Law, Artificial Intelligence, & Liability: Assessing the correspondence between Kenyan law and civil liability for autonomous artificially intelligent systems.

Submitted in partial fulfilment of the requirements of the Bachelor of Laws Degree, Strathmore University Law School

By
Beatrice Shako

137062

Prepared under the supervision of

Mr Juan Riofrio

March 2024

Word count (excluding footnotes and bibliography)

Declaration

I, [BEATRICE SHAKO], do hereby declare that this research is my original work and that to the best of my knowledge and belief, it has not been previously, in its entirety or in part, been submitted to any other university for a degree or diploma. Other works cited or referred to are accordingly acknowledged.
Signed: .....
Date: .....

This dissertation has been submitted for examination with my approval as University Supervisor.

Signed:.....
[Supervisor's Name]

Declaration Page shall be in the prescribed format (See Annex I)

Acknowledgements

List of legal instruments:

- Constitution of Kenya(2010)
- The Computer Misuse and Cybercrimes Act(2018.
- The Kenya Robotics and Artificial Intelligence Society Bill, 2023
- Competition Act, No.12 of 2010.
- Data Protection Act, No 24 of 2019.
- Consumer Protection Act, (No. 46 of 2012)
- United Nations Convention on the Use of Electronic Communications in International Contracts,
- Convention of 2 October 1973 on the Law Applicable to Products Liabilit
- Sale of Goods Act(No. 19 of 1964)
- Convention On The Law Applicable To Products Liability(1973)
- The Products Liability Directive(1985),
- Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products.
- Road Traffic Act (Straßenverkehrsgesetz) amended by the publication of 5 March 2003
- DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), 2022.
- Regulation of the European Parliament and of The Council of laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 2024
- U.S Department of Transportation, Federal Automated Vehicles Policy, September 2016
- Self Drive Act, H.R.3711 — 117th Congress, 2021

Soft Law:

- EU C: Expert Group on Liability and New Technologies New Technologies Formation, LIABILITY FOR ARTIFICIAL INTELLIGENCE AND OTHER EMERGING DIGITAL TECHNOLOGIES, 2019.
- European Commission, 'Comparative Law Study on Civil Liability for Artificial Intelligence', November 2020
- Commission Staff Working Document, Liability for Emerging Digital Technologies COM(2018) 237 final, 2018,

List of Cases:

- Garcia v. Character Techs., Inc. (2024)
- Khayigila vs Gigi & Co. Ltd & Another (1987)eKLR
- Anyanzwa vs Gasperis(1981)eKLR;

- Kenya Tea Development Authority Ltd -VS- Andrew Mokaya, HCCA No. 174 of 2006
- Cox v Ministry of Justice (2016)
- Ryland vs Fletcher(1868
- Christine Kalama v Jane Wanja Njeru & another (2021) eKLR
- Midans Services Limited & another v Ronald Kapute(2022) eKLR
- Brisley v Drotsky(2002)
- United States v. Carroll Towing Co(1947), US Court of Appeals for the Second Circuit
- Amanat v. S.E.C, United States Court of Appeals for the Third Circuit(2008)
- Jones v. W + M Automation, Inc (2006) vLex, New York Supreme Court Appellate Division
- Kenneth Maweu Kasinga v. Cytonn High Yield Solutions LLP & another(2020)

List of Abbreviations:

- ANT - Actor-Network Theory
- AI - Artificial intelligence
- CPA-Consumer Protection Act
- DL - Deep Learning
- EU - European Union
- ML - Machine Learning DL
- IoT - Internet of Things
- PLD - Product Liability Directive
- RSK - Robotics Society of Kenya

Abstract

The rise of AI throws a wrench into established civil liability rules. This paper identifies these roadblocks and proposes solutions on how to adapt legal frameworks to hold AI accountable.

The investigation carried out within this dissertation focuses on the multilayered interplay between AI and regulation. The primary goal of this dissertation is to investigate and determine who is liable for harms caused by AI systems. The paper explores the complex issues surrounding assigning blame for the defects in AI systems, aiming to offer a thorough grasp of the practical, ethical, and legal aspects of this matter. Furthermore, we examine the evolving field of artificial intelligence (AI), the current legal systems, and emerging issues via a multidisciplinary lens with the goal of illuminating the way forward for developing a fair and responsible framework for AI system accountability. The analysis moves for a proposal for a legal framework that can

accompany the development of AI in Kenya, as it is paramount for the diverse actors who operate within.

CHAPTER I
INTRODUCTION

"Intelligence is the ability to adapt to change"

Stephen Hawking

1.0 BACKGROUND

Artificial intelligence(AI) has undoubtedly revolutionised the world as we know it. Its presence permeates through almost all human activities and inclinations, it is a fact that we are living in an era of intelligent machines. Tractica projects that by 2025, the global market for AI software would have grown from $1,4 billion in 2016 to about $60 billion[1]. Applications of AI are expanding endlessly; from virtual personal assistants, commercial recruitment systems, video games and electric home appliances, to facial recognition software that can flag criminals. Law firms are already hiring artificially intelligent attorneys[2], driverless cars are transforming transportation[3], AI systems execute complex financial transactions, and events such as the Project Maven expose the use of AI in military operations[4].

This is also true for the medical field as AI is adopted in various significant areas[5]. In Austria, it aids radiologists in more precise tumour detection by swiftly comparing X-rays with a substantial dataset of other medical information[6]. Meanwhile, in Denmark, AI plays a crucial role in life-saving efforts, enabling emergency services to diagnose conditions like cardiac arrests based on the caller's voice[7]. Additionally, farms across Europe have embraced AI to monitor and regulate their animals' movement, temperature, and feed consumption. This AI system can automatically adjust heating and feeding equipment, empowering farmers to oversee their animals' well-being while allocating time for other responsibilities[8].

---

[1] Wheelock C, 'Artificial Intelligence Software Revenue to Reach $59.8 Billion Worldwide by 2025', BusinessWire, < Artificial Intelligence Software Revenue to Reach $59.8 Billion Worldwide by 2025, According to Tractica | Business Wire> on May 2nd 2017.

[2] Jesus C,'AI Lawyer "Ross" Has Been Hired By Its First Official Law Firm', *Futurism*, < AI Lawyer "Ross" Has Been Hired By Its First Official Law Firm (futurism.com)> on 25th October 2017.

[3] Gordon C,'Driverless cars market leaders innovating the transportation industry', *Forbes,* <Driverless Car Market Leaders Innovating The Transportation Industry (forbes.com)> on December 29th 2021

[4] Manson K,'AI Warfare Becomes Real for US Military With Project Maven', *Bloomberg*, <AI Warfare Becomes Real for US Military With Project Maven (bloomberg.com)> on February 28th 2024

[5] European Commission,Artificial Intelligence for Europe, *SWD(2018) 137 final*, 2018, 1.

[6] The Royal Marsden Foundation Trust, AI Imaging Hub, <AI Imaging Hub | The Royal Marsden>.

