سائد كراجة ومشاركوه
Sa'ed Karajah & Partners LLP
Attorneys at Law & Legal Counselors

Assessing the Effectiveness of Traditional Tort Law in Regulating Liability Related to Generative Artificial Intelligence

Research paper by Hashem Al Sarayreh & Zaid Al Zaghal

Sa'ed Karajah & Partners LLP

November 2024

(962) 6 5620132
info@karajahlaw.com
www.karajahlaw.com
Salem Al Hindawi St. 23
PO BOX 9431 Amman
11191 Jordan

Table of Contents

Introduction 3
General Principles of Liability Applied to Generative AI 4
A Lex Specialis for Generative AI Tort Liability 6
Who can be held liable? 11
Conclusion 15

Introduction

As a concept, Artificial Intelligence (AI) is not particularly novel, as the foundations for AI are generally credited to Alan Turing's work back in 1935[1]. However, the field has undergone substantial developments in the last decade, which gave way to Generative Artificial Intelligence's inception (GAI).

On the 13th of March 2024, the EU Parliament passed the EU Artificial Intelligence Act for the 'development, placing on the market, and use of artificial intelligence'[2] "AI EU Act". The AI EU Act does not define AI, but rather gives a definition of the 'AI System', which is similar to the definition given by the Organization for Economic Cooperation and Development "OECD", whereas:

EU AI Act definition of an AI System: "A machine-based system that is 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;" [3]

OECD's definition of an AI System: "An AI system is a machine-based system 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. Different AI systems vary in their levels of autonomy and adaptiveness after deployment."[4]

Both definitions provide a general, comprehensive, and holistic account of AI systems, wherein they highlights the core unique features that AI offers, namely its ability to adapt after deployment, inferring from its received inputs, and generating afterwards. This paper's main concern is AI systems that are solely designed to generate content. Content in this case includes new meaningful text, images, and audio, the output of what is commonly known as GAI.[5] Examples of such models are OpenAI's ChatGPT and

[1] Darrell M West and John R Allen, Turning Point: Policymaking in the Era of Artificial Intelligence (Brookings Institution Press 2020)
[2] EU Artificial Intelligence Act https://www.euaiact.com/#:~:text=The%20EU%20AI%20Act,intelligence%20in%20the%20European%20Union
[3] Ibid, Art (3)
[4] OECD, 'Definition of Artificial Intelligence' https://oecd.ai/en/wonk/definition
[5] S Feuerriegel, J Hartmann, C Janiesch et al, 'Generative AI' (2024) 66 Bus Inf Syst Eng 111 https://doi.org/10.1007/s12599-023-00834-7

DALL-E, Microsoft's Bard, and Google's Gemini. GAI generates vast amounts of content, raising important issues such as intellectual property considerations, ethical use of GAI, and tort liability[6].

On the flipside of GAI's many benefits, such as its use in drug development and music composition, it begets risks that are arguably inherent to its developing nature. These risks pose serious considerations for a technology that has gained rapid enthusiasm across the world. GAI poses risks, for instance, in relation to bias and fairness issues, ethical concerns about the data it produces and learns from, security risks, privacy concerns, and liability challenges[7].

This paper discusses whether traditional legal torts' frameworks are sufficient to resolve tortious contentions matters arising out of the use of GAI or are related to it. Specifically, this paper examines the liability related to GAI from a Jordanian law perspective while shedding light on other legal systems, namely English Law and European Law. This paper does not address contractual liability, which arises pursuant to a party's breach of its contractual obligations. Rather, this paper only explores the implications affecting the establishment of tort claims, with a particular focus on the requirement of causation, under Jordanian Law and other legal systems.

Currently, there are minimal regulations or legislation specifically addressing tort claims related to GAI. Existing laws primarily focus on violations of fundamental human rights, leaving tort claims unaddressed[8]. Therefore, courts would need to apply traditional tort rules to liability claims, which highlights the necessity of understanding the process GAI systems undergo to deliver their outcomes in order to determine "Who is Liable".

