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023> [Accessed 14 August 2024].

[217] Eshan Navabi, Katherine Daniell, Elizabeth Williams and Caitlin Bentley, 'AI for Sustainability: A Changing Landscape', in Artificial Intelligence: For Better or Worse, Future Leaders, 157-176 (2019) (p. 157).

[218] Dignum, Responsible Artificial Intelligence, p. 54.

[219] Alejandro Barredo Arrieta et al., 'Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI', Information Fusion, 58 (2020), 82-115 (p. 83).
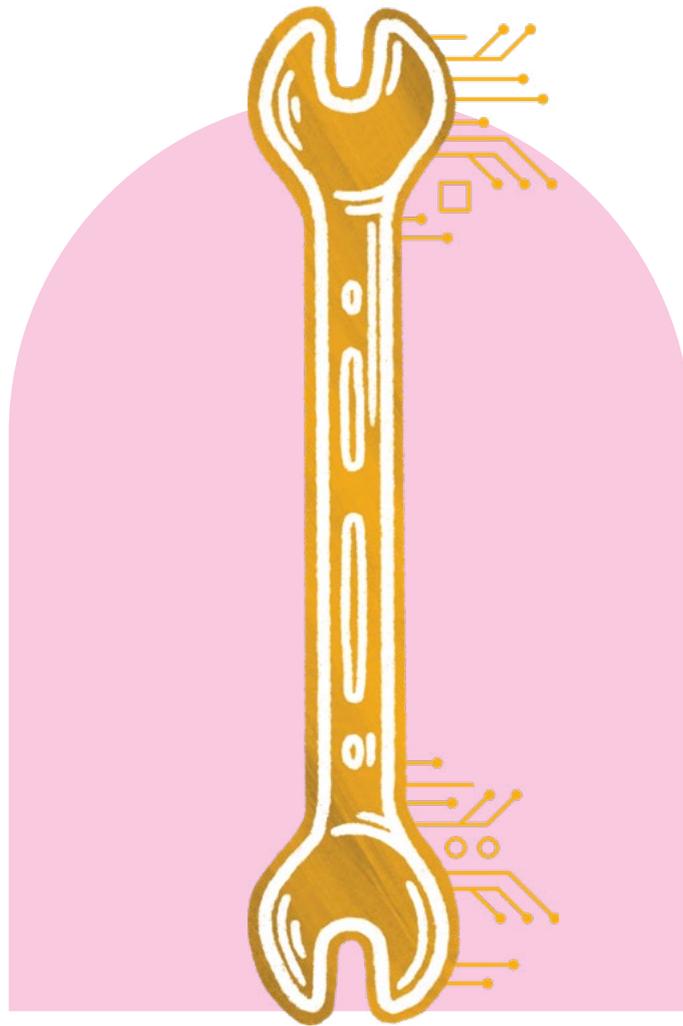
Appendix 2:
AI Tools Directory

Appendix 2.1: Creative Writing/Screen-writing

Functionality	Transparency	Ethical Considerations
Novelcrafter		
AI-powered writing platform which integrates multiple AI models, such as GPT-3, GPT-4, and Claude, among others. Users can choose which AI model to connect with. Main functionalities include: generating story ideas, assisting with character development, and writing portions of the text based on user prompts. Also offers a personal wiki (Codex) for organising information about characters, locations, and lore.	The platform clearly advertises its AI capabilities and provides detailed documentation on how to connect and use different AI models for users who choose to seek out this information. This includes guidance on how to set up and use models from providers like OpenAI and OpenRouter, as well as the benefits and limitations of each. Does not reference the existing texts it draws from when generating content.	Could potentially remove work for human editors. Draws on existing literary content without attribution or direct financial compensation. Potential for unintentional plagiarism. A writer could publish work generated by this programme without flagging it as such, as there are no explicit frameworks in place regarding how generated work should be credited. The generative aspect of this program could reinforce biases and/or stereotypes.
Squibler		
AI-powered writing assistant which utilises AI to generate real-time suggestions as the user is writing. Provides grammar corrections, style enhancements, and readability improvements, ensuring that writers can refine their prose while maintaining their unique voice. Also generates writing suggestions and prompts to aid with ideation. AI analyses the content of the text, recommending plot points and ideas for character development.	Open in its status as an AI-powered programme. Features powered by AI are clearly labelled and explained, allowing users to understand the benefits and functionalities. This approach helps writers make informed decisions when utilising AI tools within the platform. However, Squibler does not cite the exact sources for its AI capabilities, nor its algorithms or training processes.	Squibler lacks clear information regarding who owns the generated output. This is an issue, especially regarding texts in which the AI provides substantial creative input. The AI in Squibler is trained on large datasets, which may contain biases, which could result in character/plot suggestions which perpetuate prejudices. Writers are advised to be mindful that their use of Squibler does not misrepresent cultures. Accidental plagiarism is possible if the user relies heavily on Squibler for generation. May contribute to fewer employment opportunities for human editors.
Sudowrite		
AI-powered writing assistant. Offers real-time grammar corrections, style improvements, and vocabulary suggestions to polish text. The platform also aids in idea generation, providing writing prompts, story ideas, and plot outlines based on user input. Additionally, Sudowrite incorporates research tools to assist with summarising relevant information and offering resources to support writing projects.	Openly promotes its AI-driven features and capabilities in its marketing materials and website, highlighting how it supports various aspects of the writing process. Although it does not specifically disclose the sources of its AI algorithms or datasets in publicity materials, Sudowrite generally clarifies the role of AI in analysing text, generating suggestions, and enhancing writing quality.	The handling of user data by Sudowrite raises concerns about how securely personal and potentially sensitive information is stored and used. As not all original data is cited, there's a risk of users inadvertently producing work that closely resembles existing texts, potentially leading to plagiarism issues. The AI algorithms may reflect biases present in the training data, which can influence the suggestions and outputs generated by the tool. While Sudowrite is generally transparent about its use of AI, there may still be concerns about the specific details of how its algorithms work and the sources of its training data.

Functionality	Transparency	Ethical Considerations

Hyperwrite

Platform designed to enhance the writing process for various types of content, including articles, blog posts, and essays. Provides AI-powered assistance by offering real-time suggestions for grammar, style, and readability improvements, generating content ideas and summaries, and assisting with research by summarising external information and suggesting additional resources. HyperWrite AI also features analytics tools that provide feedback and help refine writing skills, and it continuously learns from user interactions to tailor its responses specifically to individual user needs.

HyperWrite AI is transparent about being an AI-powered writing tool, prominently showcasing its use of advanced natural language processing (NLP) technology in its website and marketing materials. It clearly labels AI-driven features such as real-time suggestions, content generation, and feedback mechanisms, ensuring users know they are interacting with an AI system. The platform allows users to control AI-generated suggestions by accepting, rejecting, or modifying them, thus allowing a significant amount of agency over the writing process. Additionally, HyperWrite AI includes features like 'Scholar AI', which provides citation-backed results from scholarly articles, making it easier for users to trace the sources of the information it generates. However, not all features of Hyperwrite generate citations.