[7] Corti, 'How the Capital Region of Denmark leverage Corti's AI to improve patient care', <corti.ai/stories/capital-region-of-denmark-using-ai-to-improve-patient-care-on-the-national-healthcare-system>, on 30th August 2021.

[8] Owkzarek D, 'AI-Based Smart Farming. The Rise of Machine Learning in Livestock Farming',*Nexocode*, <AI-Based Smart Farming. The Rise of Machine Learning in Livestock Farming (nexocode.com)> on October 11th 2022.

Various sectors in Kenya are actively embracing AI and machine learning to enhance business processes, analyse user behaviour, and forecast potential purchasing patterns. The utilisation of AI presents notable advantages that can be harnessed to address Kenya's prevalent challenges across diverse fundamental societal sectors[9]. It is currently observed in multiple industries including agriculture, health, education, fintech, and transportation, among others[10].The nation has been placing itself in a position to benefit from Fourth Industrial Revolution (4IR) [11] technology like AI. The exploration of AI in Kenya is not only broad with examples such as the KEMSA AI pilot project[12] and the 'cervical selfies' program[13] but concerns other matters of great significance. The future will witness a surge in conflicts involving AI due the expanding usage of AI technology in our daily lives and the potential for harm that might result from its failure. Policies and regulation must therefore balance with innovation. The goal is promoting the private sector vis a vis the protection and security of Kenyans.

The recent progress in AI technology has led to the emergence of systems capable of autonomous and unexpected actions without human supervision. Through unsupervised learning, machine learning algorithms can learn from previous actions and independently create novel behaviour patterns. While this may be good[14], these methods open the possibility for algorithmic misbehaviour without human influence. It becomes apparent that while the increasing autonomy facilitated by AI brings numerous benefits, it also introduces unknown risks. AI brings with it various potential hazards, including obscure decision-making processes, bias, unlawful content moderation practices, invasion of privacy, and potential exploitation for criminal activities, among others. The issue of AI liability, specifically, has become a tangible concern in contemporary societies. Instances of flawed reasoning and decision-making have been increasing over the years[15], common occurrences are cases of self-driving cars[16] killing pedestrians or Microsoft's 'Tay' AI bot that severely exposed it's unfounded racial bias among other negative

[9] Kenya AI, *AI in Kenya Research* <https://kenyaai.ke/research/ai-in-kenya> August 8th 2024

[10] Akello J, Policy Brief: Artificial Intelligence in Kenya, *Paradigm Initiative*, page 6<https://paradigmhq.org/wp-content/uploads/2022/02/Artificial-Inteligence-in-Kenya-1.pdf > January 2022.

[11] Crameri A, Artificial Intelligence: The fourth industrial revolution,MMAN, <Artificial Intelligence: The fourth industrial revolution (information-age.com)>, 3rd October 2019.

[12] Kabuchi J, "*KEMSA Banking On Tech To Boost Its Supply Chain System"* <KEMSA Banking On Tech To Boost Its Supply Chain System | CIO Africa> April 5 2022

[13] Global Citizen, "*Yes, cervix selfies are happening in Kenya— for a good cause* <Yes, cervix selfies are happening in Kenya— for a good cause (globalcitizen.org)> November 10, 2015.

[14] European Parliament, REPORT on saving lives: boosting car safety in the EU,*A8-0330/2017*, Section 16.

[15]Ravi R, *"From Tesla's fatal car crash to false facial recognition matches, here are five times AI failed to deliver*, <AI Gone Wrong 5 Biggest AI Failures Of All Time (jumpstartmag.com)> June 28,2021.

[16] Inc Africa, "*The 1 Fatal Flaw of Self-Driving Cars Industry experts say plans for self-driving cars depend on one unreliable assumption*." <The 1 Fatal Flaw of Self-Driving Cars (incafrica.com)> July 19, 2016.

connotations.[17]Although current AI technology is far from capable of substantial self-redesign, AI-driven machines and algorithms are integral to our reality today, capable of causing both physical and non-physical harm to individuals and society[18].

Amid intense global competition, it is imperative to establish a robust national strategy, leveraging the insights from the Blockchain and Artificial Intelligence Taskforce[19] initiated by the Kenyan government in 2018. This taskforce was mandated to provide guidance on the efficient utilisation of AI. Furthermore, Kenya is considering the adoption of the Kenya Robotics and Artificial Intelligence Society Bill[20], aimed at promoting the growth and development of robotics and artificial intelligence within the country. In Kenya, liability for defective goods is placed on the manufacturer who has a duty to compensate for harm or injury.[21]However this application is inconsistent with the nature of AI, as well as the recognition of AI as a service. Additionally, the legislative delineation of 'goods' within the Act fails to encompass AI.

The Data Protection Act(Rev. 2022) is similar as the onus is strictly on the data controller or data processor. However, clause (3) of Section 65 provides for a prominent loophole on the subject of liability if the accused(data controller) successfully proves they are not in any way responsible for the event giving rise to the damage[22];thus the plaintiff may be left uncompensated, a shortfall within the legal framework. The current legal atmosphere does not go so far as to recognize the peculiar and innovative nature of AI systems, whereby defects are not always the fault of the manufacturer. The Consumer Protection Act(2012)[23] aims to safeguard customers against dangerous, subpar, or unsafe products, unsatisfactory services, false or misleading information, or unfair trade practices. At any rate, the Act is based on a conventional producer-consumer paradigm. The subtleties of AI systems, which frequently include several players (developers,manufacturers, and users) contributing to the finished product or service, may be difficult for this model to fully capture.As such, the existing national framework may not sufficiently handle the aspect of unanticipated harm arising from the autonomous learning of AI.

Defining AI

The law and its principles are constrained because we don't know how far we can stretch human responsibility over the actions of AI. Thus, in pursuit of this objective, it is imperative to begin at

[17] Reese H, "*Why Microsoft's 'Tay' AI bot went wrong*", <Why Microsoft's 'Tay' AI bot went wrong | TechRepublic> March 24, 2016,
[18] Garcia v. Character Techs., Inc. (2024): A new lawsuit claims an AI chatbot responsible for minor's suicide.
[19] Makubi C*, Kenya Blockchain and AI taskforce report focuses on Government's Big Four Agenda,* <Kenya Blockchain and AI taskforce report focuses on Government's Big Four Agenda - coinweez> July 26, 2019.
[20] The Kenya Robotics and Artificial Intelligence Society Bill, 2023
[21] Section 64*, Competition Act*, No.12 of 2010.
[22] Section 65, *Data Protection Act*, No 24 of 2019.
[23] S*ection 12-16*, Consumer Protection Act, (No. 46 of 2012)

a definition, in order to comprehend the scope of the question at hand. It is widely acknowledged that there is no universally agreed-upon definition of AI, with variations in perspectives even among experts. This lack of consensus poses significant challenges in devising a definition suitable for regulatory frameworks.