General Principles of Liability Applied to Generative AI

To establish negligence in English tort law, the plaintiff must prove the existence of four different elements. The first element is the 'Duty of Care', where the defendant must have had a duty of care towards the plaintiff according to the 'neighbor principle', which is owed to "persons who are so closely and directly affected by my act that I ought reasonably to have them in contemplation as being affected when I am directing my mind to the acts or omissions which are called in question."[9] The second element that must be

[6] P. Zhang, A. R. Smith, and J. T. Lee, 'Sustainable Artificial Intelligence: Challenges and Future Directions' (2023) arXiv https://arxiv.org/abs/2304.06632 accessed 7 November 2024
[7] S Feuerriegel, J Hartmann, C Janiesch et al, 'Generative AI' (2024) 66 Bus Inf Syst Eng 111 https://doi.org/10.1007/s12599-023-00834-
[8] Council of Europe, 'Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law' in Council of Europe Treaty Series No. 225, Chapter IV (2024) https://rm.coe.int/1680afae3c
[9] Donoghue v Stevenson [1932] AC 562.

established is the 'Breach of Duty', which requires the proof that the duty of care owed by the defendant towards the claimant was breached by the defendant. The latter refers to the failure to meet the standard of care, which is the level of care and skill that a reasonable person would exercise in a similar situation. Upon establishing both Duty of Care and Breach of Duty, the claimant must prove 'Causation'. This third element consists of proving that the defendant's breach of duty caused the claimant's injury or loss in a direct and foreseeable manner, namely by applying the 'but for' test.[10] Finally, the claimant must prove 'Injury' which necessitates the proof of certain loss or injury resulting from the defendant's breach.

Under Jordanian law, the tort of negligence corresponds to one of the primary sources of legal obligations, specifically referred to as the 'harmful act'. In Jordanian Law, the plaintiff bears the burden of proving the existence of three essential elements. First, they must prove the 'Harmful act'[11], which is committing an act that caused damage or injury to others without the said act being a legitimate exercise of a right[12]. Secondly, the plaintiff must prove 'Damages'[13], where the plaintiff must be able to prove that damages or injury had taken place following the defendant's 'harmful act'. Finally, the 'Causal Link'[14] must be established, drawing a connection between the damages or the injury incurred by the plaintiff as a direct or eventual result of the 'harmful act' committed by the defendant.

Importantly, both Jordanian Law and English Law necessitate the establishing of 'causation' or a 'causal link' between the act/breach and the loss or injury. The latter is particularly relevant as GAI systems tend to not offer much explainability and transparency due to its current 'black box' nature. Explainability and transparency are crucial not only for legal accountability, but also for meeting ethical standards in AI deployment. The 'black box' nature of GAI systems refers to the inability to ascertain how exactly a GAI system draws out its outputs considering that there are numerous processes involved in generating the outcome, like data collection and labeling, as well as the formulation of the algorithms. The decision-making processes of GAI are not easily interpretable, making it difficult to trace how specific outcomes are produced. This opacity poses challenges for establishing causation in legal contexts, as it complicates the process of demonstrating a direct link between the system's actions and the incurred harm. What further exacerbates the need for explainability and transparency is the vast array of fields that GAI would impact, in addition to the fact that professionals, like attorneys and physicians, may soon be required to use explainable

[10] Cork v Kirby MacLean Ltd [1952] 2 All ER 402.
[11] Jordanian Civil Law, s 256.
12 Jordanian Cassation Court, Case No 6627/2023.
13 Jordanian Civil Law, s 266.
14 Jordanian Civil Law, s 257.

GAI systems[15]. The latter is significant as the use of GAI systems in various professions might become a significant component of the processes adopted by practioners of those fields[16]. If these systems do become more prevalent, consumers may increasingly expect and be entitled to the most advanced and effective services available in each industry. This expectation would entail that GAI systems are explainable and transparent, ensuring that professionals can provide high-quality accountable services to their clients.

