

Large Language Models and EU Intermediary Copyright Liability: Quo vadis?

Zoi Krokida¹

I. Introduction

Large language Models (LLMs),² a subset of generative AI tools more usually referred to as AI foundation models or general-purpose AI models,³ are gaining momentum. Their services, which aim to generate direct text or images on a topic in response to users' prompts or to produce textual summaries of a search inquiry, have skyrocketed their popularity. For instance, ChatGPT, developed by OpenAI, has attracted more than 100 million users since the beginning of its operation in 2022,⁴ while Google's Bard has attracted 142.6 million users since May 2023; it has been estimated that by 2025 this figure will reach approximately 1 billion.⁵

However, LLMs appear to attract severe criticism with regard to their legal implications in the field of copyright law. Indeed, at judicial level, a number of lawsuits have been filed against LLMs alleging copyright infringements. In particular, in June 2023, two authors, Mona Awad and Paul Tremblay, initiated legal proceedings against OpenAI, arguing that the company used their copyrighted books in datasets which train LLMs.⁶ Another example is found in the Getty Images lawsuit against Stability AI where it is claimed there is a link between the unauthorized use of Getty Images in the training process of the LLM and the output that has been delivered.⁷ Finally, a class action against Copilot, an internet hosting repository for software development, has been filed which challenges its LLM training, arguing that 'the output is often a near-identical reproduction of code from the training data'.⁸

¹ Lecturer in AI, Innovation and Law at Brunel University London.

² Hereinafter LLMs.

³ General-purpose AI models is the terminology used in Article 3 (44b) of the EU AI Act

⁴ Reuters, <https://www.reuters.com/technology/openai-says-chatgpt-can-now-browse-internet-2023-09-27/>, last accessed on 29.9.2023

⁵ Google Bard Statistics & facts [July 2023], <https://www.mlyearning.org/google-bard-statistics-facts/>, last accessed on 29.9.2023

⁶ E. Martin, 'Bestselling authors Mona Awad and Paul Tremblay sue OpenAI over copyright infringement', <https://www.latimes.com/entertainment-arts/books/story/2023-07-01/mona-awad-paul-tremblay-sue-openai-claiming-copyright-infringement-chatgpt>, last accessed 18.3.2024

⁷ J. Vincent, 'Getty Images sues AI art generator Stable Diffusion in the US for copyright infringement' (Verge, 6 February 2023), <https://www.theverge.com/2023/2/6/23587393/ai-art-copyright-lawsuit-getty-images-stable-diffusion>, last accessed 18.3.2024; The Getty Images case are currently preoccupying the US and UK courts respectively. With regard to the UK court, the High Court in London accepted the copyright claims by Getty Images and the case proceeds to trial.

⁸ Complaint class action v Github, Microsoft, OpenAI in front of the US District Court of California, para. 46.

At policy level, in 2022, a study conducted on behalf of the European Commission revealed that creators and artists recognise the benefits that revolve around LLMs, but expressed concerns with regard to the risks associated with copyright law that might occur through their use.⁹ In addition, the European Guild for Artificial Intelligence Regulation voiced their concerns through a manifesto for the regulation of AI companies in Europe, pointing out that ‘... our data and intellectual properties are being exploited without our consent, on a scale never seen before’.¹⁰ In addition, French lawmakers recently proposed a number of reforms to the existing copyright law framework in order to address the legal issues surrounding LLMs,¹¹ meanwhile Spain has announced the establishment of the first supervisory agency for Artificial Intelligence.¹² Finally, on 13 March 2024, the EU AI Act has been voted by the European Parliament and it is expected to come into force in two years’ time.¹³

From a copyright law perspective, growing academic scholarship deals with copyright implications on generative AI and, in particular, on questions about AI authorship and ownership as well as copyright infringements in the context of the training datasets for LLMs.¹⁴ However, so far, limited scholarship has addressed the issue of liability for copyright infringement in the context of AI generative tools¹⁵ and, more specifically, whether LLMs could be held liable for copyright infringing outputs, and if so, what type of liability should be imposed.¹⁶

⁹ EU Commission, ‘Study on copyright and New Technologies: Copyright data management and artificial intelligence’ (February 2022).

¹⁰ <https://www.egair.eu/>

¹¹ Assemblée nationale, position de loi n°1630, https://www.assemblee-nationale.fr/dyn/16/textes/l16b1630_proposition-loi#, last accessed at 18.3.2024

¹² P. Solimano, ‘Spain Just Created the First European AI Supervision Agency’, (Decrypt, 23.08.2023), <<https://decrypt.co/153482/spain-just-created-the-first-european-ai-supervision-agency>, last accessed at 18.3.2024

¹³ EU Parliament, ‘EU AI Act: first regulation on artificial intelligence’ (19.12.2023), available at < [EU AI Act: first regulation on artificial intelligence | News | European Parliament \(europa.eu\)](https://www.europarl.europa.eu/news/en/artificial-intelligence/2023/12/19/eu-ai-act-first-regulation-on-artificial-intelligence) last accessed 18.3.2024

¹⁴ <https://www.forbes.com/sites/bernardmarr/2019/11/11/13-mind-blowing-things-artificial-intelligence-can-already-do-today/>

¹⁵ A. Guadamuz, ‘A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs’ (2024) 2 GRUR International 111; B. Arcila Botero, ‘Is it a Platform? Is it a Search Engine? It’s Chat GPT! The European Liability Regime for Large Language Models’ (2023) 3 Journal of Free Speech Law 2023; C. Novelli, F. Casolari, P. Hacker, G. Spedicato, L. Floridi, ‘Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity’ (15 March 2024) arxiv Computers and Society < <https://arxiv.org/abs/2401.07348v4>, last accessed 18.3.2024.

¹⁶ The seminal importance of the liability for copyright infringement is demonstrated by the recent initiatives decided by AI companies with the aim to protect their customers from being sued for copyright infringements. For example, Microsoft announced its Copilot Copyright commitment to support its commercial customers in face of copyright infringement litigation for using Microsoft’s Copilot for the output that they generate in B. Smith, ‘Microsoft announces new Copilot Copyright Commitment for customers’ (7 September 2023), <https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/#:~:text=Specifically%2C%20if%20a%20third%20party%20sues%20a%20commercial,content%20filters%20we%20have%20built%20into%20our%20products>, last accessed 18.3.2024; meanwhile, Getty Images announced that its new AI generative tool has not been trained on copyrighted content and any lawsuits on intellectual property disputes against their customers will be financially covered by the company in M. Heikilla, ‘Getty images promises its new AI doesn’t contain copyrighted art’ (MIT Technology Review, 25 September 2023), <https://www.technologyreview.com/2023/09/25/1080231/getty-images-promises-its-new-ai-doesnt-contain-copyrighted-art/>, last accessed 18.3.2024.

Moreover, in order to reflect on potential liability, it is important to define the legal classification of LLMs. While LLMs are considered general purpose AI models, as per the EU AI Act, there is a growing voice in academic scholarship that argue that they could also be considered intermediaries.¹⁷ This implies that intermediary liability provisions might be applicable in the context of LLMs. Therefore, one might wonder how copyright holders could enforce their rights.

Against this background, this article aims to shed light on the liability of LLMs for copyright infringement at the European level, and to reflect on a set of recommendations in order for copyright holders to safeguard their rights. In doing so, the article explores the rise of LLMs and discusses their different types based on their functional and technical aspects (Section II). Secondly, it examines the operation of LLMs with reference to copyright law and, in particular, it addresses the dilemma of copyright infringement, namely to what extent the content that is provided as an input for the training of the LLMs and the content that is generated as an output could be considered copyright infringing (Section III). Following that, the article deals with the legal classification of LLMs through the lens of EU AI Act, but also through the EU intermediary liability regime (Section IV). With regard to the latter, normative considerations of LLMs as intermediaries and relevant definitions will be discussed. Finally, the article delves into an analysis of potential types of liability, taking into consideration the EU AI Act, the Digital Copyright Directive and Digital Services Act (Regulation (EU) 2022/2065) and relevant case law from the CJEU as well as reflects on solutions that could accommodate the liability framework of LLMs for copyright infringement (Section IV).