The term artificial intelligence was first coined by John Mcarthy, otherwise known as the father of AI. According to him, intelligence is "the computational part of the ability to achieve goals in the world". On the other hand AI, or 'cognitive technologies', is simply "the science and engineering of making intelligent machines, especially intelligent computer programmes."[24]
Stuart Russell and Peter Norvig's *Artificial Intelligence: A Modern Approach* is one of the most influential textbooks concerning the subject matter. The authors adopt a human-centric viewpoint, organising diverse scientific definitions of AI into four categories based on human thought processes and behaviour: thinking humanly, acting humanly, thinking rationally, acting rationally.

Utilising the cognitive modelling approach, thinking humanly entails developing a detailed theory of the mind and translating it into a computer program. This approach aims to automate activities typically associated with human cognition, including decision-making, problem-solving, and learning. Acting humanly is seen from the lense of the Turing Test, wherein a computer passes the test if, following the administration of written questions, a human interrogator is unable to distinguish between written replies coming from a person and those from an automated system[25].
Rational thought entails leveraging existing knowledge to achieve a desired outcome. It emphasises the application of logical reasoning, sound judgement, and critical thinking skills to arrive at well-founded conclusions and decisions. Finally, in acting rationally, the approach taken views a rational agent as one 'that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome'[26].
While Norvig and Russell's classification system encompasses a wide range of attributes associated with potential AI capabilities, it falls short of providing a concise definition. This inconsistency in definition is not deliberate; rather, it reflects the evolving nature of technology. The increasing capabilities of computers enable them to perform tasks that were previously beyond their reach, thereby influencing the terminology and conceptualization of AI.
In reality, artificial intelligence is an umbrella term consisting of multiple subfields. For the purposes of this paper. this thesis concurs with Matthew Scherer's concept, according to which

[24] McCarthy,J (1979*) "Ascribing mental qualities to machines. In Martin Ringle (ed.), Philosophical Perspectives in Artificial Intelligence.* Humanities Press.*:* Mccarthy J, *'WHAT IS ARTIFICIAL INTELLIGENCE?,* Stanford University, November 2017.
[25] Turing A, 'COMPUTING MACHINERY AND INTELLIGENCE', Oxford University Press on behalf of Mind Association, Mind, New Series, Vol. 59, No. 236,2010, at 434.
[26] Stuart Russell, Peter Norvig, '*Artificial Intelligence: A Modern Approach',* Third Edition, Pearson Education, New Jersey, 2010, at 20.

artificial intelligence may be defined as any expert system that possesses the systemic capacity to carry out activities that call for human intellect[27].

From a national standpoint, The Kenya Robotics & Artificial Intelligence Society Bill(2023) dictates that artificial intelligence means 'the ability of machines to perform tasks that are typically associated with human intelligence, such as learning and problem-solving'[28].
This is similar to the EU definition in the AI Act, whereby an 'AI system' means "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"

1.2 Statement of the Problem

Artificial intelligence is increasingly being integrated and adopted in Kenya transcending almost all industries,' it is virtually a component influencing our daily lives[29]. Emerging technology such as AI are envisioned in the Kenya National Digital Master Plan(2022-2032)[30] , as they offer substantial benefits that may be utilised to address the most pressing issues facing Kenya, spanning across several fundamental socio economic sectors[31]. As it continues to flourish, there comes a significant risk of the inevitability of faults or flaws made by the autonomous AI systems inferring disputes in this regard, as with any product or service. Internationally, AI has already demonstrated its massive potential for significant defects to which the law must regulate. This is due to it's Blackbox nature and exacerbated autonomy through deep learning structures; an AI system may solely and certainly arrive at wrong or unethical decision which causes harm or injury. Concomitantly, there are other stakeholders present such as programmers and designers that may be involved in the matter. The inputs and operations of the impenetrable systems remain opaque to the user or any other interested party, this implies a huge obstacle in proving causative links for claimants thus evading justice and responsibility. The applicability of the common law tort causation test has always been founded upon human actions, yet it would be illogical to utilise this test when determining liability in AI. It would not be prudent to enact legislation that exclusively imposes strict liability on manufacturers within the existing legal framework in Kenya.

---

[27] Scherer M, 'REGULATING ARTIFICIAL INTELLIGENCE SYSTEMS: RISKS, CHALLENGES, COMPETENCIES, AND STRATEGIES', Harvard Journal of Law & Technology Volume 29, Number 2, 2016, at 362.
[28] *Sec 2*,The Kenya Robotics and Artificial Intelligence Society Bill, 2023
[29] Kirwa M, "Kenya 5th in Africa on readiness to adopt AI in public services" <Kenya 5th in Africa on readiness to adopt AI in public services (the-star.co.ke)> on 23 February 2023.
[30] Kenya Digital Master Plan, Kenya - Digital Master Plan.pdf (kippra.or.ke),71
[31] Paradigm Initiative, Artificial Intelligence in Kenya <https://paradigmhq.org/wp-content/uploads/2022/02/Artificial-Inteligence-in-Kenya-1.pdf> page 6.

This research will therefore seek to determine how liability should be attributed for civil faults made by AI systems. This study will be limited as within the Kenyan context concerning propositions for change and transformation with regards to AI.

1.3 RESEARCH OBJECTIVES

1. This study aims to determine 'Who should be liable for faults or harm(civil liability) caused by AI systems in Kenya?'
2. The research intends to assess the forms of civil liability that exist in relation to AI.
3. The study will then attempt to assess the challenges surrounding attribution of liability.
4. An investigation will be done to evaluate who are the agents that may be held liable.
5. Next, the study will explore the international legal framework governing civil liability of AI related harm.
6. The research will then pursue a discussion on whether the current Kenyan law is appropriate to address liability claims for emerging technologies such as AI.
7. Subsequently, the research will review how courts are currently resolving this issue, aiming to finally examine the most suitable legal framework to regulate defective AI in Kenya.

1.4 RESEARCH QUESTIONS

The main ORQ is 'Who should be liable for faults or harm(civil liability) caused by AI systems in Kenya?'

In order to solve this, we must examine the various forms of civil liability that ensue. It is also paramount to consider the complexities surrounding the attribution of liability, taking into account factors such as causation, control, and autonomy. Essential to the examination is a scrutiny of the principal agents and actors involved, evaluating their potential liability in AI-related incidents. Furthermore, the study extends its purview beyond national boundaries to scrutinise the international legal framework, encompassing both established norms at the international level and pertinent case law. It is imperative to critically assess whether the current national legal framework in Kenya is adequately equipped to address claims of culpability arising from emerging technologies such as AI. Moreover, aiming to comprehend how courts currently grapple with this intricate matter and explore potential avenues for the development of a more fitting legal framework to govern defective AI systems in Kenya. This endeavour entails considering whether an expansion of the prevailing liability regimes would suffice, or if fundamental enhancements to the existing legal framework are imperative.