However, a major concern in using pre-existing sources of liability to address GAI technology is that such provisions were not set in account of technology that is as complex as GAI technology. It might be worth considering that GAI requires specific *lex specialis* rules that account for its unique nature, as discussed in the next section.

Tort liability is established through components that are not easily adaptable to GAI technology due to the lack of explainability and transparency. Tort liability relies on causal link. It must be proven that the harmful act has led to the damage, which includes demonstrating a breach of the duty of care. However, with regard to GAI, the causal link is difficult to establish, as it is not possible to identify what component of the GAI system or part of the processes involved 'caused' the injury. This lack of clarity complicates proving not only the standard of care but also the breach and the direct cause of the harm, making the traditional tort framework less effective in regulating and compensating damages arising from GAI technologies. Further, to establish liability under tort law, a standard of care must be determined, typically based on how a reasonable person would act in similar circumstances. However, the complexity of GAI makes it challenging to define what constitutes "reasonable" or standard behavior for such technology.

A Lex Specialis for Generative AI Tort Liability

Following its November 2021 General Conference, the United Nations Educational, Scientific and Cultural Organization "UNESCO" has adopted the "Recommendation on the Ethics of Artificial Intelligence", outlining guidelines and principles to be adopted by the global community when approaching national

15 P Hacker, R Krestel, S Grundmann, & F Naumann, 'Explainable AI under contract and tort law: legal incentives and technical challenges' (n.d.).
16 M Goto, 'Accepting the Future as Ever-Changing: Professionals' Sensemaking about Artificial Intelligence' (2022) 9 Journal of Professions and Organization 77.

legislation concerning AI. While not binding on Member States, the Recommendations offer clarity as to how this area of technology must be addressed, as well as focal points to consider in legislation[17].

The Recommendations draw a focus on AI actors and define them "as any actor involved in at least one stage of the AI system's life cycle and can refer both to natural and legal persons"[18]. This definition establishes a cornerstone in any AI legislation, especially for its identification of who AI actors are.

Further, by understanding who AI Actors are, the Recommendations hold that such actors must be responsible for AI through ensuring the transparency of their AI products and that algorithms remain explainable. The Recommendations highlight that transparency and explainability are essential not only for creating trust between the producers of this technology and its consumers, but also to provide them with the means to identify who must be held accountable whenever AI results in damages.[19]

It is a truism that accountability is a key element upheld across several jurisdictions and bodies concerned with AI regulation[20]. In this regard, section 1.5 of the OECD's "Principles on Artificial Intelligence" provides that AI actors must remain accountable for the AI product throughout the whole of its life cycle as well. The OECD Principles do not outline how accountability must be upheld on a national level, but they do emphasize the importance of it when addressing AI regulation[21]. Therefore, to reinforce accountability in GAI on a national level it must be enshrined that the automated nature of any decision should not relinquish AI actors from the roles they may have played[22].

To illustrate the unique nature of GAI systems' liabilities, the concept of Transparency and Accountability will be explored. Transparency is a multifaceted concept where it may relate to the users of GAI systems being aware of how these systems operate and make decisions.[23] By the same token, it is also crucial to highlight the importance of algorithmic transparency. The latter refers to users' ability to determine how the

[17] UNESCO, 'Recommendation on the Ethics of Artificial Intelligence' https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence
[18] Ibid.
[19] Ibid.
[20] 'AI Technology' https://www.cell.com/action/showPdf?pii=S2666-3899%2822%2900233-1
[21] OECD.AI, 'AI legal cases are increasing: how can we prepare?' https://oecd.ai/en/wonk/increasing-legal-cases
[22] 'Artificial Intelligence and Law' https://link.springer.com/chapter/10.1007/978-94-6265-531-7_4
[23] George Benneh-Mensah, 'Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems' (2024) https://www.researchgate.net/profile/George-Benneh-Mensah/publication/375744287_Artificial_Intelligence_and_Ethics_A_Comprehensive_Review_of_Bias_Mitigation_Transparency_and_Accountability_in_AI_Systems/links/656c8e46b86a1d521b2e2a16/Artificial-Intelligence-and-Ethics-A-Comprehensive-Review-of-Bias-Mitigation-Transparency-and-Accountability-in-AI-Systems.pdf