II. The rise and categories of Large Language Models

The history of LLMs traces back in 1960s with the development of ELIZA, the first ever Chatbot, at MIT. It was an early natural language processing programme that mimicked a discussion with a therapist. Following that, natural language processing technologies were developed and provided more complex answers due to a greater amount of training data. In 2011, Google founded the brain project that enabled researchers to conduct sentimental analysis and which subsequently led to the development of deep learning technologies in 2017. The latter heralded the rise of the era of the more sophisticated

¹⁷ B. Botero Arcila, 'Is it a Platform? Is it a Search Engine? It's Chat GPT! The European Liability Regime for Large Language Models' (2023) 3 Journal of Free Speech Law 2023; where the author argues that ChatGPT could be seen as a search engine in the context of defamatory outputs; S. Stalla-Bourdillon, 'What if ChatGPT was much more than a chatbox? What if LLM-as-a-service was a search engine? (3 April 2023, Beep Peep!)', available at < <https://peepbeep.blog/2023/04/03/what-if-chatgpt-was-much-more-than-a-chatbox-what-if-llm-as-a-service-was-a-search-engine/>, last accessed 18.3.2024; J. Quintais, 'Generative AI, Copyright and the AI Act' (Kluwer Copyright blog, 9 May 2023), available at < https://copyrightblog.kluweriplaw.com/2023/05/09/generative-ai-copyright-and-the-ai-act/#_ftn1, last accessed 18.3.2024 where the author argues that LLMs might be fall under the definition of 'online interface of a platform'.

language LLMs and, therefore, the high-quality generated outputs we are now experiencing.¹⁸ For instance, a computer-generated novel was shortlisted for a literary competition in Japan in 2016,¹⁹ while a sci-fi film generated by AI tools participated in the annual film festival in London.²⁰ Also, the IBM's project debater is but proof that AI can shape stronger arguments and rebuttals in complex subjects to debate against human participants.²¹

LLMs operate on a probabilistic basis and generate new data. This means they can accurately predict the next word in a sequence based on the preceding words.²² This understanding has been reinforced by Franceschelli and Musolesi who explain in the context of Generative Deep Learning, 'that GDL involves the creation of a probabilistic model describing data of interest, from which we obtain new works through sampling'.²³ Meanwhile, Foster argues that, "sampling from this distribution allows us to generate new observations".²⁴ The operation of LLMs is based on their training with a large set of data that teaches the probability of word occurrences and patterns in the use of the language.²⁵ The set of data can be structured and unstructured, depending on the standards and parameters that have been used, or not used, during its collection. For instance, the dataset that has been used for the training of the GPT-3 stems from various sources, namely 60% from Common Crawl, 22% from WebText2, 8% from Books1, 8% from Books2, and 3% from Wikipedia.²⁶ The larger, or faster, the amount of data is, the better it can improve the operation of an AI generative tool. As O'Leary notes, 'under situations of large volumes of data, AI allows delegation of difficult pattern recognition, learning, and other tasks to computer-based approaches'.²⁷

Interestingly, the evolving nature of LLMs makes it difficult to divide them into separate categories, but broad classification could comprise the following models, namely the pre-trained, the multimodal, the fine-tuned and the multi-language ones. Pre-trained LLMs constitute state-of-the-art language

¹⁸ T.J. Sejnowski, 'Large Language Models and the Reverse Turing Test' (2023) 35 *Neural Comput* 312.

¹⁹ D. Lewis, 'An AI-Written Novella Almost Won a Literary Prize: A short novel co-written by humans and AI passed the first round of a Japanese literary contest' (Smithsonian Magazine, 28 March 2016), available <
<https://www.smithsonianmag.com/smart-news/ai-written-novella-almost-won-literary-prize-180958577/>, last accessed 18.3.2024; C. Craig and I. Kerr, 'The Death of the AI Author' (2019) Osgoode Legal Studies Research Paper 32.

²⁰ D. Lewis, D. (2016, 28 March). An AI-Written Novella Almost Won a Literary Prize. Smithsonian Magazine (28 March 2016), <https://www.smithsonianmag.com/smart-news/ai-written-novella-almost-won-literary-prize-180958577>, last accessed on 13.10.2023.

²¹ IBM, Project Debater, <https://research.ibm.com/interactive/project-debater/>

²² Z. Ghahramani, 'Probabilistic machine learning and artificial intelligence' (2015) 521 *Nature* 452.

²³ G. Franceschelli & M. Musolesi, 'Copyright in generative deep learning' (2022) *Data & Policy* e 17- 8.

²⁴ D. Foster, *Generative Deep Learning: Teaching Machines To Paint, Write, Compose, and Play* (O'Reilly, 2023)

²⁵ Ibid.

²⁶ T. Brown and others, 'Language Models are Few-Shot Learners' (22 July 2020) arXiv Computers and Society, available at <https://arxiv.org/pdf/2005.14165.pdf>; see also M. Diaz, 'How to block OpenAI's new AI-training web crawler from ingesting your data (ZDNET, 2023), available at <https://www.zdnet.com/article/how-to-block-openais-new-ai-training-web-crawler-from-ingesting-your-data/>, last accessed on 18.3.2024.

²⁷ D. E. O'Leary, 'Artificial Intelligence and Big Data' (2013) 28 *IEEE* 97.

processing systems that generate human-like texts. A representative example can be found in ChatGPT which has been trained with 499 billion tokens of data. As Guadamuz explains, a token in this model is a sequence of characters that are found in the body of the training data and consist of four letters or numbers.²⁸ Another example of a pre-training model is Jasper, an AI tool that can generate content from a headline of an article and that may be comparable to the content produced by professional writers.²⁹ Its purpose is to generate Facebook ads, marketing emails, and descriptions of products.³⁰

As far as the multimodal models are concerned, they constitute a combination of the generation of text with images or videos in order to perform multimodal tasks.³¹ Representative examples are the DALL* E- 3 or Midjourney. These models are trained to understand the relationship between images and texts and to generate images as a response to textual prompts by users, and vice versa. For instance, DALL-E is based on the portmanteau of the surrealist Salvador Dali and Pixar's WALL-E, the robot protagonist of an animation film.³² It was released in January, 2021 and generates images from textual prompts initiated by users.³³ The second version of the DALL-E2 was released in April, 2023 and seems to be more sophisticated while DALL-E3 was released in October, 2023.³⁴ Another example is Midjourney which generates images from a textual description made by users.³⁵ Its operation is based on machine learning technologies that can understand textual descriptions and translate them into numerical versions. At the same time, the use of diffusion models, which are a type of deep generative model, can add noise into the dataset of images and generate images by reversing the noise. In that vein, MusicLM aims to generate high fidelity music from textual descriptions, such as songs with a textual description of "soulful jazz for a dinner party".³⁶

²⁸ A. Guadamuz, 'A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs' (2024) 2 GRUR International 111.

²⁹ G. Frosio, 'The artificial creatives: the rise of combinatorial creativity from DALL-E to GPT-3' in Martha Garcia-Murillo, Ian MacInnes, and Andrea Renda (eds), Handbook of Artificial Intelligence at Work: Interconnections and Policy Implications (Edward Elgar, forthcoming).

³⁰ T. Simonite, 'The Future of the Web Is Marketing Copy Generated by Algorithms: The killer app for GPT-3 could help marketers lure clicks and game Google rankings.' (18 April 2022, Wired) is available at <<https://www.wired.com/story/ai-generated-marketing-content/>> last accessed 18.3.2024.

³¹ Shukang Yin, Chaoyou Fu, Sirui Zhao, Ke Li, Xing Sun, Tong Xu, Enhong Chen, 'A Survey on Multimodal Large Language Models' (2023) arxiv Computer Vision and Pattern recognition <[arXiv:2306.13549](https://arxiv.org/abs/2306.13549)>, last accessed 18.3.2024.