1.5 HYPOTHESIS

Kenya is swiftly integrating emerging technologies such as artificial intelligence within its national systems and programs. The lack of legal clarity and soundness in view of civil liability thus becomes a pressing issue.The primary challenge here is the inherent ambiguity surrounding the

decision-making processes of AI systems. The traditional legal approach adopted is one of strict liability on the manufacturer for any defects or harm caused by such AI systems. This practice pursues the legitimate aims of justice and equity[32] through the provision of guaranteed protection. However, the law is suitable only to the extent that liability for defective goods was actually the fault of or rather caused by the manufacturer[33]. Furthermore, Section 66(3) of the Competition Act (No.12 of 2010)[34] stipulates that it is a defence to establish that the state of scientific or technical knowledge at the time when the goods were supplied by their actual manufacturer was not such as to enable that defect to be discovered. Seeing as AI is founded upon machine learning, this can easily be proven and responsibility evaded. Hence, depending solely on conventional models to assign liability for consumer goods and services, including AI, would ultimately undermine the pursuit of legal justice, as these approaches prove inadequately equipped for the task.

Therefore, this study postulates that a comprehensive framework that considers the distributed responsibilities within the AI ecosystem, including developers, users, and regulatory bodies, is necessary to ensure fair, ethical, and effective redress mechanisms. The hypothesis posits that a nuanced approach that encompasses legal, ethical, and technological considerations will contribute to a more robust and adaptive system for AI liability in Kenya, fostering responsible AI development and usage in alignment with global standards and emerging best practices.

1.6 JUSTIFICATION

Kenyan law concerning emerging technologies such as AI is vague and rather novel[35]. Inferences of this is legal insufficiency, parallel to the significant degree of AI implementation within the country[36]. The law, or a lack of would thus influence the society at large as there remains a void that requires substantial fulfilment. The current law as it exists encompasses traditional common law principles that exclude the AI aspect of machine learning and independent AI decision-making as well as the possibility of liability on other stakeholders. Therefore, a plaintiff's case may be doomed if they are unable to satisfy the fundamental requirements of evidentiary sufficiency regarding causation. This is because they will not be able to reconstruct the circumstances of AI's reasoning process and trace the chain of events that led to a specific (faulty) output. Additionally, continuing to place strict liability on the manufacturer would be unjust. In order to promote technological innovation in the private sector, the law must facilitate security and guidelines for the participating agents.

Consequently, this research will be useful to vividly address the minimal legal framework surrounding AI civil liability to ensure that claims for compensation for harm caused by AI systems

[32] Ubi Jus Ibi Remedium: where there is a right there is a remedy

[33] Section 64*, Competition Act*, No.12 of 2010

[34] Section 66, *Competition Act*, No. 12 of 2012.

[35] Kenya National Digital Masterplan(2022-2032), 72.

[36] Lukhanyu V, Challenges in Implementing AI Regulations in Kenya, *Bake Bloggers Association of Kenya*, 2025 <https://blog.bake.co.ke/2025/02/18/challenges-in-implementing-ai-regulations-in-kenya > on February 18th 2025.

are more successful. This aspect of AI application is critical yet has not been accorded sufficient weight; it is the mandate of the law to not only progress and improve with the changing realities but regulate and control as well. This research must therefore assess the most suitable legislation on the matter of attributing liability within the Kenyan context, giving substantial weight to the nature of AI systems. An in-depth analysis of the amplification of the element of causation in liability claims would aid legislators and adjudicators in the pursuit of evolving the law, as well as the tech industry and researchers in addressing AI and liability.

1.7 THEORETICAL FRAMEWORK

A Network Theory Analysis of AI Liability: (Actor-Network Theory)A framework to ascribe responsibility

The significance of network theory is frequently overlooked while analysing and resolving unique legal dilemmas brought on by developing technology[37]. This notion has the potential to aid regulators, judges, and insurers in comprehending intricate legal issues more effectively. By providing a visual framework, it helps them grasp the complexity of novel legal inquiries and discern more suitable policy interventions[3839]. Within the legal domain, scholars have extensively explored and applied network theory across various legal analyses, marking a notable increase in its presence within legal academic discourse in recent years. Examples of its applications include the proposal of models to address internet jurisdiction challenges, enhancing the regulation of the intersection between intellectual property and antitrust laws[40]. Additionally, network theory has been employed to examine the structural relationships formed by case law citations[41], aiding in a more profound understanding of the violation of privacy rights concerning information circulating within a network, distinguishing between public and private realms[42]. It has also served as a

[37] Law J, Actor Network Theory and Material Semiotics in B. S. Turner (Ed.),*The New Blackwell Companion to Social Theory*(2008), Wiley-Blackwell, Oxford, PG 7-10.

[38] Callon M, 'Actor-Network Theory – The Market Test',The Sociological Review 47 Volume 47 Issue 1,pg 192, <https://doi.org/10.1111/j.1467-954X.1999.tb03488.x> May 1999.

[39] Strahilevitz L, 'A Social Networks Theory of Privacy, U Chicago Law & Economics, Olin Working Paper No. 230; U university of Chicago, Public Law Working Paper No. 79, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=629283#> 3rd December 2004.The author places great emphasis on the application of network theory in the field of law in the future. He posits that our assessment of the necessity for private actor legal regulation is aided by our understanding of the pertinent social networks. In summary, social network theory might serve as a basis for intriguing new societal norms in the future. can be regarded as responsible, they are not autonomous nodes, as this article's remaining sections will make clear.

[40] McGowan D, 'Networks and Intention in Antitrust and Intellectual Property', *Hein Online*, 1999, , <https://www.semanticscholar.org/paper/Networks-and-Intention-in-Antitrust-and-Property-McGowan/2880a3da4bf28c261f56fe45bb3e315d127b5872>, on 1st April 1999

[41] Howler J et al, 'Network Analysis and the Law: Measuring the Legal Importance of Supreme Court Precedents, Political Analysis', Vol. 15, No. 3, pp. 324-346, 2007,9 , <Network Analysis and the Law: Measuring the Legal Importance of Supreme Court Precedents by James H. Fowler, Timothy R. Johnson, James F. Spriggs II, Sangick Jeon, Paul J. Wahlbeck :: SSRN>on 29th September 2009.

[42] Strahilevitz L, 'A Social Networks Theory of Privacy, U Chicago Law & Economics, Olin Working Paper No. 230; University of Chicago, Public Law Working Paper No. 79, pg 25, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=629283#> 3rd December 2004

valuable tool in determining the secondary liability of a contact supplier, within contract law, assessing the non-obvious requirement for patents[43], and creating models for studying the enforcement networks of governmental agencies[44]. This multifaceted application underscores the versatility and relevance of network theory within the evolving landscape of legal scholarship and analysis.