GAI systems' algorithm's function – and how they were set out to function.[24] This is important because the use of GAI systems may naturally result in harmful or unfair outcomes.[25] Even if algorithms and their codes were not intentionally designed to discriminate, the issue runs deeper. The interaction between the data and the algorithms can still produce harmful and discriminatory outcomes. Moreover, it is highly unlikely that a review of the code alone would reveal the source of the error, making it difficult to identify and address the underlying bias.[26] Negligence may occur when developers develop algorithms without properly considering the data, leading to harmful outputs. Under general tort principles, developers could be held liable for breaching their duty of care. However, these breaches are not easily identifiable through conventional methods. Thus, the need for Transparency and Accountability – a matter that legislators must certainly ensure when attempting to regulate GAI, its use, and its design. For instance, legislators may require deployers to take necessary steps to trace liability and ensure transparency. The AI EU Act focuses more on the obligations of AI Actors in ensuring safety and fundamental human rights rather than on introducing new mechanisms to address directly the liability that may result from AI[27].

When it comes to liability, the Act relies on the pre-existing Product Liability Directive (PLD)[28] and national tort laws. Liability tort laws in the EU, such as those in the French Civil Code[29], German Civil Code[30], and Italian Civil Code[31], are established on a fault-based liability system, where the claimant must prove both fault and causation that have led to actual damage. Alternatively, many EU Member States, including Germany (§ 1 ProdHaftG)[32] and the UK (Consumer Protection Act 1987)[33], recognize the notion of strict liability in product-related cases. Under strict liability, a victim must prove that damage has occurred due to the product or responsible person's actions, regardless of whether the responsible party was in compliance with the law or not, without needing to prove fault or breach of duty of care.. The PLD, on the other hand,

[24] 'Policy Review' (2020) https://policyreview.info/pdf/policyreview-2020-2-1469.pdf
[25] Ibid.
[26] Ibid.
[27] European Parliament, 'EU Artificial Intelligence Act | Up-to-date developments and analyses of the EU AI Act' (1 June 2023) https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence
[28] European Commission, 'Liability for Defective Products' https://single-market-economy.ec.europa.eu/single-market/goods/free-movement-sectors/liability-defective-products_en
[29] French Civil Code.
[30] German Civil Code
[31] Italian Civil Code
[32] Product Liability Act (Produkthaftungsgesetz, ProdHaftG) § 1.
[33] Consumer Protection Act 1987.

is an instrument that highlights possible routes for claiming damages regarding specified products existing within the EU[34].

The challenges of relying solely on pre-existing notions of liability are directly linked to the complex nature of AI technology. Firstly, in a fault-based liability system, the victim of an AI-related mishap must prove that 1) the relevant AI actors have actively breached or failed to comply with specific obligations under the EU AI Act, 2) damage has in fact occurred, and 3) the damage is a direct result of that breach. Secondly, even under strict liability, while the victim does not need to prove fault, they still bear the burden of proving the damage and demonstrating a causal link between the damage and the AI product or system. Thirdly, under the Product Liability Directive (PLD), the victim must prove that the AI product was defective at the time of use. However, given the inherent complexity and opacity of AI systems, identifying non-compliance or linking specific technological failures to the damage can be exceedingly difficult, making it harder to demonstrate that the product is effectively defective.

The 'black box' nature of AI technology is in and of itself a veil between the AI actors and the technology, let alone external users. GAI systems employ neural networks, which are closed systems that receive and input and produce an output without any justification or reasoning behind the output[35]. To illustrate with an example, in 2014, an internet bot tasked with buying random items online had bought illegal drugs from the deep web[36]. Eliminating the reoccurrence of the latter or similar unwarranted occurrences will either require the root cause of its occurrence or will compel developers to seek different means of prevention. Importantly, and with emphasis on the fact that the AI system does not provide an answer as to how an output was reached, the need for special regulation grows further, since it is technically impossible to ascertain causation.