HyperWrite AI allows users to conduct research through Scholar AI, which implements a clearer sourcing and referencing system that helps users trace the origins of the information. However, other functions of the cite generate content without clear referencing. Users are also advised to verify the accuracy of these sources and the generated content. There is a lack of transparency regarding how the user's data will be handled – who owns it, and will it be used as the basis for further training.

AI screenwriter

A tool specifically designed for screenwriters. Features include: generating realistic dialogue, assisting with plot development, helping with character evolution, and ensuring proper script formatting. The tool supports collaborative writing, provides editing and feedback, and allows for export in various script formats. Utilising advanced AI technologies like natural language processing (NLP) and machine learning, AI Screenwriter analyses user inputs and offers suggestions based on a large dataset of existing scripts. It provides real-time feedback, learns from user interactions to improve its suggestions, and adapts to individual writing styles over time.

AI Screenwriter is moderately transparent about its use of AI. It offers basic explanations of its AI technology, highlighting the use of natural language processing (NLP) and machine learning. It informs the reader that its content is trained on a wide and diverse range of screenplays. However, it does not provide specific information about its datasets, and does not cite sources. Additionally, it includes terms of service and privacy policies that outline the collection and use of user data.

While AI Screenwriter offers basic insights into its algorithms, the lack of detailed information about specific datasets could result in accidental plagiarism on the part of the user, original authors not being compensated/credited, and the user not being able to make informed decisions regarding potential biases/ prejudices in the programme's suggestions. However, the terms of service sets out quite robust policies in regards to privacy and ownership in comparison with many similar AI tools.

Charisma.ai

Storytelling platform that allows users to create interactive and immersive experiences with virtual characters. Designed for a variety of users, including content creators, game developers, educators, and marketers, it features a no-code environment enabling the creation of dynamic characters with emotions and memories capable of engaging in natural, freestyle conversations. Charisma.ai integrates with popular game development tools like Unity and Unreal Engine, facilitating the inclusion of its interactive characters into larger projects. It also provides real-time analytics to help creators understand audience engagement and make data-driven improvements.

Charisma.ai demonstrates a moderate level of transparency regarding its use of AI. The platform openly advertises its AI-driven capabilities, including proprietary natural language processing (NLP) and machine learning technologies. This transparency is reflected in its marketing materials, website, and user documentation. Yet, it does not cite its sources or explicitly provide details about its dataset.

Charisma.ai tracks user interactions and provides real-time analytics. Ensuring that user data is collected, stored, and used securely, and that privacy policies are transparent and comprehensive, is crucial, yet the programme does not provide comprehensive information regarding this. The AI's responses and character interactions are shaped by its training data, which may contain biases. While Charisma.ai allows users to create and refine content, there is a need to ensure that users retain control and can make ethical decisions about how AI-generated content is used and presented.

Functionality	Transparency	Ethical Considerations

Poe

AI chatbot aggregator that provides users with access to a range of AI-powered chatbots from leading providers such as OpenAI, Anthropic, Meta, and Stability AI. The platform aims to centralise various AI models, allowing users to engage with different technologies to create individualised chatbots for diverse purposes, including casual conversations, specialised advice, and text generation. By integrating models specialising in natural language processing and image generation, Poe facilitates a wide range of applications, from text-based tasks to image creation.

Poe is transparent about its identity as an AI-powered platform and its integration with multiple well-known AI models. It highlights the use of these advanced models for various tasks, including text-based interactions and image generation. However, Poe generally does not provide specific citations for the information generated by these pre-existing AI models. The responses are based on the models' training data, which comes from extensive datasets, but the exact sources of the information are often not disclosed. This lack of citation can be a limitation for users who need to verify the accuracy and reliability of the information provided. However, users who create their own AI models through Poe and design the knowledge base to draw upon specific texts may notice clear citations to written source material, which can help users keep track of both copyright and artistic inspiration.

The custom chat models on Poe are customisable by the user, and the user can upload specific data sources to serve as a foundation of knowledge. This may help to circumvent any potential issues relating to bias and representation. However, the responses provided by AI models are based on extensive datasets, yet largely lack clear citation. This may lead to accidental plagiarism, as well as issues with the accuracy and reliability of the information. As users interact with different AI models and create their own chatbots, there are questions about the ownership of content generated. This is not made clear on the site. Given that Poe aggregates data from multiple AI models and handles user interactions, it is important to ensure that user data is collected, stored, and managed securely. Clearer and more robust frameworks need to be put in place to clarify how this data is collected and handled.

ChatGPT

ChatGPT is an AI tool developed by OpenAI that processes large amounts of text data to understand and produce coherent and contextually relevant responses. Uses machine learning, specifically a variant of the Transformer architecture, to generate human-like text based on user inputs. The main functions of ChatGPT include answering questions, providing explanations, engaging in conversation, and assisting with a variety of writing tasks.

ChatGPT is relatively transparent about its use of AI, as it is marketed as an AI-powered tool and provides clear information regarding the technical process through which it draws on datasets to produce responses. However, it lacks clear information regarding the specific algorithms and training processes used, and it does not directly cite sources for the information it provides.

ChatGPT could be improved in relation to AI ethics by enhancing transparency about how it generates responses, implementing more robust mechanisms to detect and mitigate biases, ensuring greater accuracy to reduce the spread of misinformation, strengthening privacy and data security measures, and establishing clearer guidelines for content moderation to prevent harmful outputs.

Appendix 2.2: Image Generation

Functionality	Transparency	Ethical Considerations

Runway

RunwayML is an AI-powered image generation and modification tool aimed at visual artists and filmmakers. Designed to help streamline creative workflows, allowing users to either generate still images/videos, or apply effects, motion graphics, and enhancements to existing footage. Users can generate new images from text prompts, edit an image based on user input, or apply algorithmic filters. AI can also be used to recognise and manipulate objects within images. This allows creatives to generate complex visual content based on simple inputs, such as transforming rough sketches into detailed visuals. Inputs can be text-to-image or image-to-image.

RunwayML has taken several steps to provide a level of transparency around the AI models it employs. RunwayML allows users to access a wide variety of pre-trained AI models, and gives detailed information regarding performance metrics, what they are intended for, the type of data they were trained on, and how users can apply them to their creative projects. However, RunwayML could improve transparency by offering more detailed information on the specific datasets and algorithms behind all its pre-trained models. One of RunwayML's standout features is its ability to allow users to train custom models using their own datasets. While this gives users full control over their AI applications, the platform is somewhat opaque in explaining how the underlying algorithms might be optimised, the trade-offs in model training, and the potential biases that could arise depending on the datasets used.

RunwayML offers users the choice to can upload their own datasets, adjust parameters, and create AI models tailored to their unique aesthetic or project requirements. This level of individualisation can help the user overcome issues with training bias and avoid committing copyrighting infringement. However, the pre-set models do not directly cite the data used for training. RunwayML could also implement more detailed tracking of data provenance, ensuring that users know whether the datasets used were ethically sourced, and if any user-generated content is used in model training, how it is protected. If users employ AI tools to generate or alter images or videos of real people without their consent, this raises significant ethical concerns, particularly around privacy and consent. There are currently no formal restrictions on doing this in Runway.