³² <https://openai.com/research/dall-e>, last accessed 18.3.2024

³³ C. Newton, 'How DALL-E could power a creative revolution: Thoughts on my first week with OpenAI's amazing text-to-image AI tool' (The Verge, 10 June 2022), available at <https://www.theverge.com/23162454/openai-dall-e-image-generation-tool-creative-revolution>> last accessed on 18.3.2024.

³⁴ P. Delobelle and others, 'RobBERT: a Dutch RoBERTa-based Language Model' (2020), available at <<https://arxiv.org/pdf/2001.06286.pdf>>, last accessed 13.10.2023

³⁵ C. Wankhede, 'What is Midjourney AI and how does it work? How does a state-of-the-art AI image generator work? We break it down.' (Android Authority, 29 September 2023), available at <https://www.androidauthority.com/what-is-midjourney-3324590/>, last accessed 18.3.2024.

³⁶ Turn ideas into music with MusicLM, <https://blog.google/technology/ai/musiclm-google-ai-test-kitchen/>; Also see <https://musiclm.github.io/>, it also states that 'MusicLM casts the process of conditional music generation as a hierarchical sequence-to-sequence modeling task, and it generates music at 24 kHz that remains consistent over several minutes'.

Fine-tuned LLMs incorporate a large training dataset but are mainly used for specific tasks, such as to conduct sentiment analysis and text classification. Telling examples are the Bert, Roberta, and Albert models. Roberta conducts sentimental analysis—otherwise known as opinion mining—extracting and identifying the emotional tone of a text by analysing each word; a case study on Twitter revealed that Roberta can identify whether a tweet is emotively positive, negative, or neutral.³⁷ Finally, multi-language models are conducting textual analysis in multilingual datasets. For instance, XLM- Roberta is trained on data consisting of 100 languages and it is mainly used for the analysis of non-English texts.³⁸

Overall, it can be concluded that there are different types of LLMs with highly sophisticated performances and services to users. Yet, it seems that a number of lawsuits have been filed against the LLMs, arguing copyright infringement. Should the judiciaries confirm that there is a copyright infringement, LLMs might be subject to claim damages, and therefore liability.

Nevertheless, before addressing their potential liability, it is important to assess whether a copyright infringement might occur in the context of their operation. Such discussion is of seminal importance to the extent that finding copyright violations would lead to the liability of LLMs; otherwise, lack of copyright infringement would not establish liability on the part of LLMs.

Given there are different categories of LLMs, this article will mainly focus on the operation of LLMs that generate outputs in response to users' prompts for multiple occasions and therefore address their legal classification and liability rules for copyright infringing outputs.

III. The dilemma of copyright infringement in the context of LLMs

From a copyright law perspective, any reproduction of the work or communication to the public of the work or distribution of the work without the authorization of the copyright holder constitutes copyright infringement. At the European level, this understanding has been set forth in Articles 2, 3, and 4 of the InfoSoc Directive.³⁹ In the context of LLMs, a potential copyright infringement might take place within their operation. This means that it is necessary to examine the content that is provided as an input for the training of the LLMs, but also the content that is generated as an output.

³⁷ M. Rotulo, ‘

Crucially, while the main focus of this article is to address whether there is a copyright infringement in the output generated by LLMs, it seems necessary to discuss stances of potential copyright infringement in the datasets that have trained the AI generative tool. This is because a causal connection between input and output might exist.⁴⁰ In simple terms, if a copyright infringement occurs in the training dataset, it is highly likely that a copyright infringing output would be generated by the LLMs. This understanding has been reinforced by a number of scholars. For instance, Samuelson notes that ‘when many ingested works depict the same image (e.g., Mickey Mouse), the LLM may memorize that image and produce infringing outputs’.⁴¹ In addition, Guadamuz reports that AI generative tools can reproduce outputs of famous characters after textual prompts from users, and therefore copyright violations might take place. This is because the output could be infringing the character’s copyright as well as reaching the ‘well-known’ mark.⁴²

Moreover, in order to assess whether a work infringes copyright, a comparison between the original work and the copyright infringing work must be undertaken. Consider, for instance, the seminal case of *Temple Islands Collection* where Judge Birss concluded that two images in the case exhibited striking similarities and noted, ‘there are a myriad of ways in which a bus could be portrayed in front of the Houses of Parliament that would not have been inappropriately based upon the claimant’s work yet the defendants have done so in a way which is very similar indeed to the claimant’s work’.⁴³ Therefore, the evaluation of output which contains a potential copyright infringement presupposes a pre-existing work so as to enable the comparison between the allegedly infringing work and the pre-existing work.

A. Training data as input

The input of LLMs is about the data that is collected from different sources in order to train the AI tool. Such training data consists of books, articles, and websites that either belong to the public domain or are copyright protected. For instance, a report conducted by the Washington Post examined the dataset C4, which has been used to train some of Google’s and Meta’s AI generative models.⁴⁴ The

⁴⁰ See also the complaint by Getty Images against Stability AI in para. 52 to the US District Court. ‘In many cases, and as discussed further below, the output delivered by Stability AI includes a modified version of a Getty Images watermark, underscoring the clear link between the copyrighted images that Stability AI copied without permission and the output its model delivers.’

⁴¹ P. Samuelson, ‘Generative AI meets copyright’ (26.4.2023, YouTube), available at <https://www.youtube.com/watch?v=6sDGlrVO6mo&t=1058s>, last accessed 18.3.2024.

⁴² A. Guadamuz, ‘A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs’ (2024) 2 GRUR International 111.

⁴³ *Temple Island Collections Ltd v New English Teas Ltd & Anor* [2012] EWPCC 1 (12 January 2012), para. 11

⁴⁴ N. Diakopoulos, Finding Evidence of Memorized News Content in GPT Models, (Medium, 5 September), available at < <https://generative-ai-newsroom.com/finding-evidence-of-memorized-news-content-in-gpt-models-d11a73576d2>, last accessed 18.3.2024.

report demonstrated that business and industrial websites held the first place in terms of the amount of data in the dataset, while the “News and Media” category of sites held the third place. Similarly, Anthropic, a recently launched LLM, collects data from publicly available information via the Internet, datasets that the LLM license from third party businesses, and data that LLM’s users or crowd workers provide.⁴⁵

Crucially, some of the data is copyright protected and a risk of copyright infringement might occur. This is mainly because collecting data for training involves mining text and other data sources. The mining method presupposes that a copy of the data has been made in order that information may be accessed, read, stored, analysed, and extracted. Lack of authorization from the rights holder to make a copy of the work could trigger accusations of copyright infringement as per Article 3 of the InfoSoc Directive.

At judicial level, an increasing number of lawsuits against AI providers reveal the copyright risks that might exist in the use of data for training LLMs. More specifically, in November 2022, a programmer filed a class action against Microsoft and OpenAI for designing the Copilot platform. Copilot’s aim is to provide autocomplete suggestions after analysing billions of lines of computer code. In this light, the programmer argued that this training process violates his right to authorship and his contribution to the development of the original code.⁴⁶ Further, in July 2023, the authors Sarah Silverman, Christopher Golden, and Richard Kadrey sued OpenAI and Meta in the US District Court arguing that their copyrighted books had been used as training material for ChatGPT without their consent.⁴⁷ While the case is still pending, it is worthy of note that in the motion to dismiss the authors’ claims, OpenAI and Meta never asserted that they did not use the works of authors for training their data. Finally, in March 2024, the French Regulator fined Google with €250m for using copyrighted content from publishers and media agencies without their permission in order to train its AI Chatbot.⁴⁸

At policy level, more than 10,000 authors signed an open letter from the Authors Guild calling on all AI Industry Leaders to protect writers. More specifically, the president of the Authors Guild said,

the output of AI will always be derivative in nature. AI regurgitates what it takes in, which is the work of human writers. It’s only fair that authors be compensated for having ‘fed’ AI and

⁴⁵ Anthropic, <https://support.anthropic.com/en/articles/7996885-how-do-you-use-personal-data-in-model-training>, last accessed 18.3.2024.