The EU AI Act is thus lacking in its reliance on pre-existing notions of liability, and while the EU is currently re-drafting the PDL to adapt it to new technology[37], these instruments were not established in consideration of the complex nature of AI technology and the effects the latter has on tort instances and

[34] European Parliament. (2023, June 1). EU Artificial Intelligence Act | Up-to-date developments and analyses of the EU AI Act. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence
[35] Karen Yeung, 'What's Inside the Black Box? AI Challenges for Lawyers and Researchers' (2019) Legal Studies https://www.cambridge.org/core/services/aop-cambridge-core/content/view/8A547878999427F7222C3CEFC3CE5E01/S1472669619000021a.pdf/whats-inside-the-black-box-ai-challenges-for-lawyers-and-researchers.pdf
[36] Ibid
[37] European Commission, 'Liability for Defective Products' (Single Market, Economy and Industrial Policy) https://single-market-economy.ec.europa.eu/single-market/goods/free-movement-sectors/liability-defective-products_en

claims. To address this serious concern, the EU has also begun to draft the AI Liability Directive, a tool separate from, but parallel to, the AI Act, introducing new means of addressing the technical aspect of AI Actors as defined in the Act in coherence with the pre-existing tools.

On the 28th of September 2022, the European Commission issued the Proposal of the EU AI Liability Directive (AILD)[38], which the Council of the European Union has proposed an amendment to update it as to adapt non-contractual civil liability rules to Artificial Intelligence[39]. Currently, the AILD assumes that the AI Act shall ensure that all AI technologies within the EU are deemed acceptable for use[40]. It also incorporates a notion of fault-based liability[41] and introduces rebuttable 'presumption of causality'[42], easing the burden of proof on the victims of AI harm and damage. A presumption of causality means that a victim simply may claim that the AI has resulted in harm due to a fault, such as non-compliance with the Act's requirements of transparency and cybersecurity, but without having to explain how that fault directly resulted in the damage, but that it was 'reasonably likely', hence causality is assumed. Yet, its rebuttable nature allows the AI Actors, as defendants, to prove that the fault either does not exist, that it did not lead to the damage, or that the claimant has the resources necessary to establish and thus prove causation and thus must do so. The latter may effectively deter victims from claiming compensation all-together if they do not have sufficient resources to proceed with their claims. The proposal to update the AILD has addressed the latter concern by introducing a number of options that involve easing the burden of proof for victims trying to prove their liability claims, reassessing the need to harmonise strict liability rules with AI use cases, and most importantly by introducing the 'targeted rebuttable presumption of causality' whereby the defendants, namely GAI providers, are then required to rebut the causal link between their non-compliance and the damage that had taken place[43]. The latter aims to portray a clearer image of liability for businesses in terms of compliance, while providing claimants with a more equitable framework to prove their claims given the complexity of drawing out causation.

[38] European Parliament, 'Cybersecurity in the EU: Current State of Play and Challenges' (2023) EPRS Briefing Paper https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)739342_EN.pdf

[39] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space' (COM(2022) 496 final) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022PC0496 accessed [5/10/2024].

[40] European Union, 'Artificial intelligence liability directive' https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0496

[41] European Parliament, 'Cybersecurity in the EU: Current State of Play and Challenges' (2023) EPRS Briefing Paper https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)739342_EN.pdf 2

[42] Ibid, 4

[43] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space' (COM(2022) 496 final) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022PC0496 accessed [5/10/2024], p. 13.

The AILD further allows courts to order the disclosure of evidence regarding high-risk AI systems when necessary and proportionate[44]. A positive aspect of the AILD is that it does provide novel notions for liability within the EU but is rather a case-specific route of action considering the complex nature of AI technology.

It is argued that the EU AILD provisions on explainability are insufficient for ensuring meaningful transparency[45]. The Directive is also critiqued for its preliminary nature, where it leaves many critical areas, such as accountability and ethical considerations, inadequately addressed, thus calling for a more robust framework to manage the intricate dynamics of AI systems[46].