Kive.ai

AI-driven creative platform that enhances visual content creation and collaboration through a range of innovative tools. With features like AI-powered image generation, dynamic mood boards for visual storytelling, and intelligent asset management, Kive is designed to streamline the entire creative workflow. Users can generate unique images and visuals, while AI-assisted organisation can be used as a tool to tags and categorise assets.

Kive is upfront about its integration of AI, clearly marketing itself as an AI-powered platform. The platform could be more transparent about the origins of the data used to train its AI models. Currently, it does not provide explicit details about the sources or datasets that inform the AI's outputs. For users concerned with the ethical implications of data usage (such as whether the data is free from biases or copyright issues—greater clarity is needed). While Kive's tools are highly functional, a more transparent system that allows users to understand how the AI arrived at a particular result—such as a model's training data or decision-making processes—would improve transparency.

It can be unclear who holds the rights to images or designs produced by AI on Kive. Users may face challenges in claiming copyright or ensuring the originality of AI-generated works. Offering more detailed information to users on the datasets and the methods used to ensure ethical standards would also enhance transparency and accountability. If Kive's models are trained on biased or unrepresentative datasets, the AI-generated images or content could perpetuate stereotypes or present skewed perspectives. Regular assessments of diversity and representation in datasets could mitigate such risks. here is currently a lack of safeguards in place to prevent misuse, including clear policies, user guidelines, and detection systems for manipulated content.

Midjourney

Primary generative AI imaging software that pre-trained machine learning models that perform tasks such as image and video generation, style transfer, segmentation, and object detection. These models allow users to create or enhance media by generating unique images, applying artistic styles to photos or videos, and replacing backgrounds in real-time.

Midjourney is relatively transparent about its usage policies and licensing. They provide information on how users can utilise the generated images, including commercial use and attribution requirements. Midjourney also provides some insight into the general workings of its AI, such as the fact that it uses machine learning algorithms to generate images based on textual descriptions. However, the specifics of the model architecture, training data, and algorithms are not fully disclosed.

The company has guidelines for content creation and usage to prevent the generation of harmful or inappropriate content. They outline what is and isn't allowed, and they take measures to enforce these guidelines, but the specifics of their moderation processes are not always clear.

Functionality	Transparency	Ethical Considerations

Adobe Firefly

Adobe Firefly primarily uses generative AI as one of its core functionalities. The main purpose of Firefly is to leverage generative AI techniques to create, enhance, and manipulate visual content based on user inputs. This includes generating images from textual descriptions, creating design elements, and automating various aspects of visual content creation and editing.

Adobe provides clear information on how Firefly utilises AI, including its capabilities for generating and enhancing visuals based on user inputs. They describe the AI-driven features, such as generative content creation, smart editing, and automated design suggestions, in user-facing documentation and marketing materials. Adobe also makes a concerted effort to provide some level of transparency about the data used to train their AI models. For Firefly, they emphasise that their AI is trained on a diverse dataset, but the specific details about the data sources are generally not disclosed in depth. Adobe also provides relatively clear information on how user data is handled and protected when using Firefly. They have privacy policies in place that outline data collection, storage, and usage practices.

Adobe Firefly offers guidance discouraging the generation of harmful, inappropriate, or misleading content. Adobe's policies aim to prevent misuse of the AI, such as creating discriminatory or offensive imagery. Adobe has privacy policies that outline how user data is collected, stored, and used, focusing on protecting user information and maintaining confidentiality. Adobe also takes measures to address intellectual property concerns, aiming to ensure that generated content does not infringe on existing copyrights or trademarks. However, Firefly could cite more clearly the data used to train its models. It could also improve its ethical standing by offering users more agency over the way in which AI generates content. For instance, allowing users to set ethical parameters or filters for content generation can help prevent misuse and ensure adherence to individual or organisational standards.

Stable Diffusion

AI program designed for generating and manipulating images based on textual descriptions. Primarily a generative AI model designed for producing still images. Stable Diffusion also allows users to influence the generated images through various parameters and settings. This includes adjusting aspects like style, colour, and composition, giving users greater control over the final output.

Stable Diffusion provides a broad overview of its AI model and methodology on its website and related documentation. Users can learn that it employs diffusion models to generate images from text prompts and get a sense of how the process works at a high level. There are guidelines available regarding the responsible use of the technology. These include recommendations to avoid generating harmful or inappropriate content and respecting intellectual property. However, providing more detailed information about the training data and processes used to develop the diffusion model would enhance transparency.

Greater transparency about the measures taken to address biases in the AI model could improve user trust. Users would benefit from knowing more about the dataset's composition, how biases are managed, and how the model's performance is evaluated. Detailed information on how the model is tested for fairness and how biases are mitigated would be valuable. More information on how user data is handled, particularly if the platform collects or stores any data from user interactions, would also be beneficial.

Imagine.art

AI-powered platform designed to create both still images and videos from text prompts. It allows users to describe their vision in words, which the AI then transforms into visuals. Art's core functionality focuses on text-to-image generation. It also offers a text-to-video feature, enabling users to fully generate HD videos from written scripts. Imagine.art also offers advanced tools like upscaling images, maintaining character consistency in visual narratives, and remixing images to experiment with different styles.

Imagine.art provides a moderate level of transparency about its use of AI. It clearly states that it uses advanced AI algorithms to transform user inputs into visual outputs. However, the platform does not offer detailed technical explanations about how the AI operates, including specifics about the models it uses (e.g., neural network architectures or third-party AI tools), the data sources for training, or the ethical guidelines that govern AI-generated content. Providing more specific information about the AI models and techniques used would give users a clearer understanding of the technology behind the platform. Disclosing the datasets used to train the AI would also help to address issues regarding copyright and bias.

It's unclear who owns the rights to the final work and how the platform ensures it doesn't infringe on existing copyrights. Additionally, the platform's bias in training data is a concern, as the AI may unintentionally favour certain styles or demographics, reinforcing stereotypes or limiting diversity in its outputs. Misuse of the tool is another ethical challenge, as Imagine.art could be exploited to create harmful, misleading, or deepfake content. There are no clear guidelines in place to prevent this. It is also uncertain how the platform handles and stores user-uploaded images and personal data.

Appendix 2.3: Editing

Functionality	Transparency	Ethical Considerations

Vimeo – 'One Take Video'

Vimeo's 'One Take Video' program uses AI both as a generative tool and for automating key video creation processes. Generative AI is applied predominantly in the scriptwriting phase, where users can use an AI interface to generate a full script based on text prompts. Beyond generative AI, the program also offers AI-powered tools for video editing and performance support. The teleprompter feature helps users deliver scripts naturally, while the AI-based editor automatically identifies and removes filler words, pauses, or awkward moments from the video. These editing features do not involve generative AI but use AI to streamline the post-production process.

Vimeo explicitly outlines the features of its AI-powered suite, such as the generative AI scriptwriting tool, the teleprompter for recording and the AI-powered image editor. However, the platform does not provide in-depth details about the underlying AI models, data sources, or potential limitations of these tools. While it promotes ease of use and accessibility, Vimeo does not elaborate on the technical specifics of how the AI functions, such as the datasets used for training or whether any AI-related biases are being addressed. Additionally, there are also no citations for the specific datasets used to generate scripts.