⁴⁶ A. Wilkin, ‘Microsoft’s Copilot code tool faces the first big AI copyright lawsuit’ (NewScientist, 8 November 2022), [Microsoft, GitHub and OpenAI face the first big AI copyright lawsuit over Copilot | New Scientist](https://www.newscientist.com/article/mg25631010010-microsoft-github-and-openai-face-the-first-big-ai-copyright-lawsuit-over-copilot/), last accessed 18.3.2024.

⁴⁷ W. Davis, ‘Sarah Silverman is suing OpenAI and Meta for copyright infringement’ (The Verge, 9 July 2023), <https://www.theverge.com/2023/7/9/23788741/sarah-silverman-openai-meta-chatgpt-llama-copyright-infringement-chatbots-artificial-intelligence-ai>, last accessed 18.3.2024.

⁴⁸ Angelique Chrisafis, ‘Google fined €250m in France for breaching intellectual property deal’ (The Guardian, 20 March 2024) is available at <https://www.theguardian.com/technology/2024/mar/20/google-fined-250m-euros-in-france-for-breaching-intellectual-property-rules>

continuing to inform its evolution. Our work cannot be used without consent, credit, and compensation. All three are a must.⁴⁹

However, while the use of data in order to train AI models might have a detrimental effect on creators' rights, it has been argued that the use of in-copyright works to train LLMs might fall under the Text and Data Mining exception as set forth in Article 4 of the Digital Copyright Directive.⁵⁰ More specifically, Article 4 (3) notes that,

the exception or limitation provided for in paragraph 1 shall apply on condition that the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online.

This means that AI commercial businesses can reproduce and extract data in order to use them for the training purposes of the AI tools under the requirement that they have lawful access to the data.

Nevertheless, as per Article 4 (3), if rights holders do not wish for their data to be used in order to train the AI generative tools, they can evoke the opt-out mechanism. More specifically, this opt-out option gives rights holders the opportunity to withdraw their data from the training datasets. The option has already been used by a number of AI tools, such as spawning.ai which created the HavelBeenTrained website with more than one billion requests for the removal of pieces of work from the Stable Diffusion training dataset.⁵¹ Indeed, the website's purposes state that,

With HavelBeenTrained, artists can search these databases for links to their work and flag them for removal. We're incorporating new datasets as they are released and we're also partnering with other organizations who collect and use image links, so HavelBeenTrained can serve as a once only opt-out tool that applies to every dataset used to train generative AI Art tools.

Yet, this opt-out option does not come without limitations⁵² because it is only applicable to training datasets that are publicly available. This means that rights holders would need to be aware that their

⁴⁹ Authors Guild, 'More than 15,000 Authors Sign Authors Guild Letter Calling on AI Industry Leaders to Protect Writers' (18 July 2023), <https://authorsguild.org/news/thousands-sign-authors-guild-letter-calling-on-ai-industry-leaders-to-protect-writers/>, last accessed 18.3.2024.

⁵⁰ Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market.

⁵¹ K.Wiggers, 'Spawning lays out plans for letting creators opt out of generative AI training' (Techcrunch, 3 May 2023), available at <https://techcrunch.com/2023/05/03/spawning-lays-out-its-plans-for-letting-creators-opt-out-of-generative-ai-training/>, last accessed 18.3.2024.

⁵² J. Quintais, 'Generative AI, Copyright and the AI Act' (KluwerCopyright blog, 9 May 2023), available at < https://copyrightblog.kluweriplaw.com/2023/05/09/generative-ai-copyright-and-the-ai-act/#_ftn1, last accessed 18.3.2024.

data are being used in the training datasets. Otherwise, lack of knowledge would deprive them of the opportunity to opt their data out of training the LLMs.

B. AI-generated outputs

It has been argued that a copyright infringement might exist in output generated by LLMs. This is because there might be a risk of replication between the data that has been used in order to train the AI model and the output.⁵³ Consider, for instance, cases of identity, or cases of similarity between the output and one, or several, of the inputs, might occur.⁵⁴ Such cases could trigger copyright infringement. More specifically, under national laws, a copyright infringement could be established if a work is substantially similar to a work that forms part of the training data. Telling examples can be found in the UK and Germany. In the UK, as per S. 16 (3) of the CDPA 1988, one of the requirements to assess copyright infringement is whether the defendant has copied the whole work or used a substantial part of the protected aspects of the work. In Germany, the assessment of copyright violations is left to the courts' examination, but the German courts recognize a presumption of copying if there are striking similarities between the two works.⁵⁵

However, it has been argued that the risk of replication is limited. Indeed, Lemley notes in his seminal article that 'the circumstances in which the output is similar to any given input will be rare'.⁵⁶ Likewise, Guadamuz argues that '... the owners of a replicated image... could sue for copyright infringement, but the rates described in the papers would make it difficult for any random rightsholder to find some replication of their work.'⁵⁷

Yet, Elkin-Koren, citing a number of studies, points out that generative AI tools can reconstruct specific copyright works that form part of the generative AI dataset.⁵⁸ More specifically, a case study on Stable Diffusion, conducted by computer engineers, reveals that strong replication rates can be performed in smaller datasets with limited likelihood of duplicated images.⁵⁹ As the experiment has indicated, a random image from the dataset is duplicated 11.6 times, while a typical matched image is only

⁵³ See footnote 25.

⁵⁴ See footnote 47.

⁵⁵ Albrecht Conrad and Fabian Seip, 'Copyright litigation in Germany: overview', available at <[https://uk.practicallaw.thomsonreuters.com/w-011-3790?transitionType=Default&contextData=\(sc.Default\)>](https://uk.practicallaw.thomsonreuters.com/w-011-3790?transitionType=Default&contextData=(sc.Default)>), last accessed 13.10.2023

⁵⁶ M. A. Lemley, 'How Generative AI Turns Copyright Upside Down' (2023) 3, available at SSRN: <https://ssrn.com/abstract=4517702>, last accessed 18.3.2024.

⁵⁷ See footnote 25.

⁵⁸ H. Wasserman Rozen, N. Elkin- Koren and R. Gilad- Bachrach, 'The case of explainability' (20 May 2023) arxiv Artificial Intelligence <[arXiv:2305.12167](https://arxiv.org/abs/2305.12167)>, last accessed 18.3.2024.

⁵⁹ G. Somepalli, 'Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models' (12 December 2022) arxiv Machine Learning < <https://arxiv.org/abs/2212.03860>, last accessed 18.3.2024.

duplicated 3.1 times. Therefore, as the scientists have stated in relation to generative AI tools, they can 'blatantly copy from their training data'.⁶⁰

What is more, a potential replication might be triggered due to risks of memorization of inputs. From a legal perspective, whilst memorization is not unlawful it might lead to copyright infringement. Indeed, a number of studies have demonstrated the risks of memorization of training data in the dataset. For instance, a study conducted by computer scientists on neural language models revealed that a six billion parameter GPT-J model memorized at least 1% of its training dataset.⁶¹ In similar fashion, another study evaluated the training dataset in ChatGPT and GPT-4 and concluded that '... OpenAI models have memorized a wide collection of copyrighted materials, and that the degree of memorization is tied to the frequency with which passages of those books appear on the web'.⁶² Finally, this understanding has been reinforced by a recent experiment conducted by Diakopoulos. To test the risks of memorization, Diakopoulos used unique text that appears alongside every opinion that is published in the New York Times.⁶³ It reads as follows:

The New York Times journal include a unique text at the end of every opinion article that is published on the journal, "The Times is committed to publishing a diversity of letters to the editor. We'd like to hear what you think about this or any of our articles. Here are some tips. And here's our email: letters@nytimes.com."