Jordan's approach to legislation around AI is promising due to the government's issuing of the National Artificial Intelligence Code of Ethics in February 2022[47]. Released under the competence of the Ministry of Digital Entrepreneurship, Jordanian legislators expect that any regulation surrounding AI must consider identifying the different parties involved in the life cycle of the AI, mainly to attribute responsibilities and liability[48]. It is emphasised within the Charter that it is these AI Actors that must be held liable for any harm caused by the technology, and not to spread out liability so as to limit the benefits of the technology itself. The Charter further highlights that any negative effects of AI must be examined in order to be remedied through reliable technical measures, entailing the necessity for record keeping and traceability when establishing any relevant regulations. Jordan wishes to ensure transparency and accountability in its legislative process surrounding AI technology, establishing relevance to the UNESCO's guiding Recommendations[49].

Who can be held liable?

An evident key component common in draft regulations surrounding AI is the emphasis on identifying the parties responsible for AI technology once it is deployed for public usage. The purposes for which AI Actors must be defined are several, including clarity as to AI Systems' life cycles, and the fair distribution of liability amongst the different involved parties. AI Actors must remain accountable for the product that is

[44] Ibid.
[45] M Veale and FZ Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) 39 Computer Law & Security Review 105615.
[46] VL Raposo, 'Ex machina: Preliminary Critical Assessment of the European Draft Act on Artificial Intelligence' (2022) 30(1) International Journal of Law and Information Technology 88.
[47] Ministry of Digital Economy and Entrepreneurship (MoDEE), 'MoDEE presents the draft of the national charter on the ethics of AI for general consultation' (16 February 2022) https://www.modee.gov.jo/EN//NewsDetails/MODEE_presents_the_draft_of_the_national_charter_on_the_ethics_of_AI_for_general_consultation accessed [date you accessed the document].
[48] Ibd, 5.
[49] Ibid, 18.

released and made available for use. While AI actors must comply with the legislation and guarantee fundamental human and consumer rights, importantly they must be able to explain the choices they made while carrying out their role through involving experts and external stakeholders. Foremostly, AI Actors must be held responsible in case the technology causes harm.

Yet, an issue with defining AI Actors is that definitions are often broad, encompassing developers, engineers, vendors, developers, deployers, and data providers[50]. However, the harm that an AI may cause, or the fault which would have led to such harm, may be difficult to pinpoint. The presence of several parties involved in the process of creating functional AI creates an accountability gap. Even if one was able to determine what went wrong with the technology, then the problem would reside in the fair distribution of fault and consequent liability.

What further complicates the issue is the 'black box' nature of AI, where the algorithm behind the technology becomes extremely complex and opaque, so much so that even the parties were responsible for creating it find difficulty in understanding the AI's reasoning[51]. Simply put, the prompt is provided to the machine, and the machine provides the output. We may understand what knowledge and data the machine had learned from, but we might not be able to identify how the machine used that data to provide the output, or what data it even used for that specific output.

The black box problem paired with the accountability gap establishes serious concerns into how liability may be assigned. As such, regulation should account for features that must be incorporated into the entire life cycle of technology. By emphasizing certain aspects of regulation, the accountability gap may be bridged, at least partially, and the black box may become simpler to approach. Traceability and explainability must remain at the forefront of AI legislation, ultimately opening the door to accountability and transparency. Traceability is a feature by which the entire process from input to output is exposed and made available to users, to the extent possible. Such traceability would allow users to understand what data is being used in creating the desired outputs. Traceability may help in reducing infringement on data by identifying its sources, allowing property owners more access to how their data and information is being used[52]. On the other hand, explainability in AI technology pertains to uncovering as many aspects of the black box as

[50] Wikborg Rein, 'Who is liable when the use of AI leads to harm?' (2023) https://www.wr.no/en/news/who-is-liable-when-the-use-of-ai-leads-to-harm.
[51] Ibid.
[52] Adesso, 'Why Traceability is Important in Artificial Intelligence' (2023) https://www.adesso.de/en/news/blog/why-traceability-is-important-in-artificial-intelligence-2.jsp#:~:text=Traceability%20is%20critical%20to%20increasing,development%20of%20an%20AI%20model.