The program collects user inputs such as video recordings and potentially personal information through its teleprompter and AI-based editor. Vimeo needs to be transparent about how this data is stored, processed, and protected from unauthorised access. The tool could also potentially result in less work being available for human editors and editorial assistants. Additionally, it's important to clarify who holds the rights to the final video. This is particularly relevant when using generative AI for scriptwriting— does the content fully belong to the user, or does Vimeo retain some rights? To improve transparency, Vimeo could provide more information on the AI technologies and ethical considerations behind the program, including data privacy and bias mitigation efforts. The generative AI scriptwriter might produce content influenced by biased data sets.

Final Cut Pro

AI features in Final Cut Pro are deeply integrated into the editing workflow. They are designed to be intuitive and work seamlessly with the software's existing tools. Users can access AI-powered functionalities through standard editing interfaces, making advanced technology accessible without requiring technical expertise. AI is implemented in functions such as: altering aspect ratio; organising and searching through footage; improving audio quality; colour matching; automatic transcription and subtitles; facial recognition and tracking; and object/motion tracking.

Final Cut Pro does not prominently advertise itself as an editing platform founded on AI technology. However, within the system, the features that utilise AI or machine learning are often highlighted and described in terms of their advanced technological capabilities. The program also offers further information about how the technology works in techniques like smart conform and automatic transcription. However, because these features are designed to be seamlessly integrated into the editing interface, users can use them without consulting this further information.

AI features like facial recognition and object tracking could potentially be used to infringe on privacy. If not properly managed, these tools might collect or analyse sensitive information without proper consent. AI algorithms can inadvertently introduce bias, particularly if they are trained on skewed or unrepresentative data. For example, facial recognition or scene detection algorithms might perform unevenly across different demographics, leading to less accurate results for certain groups. The automation of certain editing tasks through AI might reduce the need for human editors, potentially leading to job displacement or shifts in the industry.

Kapwing

Kapwing is an online video editing and content creation platform that leverages AI to simplify and enhance various aspects of the creative process. Kapwing integrates AI seamlessly into its user interface, allowing users to access powerful features through simple, intuitive tools. The AI performs complex tasks in the background while users interact with straightforward controls. Functions include: automatic subtitles/captioning; isolating and removing backgrounds; and enhancing image and video quality (i.e. resolution and colour changes); suggesting templates and design elements to fit video content; applying filters and effects. The use of AI here is primarily tools-based rather than generative, as these features augment and streamline existing editing processes, rather than generating content from scratch.

Kapwing offers help resources with basic explanations about how its AI technology benefits users. It provides clear descriptions of AI-powered features like automatic subtitles and background removal. While the platform is transparent about the presence and functionality of AI, it generally does not delve deeply into the technical details of the underlying algorithms, including data training.

Kapwing should provide more detailed information about how AI algorithms work and how user data is processed and stored; Implement and regularly update mechanisms to identify and reduce biases in AI tools; offer more granular controls for users to manage how AI features interact with their content, such as options for reviewing and adjusting AI-generated outputs to ensure accuracy and relevance and develop and enforce guidelines for the ethical use of AI features, including clear policies on acceptable content and practices.

Functionality	Transparency	Ethical Considerations

Frame.io

A collaborative video review and approval platform. It utilises AI to enhance various aspects of the video production and review process. Functionalities include: automated meta-data tagging; automated transcription; and providing suggestions based on content analysis. The AI in Frame.io is primarily tools-based rather than generative. It focuses on enhancing existing processes through automation and improved searchability, rather than creating new content from scratch.

The platform provides clear descriptions of AI-powered features, such as automated meta-data tagging and smart transcription. Users are informed about how these features enhance the video review and approval process. Frame.io offers documentation and support resources that explain the functionality of its AI features. This helps users understand the benefits and operation of the tools.

Frame.io could improve its transparency by providing detailed technical documentation about its AI algorithms, explaining how AI makes specific decisions, and clearly communicating the limitations and potential biases of its AI features.

Descript

Descript uses AI to streamline audio and video editing by offering features like automatic transcription, which allows for text-based editing of media; Overdub, which creates synthetic voice clones for generating new content without re-recording; automated clean-up tools to remove filler words and background noise; and speaker identification for organising multi-speaker content. These AI-driven functionalities simplify the editing process, and are not primarily generative.

Descript is relatively open and transparent about its use of AI in its functions. It clearly describes how AI is utilised for automatic transcription, text-based editing, and features like Overdub for voice cloning. The platform provides detailed information about these functionalities on its website and in its help resources. Descript could enhance its transparency by providing detailed technical documentation on its AI algorithms, publishing regular transparency reports on updates and issues, and explaining AI decision-making processes, offering user feedback mechanisms for adjusting AI results.

Descript could improve its ethical standing by enhancing privacy measures to protect user data, ensuring explicit consent and control for voice cloning, improving accuracy and transparency in AI outputs, providing user education on ethical AI use, and implementing monitoring systems to detect and address misuse.

Veed

Veed is an AI-powered video editing platform that uses artificial intelligence to streamline and enhance various aspects of video creation and editing. Functions include: automatic subtitles and transcription; cutting, trimming, and editing videos with AI suggestions; background alteration/removal; algorithmic effects/filters; video enhancements like colour and resolution alteration. Not generative.

Veed is reasonably transparent about its use of AI, clearly describing how AI features like automatic subtitles, background removal, and video enhancements work on its website and in user guides. However, it could improve transparency by providing more detailed technical documentation about the underlying AI algorithms and models, explaining how they process data and their limitations.

Veed needs to ensure that user data, including video content, is securely managed and protected from unauthorised access or misuse. It is important for users to check the outputs of Veed such as subtitles/captions for accuracy. Users should be informed and give consent regarding how their content is processed by AI, including features like background removal and transcription. Veed should also take greater steps to ensure that user data, including video content, is securely managed and protected from unauthorized access or misuse. Addressing these ethical considerations could further ensure that its AI is used responsibly and transparently.

Appendix 2.4: Music/Sound

Functionality	Transparency	Ethical Considerations

Suno

Suno is an AI-powered music generation tool designed to democratise music creation by allowing users to generate music, regardless of skill level. Users can input text prompts to generate full songs, complete with vocals, lyrics, and instrumental tracks. Users can specify genres, moods, and other musical elements to tailor the output to their preferences. For more control over the music creation process, users can use the custom mode to tweak inputs and make real-time adjustments to lyrics, song structure, and styles. This allows for a more refined and personalised output. Suno AI also supports the editing of generated tracks through processes like extending full tracks, or isolating and fragmenting different segments.

The transparency of Suno could be improved, as it does not explicitly detail the underlying AI mechanisms, data usage policies, or the specifics of intellectual property management. While the website offers an overview of its vision, it lacks comprehensive transparency about how user data is handled or the ethical boundaries around content creation. More detailed policies and clearer explanations of the algorithms and technologies used would provide users with a better understanding of the platform's inner workings.

Suno faces potential issues regarding intellectual property, especially when users create covers or remix existing music. The platform needs clear guidelines on how copyright material is managed to avoid legal disputes. Furthermore, it should address how personal data is stored and shared, as there is currently limited information about these aspects.