When entering the prompt, 'The Times is committed to publishing', the AI model generated the rest of the text as an output which matched the original text. Therefore, one might conclude that a copyright infringing work used as an input would run the risk of a potential copyright infringing work as an output.

Against this background, it remains to be seen how the EU courts approach the dilemma of copyright violations in LLMs, in particular in the context of the output generated by the AI tools. Should the judiciaries confirm copyright infringement in the output, potential liability for LLMs might arise. Yet, in order to address the potential liability that might be ascribed to LLMs, it is warranted for identification to be made of their legal classification—mainly because different frameworks of liability are contingent upon their legal nature. To do so, the following section provides a critical account of the legal classification of LLMs.

⁶⁰ Ibid.

⁶¹ N. Carlini and others, 'Quantifying memorization across neural language models' (6 March 2023) arxiv Machine Learning <<https://arxiv.org/abs/2202.07646>, last accessed 18.3.2024.

⁶² K. Change and others, 'Speak, Memory: An Archaeology of Books Known to ChatGPT/GPT-4' (20 October 2023) arxiv Computation and Language <<https://arxiv.org/abs/2305.00118>, last accessed 18.3.2024.

⁶³ N. Diakopoulos, Finding Evidence of Memorized News Content in GPT Models, (Medium, 5 September), available at <<https://generative-ai-newsroom.com/finding-evidence-of-memorized-news-content-in-gpt-models-d11a73576d2>, last accessed 18.3.2024.

IV. Demystifying the legal classification of LLMs

LLMs generate texts, music, or images in response to textual prompts from users. However, their legal classification is subject to debate, mainly due to their training operation and the generative services they provide their users. In this light, LLMs could be seen as general-purpose AI models, but at the same time it has been argued that LLMs might also be considered as intermediaries between users and the output provided.

A. LLMs as foundation Models

To begin with, LLMs' services⁶⁴ seem to be covered by the definition of general-purpose AI models⁶⁵ as set forth in the EU AI Act.⁶⁶ As per Recital 99,

'Large generative AI models are a typical example for a general purpose AI model, given that they allow for flexible generation of content (such as in the form of text, audio, images or video) that can readily accommodate a wide range of distinctive tasks.'

Recital 105 further explains that,

'General-purpose models, in particular large generative models, capable of generating text, images, and other content, present unique innovation opportunities but also challenges to artists, authors, and other creators and the way their creative content is created, distributed, used and consumed. The development and training of such models require access to vast amounts of text, images, videos, and other data.....'

Indeed, a number of opportunities and risks have been declared by different stakeholders in light of the use of LLMs. For instance, a report commissioned by the UK Competition and Markets Authority demonstrates that LLMs might trigger opportunities as well as challenges for consumer protection, labour market, misinformation, intellectual property laws, data protection and investments in the public cloud infrastructure.⁶⁷ What is more, LLMs' training datasets consist of different sources. A

⁶⁴ Large Language Models (LLMs) constitute a sub-group of Large Generative AI models.

⁶⁵ As per Article 3 (44b), 'general purpose AI model' means an AI model, including when trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable to competently perform a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications.' This is also in line with the OECD's definition on Generative AI models, "An AI system is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments'.

⁶⁶ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts 2021/0106(COD) Draft (Final draft as updated on 21/01/ 2024).

⁶⁷ CMA, 'AI Foundation Models: Initial review Competition and Markets Authority' (May 2023) 7.

telling example can be found in LamDa which comprises of 12.5% C4-based data, 12.5% English language Wikipedia, 12.5% code documents from programming Q&A websites, tutorials, and others, 6.25% English web documents, 6.25% Non-English web documents and 50% dialogs data from public forums.⁶⁸

B. LLMs as intermediaries

Apart from general purpose AI models, it has been argued that LLMs might perform a mediatory function between third parties and therefore could be seen as intermediaries.⁶⁹ This is because they might be viewed as middle entities between their users and the third-party content that is displayed in response to textual prompts. In this light, the following section explores a set of reasons that justify the consideration of LLMs as intermediaries.

i. Normative considerations

Firstly, intermediaries could be seen as a source of deterrence of the wrongdoings. This is summarized under the umbrella of gatekeeper theory developed by Kraakman,⁷⁰ which views intermediaries as those entities that can control the information and content that pass through their gates as well as design the world by choosing which information will reach or not the end-users. Drawing parallels with the gatekeeper theory, LLMs can control the information that end-users will receive following their textual prompts. In addition, they are also responsible for deciding the content of the summaries or the answers that reach their users. For instance, the output delivered to users stems from the dataset that has been used as an input in order to train the algorithm of the LLMs.

Secondly, through the lens of the gatekeeper theory, intermediaries can detect wrongdoings at a reasonable cost. This means that intermediaries can identify illicit activity on their platforms faster and more efficiently than other parties, such as the copyright holders. This understanding has been described by Lichtman and Posner in their seminal article where they argue that intermediaries can mitigate the increasing amount of malicious software that passes through their gates.⁷¹ In similar

⁶⁸ R. Thoppilan and others, 'LaMDA: Language Models for Dialog Applications' (10 February 2022) arXiv Computation and Language < <https://arxiv.org/abs/2201.08239>, last accessed 18.3.2024.

⁶⁹ H. Ariyaratne, 'ChatGPT and Intermediary Liability: Why Section 230 Does Not and Should Not Protect Generative Algorithms' (2023) 23, available at SSRN: <https://ssrn.com/abstract=4422583> where it has been stated that 'there are instances—such as when ChatGPT quotes directly from a source—that make them more of an intermediary of third-party speech'.

⁷⁰ RH Kraakman, 'Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy' (1986) 2 *Journal of Law, Economics and Organization* 61.

⁷¹ D Lichtman and E Posner, 'Holding Internet Service Providers Accountable' (2006) 14 *Supreme Court Economic Review* 221-222.

fashion, Mann and Belzley argue that the technological filters seem to be the cheapest tool to detect online wrongdoings more efficiently.⁷²

Transposing the last-cost benefit approach into the context of LLMs, it appears that LLMs might be in a better position than copyright holders to detect unauthorized content at a reasonable cost. At the same time, LLMs could resort to filtering mechanisms that enable them to detect the copyright infringement and exclude the illicit content from the generated output. For instance, Azure OpenAI Service already operates a content filtering system that detects and prevents any illegal content in the input or output of the LLMs.⁷³

Finally, from a moral perspective, intermediaries can be considered morally liable due to material cooperation theory. As per that theory, intermediaries design their business model for legitimate purposes but accept infringements as unwanted side effect.⁷⁴ Similarly, LLMs have been designed to pursue legitimate aims but in the course of their operation, illicit activities, such as copyright infringements might appear as unwanted side effects.

Having discussed the normative considerations for LLMs as intermediaries, the next question concerns the potential definitions of intermediaries in light of the relevant legal provisions in the Digital Copyright Directive and the EU Digital Services Act.

ii. Digital Copyright Directive

At first glance, as far as the Digital Copyright Directive is concerned, it seems unclear whether LLMs would fall into the scope of online content sharing providers. More specifically, Article 3 (6) defines them as ‘online content-sharing service providers’, meaning;

a provider of an information society service of which the main or one of the main purposes is to store and give the public access to a large amount of copyright-protected works or other protected subject matter uploaded by its users, which it organises and promotes for profit-making purposes.

⁷² RJ Mann and SR Belzley, ‘The Promise of Internet Intermediary Liability’ (2005) 47 William and Mary Law Review 268.

⁷³ See also <https://learn.microsoft.com/en-us/azure/ai-services/openai/concepts/content-filter?tabs=warning%2Cpython>, last accessed 18.3.2024.

⁷⁴ R.A. Spinello, ‘Intellectual Property: Legal and Moral Challenges of online file-sharing’ in K. Einar Himma and H. Tavani (eds.), Handbook of Information and Computer Ethics (2008) 565.