As such, this theory suggests an analysis of network theory as one possible instrument to enable stakeholders, such as legislators, regulators and the judicial system, to obtain tangible measurements to better identify the human entity that should be held liable once an AI entity has caused damages[45]. The theory sees legal issues as linked systems with different players acting as "nodes" and influencing one another through relationships called "edges". In simple terms edges are the interactions and actions that link human nodes and non-human nodes in the form of AI units. The position hinges on the knowledge that the interactions between individual components have a significant impact on their behaviour, highlighting the need of bringing the network's collective dynamics into account. Because AI systems are inherently complex, this methodology is especially well-suited for studying AI liability.

Simply put by Markus Schimer[46]Network theory involves the study of the way elements in a network interact. It is the study of symmetric or asymmetric relationships among interconnected objects, or "nodes" in a system, in order to better comprehend the operations and interactions of its many parts[47]. The first step is to identify and scrutinise the key actors or crucial nodes within the AI lifecycle, these include developers, makers, operators, users, regulatory entities, and the AI system itself. By examining the unique roles and duties of each element, the study aims to reveal how influence and accountability flow within the network structure.

These two components-edges and nodes-give an overview and comprehension of a network for every particular system that a stakeholder wants to study. There are an endless number of connections, or edges, that link nodes. The "boundary," or the pertinent nodes and edges in the network under consideration, is another crucial concept[48]. The determination of a network's boundary is typically subjective, depending on the observer's perspective. In a theoretical AI liability scenario, the observer could be a judicial or administrative authority examining an AI

---

[43] Pedraza-Fariña L, 'A Network Theory of Patentability' ,University of Hong Kong Faculty of Law Research Paper No. 2019/005, 2019, pg 137 , <A Network Theory of Patentability by Laura G. Pedraza-Fariña, Ryan Whalen :: SSRN>, on 5th May 2020.

[44] Nguyen L, 'Network Analysis in Public Policy',Encyclopedia of Public Policy, Network Analysis in Public Policy | SpringerLink> on 12 May 2023.

[45] Matwyshyn A, 'Of Nodes and Power Laws: A Network Theory Approach to Internet Jurisdiction Through Data Privacy', Northwestern University Law Review, Vol. 98, 2004, 36 ,Of Nodes and Power Laws: A Network Theory Approach to Internet Jurisdiction Through Data Privacy by Andrea M. Matwyshyn :: SSRN > 15th June 2006.

[46] Borgatti, Stephen & Lopez-Kidwell, V.. (2011). Network theory. The Sage Handbook of Social Network Analysis. 40-54.

[47] University of Southern California, *"What You Need to Know About Network Theory"* < What You Need to Know About Network Theory | USC Online MCM

[48] Lior A, 'Artificial Intelligence and Tort Law – Who Should be Held Liable when AI Causes Damages?', *Heinrich-Böll-Stiftung Logo,* part 3, Artificial Intelligence and Tort Law – Who Should be Held Liable when AI Causes Damages? | Heinrich-Böll-Stiftung | Tel Aviv - Israel (boell.org), on 24th December 2021.

incident, with the nodes and edges in the network encompassing entities like the victim, the AI system, its manufacturer, and programmer.

AI Liability & Network Theory in Practice

After an AI entity causes harm, the network boundaries broaden to encompass nodes representing the victims, human entities associated with the AI entity, and so forth. The network, designated as the "Accident Network," integrates numerous nodes linked through diverse types of connections. The primary nodes of significance are the AI entity itself and the victim or claimant, connected by an edge denoting "damage." Subsequently, the AI entity extends to incorporate additional edges, such as ownership, testing, production, programming, training, cooperation, etc. The analysis of the overall relationship between the AI injurer, the harmed parties, and the human entities associated with the AI entity takes into account the nature of these specific edges(i.e those who were hurt, and the people involved in creating and using the AI system, e.t.c).

It is probable that these nodes have varying degrees of connectivity with one another, serving different functions. The boundaries of the accident network will be determined by the victim and tortfeasor nodes based on the nature of their edges and the importance of their nearby nodes[49]. There can't be an endless accident network. Its usefulness as a tool to pinpoint the responsible party would be compromised by this. It also cannot be constrained by an exact, fixed number of edges or nodes. To make sure that only pertinent nodes and edges are taken into account, each accident network must be assessed and defined. The real advantage of applying network theory in the context of AI liability lies in its capacity to reveal the adjacent human nodes and their crucial function within the network. This ensures that these individuals are appropriately held accountable for their actions. It is crucial to remember that AI entities are not independent nodes; rather, they are directed and managed by human nodes that are connected to the AI node by various kinds of edges. This is critical information to keep in mind when determining which party should bear responsibility. These human nodes are in charge of the AI entities and are able to direct and advise them. In an accident network, the responsible human node or nodes should be determined by their level of control over the AI entity's behaviour, as determined by their capacity to monitor, direct, and pilot the AI node based on the nature and magnitude of their connection.

Immanently, AI nodes exhibit a high degree of connectivity, rendering them central entities with a considerable 'weight' in an accident network[50]. Their actions can have adverse effects on numerous human nodes directly or indirectly linked to them. Examining the characteristics of an AI node responsible for harm and the connected human nodes serves as an invaluable tool for

[49] Heath S, FullerA & Johnston B, 'Chasing Shadows: Defining Network Boundaries in Qualitative Social Network Analysis', Volume 9, *Sage Journals*, Issue 5, 2009, 645-661

[50] The smallest number of nodes or edges that can be removed in order to isolate the remaining nodes is how connectivity is determined. Analysing a system's degree of connectedness may provide a lot of information. For instance, the kind of edge (buyer-seller, trade partners, authority, physical connections, etc.) and its weight (a metric used to assess the relative relevance of a specific edge) might provide us with indirect information about the centrality, strength, and other characteristics of the connection.

establishing the requisite legal connections between damages and the wrongful acts that precipitated them. This proves crucial irrespective of the legal regime applied.

Actor-Network Theory (ANT) seeks to redefine society's connection to technology and non-human entities by treating both human and non-human actors, or nodes, equally within technology-infused networks[51]. While this dissertation acknowledges AI entities as autonomous nodes capable of causing harm, it does not endorse the notion that they should be granted the same agency as human nodes, however this will be elaborated further. This stance is taken based on the recognition that AI entities do not possess the inherent qualities and ethical considerations associated with human agency[52].

1.8 LITERATURE REVIEW

THE CURRENT LEGAL ATMOSPHERE & PROPOSED SOLUTIONS FOR THE FUTURE OF AI

The proper civil liability regime that should apply to AI entities is a topic of heated discourse within the academic and legal community.