AudioSonic

AudioSonic offers advanced AI-powered text-to-speech (TTS) technology, allowing users to convert written text into customisable, natural-sounding speech. This facilitates the generation of lyrics, which may layered over backing and instrumental tracks produced elsewhere. It supports different languages and integrates seamlessly with existing content management systems, enhancing user engagement across multiple industries.

AudioSonic provides clear information about its features, such as its use of deep learning algorithms, customisation options, and integration capabilities. However, details about its data usage, privacy, and ethical guidelines could be more explicit.

Ethical concerns include potential misuse for deepfake audio creation and the platform's impact on industries reliant on human voice work. More comprehensive information regarding ethical safeguards would enhance trust.

AIVA

AIVA is an advanced AI tool designed to generate and assist in music creation. It supports over 250 musical styles and allows users to generate personalised compositions. It offers tools for editing, downloading, and exporting music in various formats, making it useful for artists and media creators. The platform also supports collaborations, enabling users to work on projects collectively.

AIVA clearly outlines its pricing plans, licensing options, and feature details, allowing users to understand the different subscription tiers and what rights they hold over the generated music. The platform also explains how users can utilise generated compositions, including for commercial use, depending on the chosen plan.

AIVA's ethical considerations include issues around copyright ownership, the impact on human musicians and the creative industry, questions of originality in AI-generated music, and concerns about the lack of emotional and cultural depth in compositions.

Soundful

Soundful is a robust AI-powered platform designed for creators, offering quick access to royalty-free music in various genres. The platform allows users to tailor compositions to their specific needs, streamlining the process of finding original music.

Soundful is upfront about its subscription plans, pricing, and the royalty-free nature of its music. Users clearly understand the ownership rights for the tracks they generate and how they can use them, whether for personal projects or commercial purposes. The platform provides transparent guidance on licensing, ensuring users know the extent of their rights.

Soundful raises questions about the impact on professional musicians and the creative process. There is concern over the potential displacement of human composers, especially as AI-generated music becomes more accessible and affordable. Soundful's model of generating royalty-free music could also affect traditional revenue streams for artists. However, by offering clear licensing terms, the platform mitigates some ethical concerns by ensuring fair use of its content and clarifying authorship rights.

Functionality	Transparency	Ethical Considerations

PlayHT

PlayHT offers advanced text-to-speech (TTS) and AI voice generation services, allowing users to create voiceovers in over 142 languages and accents. Its features include: voice cloning, custom pronunciations, and different speech styles. This makes it suitable for various use cases like podcasts, e-learning, gaming, and IVR systems.

PlayHT is relatively clear about the role of AI in its features, providing information regarding the rights users have over the AI-generated voices, and usage options. It specifies the types of licenses available for commercial use and provides detailed information on what users can legally do with their AI-generated voices.

The ethical concerns include the potential misuse of voice cloning, particularly regarding deepfake audio, and the impact on voiceover professionals. PlayHT addresses these concerns by emphasising responsible AI usage and offering full commercial rights for generated voices.

AutoPod

AutoPod offers the automated streamlining of various editing tasks such as cutting based on silence detection, multi-camera angle switching, and resizing videos for social media platforms. It is designed to be integrating with other editing programmes such as Adobe Premiere Pro, providing tools like auto frame adjustment and aspect ratio changes that save editors significant manual effort. The multi-camera tool is especially useful for podcast creators working with multiple angles, enabling smoother workflows.

AutoPod is transparent about what it offers. It clearly outlines its key features like automatic editing and provides demo videos, including detailed information about the use of AI in the running of these functions, allowing potential users to fully understand how the software will work within their setup.

AutoPod's focus lies in streamlining the editing process, rather than generating new content from scratch – as such, it avoids some of the ethical issues relating to copyright infringement and intellectual ownership of content. Additionally, AutoPod appears to respect user privacy, not engaging in unethical practices such as claiming ownership over user content or data.

Functionality	Transparency	Ethical Considerations
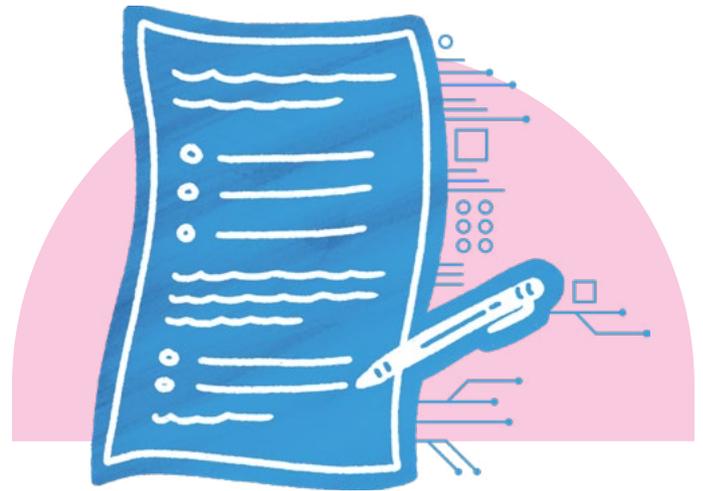
Appendix 3: Workshop Participant Surveys

Appendix 3.1: Pre-Workshop Survey

Please indicate your agreement for the Research Team to access and use your recorded responses to this questionnaire before continuing.

• I consent to take part in the survey on the basis of the Participant Information Sheet.

How old are you?

What gender do you identify with?

Choose one or more races that you consider yourself to be:

• White
• Black or African American
• American Indian or Alaska Native
• Asian
• Native Hawaiian or Pacific Islander
• Other
• Please specify if other

In order to ensure that everyone's needs are accommodated, please describe any neurodivergence you identify with and/or any accessibility needs you have.

Which of the following best describes your sector or profession?

• Tech developer
• Policymaker
• Academic
• Media Production
• Other

What is your role within media production?

How familiar are you with Generative AI?

• Not familiar at all
• Slightly familiar
• Moderately familiar
• Very familiar
• Extremely familiar

How do you feel about the use of generative AI in media production?

Please rate on a scale from 0 to 100, with 0 representing a completely negative opinion and 100 representing a completely positive opinion.

Please explain why you gave your answer.

What potential opportunities and risks do you see with the use of Generative AI?

How frequently do you use generative AI in your role in media production?

- Daily
- 2-3 times a week
- Once a week
- Once a month
- Less than once a month

What tools do you use?

How competent do you feel you are in using generative AI?

- Very competent
- Competent
- Somewhat competent
- Incompetent
- Very incompetent

How would you describe your experience using generative AI tools for your media production tasks?

Positive

- Strongly agree
- Agree
- Somewhat agree
- Disagree
- Strongly disagree

Easy

- Strongly agree
- Agree
- Somewhat agree
- Disagree
- Strongly disagree

Useful

- Strongly agree
- Agree
- Somewhat agree
- Disagree
- Strongly disagree

Reliable

- Strongly agree
- Agree
- Somewhat agree
- Disagree
- Strongly disagree

What does a 'reliable' generative AI tool mean to you?

What do you understand by the term 'responsible AI', considering its implications for your work?