Crucially, from a formalistic interpretation, the services of AI models do not fit the definition of online content sharing service providers. This is because they do not host user generated material. Rather, their services host content generated by AI models.⁷⁵

Yet, from a purposive approach, one could argue that LLMs might fit within the definition of online content sharing service providers. Indeed, the Digital Copyright Directive refers to one of its aims, in Recital 2, which is to achieve a well-functioning and fair marketplace for copyright. Given that the operation of LLMs is mainly based on training datasets that might consist of copyright content, a purposive interpretation would be warranted. More specifically, it could be argued that LLMs might communicate the works to the public after textual prompts from users.

However, even if the latter interpretation is followed, it seems difficult to satisfy the requirement of the ‘public’. This is because AI generated content is not accessible to the public, rather it is only accessible to the user who prompted the output. In this way, this understanding fails to meet the threshold for the concept of public as concluded in the EU landmark case of *Rafael Hotel*.⁷⁶ In that case, the CJEU examined the preliminary reference as to whether hotel guests watching TV in their rooms fall within the scope of the right of communication to the public. The CJEU offered an interpretation of the concept of ‘public’; by setting the threshold in paragraph 30 of the judgement, ‘as a general rule, a fairly large number of persons are involved, so that they may be considered to be a public, having regard to the principal objective of Directive 2001/29’. In the context of LLMs, the limited number of users might not satisfy the threshold of the concept of ‘public’, unless that content is posted in an online platform, and therefore gains greater visibility. On the other hand, there is a risk where a great number of users might input the same or similar prompt and as a result the same or similar output might be delivered, thus reaching the threshold of ‘fairly large number of persons’.

iii. Digital Services Act

Finally, whether LLMs might fall under the provisions of Digital Services Act would also be subject to debate. This is because the services provided by LLMs appear to be difficult to fit within the scope of hosting providers. More specifically, as per Article 6, ‘a hosting service provider is an information society service ... that consists of the storage of information provided by a recipient of the service ...’. Under a formal interpretation, in the context of LLMs, this definition might not be applicable. This is because LLMs do not host content provided by the recipients of the service. In similar fashion, Hacker and others argue that the provisions on Digital Services Act do not fit with the services provided by

⁷⁵ P. Hacker, A. Engel, and M. Mauer, ‘Regulating Chat GPT and other Large Generative AI Models’ (2023) ACM Conference on Fairness, Accountability, and Transparency 1118.

⁷⁶ S. Karapapa, ‘The requirement for a “new public” in EU copyright law’ (2017) 1 European Law Review 70.

hosting services because ‘while users do request information from LGAIMs via prompts, they can hardly be said to provide this information. Rather, other than in traditional social media constellations, it is the LGAIM, not the user, who produces the text’.⁷⁷

However, it could be argued that the services of LLMs would be analogous to online search engines. This is mainly due to the way they operate. More specifically, as per Article 3 (j) of the Digital Services Act, a search engine is ...

an intermediary service that allows users to input queries in order to perform searches of, in principle, all websites, or all websites in a particular language, on the basis of a query on any subject in the form of a keyword, voice request, phrase or other input, and returns results in any format in which information related to the requested content can be found.

This means that when online users input a query, the search engine generates a list of results. This is also described in the OECD report on the economic and social role of internet intermediaries which states that ‘Internet search engines and portals operate websites that use a search engine to generate and maintain extensive databases of Internet addresses and content in an easily searchable format’.⁷⁸

Likewise, LLMs generate content in response to users’ textual prompts. This understanding has been reinforced by Botero Arcila who argues that certain LLMs could be seen as search engines by analogy.⁷⁹ This is because LLMs can fit within the definition of the hosting providers as evidenced in a line of case law. More specifically, in *Google France/Inc. v. Louis Vuitton Malletier*⁸⁰ the Advocate General noted that search engines fall into the scope of hosting ISSPs as per Article 14 of the ECD and held that,

... in any case, the provision of hyperlink services and search engines falls squarely within the notion of information society services and, most importantly [as I will argue next] their inclusion is consistent with the aims pursued by Directive 2000/31.

However, given the size and the large number of users, LLMs could draw parallels with the operation of very large online search engines. More specifically, Stalla-Bourdillon outlines the correlation of very large online search engines with LLMs and also discusses the extension of the provisions of the EU Digital Services Act (DSA) to LLMs.⁸¹ In particular, Article 33 of the DSA states very large online search

⁷⁷ See footnote 67.

⁷⁸ OECD, ‘The economic and social role of internet intermediaries’ (2010) 12.

⁷⁹ B. Botero Arcila, ‘Is it a Platform? Is it a Search Engine? It’s Chat GPT! The European Liability Regime for Large Language Models’ (2023) 3 *Journal of Free Speech Law* 2023.

⁸⁰ Joined Cases C-236/08, C-237/08 and C-238/08 of *Google France/Inc. v. Louis Vuitton Malletier* (2009) ECLI:EU:C:2009:569 para. 134.

⁸¹ S. Stalla-Bourdillon, ‘What if ChatGPT was much more than a chatbox? What if LLM-as-a-service was a search engine? (3 April 2023, Beep Peep!), available at < <https://peepbeep.blog/2023/04/03/what-if-chatgpt-was-much-more-than-a-chatbox-what-if-llm-as-a-service-was-a-search-engine/>, last accessed 18.3.2024.

engines are considered those that reach the minimum threshold of 45 million users per month. Interestingly, in the context of LLMs, it has been evidenced that ChatGPT reached an estimated 123 million active users per month within the first three months of its launch.⁸²

Following that, Article 34 notes that very large online search engines might cause systemic risks. Such systemic risks include the dissemination of illegal content online and the any actual or foreseeable negative effects for the exercise of fundamental rights, including the right to freedom of expression and information as per Article 11 of the EU Charter of Fundamental Rights. Similarly, the operation, design and functioning of the LLMs might also trigger systemic risks. This understanding is supported by a number of lawsuits alleging copyright infringement in the context of inputs and outputs of LLMs. What is more, the use of filtering tools in order to prevent the dissemination of illegal content in the outputs might jeopardize right of users to receive information and thus a risk of censorship might arise.

Moreover, it could also be argued that LLMs which are embedded in hosting providers could fall under the definition of an online interface of an online platform. As per Article 3 (m) of the DSA, ‘online interface’, means ‘any software, including a website or a part thereof, and applications, including mobile applications’. Telling examples can be found on Amazon with its new generative AI service, Amazon HealthScribe, which summarizes and develops clinical summaries of patients on behalf of healthcare companies.⁸³ Similarly, eBay has created a tool based on Open AI’s ChatGPT which provides a product description following sellers’ information about a product’s category, condition, and brand.⁸⁴ In addition, Recital 70, with regard to the online interface of a platform, states that ‘... this is done, for example, by algorithmically suggesting, ranking and prioritising information, distinguishing through text or other visual representations, or otherwise curating information provided by recipients’. For instance, Google integrates AI tools into its data and analytical tools,⁸⁵ while it has been reported that Amazon plans to add an LLM similar to the ChatGPT tool to its online store.⁸⁶

⁸² D. Milmo, ‘ChatGPT reaches 100 million users two months after launch’ (2 February 2023, The Guardian), available at <<https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app>>, last accessed 18.3.2024.

⁸³ A. Ghoshal, ‘AWS updates Amazon Bedrock service with new large language models’ (27 July 2023, Infoworld), available at <<https://www.infoworld.com/article/3703568/aws-updates-amazon-bedrock-service-with-new-large-language-models.html>>, last accessed 1.12.2023

⁸⁴A. Berthene, ‘Ebay rolls out its generative AI listing tool to all marketplace sellers in app’ (7 September 2023, Digital Commerce 360), available at <<https://www.digitalcommerce360.com/2023/09/07/ebay-generative-ai-listing-tool-marketplace-sellers-app/>>, last accessed 18.3.2024.