Scholars such as Stanley Chukwubueze have approached the question by first addressing the place or legal status of AI[53]. The pursuit of legal personhood is viewed as a possible route to securing certain rights and safeguards for AI. Liability employs the process of "determining who is to blame with system failures—or, more precisely, who society may take legal remedy from—when anything goes wrong, in the words of Barfield.[54]If an AI system were to be given legal rights, the inference is that it would be held personally responsible for everything it did. Mostamai Sebabatso[55] considers this matter, suggesting that while the supervisory responsibility currently falls on humans, it's not the optimal solution for addressing liability concerns. However his idea of assigning criminal liability to AI seems defective as AI lacks the legal and moral attributes necessary for such responsibility. In contrast, Claudio Novelli considered the proper legal remedy to be the conferral of legal personhood, primarily justified by the necessity to address the resulting liability gap in the most effective manner. The author explores various manifestations of legal

---

[51] Latour B, *Reassembling the social: An introduction to Actor-network-theory*, Oxford University Press New York, 2005, pg 219

[52] Winner L, Upon Opening the Black Box and Finding It Empty: Social Constructivism and the Philosophy of Technology, Science, Technology & Human Values Volume 18, No. 3, 2008, 365, (PDF) Upon Opening the Black Box and Finding It Empty: Social Constructivism and the Philosophy of Technology (researchgate.net), on 9th December 2008.

[53] Chukwebueze O, Artificial Intelligence Vis-à-vis Its Prospects And Challenges:Legal Rights And Liabilities, *Sabi Law*<Artificial Intelligence Vis-à-vis Its Prospects And Challenges:Legal Rights And Liabilities (sabilaw.org) on 28 September 2022

[54] Barfiled W, 'Legal personhood in the age of artificially intelligent robots', Research handbook on the law of artificial intelligence,pg 213-250, on 28 Dec 2018.

[55] Motsamai S, 'Criminal and Civil Liability of an Artificial Intelligence (AI) for Cybercrimes, Machine or Robot for an Act or Conduct Committed Independent of Human Intervention or Control', Journal of Civil & Legal Sciences, Volume 11 • Issue 6 2022, pg 4, <Criminal and Civil Liability of an Artificial Intelligence (AI) for Cybercrimes, Machine or Robot for an Act or Conduct Committed Independent of Human Intervention or Control (omicsonline.org)> 2022.

personhood, alongside considerations regarding the financial or asset structure upon which the hypothetical new legal entity could be founded[56]. The hypothesis contemplated assumes that an AI possesses its own assets, enabling victims and creditors to seek compensation in the event of civil wrongdoing. Endowing artificial agents with their own assets would facilitate liability limitation and asset segregation, thereby mitigating risks akin to the organisational structure of legal entities. Depending on the design of the legal personality status, it would be feasible to distribute internalisation costs across multiple stakeholders while separating their individual assets, thereby incentivizing economic production and innovation.

Anat Lior[57] presents a different approach to legal personality in discussing the notion of a negligent liability regime, and infers the creation of a new standard of a "reasonable computer." Supporting this is the work of Ryan Abbott[58] who presents an argument that strict liability laws would inhibit innovation. In essence, strict liability should not apply if a provider can demonstrate that an autonomous computer, is safer than a reasonable person. The negligence test would consider a computer tortfeasor more like a person than a product since it would put more emphasis on the computer's actions rather than its design. Suppliers must thus demonstrate in a cost-benefit analysis that an AI entity is safer than an ordinary person. Responsibility will be determined by assessing negligence(standard) in cases where an autonomous computer is occupying the position of a reasonable person in the traditional negligence regime through focusing solely on the action of the AI rather than it's form.[59] A reservation however would perhaps be that the "reasonable person" standard will cease to exist; rather, our human acts will be held to a "higher" reasonable standard – that of a "reasonable computer.

Gerstner[60] however has observed the reluctance of courts to utilise the negligence test due to the ambiguity in identifying the source of the flaw to prove causation. A key diversion from this is the work of Jean-Sebastien Boghett and Gerhard Wagner[61] who propose a fait générateur[62] .This would mean considering liability based on circumstance. This implies an empirical and gradual development of law on the subject of liability as it would occur on a case to case basis.

---

[56] Novelli C, 'AI AND LEGAL PERSONHOOD: A THEORETICAL SURVEY', Universite du Luxembourg & Universita di Bologna, Published PhD Thesis, pg 131 ,16th June 2022

[57] Lior A, *Artificial Intelligence and Tort Law – Who Should be Held Liable when AI Causes Damages?* <Artificial Intelligence and Tort Law – Who Should be Held Liable when AI Causes Damages? | Heinrich-Böll-Stiftung | Tel Aviv - Israel (boell.org)> 24 December 2021

[58] Abbott R, The Reasonable Computer: Disrupting the Paradigm of Tort Liability The Reasonable Computer: Disrupting the Paradigm of Tort Liability - (gwlr.org)> January 2018 Vol. 86 No.

[59] Abbott R, The Reasonable Computer: Disrupting the Paradigm of Tort Liability The Reasonable Computer: Disrupting the Paradigm of Tort Liability - (gwlr.org)> January 2018 Vol. 86 , 29

[60] Gerstner M, *Liability Issues with Artificial Intelligence Software,* Vol 33, No.1, Santa Clara Law Review.

[61] Gerhard W, Robot Liability https://ssrn.com/abstract=3198764 > on 14 July 2018

[62] literally: the generating fact, but the expression is often translated into English as the 'event giving rise to liability'.This implies that depending on the circumstances and on the legal systems, different regimes may be applicable, such as product liability or liability for fault.

This is similar to the proposition of Yaniv Benhamou and Justine Ferland[63] who propose the adaptation of current fault-based liability regimes over novel principles. So, instead of considering solutions that would require amendments, simply adopting regimes with enhanced duties of care and precisions regarding shared liability and solidarity between tortfeasors potentially through case-law, is sufficient

Perhaps there ought to be an introduction of an e-personhood status as propounded by Tany Bonfim. In her analysis, liability is directed at the AI itself , endorsing it as a legal subject with rights and duties similar to that of corporate personhood. While this would not grant human rights to the AI, it would bring AI entities within the purview of legal liability[64]. Contributing to this discourse is Susanne Beck as she highlights a crucial differentiation between legal personhood attributed to corporations and the concept of electronic personhood: electronic personhood could potentially evolve to possess empathetic capacities and establish connections with humans beyond those of traditional legal entities[65]. She contends that a novel form of personhood tailored specifically for electronic entities is necessary to address the gap currently existing between conventional legal personality and the complexities arising from the actions of electronic entities not encompassed by traditional personality frameworks.
Adding to this discussion is the work of Filipe Alexandre[66] as he grapples on this question; *should robots pay taxes?*. This is disagreeable seeing as the utilisation of public services or infrastructure by an artificially intelligent agent does not result in a benefit for the agent itself, but rather for the user or designer who directed it to engage in activities requiring such public services or infrastructure. Given that AI agents are programmed to directly or indirectly serve human welfare, it follows that a human will inevitably be the ultimate recipient of the services or infrastructure utilised by the agent to fulfil its objectives. Consequently, he suggests a regulatory framework that encompasses two potential scenarios concerning autonomous actions: If the action stems from a coding deficiency or product defect, liability would be governed by negligence laws, holding the designer accountable. Alternatively, if the action arises solely from the evolving behaviour of the agent, the designer could absolve themselves by procuring insurance on behalf of the agent, otherwise assuming direct liability. By ensuring the agent possesses the means to compensate third parties for any damages it incurs, the designer effectively transfers their liability to the agent.