Please indicate the importance of the following items to you and your role:

Collaboration

- Not at all important
- Slightly important
- Moderately important
- Very important
- Extremely important

Bias

- Not at all important
- Slightly important
- Moderately important
- Very important
- Extremely important

Labour

- Not at all important
- Slightly important
- Moderately important
- Very important
- Extremely important

Copyright

- Not at all important
- Slightly important
- Moderately important
- Very important
- Extremely important

Productivity

- Not at all important
- Slightly important
- Moderately important
- Very important
- Extremely important

Are you familiar with any (proposed) laws/policies/industry guidelines about the use of generative AI tools in media production?

- Yes
- No

Please specific and/or give an example.

What kind of developments would you like to see in Generative AI in the future?

Appendix 3.2: Post-Workshop Survey

Please indicate your agreement for the Research Team to access and use your recorded responses to this questionnaire before continuing.

- I consent to take part in the survey on the basis of the Participant Information Sheet.

Which days did you attend the workshop(s)?

- Workshop 1 (Screen-writing with AI) – Day 1 (Friday 26 April)
- Workshop 1 (Screen-writing with AI) – Day 2 (Saturday 27 April)
- Workshop 2 (Image Creation with AI) – Day 1 (Friday 24 May)
- Workshop 2 (Image Creation with AI) – Day 2 (Saturday 25 May)
- Workshop 3 (Editing with AI) – Day 1 (Thursday 13 June)
- Workshop 3 (Editing with AI) – Day 2 (Friday 14 June)
- Workshop 4 (Sound and Music with AI) – Day 1 (Friday 28 June)
- Workshop 4 (Sound and Music with AI) – Day 2 (Saturday 29 June)

After the workshop, how do you feel about the use of generative AI in media production?

Please rate on a scale from 0 to 100, with 0 representing a completely negative opinion and 100 representing a completely positive opinion.

After the workshop, has your attitude toward generative AI changed?

- Improved significantly
- Improved somewhat
- Remained the same
- Worsened somewhat
- Worsened significantly

Could you please explain why you gave your answer?

What new AI tools have you used since the workshop and in which context?

After the workshop, how has your competency in using generative AI tools changed?

- Improved significantly
- Improved somewhat
- Remained the same
- Worsened somewhat
- Worsened significantly

Did the workshop contribute to your understanding of 'responsible AI' in media production?

- Yes
- No

How did your understanding of 'responsible AI' in media production change after the workshop?

Did the workshop contribute to your understanding of 'creativity' in media production?

- Yes
- No

How did your understanding of 'creativity' in media production change after the workshop?

Did the workshop contribute to your understanding of 'bias and representation' in media production?

- Yes
- No

How did your understanding of 'bias and representation' in media production change after the workshop?

Did the workshop contribute to your understanding of 'collaboration' in media production?

- Yes
- No

How did your understanding of 'collaboration' in media production change after the workshop?

After the workshop, are you now more familiar with any (proposed) laws/policies/industry guidelines about the use of generative AI tools in media production?

- Yes
- No

Could you please explain why you gave your answer?

Did the workshop change your expectations of the future development of Generative AI in media production?

- Yes
- No

How have your expectations for the future development of Generative AI in media production changed since attending the workshop?

Please share any other thoughts or feedback you have.

Appendix 4: Expert Bridging Group Recommendations for Responsible AI

22nd July 2024 – 10:00-16:00 (BST), Bournemouth University

GAI specificity, transparency and scalability

- Develop a clear set of distinctions when talking about and defining GAI, since the term encompasses a broad range of tools and models produced under different circumstances.
  - For example, policy solutions designed to mitigate the impact of 'scraping' in LLM training processes may need to apply to popularised models like ChatGPT and Suno, but not necessarily to models created by individual people from scratch, trained on privately-owned data in which the developer owns the copyright of any and all material fed into the model.

- Greater transparency in how producers from big studios are using GAI in the decision-making process for commission or production, enabling audiences to challenge or ask questions if this does not align with their ethical values.

- Support for the publication of pre-set GAITs according to questions of cultural specificity and age restrictions e.g. Meta in the process of rolling out different models of its GAI for child appropriate use and differences for specific countries.

- Clearer guidance and guidelines to producers on how they can use GAI regarding current and future workload practices.

-
- The suggestion of legal agencies seeking permission from individuals before their work is ingested into AI models and creating a transparent framework for licensing their data.

- Clearer industry communication about the actual impact of AI on workflows and ensuring job security.

- A general need for more transparency about the sources of training data and whether users can attribute particular outputs to specific contributors.

- The requirement for a more comprehensive system of tracing training data back to its original sources. This would ensure greater transparency and help users make more informed decisions relating to their use of AI and distribution of AI-generated images.

Diversification of stakeholders

- Increased diversity of workforce at every stage in the GAI production process, but particularly for those involved in the training of LLMs, so as to combat bias and stereotype at its source, rather than relying on corrections after the release of the model.

- Facilitate and promote an open-source development community, whereby developers can engage with creatives so as to be made more aware of what risks exist within the GAI landscape for artists, writers and practitioners.

- Inclusion of various stakeholders (government, private sector, researchers, civil society) in discussions about the role of AI in decision-making processes and governance.

- Proposals to involve diverse risk management experts from multiple fields (e.g., social science, philosophy, medical fields) to help companies identify and mitigate risks in AI technology

- One speaker emphasized the importance of involving diverse expert stakeholders and ensuring public accountability in the process of clarifying intellectual property and data protection rules. Government cross-sector collaboration was mentioned as crucial.

Regulation and accreditation

- The establishment of a third-party organisation, interdisciplinary in nature (developers, academics, lawyers, philosophers, users), designed to evaluate and certificate 'responsible AI'.
  - This could also be a contractor where companies could go to seek guidance

on RAI practices, how to effectively manage AI risk, and further embed socio-cultural concerns into pre-existing technical expertise.
- An offshoot of this could be a 'Responsible AI' accreditation for feature and series productions, perhaps following the lead of BAFTA albert's environment and sustainability accreditation procedures. This could even be in partnership with a body like BAFTA, in which assessors provide 'bronze', 'silver' and 'gold' level accreditations to productions based on their adherence to responsible AI guidelines. However, this would need to be weighted differently for independent productions and big-budget studios.

- Take the lead from the EU AI Act, investigating the possibility of establishing a Risk Regulatory Framework (currently, the EU are only proposing regulations for 'High Risk' AI systems, and they have labelled GAI chatbots and GAI-produced content as 'Limited Risk').

- A recognition that, in every stage of the GAI production process, human collaboration and oversight is required to combat algorithmic decision making that risks perpetuating bias – a 'right to human review' embedded into legislation.

- A motion to introduce more AI into council policies, indicating the need for AI governance in public systems.

- Regulatory bodies currently lack sufficient enforcement capacity to effectively manage AI governance, which was identified as a critical issue. Strengthening these capabilities would enable regulators to enforce laws more robustly.

- Participants discussed the lack of formalized image rights in the UK, contrasting it with the US, where image rights are more structured. They hinted at the need for clearer regulations around the use of posthumous digital identities and how AI uses them, such as celebrity likenesses.