possible. This feature entails that users are able to know and understand how the AI has processed the data and what programs and algorithms it utilized. Explainability would help reduce biases in AI outputs, by identifying built-in components that have led to such biases. Such features would be beneficial in understanding how the machine creates the outcome and where things 'have gone wrong', and who was responsible, even if liability were joint amongst several parties. However, policymakers may hesitate to implement such requirements due to concerns around creativity and intellectual property.

Along with a fault-based liability and a product liability, it ought be stressed that liability may certainly arise out of users of GAI actions. That is, one may use the GAI's outputs in a way that may cause harm, injury, or loss to others. In such cases, one must inquire about the extent to which the GAI must be held accountable for its contribution to the harm or injury that has taken place. For the purposes of examining the latter notion, one will explore the issue of 'deepfakes', which is a relatively 'not-so-novel' application of GAI following its inception on Reddit in 2017[53].

A combination of 'Deep Learning' and 'fake', deepfakes are synthetic media in which a person in an existing image or video is replaced with someone else's likeness using advanced artificial intelligence techniques[54]. Therefore, using deepfake technology allows one to mimic certain persons in unlimited potential scenarios and uttering priorly chosen speech so as to lead (or more appropriately 'mislead') viewers of the media into believing a certain state of affairs. The impact of such use of GAI can reasonably be described as catastrophic given the vast array of applications it may have. For instance, political agendas may be either accelerated or halted using deepfakes, wherein videos may emerge as supposedly 'leaked' for politicians and other public figures who may exercise at least some public influence. Earlier this year, both Sadiq Khan and Keir Starmer, both UK politicians, have been subjects of deepfake technology, which could just tilt the party's voting standing prior to the general elections. Notwithstanding, the dire results of misusing deepfake technology spans across almost all areas of human endeavor[55].

[53] Felipe Romero Moreno, 'Generative AI and Deepfakes: A Human Rights Approach to Tackling Harmful Content' *International Review of Law, Computer & Technology* https://www.tandfonline.com/doi/epdf/10.1080/13600869.2024.2324540?needAccess=true [02/05/2024].

[54] Mika Westerland 'The Emergence of Deepfake Technology: A Review' (2019) [Volume 9(Issue 11)] *Technology Innovation Management Review* https://timreview.ca/sites/default/files/article_PDF/TIMReview_November2019%20-%20D%20-%20Final.pdf accessed [06/10/2024].

[55] Marianna Spring, 'Sadiq Khan Says Fake AI Audio of Him Nearly Led to Serious Disorder' BBC News (14 February 2024).

To illustrate, in February 2024, Arup, a British engineering firm, suffered an amount of 20 million pounds sterling due to a fraudulent attack utilizing deepfake technology. A financial officer at the firm was part of a supposed online meeting with the firm's Chief Financial Officer and other members of staff, wherein he was convinced by the "CFO" to transfer an amount of 20 million pounds sterling to a given account. It was later discovered that the images and voices produced in the 'meeting' were fake and created using AI technology, and no meeting in fact had taken place with the CFO[56]. Such an incident ought to work towards raising awareness about the sophistication of new and advanced technologies, especially that the technology is fairly accessible to virtually anyone. In light of the foregoing, the question about GAI systems' responsibility is once again asked; if presumably the persons behind Arup's loss had utilized a GAI model to produce the fake images and voices, then to what extent is the said GAI model (or the persons behind its development and deployment) liable for providing such a facility?