Ville Rautanen advances a different perspective, asserting that providing a definitive answer to the question of liability would entail stating: "The liable party hinges upon the circumstances of the

[63] Yaniv B & Ferland J, : Leading Legal Disruption: Artificial Intelligence and a Toolkit for Lawyers and the Law, Pina D'Agostino / Carole Piovesan / Aviv Gaon (éd.), Thomson Reuters Canada 2020
[64] Bonfim T, 'Criminal liability of artificial intelligent machines: eyeing into AI's mind',Lund University, Published LLM Thesis, pg 48, Spring 2022.
[65] Susanne Beck, 'Intelligent Agents and Criminal Law—Negligence, Diffusion of Liability and Electronic Personhood', Leibniz University, Unpublished pg 142, Spring 2016.
[66] Alexandre F, The Legal Status of Artificially Intelligent Robots, Tilburg University, pg 32, <(PDF) The Legal Status of Artificially Intelligent Robots: Personhood, Taxation and Control (researchgate.net)> 12th June 2017.

incident and the specific AI application involved." Therefore, a more pertinent inquiry might be: "What specific rationale renders party X liable in this particular case?"[67] Additionally, the author submits that an owner's liability approach would be rather suitable, comparable to an employer's liability when acting under the capacity of an agent, or parental responsibility.
Schaerer, Kelley, and Nicolescu draw a comparison regarding owner liability, suggesting that semi-autonomous machines bear a closer resemblance to animals, which exhibit independent behaviour, than to typical products, which lack autonomy and do not act autonomously[68].

In her analysis, Matilda Claussén-Karlsson deems that the befitting solution would be the imposition of a supervisory duty on the owner.[69]Establishing a civil law obligation for overseeing and intervening in AI operations to mitigate potential harm, even in cases where the action was unforeseeable to the supervisor. The rationale behind this approach is that implementing such a solution would not impede the development of AI technologies beneficial to humanity. Instead, it would curtail the likelihood of individuals engaging in risky behaviour with the knowledge that their actions or inaction would likely not result in criminal liability. Jomon Jose relies heavily on a general principle laid down in Article 12 of the United Nations Convention on the Use of Electronic Communications in International Contracts[70]. Considering the preceding points, he adopts the notion of AI-as-tool, wherein strict liability principles dictate the actions of the machine. This holds accountable the natural or legal entity on whose behalf the machine operates, irrespective of whether such behaviour was intended or anticipated.This exemplifies the use of respondeat superior liability doctrine, which holds that an individual bears accountability not for his own wrongdoing but rather for his association with the tortfeasor AI. The author continues by saying that the AI user or operator may sue the designer or manufacturer for damages if the AI behaves erratically and causes harm to a third party. According to him, regulatory requirements must aim to guarantee that AI systems have developed subsystems that provide an explanation of the decision logic for interrogation purposes[71].
Researchers Danila Kirpichnikov et al. advocate for the recognition of artificial intelligence as a potential subject of crime, asserting that such acknowledgment is both scientifically and practically warranted. This necessity arises from the objective imperative to safeguard the security of individuals, society, and the state amid the rapid advancement of technology and widespread

[67] Rautanen V, 'Liability issues with Artificial Intelligence in the national and international context', Published LLM Thesis, University of Turku, Pg 84, June 2020..
[68] Schaerer E et al, 'Robots as Animals: A Framework for Liability and Responsibility in Human-Robot Interactions', 18th IEEE International Symposium on Robot and Human Interactive Communication,2009, , <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2271466> 27th September 2009.
[69]Claussén-Karlsson M,'Artificial Intelligence and the External Element of the Crime An Analysis of the Liability Problem',Orebro Universtity, pg 45, <https://www.diva-portal.org/smash/get/diva2:1115160/FULLTEXT01.pdf> Spring 2017.
[70] *Article 12(2)(212),*United Nations Convention on the Use of Electronic Communications in International Contracts, United Nations Commission on International Trade Law.
[71]Jose J, 'Liability Issues and Regulation of Artificial Intelligence', National Law School of India University Bengaluru, pg 64, 2017-2018.

digitalization[72]. One may also argue, like Panagiota Stamatiou, that there is no necessity to formulate an entirely new AI-specific legal framework to handle product liability cases. However, some minor adjustments may be required to clarify certain concepts, such as 'product', 'producer', 'defect', 'damage', and the 'burden of proof'.These clarifications could be achieved through guidelines incorporated into legislative texts, which would provide specific definitions for these terms. That conventional frameworks(such as the Product Liability Directive) encompass a diverse array of defective products, and the maintenance of such legislative texts is crucial to respond to the demand for legal certainty[73].

Lena Anna Kuklińska distinguishes among three primary forms of liability applicable to AI. The merits and drawbacks of producer, operator, and AI's own (algorithm) liability are examined to determine which option might be most appropriate for potential application.It is recommended that the law should opt for a distinguished legal personality for AI, backed by approximation(within the EU) to allow for uniform tackling of the situation. In essence, this approach maintains that thee creation of a functional legal personhood for AI is key, with rights and obligations that could be outlined in the document, as well as an explanation of what AI is, the circumstances under which AI liability might apply, and the manner in which other frameworks might coexist with a new legal actor. Additionally, it is crucial to offer an explanation of the damage payment guidelines[74].
Numerous legal scholars emphasise the imperative of granting legal personality to AI, primarily due to the serious criminal implications resulting from its absence. This approach could facilitate the alignment of various legal domains with technological progress, preempt unforeseen legal complications stemming from innovative directions, and enhance clarity and certainty regarding the roles of distinct actors[75].

Evidently, there are currently multiple contrasting schools of thought on the subject of culpability. In line with the prevailing view, Charikleia Bertsia posits that the focus of the law should be on the legal persons(companies and individuals) who operate behind the AI systems. As AI cannot bear fundamental human characteristics such as morality or free will, liability for compensation falls on the party directly causing the damage or, in certain situations, on a party responsible for the actions of the damager, such as a parent for their child's actions.The tort system is portrayed by the author as the best option; backed by further state initiatives, it represents a mode of

[72] Kirpichnikov D, Pavlyuk A, Grebneva Y, Okagbue H, 'Criminal Liability of the Artificial Intelligence',E3S Web of Conferences 159,pg 8 , 2020.
[73]Stamatiou P, 'Artificial Intelligence and Product Liability', National and Kapodistrian University of Athens,Published LLM Thesis, pg 42, Athens, October 2021.
[74]Kuklińska L, 'Establishing liability for the actions of Artificial Intelligence; Recommendation for the European Commission',Institute of New Europe, page 17, September 2021.
[75] For example: Hallevy G, 'The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control', 199-201, 2010. ;Freitas P & others, 'Criminal Liability of Autonomous Agents: from the unthinkable to the plausible' (Universidade do Minho ,5 May 2021

regulation whose advantages and disadvantages are determined case-by-case as judges gradually familiarise themselves with expertise[76].