- Regulations should put in place to prevent large tech companies from exploiting workers, predominantly in the Global South, with cheap, outsourced labour used for content moderation, labelling and detecting bias and hate speech.

- Increased regulatory capacity and enforcement of any future laws pertaining to GAI, copyright and misuse, and make politicians accountable.

- Tighter restrictions and punishments for the use of voicemods or voice cloning outside of the parameters of initial purposes set out in any given contract.

- The idea that AI models should, by default, exclude non-copyright-controlled material unless explicit permission is granted, addressing AI-generated content through stricter legal frameworks.

- The need for insurers to accept AI-generated content only if it's trained on legally approved, copyright-controlled datasets.

Training and copyright

- A proposal that GAI media outputs cannot be legally considered as 'source material', meaning that any remuneration for intellectual property or 'format fees' for writers must be traced back to the original human inputs and human labourers behind said inputs.

- Calls for developers of popularised GAIT to publicise training databases and remove the 'black box' in LLM training processes, allowing end users to acknowledge copyrighted materials and helping facilitate transparency in the GAI media production process.

- Calls on large tech companies to retrain LLMs – this would require that we agree collectively on a responsible AI training rubric.

- A set of contractual regulations regarding the use of creative media in LLM training, including, but not limited to; sought permission from CMOs representing a group's democratically expressed position on the use of creative media in LLM training; remuneration for artists whose work has been trained on LLMs based on a percentage of revenue made as a result of AI-produced media; and a voluntary, opt-in licensing framework that is renegotiated every X years, so as to prevent assumed consent.
  - One potential issue with this is that CMOs could be said to set the agenda for the industry, making it harder for individual artists not part of said CMO to exert their own individual agency

- A clear set of contractual guidelines concerning identity and performative labour for actors that allows for authoristation and commercialisation where desired. This also needs to include pathways for performer agency and co-design in the AI media production process.

- Implementation of fair use policies regarding copyright in AI training models and clearer guidelines on how AI systems interact with copyrighted material in creative outputs.

- The concept of protocolism was raised, which would allow for artists to define new ways to use AI and receive rewards for their contributions. This calls for rethinking copyright laws to account for the role of AI in creative outputs.

- The necessity for GAI systems to provide sources for the data it utilises to generate synthesised content.

- The development of training to provide guidance on how to craft better AI prompts and encouraging human collaboration to ensure higher-quality outputs.

Awareness and education

- Greater awareness of the entrenchment of harmful gender, race, disability, sexuality, class bias and stereotype that can, and does, occur as a result of voice-, text-, image- and video-based GAI.

- Increased media literacy of labourers in the process of developing GAI models, including calls for mandatory education courses and awareness training on the potential impacts of biased GAI.

- The EBG raised the idea that technical experts may know how to correct bias in AI systems, but often lack the necessary understanding of how bias manifests in specific contexts. Further training could be implemented to instruct technical developers on issues relating to bias and representation and how they relate to the context of GAI.

- The importance of educating users and developers about the nuances of consent in AI-related contexts, especially when creating avatars or digital representations.

- Providing guidance to users on the responsible use of AI systems, including privacy issues, the scope of AI capabilities, and ethical concerns. This information should be featured prominently on sites which utilise GAI and allow users to generate outputs with GAI, thus allowing users to make informed decisions regarding how they use AI.

- Initiatives like CDT programs and the Safe and Trusted AI Centre could be integrated into higher education environments to educate students on responsible AI use.

- Embed education on GAI media production into primary education, teaching young people how to prompt and get better responses from AI with an emphasis on human collaboration.

- Concerns about the AI models being trained on biased datasets that reflect harmful historical and societal stereotypes were raised. This suggests the need for regulations around what kind of data is used for training AI systems and ensuring that diverse, representative, and accurate data is included.

Further research needed

- Further research into the environmental impact of GAI media production, since this is currently unknown. At present, the lack of clarity on this is placing the onus on end users to be 'conscious' without any set of parameters or data with which to compare their environmental impact.

- The concept of how AI deals with posthumous data (like the images or digital presence of deceased celebrities) is a new area that requires further research to understand its implications on ethics, law, and culture.

- Further exploration is needed on how consent can be withdrawn after it has been given in AI applications, particularly when individuals become uncomfortable with the uses of their likeness, movements, or voices in AI-generated content.

- Further exploration of meta-data for AI-generated content and how it could play a role in transparency and accountability. This indicates a need for research into how to track AI outputs, attribute content sources, and ensure ethical use of AI tools.

- A need for further research into how AI systems interact with labor rights and economic models within creative industries.

Appendix 5: Participant Agreement Form

The title of the research project
Shared Post-Human Imagination: Human-AI Collaboration in Media Creation

Invitation to take part
You are being invited to take part in a research project. Take time to decide whether you wish to take part. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

Who is organising/funding the research?
This research is funded by the Arts Humanities Research Council (AHRC) and is part of the Bridging Responsible AI Divides (BRAID) programme.

What is the purpose of the project?
The proposed research aims to investigate responsible AI (RAI) in the context of media production, focusing on collaboration, creativity and representation. The research outputs seek to foster feedback loops among developers, users and policymakers in order to develop RAI practices.

The main questions of our research are:

- How can we make sure that there is justice, transparency, accountability and safety regarding the large training databases based on which AI tools are built? Instead of algorithmic colonialism, these databases should be intersectional in all social categories such as race, class, gender and ability. Moreover, there should be a discussion between developers, industry and policymakers about ways in which feedback can be easily circulated and turned into action to lessen bias for a more responsible AI.

- How can we make sure that the collaborative work including AI tools is accountable and just? In this context, questions of copyright, IP and job security are essential. Moreover, questions of accessibility and responsible use are equally important. Who gets marginalised in these processes and how much awareness and flexibility is granted to users to recognise and supersede biases?

- How can we make sure that the augmentation of human creativity happens in a just and safe way? In other words, while exploring the creative possibilities that generative AI offers, we need to ask what and who gets excluded by computational logic.

In order to address these research aims, we will apply a hybrid, practice-theory methodology. In four workshops, we will simulate the process of media production (script-writing, image generation, audio/music generation and post-production), experiment with significant generative AI and subsequently reflect on our experience within the context of our research questions.

Why have I been invited?
We are inviting participants to take part who are experts in generative AI/or and media creation, largely drawing on these four main areas of media production, such as scriptwriting, (moving) image generation, audio/music generation, and editing and post-production.

Do I have to take part?
It is up to you to decide whether or not to take part. We want you to understand what participation involves, before you make a decision on whether to participate. If you or any family member have an on-going relationship with BU or the research team, e.g. as a member of staff, as student or other service user, your decision on whether to take part (or continue to take part) will not affect this relationship in any way.

Can I change my mind about taking part?
Yes, you can stop participating in study activities at any time and without giving a reason. After you decide to withdraw from the study, we will not collect any further information from or about you. As regards to the information we have already collected before this point, your rights to access, change or move that information are limited. This is because we need to manage your information in specific ways in order for the research to be reliable and accurate. Further explanation about this is in the Personal Information section below.

What would taking part involve?