⁸⁵ E. Avidon, ‘Google unveils generative AI integrations for data tools’ (8 August 2023, Techtargget), available at <<https://www.techtargget.com/searchbusinessanalytics/news/366550019/Google-unveils-generative-AI-integrations-for-data-tools>>, last accessed 18.3.2024.

⁸⁶ M. Day, ‘Amazon Plans to Add ChatGPT-Style Search to Its Online Store’ (15 May 2023, Bloomberg), available at <<https://www.bloomberg.com/news/articles/2023-05-15/amazon-plans-to-add-chatgpt-style-search-to-its-online-store?leadSource=uverify%20wall>>, last accessed 18.3.2024.

Overall, one might extrapolate that the legal classification of LLMs heavily relies on the interpretation of the relevant legal provisions. Such interpretation could either favour or reject the application of rules of the EU AI Act, the Digital Copyright Directive, or the Digital Services Act in the context of LLMs. Crucially, different classifications might trigger different liability schemes. For this reason, the following section critically addresses the different liability regimes that might be adopted in order to regulate the liability of LLMs for copyright infringing outputs.

V. Different EU regulatory perspectives for Large Language Models

Should the judiciaries confirm copyright infringing outputs, LLMs might be subject to different liability rules for copyright infringements. Such rules stem from the relevant provisions in the EU AI Act, Digital Copyright Directive, and the Digital Services Act.

A. EU AI Act

The EU AI Act is applicable in the copyright law context.⁸⁷ This means that LLMs' services are regulated by the provisions of the EU AI Act. However, the EU AI Act does not envisage a liability framework for LLMs, but ascribes to them a set of responsibilities. In particular, as per Recital 106,

'...For this purpose, providers of general purpose AI models should put in place a policy to respect Union law on copyright and related rights, in particular to identify and respect the reservations of rights expressed by rightholders pursuant to Article 4(3) of Directive (EU) 2019/790. Any provider placing a general purpose AI model on the EU market should comply with this obligation, regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of these general purpose AI models take place. This is necessary to ensure a level playing field among providers of general purpose AI models where no provider should be able to gain a competitive advantage in the EU market by applying lower copyright standards than those provided in the Union.'

This means that LLMs are required to obtain the authorization of rights holders if they wish to carry out text and data mining over their works. As per Article 4 (3) of the Digital Copyright Directive, rights holders may select to reserve their rights over their works or other subject matter to prevent text and

⁸⁷ On 13 March 2024, the EU AI Act has been voted by the EU Parliament. The agreed draft text has been released.

data mining, unless this is made for scientific purposes. For instance, GPTBot crawler provides the opportunity for website owners to opt-out of having their data scraped for training datasets of LLMs.

What is more, Recital 107 states

‘In order to increase transparency on the data that is used in the pre-training and training of general purpose AI models, including text and data protected by copyright law, it is adequate that providers of such models draw up and make publicly available a sufficiently detailed summary of the content used for training the general purpose model...’ and ‘...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...’

This means that LLMs are required to issue transparency reports that contain information on the functions of AI generative tools, on the decision-making process and the sources of datasets that have been used in order to train the algorithms. This understanding is reinforced in Article 53 (1) c which requires from LLMs to keep documentation relating to testing and training and evaluation outcomes of the training datasets.

In addition, as per Recital 108,

‘With regard to the obligations imposed on providers of general-purpose AI models to put in place a policy to respect Union copyright law and make publicly available a summary of the content used for the training, the AI Office should monitor whether the provider has fulfilled those obligations without verifying or proceeding to a work-by-work assessment of the training data in terms of copyright compliance.’

This envisages the development of an AI Office which will be in charge of overseeing the operation of LLMs and supervise their compliance with the obligations, as set forth in EU AI Act. Interestingly, the development of a supervisory authority is not a novelty. Indeed, the establishment of administrative bodies have been entailed in the digital platforms’ legal framework and in particular in the EU Digital Services Act. As per Article 49 b and Article 51, the Digital Services Coordinator is assigned to supervise and enforce the Regulation in the Member states while in terms of responsibilities, the Digital Services Coordinator has the power to request information from online platforms about infringements within their services as well as impose fines for non-compliance with transparency obligations or lack of risk mitigation measures.

Finally, as per Recital 107,

‘While taking into due account the need to protect trade secrets and confidential business information, this summary should be generally comprehensive in its 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....’ This implies that LLMs are required to disclose the datasets that have been used in order to train their services, but without revealing trade secrets and confidential information for the operation of their businesses. Yet, one might wonder how transparency reports can address the practices of training and evaluate training datasets without disclosing business confidential information.

On the top of that, the ascription of a set of obligations as well as potential responsibility for non-compliance to LLMs may be challenging. This is mainly due to the opacity of AI tools. More specifically, AI tools are seen as black boxes due to their obscure internal operation.⁸⁸ For example, it seems unclear how the input results in the output. It is this obscure nature of AI tools that gives rise to mounting concerns over their operation. For instance, it could be asked under which conditions and requirements training data has been gathered, how they are processed by LLMs, and how the outputs have been produced.

Yet, it is worthwhile to emphasise that AI generative tools have been developed by human entities. Drawing parallels with algorithms, humans have developed algorithms and therefore can influence the sorting and ranking of the results.⁸⁹ As Parasuraman and Riley have pointed out, ‘although human operators can have little impact on a completely automated system’s decision, the system’s designers play an enormous role in creating that system’.⁹⁰ Likewise, LLMs have been developed by humans who decide the methods for mining data, the material infrastructure, and the training of AI generative tools. Indeed, the OpenAI company that created the ChatGPT tool comprises top researchers and scientists who have constructed and set the parameters under which ChatGPT operates.⁹¹

⁸⁸ D. K. Citron and F. Pasquale “The Scored Society: Due Process for Automated Predictions” (2014) 89 *Washington Law Review* 19.

⁸⁹ E. Bozdog, ‘Bias in algorithmic filtering and personalization’ (2013) 15 *Ethics and Information Technology* 209–227.

⁹⁰ R. Parasuraman and V. Riley, ‘Humans and Automation: Use, Misuse, Disuse, Abuse’ (1997) 39 *Hum Factors* 230.

⁹¹ C. Smith, ‘ChatGPT-4 Creator Ilya Sutskever on AI Hallucinations and AI Democracy’ (15 May 2023, *Forbes*), available at <<https://www.forbes.com/sites/craigsmith/2023/03/15/gpt-4-creator-ilya-sutskever-on-ai-hallucinations-and-ai-democracy/>>, last accessed 18.3.2024.

B. Intermediary liability framework

i. Digital Copyright Directive

Given that this article focuses on copyrighted content, the relevant provisions from the Digital Copyright Directive need to be addressed. As discussed in Section IV, depending on the interpretation, LLMs may, or may not, fall within the definition of online content sharing service providers. Where they do, their liability would be applied and regulated under Article 17 of the Digital Copyright Directive. As per Article 17 (1),

Member States shall provide that an online content-sharing service provider performs an act of communication to the public or an act of making available to the public for the purposes of this Directive when it gives the public access to copyright-protected works or other protected subject matter uploaded by its users.

This means that LLMs would be considered to give access to the public to copyrighted works.

In order to escape liability, as per Article 17 (4) (a), LLMs would be required to demonstrate that they made their best effort to obtain permission from copyright holders in order to avoid liability for copyright infringing works. In addition, as per Article 17 (4) (b), LLMs would be required to demonstrate that they made their best effort in line with industry practices and prevented the availability of infringing works within their services. Finally, as per Article 17 (4) (c), LLMs might be required to prevent the appearance of copyright infringing works by deploying measures to terminate or prevent the reappearance of infringing content.

This means that LLMs might resort to filtering tools in order to escape liability. As already discussed in Section V. A, the use of filters might pose risks to users' fundamental rights. This understanding has been reinforced by the ruling on the *Republic of Poland v European Parliament and Council (C-401/19)*.⁹² In particular, in relation to the use of filtering tools, the CJEU has concluded, in para.86,

In that context, it must be borne in mind that the Court has already held that a filtering system which might not distinguish adequately between unlawful content and lawful content, with the result that its introduction could lead to the blocking of lawful communications, would be incompatible with the right to freedom of expression and information, guaranteed in Article 11 of the Charter, and would not respect the fair balance between that right and the right to intellectual property.