The previous question refers one to prospective obligations on GAI providers, like omitting from providing deepfake processing and outputs, which is namely to cancel and discontinue the whole facility of producing fake images and voices as an available facility all together[57]. A skeptic communitarian would argue that that is the most favorable measure that might be taken given that it protects society at large. However, the alternative must be sufficiently illustrated, wherein the production of deepfakes by GAI models to users can be stopped, but that does not account for other non-controlled uses, as well as the ability to access such a facility without referring to known and licensed parties. The view that one must eliminate the whole technology from the public domain in order to preserve order and limit harmful occurrences has been historically often suggested as regulation in response to the fears regarding technological developments.

Therefore, one must consider how to balance protecting individuals from the harmful impacts of deepfake technology while ensuring their right to access information and technology and allowing creative developments to emerge.

[56] Cheng Leng and Chan Ho-him, 'Arup Lost $25mn in Hong Kong Deepfake Video Conference Scam' Financial Times (14 February 2024).
57 MH Maras and A Alexandrou, 'Determining Authenticity of Video Evidence in the Age of Artificial Intelligence and in the Wake of Deepfake Videos' (2019) 23 International Journal of Evidence & Proof 255.

Attempts to address the deepfake include enforcing a unique ID or tag for every deepfake output produced by the GAI system, clearly indicating that the media is AI-generated and is not 'real'. However, this measure can be easily bypassed through further processing and subjecting the media to editing tools, like Photoshop. Permanent distinguishing elements could either be bypassed or disturb the media's utility. The primary threat of deepfakes lies in their ability to make 'real' media indistinguishable from fabricated content potentially affecting their admissibility as evidence in court. Therefore, it is crucial to develop a facility capable of distinguishing between real media and deepfakes. An AI model leveraging machine learning and expert systems could serve as an 'independent fact-checker' for use by law enforcement and courts[58]. Governments, public bodies, and legislators worldwide should regulate the verification and production of deepfakes by GAI to address this growing concern[59].

Conclusion

In conclusion, the rapid evolution of Generative Artificial Intelligence (GAI) presents both incredible opportunities and significant challenges. As GAI becomes increasingly integrated into various sectors, from creative industries to healthcare, it is essential to address the legal implications, particularly those concerning tort liability. Current legal frameworks, such as those of the English and Jordanian legal systems, provide a basis for addressing negligence and harm but fall short in dealing with the complexities of GAI's black-box nature and the multifaceted roles of AI actors. The main legal challenge lies in proving causation, establishing a standard of care, and demonstrating a breach due to the lack of explainability and transparency in AI systems. Traditional tort and negligence rules are not equipped to handle the opacity and intricate decision-making processes of GAI, making it challenging to attribute liability and ensure accountability. This necessitates the development of new legal approaches and regulations tailored to the unique challenges posed by GAI technologies.

The need for explainability and transparency in AI systems is paramount. Without clear traceability, it becomes nearly impracticable to assign liability accurately, thereby undermining trust and accountability. As such, regulatory bodies must prioritize the development and enforcement of standards that ensure AI technologies are transparent and explainable. This includes the establishment of comprehensive legal

[58] Congressional Research Service, 'Law Enforcement Use of Artificial Intelligence and Directives in the 2023 Executive Order' (2023) https://crsreports.congress.gov/product/pdf/IN/IN12289 accessed [10/09/2024]
[59] International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, Volume 10, Issue 05 (May 2023) <www.irjet.net>

frameworks similar to the EU's AI Liability Directive, which aims to adapt existing liability concepts to the evolving technical nature of AI.

The AILD represents a significant step towards addressing these challenges, proposing various liability rules to hold AI actors accountable. However, the directive has faced criticism who argue that it may not go far enough in addressing the complexities of AI technologies.

Furthermore, while the focus on AI actors' responsibility is crucial, it is equally important to consider the liability of AI users who may misuse GAI outputs, exemplified by the potential harm of deepfake technologies. Balancing the protection of individuals and society from AI-induced harm with the promotion of technological innovation requires nuanced legislation and international collaboration.

Ultimately, ensuring a robust regulatory environment that addresses both the general and specific challenges posed by GAI is essential for fostering responsible development and use of this transformative technology. Such an approach will help mitigate risks, protect rights, and promote trust, paving the way for a future where GAI can be harnessed for the greater good.