Workshop/round table discussion participation: You will be asked to take part in one or all of the 4 workshops and the concluding round table discussion being held at BU in relation to generative AI tools and media creation. In the course of these four workshops, we will simulate the process of media creation. The purpose of the workshops is to scope specific generative AI tools by using them, to reflect on issues that arise and to prototype artefacts. The four workshops will cover four main areas of media production, such as scriptwriting, (moving) image generation, audio/music generation, and editing and post-production. We will seek your opinions about generative AI in media creation; any concerns or worries you have about its use; any hopes and expected benefits you see from its use; and your professional experience of working with these tools. As a conclusion, we will host a round-table discussion to which we will invite stakeholders from the media industry, policymakers (such as the respective guilds), developers and researchers to present the preliminary outcomes of our research and discuss the future of media production in the context of responsible AI.

Online survey: We will ask you to respond to a survey before and after the workshop and share your experience and opinion during the workshop. Participation in this study is on the basis of consent: you do not have to complete the survey, and you can change your mind at any point before submitting the survey responses. Please note that once you have completed and submitted your survey responses, we are unable to remove your anonymised responses from the study. Deciding to take part or not will not impact upon you. The online survey will take approximately 10-15 minutes to complete.

Will I be reimbursed for taking part?
Only invited panel experts and keynote speakers will be reimbursed for their time as well as their reasonable travel, accommodation and subsistence expenses.

What are the advantages and possible disadvantages or risks of taking part?
Whilst there are no immediate benefits for those people participating in the project, it is hoped that this work will enable the development of responsible AI practices in media creation.

What type of information will be sought from me and why is the collection of this information relevant for achieving the research project's objectives?
We will collect your opinions about generative AI in media creation; any concerns or worries you have about its use; any hopes and expected benefits you see from its use; and your professional experience of working with these tools.

Will I be recorded, and how will the recorded media be used?
The workshops will be video recorded, respectively the zoom sessions recorded online. The video recordings of your activities made during this research will be used only for analysis and the transcription of the recording(s) for illustration in conference presentations and lectures. No other use will be made of them without your written permission, and no one outside the project will be allowed access to the original recordings.

How will my information be managed?
Bournemouth University (BU) is the organisation with overall responsibility for this study and the Data Controller of your personal information, which means that we are responsible for looking after your information and using it appropriately. We will use your data on the basis that it is necessary for the conduct of research, which is an activity in the public interest. We put safeguards in place to ensure that your responses are kept secure and only used as necessary for this research study and associated activities such as a research audit.

Undertaking this research study involves collecting and/or generating information about you. We manage research data strictly in accordance with:

• Ethical requirements; and
• Current data protection laws. These control use of information about identifiable individuals, but do not apply to anonymous research data: 'anonymous' means that we have either removed or not collected any pieces of data or links to other data which identify a specific person as the subject or source of a research result.

BU's **Research Participant Privacy Notice** sets out more information about how we fulfil our responsibilities as a data controller and about your rights as an individual under the data protection legislation. We ask you to read this Notice so that you can fully understand the basis on which we will process your personal information.

Research data will be used only for the purposes of the study or related uses identified in the Privacy Notice or this Information Sheet. To safeguard your rights in relation to your personal information, we will use the minimum personally-identifiable information possible and control access to that data as described below.

Publication: You will not be able to be identified in any external reports or publications about the research without your specific consent. Otherwise your information will only be included in these materials in an anonymous form, i.e. you will not be identifiable. Research results will be published in a report with information, analysis and recommendations for best practices of the use of generative AI tools in media creation in the context of RAI. The outputs will be presented in a closing event to a group of media creators, policy makers and developers to foster dialogue and concrete actions.

Security and access controls: BU will hold the information we collect about you in hard copy in a secure location and on a BU password protected secure network drive held electronically. Personal information which has not been anonymised will be accessed and used only by appropriate, authorised individuals and when this is necessary for the purposes of the research or another purpose identified in the Privacy

Notice. This may include giving access to BU staff or others responsible for monitoring and/or audit of the study, who need to ensure that the research is complying with applicable regulations.

Sharing your personal information with third parties: As well as BU staff working on the research project, we may also need to share personal information in non-anonymised form with our project co-investigators at the Columbia University, USA, and Zhejiang University, China).

Further use of your information: The information collected about you may be used to support other research projects in the future and access to it will not be restricted. You will not be able to be identified in the data without your specific consent. To enable this use, the data will be added to an appropriate research data repository such as (BU's Online Research Data Repository): this is a central location where data is stored, which is accessible to the public.

Keeping your information if you withdraw from the study: If you withdraw from active participation in the study we will keep information which we have already collected from or about you, if this has on-going relevance or value to the study. This may include your personal identifiable information. As explained above, your legal rights to access, change, delete or move this information are limited as we need to manage your information in specific ways in order for the research to be reliable and accurate. However if you have concerns about how this will affect you personally, you can raise these with the research team when you withdraw from the study. You can find out more about your rights in relation to your data and how to raise queries or complaints in our Privacy Notice.

Retention of research data: Project governance documentation, including copies of signed participant agreements: we keep this documentation for a long period after completion of the research, so that we have records of how we conducted the research and who took part. The only personal information in this documentation will be your name and signature, and we will not be able to link this to any anonymised research results.

Research results: As described above, during the course of the study we will anonymise the information we have collected about you as an individual. This means that we will not hold your personal information in identifiable form after we have completed the research activities. You can find more specific information about retention periods for personal information in our Privacy Notice.

We keep anonymised research data indefinitely, so that it can be used for other research as described above.

Contact for further information: If you have any questions or would like further information, please contact Dr Szilvia Ruszev, Project Lead: **sruszev@bournemouth.ac.uk**

In case of complaints: Any concerns about the study should be directed to Professor Scott Wright, the Deputy Dean for Research & Professional Practice in the Faculty of Media and Communication, Bournemouth University by email to: **researchgovernance@bournemouth.ac.uk**

Appendix 6: Map of Generative AI Media Production Stakeholders

Developers

Big AI developers (OpenAI etc)

Software developers (FinalDraft etc)

Power

Influence

Influence

Producers / Distributors

Studios	Lobbyists
Producers	Film funds
Distributors	Commissioners
Film financing	Publishers
Ranking agencies	Film insurance companies
	Film festivals

Researchers

Academics   AI ethicists

Alignment

Policy Makers

Guilds	Unions
Bloggers	Journalist
Critics	Policy/Law makers
Advocacy groups	

Influence

Alignment

Alignment

Power

Creators

Translators	HoDs (DP, Production Designers)
Actors & performers	Screenwriters
Directors	Editors

Alignment

Appendix 7: BRAID Methodology

Research Design

Research Questions

| Creativity | Bias | Collaboration |

Data Gathering

4 Workshops / 200 Participants / 2 Days

| Scriptwriting Workshop | Image Creation Workshop | Editing / Post-Production | Music / Audio |

| Pre- and Post-Surveys | Interdisciplinary Presentations | Discussions | Miro Board |

Issues Identified

1 Expert Bridging Group

Outputs

| Policy Recommendations | Best Working Principles and Practices |

| Creative Artefact | Gen AI Interdisciplinary Glossary | Gen AI Tools Matrix |