⁹² Case C-401/19, Republic of Poland v European Parliament and Council (2022) ECLI:EU:C:2022:297.

What is more, a number of studies have evidenced the inability of filtering tools to distinguish between lawful and unlawful content and therefore raise mounting concerns with regard to censorship.⁹³

Finally, those provisions come to be added by a line of case law where the act of communication to the public has been addressed. More specifically, in *YouTube/ Cyando*,⁹⁴ the CJEU held that in order to assess whether there is a deliberate intervention, the court needs to examine whether the online platform was aware that users were making illegal content available online, whether they had taken measures to address copyright violations, and whether they operate their business model in order to promote the dissemination of copyright infringing works, or not. In the context of LLMs, this means that the court needs to assess the above-mentioned criteria. Otherwise, LLMs would not be considered to make a deliberate intervention in sharing illegal content and, therefore, cannot be subject to liability for copyright infringement, as set forth in Article 3 of the InfoSoc Directive.

ii. Digital Services Act

As already discussed in Section IV, if the LLMs are considered very large search engines, or embedded in hosting providers, and thus treated as an online interface of an online platform, the relevant provisions of the Digital Services Act would be applicable. Given the majority of EU member states consider search engines as hosting providers,⁹⁵ Article 6,⁹⁶ reinforcing Article 14 of the E-Commerce Directive,⁹⁷ formulates host internet service provider liability in the form of secondary liability in a negative way; it states that,

Where an information society service is provided that consists of the storage of information provided by a recipient of the service, the service provider shall not be liable for the information stored at the request of a recipient of the service, on condition that the provider: (a) does not have actual knowledge of illegal activity or illegal content and, as regards claims for damages, is not aware of facts or circumstances from which the illegal activity or illegal

⁹³ Z. Krokida, 'Use of Filters by Online Intermediaries and the Rights of Users: Developments in the European Union, Mexico, India and China', in J. Coates, V. Owen and S. Reilly, *Navigating copyright for libraries* (De Gruyter Saur, 2022) pp. 405-439.

⁹⁴ C-682/18 - *YouTube and Cyando* (2021) ECLI:EU:C:2021:503.

⁹⁵ EU Commission, 'Impact Assessment Report on the DSA' (2022) 95 which states that 'For example, Austria, Hungary, Spain and Portugal have adopted specific liability exemptions for search engines according to which a company can benefit if it meets the conditions that hosting service providers are required to meet in order to secure a liability exemption.'

⁹⁶ Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) PE/30/2022/REV/1, OJ L 277.

⁹⁷ Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce') OJ L 178.

content is apparent; or (b) upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the illegal content.

This means that host providers would not escape from liability if certain conditions apply; namely, if they lack knowledge, or are not aware, of the infringing act, or if they take down the infringing content upon receiving notification.

In this light, LLMs would be subject to liability if they are aware or have knowledge that the output is copyright infringing. Crucially, liability cannot be triggered due to general awareness or knowledge. Rather, LLMs must have knowledge of the specific infringement, or awareness through the lens of the diligent economic operator. In particular, with regard to knowledge, in *YouTube/Cyando*, the CJEU held, in para. 83, that on the infringing act or ...

in the event that YouTube's role is to be classified as neutral and, therefore, that its activity falls within the scope of Article 14(1) of the Directive on Electronic Commerce, the question arises, in the third place, as to whether "actual knowledge of illegal activity or information" and "[awareness] of facts or circumstances from which the illegal activity or information is apparent", within the meaning of that provision, must relate to specific illegal activities or information...⁹⁸

In addition, apart from knowledge and awareness, LLMs might be subject to liability if they fail to expeditiously remove the allegedly infringing output upon being notified. This requirement also seems to extend a duty to LLMs to prevent the allegedly unauthorized output. This understanding is indicated in a line of case law. In *L'Oréal v eBay*,⁹⁹ L'Oréal initiated legal proceedings against eBay, alleging trade mark infringement. The CJEU concluded that the host ISP can be ordered to prevent the emergence of a specific trade mark infringement by the same infringer for the same trade mark. This means that it is the host ISPs who need to identify when the same infringement occurs by the same infringer. Likewise, in *Peterson v YouTube*¹⁰⁰ the CJEU followed the same approach. In this case, Frank Peterson filed a lawsuit against YouTube alleging that the latter failed to terminate the reappearance of unauthorised videos on its network. The CJEU handed down the ruling that host ISPs have an obligation to prevent the emergence of future infringements, considering a fair balance between different interests at stake.

Finally, LLMs might be ascribed the obligations that have been assigned to very large online search engines and very large online platforms. More specifically, as per Article 34 of the Digital Services Act, LLMs would be required to conduct a risk assessment with regard to their operation and the

⁹⁸ C-682/18 - *YouTube and Cyando* (2021) ECLI:EU:C:2021:503.

⁹⁹ Case C-324/09, *L'Oréal SA V eBay Int'l AG* (2011) E.C.R. I-6011, para. 120.

¹⁰⁰ Case C-682/18, *Peterson v YouTube* (2021) ECLI:EU:C:2021:503.

dissemination of infringing outputs. They would also be required to deploy measures in order to mitigate any risks, such as the adjustment of the features and functioning of their services, or adapting terms and conditions.¹⁰¹ What is more, LLMs would be subject to an independent audit at their own expenses at least once a year and an examination of their compliance with their obligations.¹⁰² Finally, LLMs would be assigned transparency obligations and, in particular, to issue transparency reports with regard to their operation and content moderation at least every six months, reporting to the Digital Services Coordinator and the European Commission.¹⁰³ The ascription of the above obligations would be of seminal importance for the design and use of LLMs given the number of systemic risks that might arise in the course of their operation.

VI. Conclusion

This article has critically engaged with the different types and operation of LLMs and their potential liability for copyright infringing outputs. As discussed, LLMs took the world by storm with millions of users accessing their services within the first months of their operation. Yet, since then, an increasing number of lawsuits have been filed which argue copyright infringements while writers have initiated legal proceedings against the developers of LLMs, arguing that they are generating copyright infringing outputs.

Should the judiciaries confirm the copyright infringing outputs, a potential liability of LLMs might arise. Such liability, however, is contingent upon their legal classification. LLMs that generate texts in response to users' prompts could be seen as general-purpose AI models, but at the same time, as the findings indicate, they might also perform an intermediary function, either as online content sharing service providers, very large online search engines, or as online interfaces of online platforms. Whilst the EU AI Act envisages a set of duties to which LLMs are required to comply, the Digital Copyright Directive and the Digital Services Act provide the liability framework of online platforms. With regard to search engines and online interfaces of online platforms, following a formalistic or a purposive interpretation, LLMs might be subject to different legal obligations as set forth in Article 17 of the Digital Copyright Directive or Arts 35-42 of the EU Digital Services Act, accordingly. The application of those legal obligations in the context of LLMs seem to be warranted due to the systemic risks in the course of their operation. Otherwise, risks of illicit output or censorship might arise.

¹⁰¹ Article 35 of the Digital Services Act.

¹⁰² Article 37 of the Digital Services Act.

¹⁰³ Article 42 of the Digital Services Act.

In light of the above, what is sure for now is that the arrival of LLMs heralds a novel epoch in digital intermediation and might give rise to new legal challenges in the context of intermediary liability. This is also what Perault has accurately described by reference to AI tools, i.e., ‘the emergence of products fuelled by generative artificial intelligence such as ChatGPT will usher in a new era in the platform liability wars’.¹⁰⁴

¹⁰⁴ M. Perault, ‘Section 230 won’t protect ChatGPT’ (2023) *Journal of Free speech* 1.