Tarek Nakkach [77]and Jomon Jose[78] suggest that the existing liability schemes should be adopted to the emerging market realities while considering the personhood of the AI. That liability for AI actions can be classified into: Vicarious liability, Strict Liability, Criminal Liability, or Common enterprise liability. As a result, Nakkack proposes that in order to determine attribution of liability, there should be accountability, i.e. determining why and how the AI makes the decisions it does. However, this would not seem effective as it would be too costly and redundant. This is a major advantage of the strict liability regimes already in place that would save transaction costs and complex or lengthy litigation.

Antonia Rapti also demonstrates her endorsement of the dominant viewpoint arguing that any potential harm brought about by the use of AI applications can generally be attributed to risks borne by individuals or companies, who may be its users or makers[79]. The author champions the work of the EU AI Act, encouraging its human-centric approach. It presents itself as union of both strict-liability and fault-based liability systems according to risk posed by the particular AI entity.

To sum up, investigating AI liability in light of current legal frameworks exposes a nuanced environment with dynamic changes and nascent remedies. The exploration of literature reveals a vibrant international conversation on AI liability. However, legal discussions specific to Kenya are currently limited. This gap presents a unique opportunity. It opens the door for increased scholarly focus and legal advancements in this crucial area within Kenya.

1.9.METHODOLOGY

This thesis will utilise a multifaceted legal research approach to fulfil its goals of establishing responsibility for the consequences of AI system-induced harm in Kenya. The nature of the research conducted in this study will be qualitative as it is most adequate to address the interdisciplinary nature of the study: law and technology. The study employs both primary and secondary sources of data. Primary sources include law such as the Competition Act(no.12 of 2010), Consumer Protection Act(rev. 2022), and EU AI Act(2024); while secondary data is introduced through sources such as journal articles, books, and reports etc. Deploying these

[76] Charikleia B, LEGAL LIABILITY OF ARTIFICIAL INTELLIGENCE-DRIVEN SYSTEMS ("AI"), International Hellenic University, Greece,2019, pg 50.

[77] Nakkach T, Top Ten Issues on Liability and Regulation of Artificial Intelligence (AI) Systems, ACC, <Top Ten Issues on Liability and Regulation of Artificial Intelligence (AI) Systems | Association of Corporate Counsel (ACC)> August 2nd 2021

[78] Jomon J, "Legal Liability Issues and Regulation of Artificial Intelligence", National Law School of India Bengalaru, CLCF/588/1, 2017-2018. < Legal liability issues and regulation of Artificial Intelligence (AI) Dissertation work -Post Graduate Diploma in Cyber Laws and Cyber Forensics Course Code: PGDCLCF Submitted by: Jomon P Jose | Jomon Jose - Academia.edu>

[79] Antonia R, 'Liability Issues Regarding Artificial Intelligence in a European and International Context',UNIVERSITY CENTER OF INTERNATIONAL PROGRAMMES OF STUDIES SCHOOL OF HUMANITIES, SOCIAL SCIENCES AND ECONOMICS, Greece, 2023, page 28.

sources to vividly assess and comprehend the discrepancy with the correspondence between national law and liability for autonomous artificially intelligent systems.
To commence, the research leans more towards a descriptive motion in order to truly comprehend the nature of the subject matter and the legislation regarded. This can be seen in the description of various facts and crucial concepts such as 'deep learning' and the 'Blackbox' feature, as well as in depth inquiries into the elements involved. Proceeding this and more generally is a prescriptive approach as the study aims to examine de lege ferenda, what the law ought to be in light of AI responsibility in Kenya.
The methodology incorporates elements of both inductive and deductive reasoning. The inductive approach present champions the comparative analysis evidenced through the evaluation of case studies involving AI liability in Kenya and other jurisdictions. Empirical data is gathered that seeks patterns or generalisations to derive at a standard conclusion. A critical analysis is called for as we will vividly review the subject matter in all its form in the interests of gaining comprehensive understanding and a foundation for the claim. Next, a doctrinal methodology is implemented in view of analysing primary legal sources, doctrines, and principles. The aim of this is to reinforce the objectives of how the courts are resolving this issue and thus the adequacy of the applicable law in addressing liability claims for AI. The direction inferred from the findings of this research reveal implications of the legal aperture;subsequently necessitating a comparison on contrasting legal systems in order to identify common themes and viable solutions to the problem. In determining the most suitable national legal framework, it is inevitable that a brief philosophical approach will be indispensable.

This comprehensive strategy aims to enhance the understanding of AI civil liability challenges in Kenya. Through the integration of legal doctrine research,reasoning and analysis, along with a comparative and socio-legal outlook, the dissertation intends to present a thorough investigation. Moreover, it seeks to recommend a pragmatic and legally robust framework to effectively tackle AI liability issues within the Kenyan setting.

1.9.1 CHAPTER BREAKDOWN

This study aspires to determine who should be ascribed liability for faults or harm caused by AI driven systems.
Accordingly, the first chapter of this research is simply introductory.
Chapter I presents the main RQ(Who should be liable for faults or harm(civil liability) caused by AI systems in Kenya?). . It describes summarily the problem at hand and its foundation, along with the research objectives to be studied and corresponding underlying questions. The chapter also provides for a hypothesis, theoretical framework i.e the lens through which the RQ is examined, a justification of the study, and an elaborate literature review among others.This work acknowledges the elusive nature of universally agreed-upon definition AI. Since regulation hinges on a clear understanding of what AI is, the text strives to define AI as precisely as possible

Chapter II dissects the different modes of civil liability that may be applicable in this regard.
Chapter II examines the modes of civil liability for AI-related harm. It explores key liability frameworks, including strict liability, where accountability is imposed regardless of fault, and fault-based liability, which requires proof of negligence. The chapter also considers vicarious liability, applicable when an entity is held responsible for harm caused by AI operated by third parties, among others.
Proceeding this, ***Chapter III*** explores the challenges surrounding attribution of liability, the elements that could affect identification of the root cause of the problematic action or omission of the AI unit and the responsible agent.. This section breaks down the core functionalities of AI, examining how these features make it difficult to regulate AI and determine who's responsible for harm caused by it.
Chapter IV scrutinises which actors are entangled within the web of an AI lifecycle, in order to realise why, how, and under what circumstances they may indeed be held liable.This chapter goes further to uncover why the specific components of an AI entity would affect the attribution of liability in claims for civil faults or injury as would be influenced by an actor. Additionally, consideration is given as to whether AI systems could display the prerequisites to establish liability, and to what extent this is feasible.
Chapter IV critically observes the existing international legal atmosphere in addressing liability claims for emerging technologies such as AI. Concomitantly, the chapter will delve into the different liability approaches that exist under common law giving substantial weight to the paramount examination of the element of legal causation.This chapter engages in discourse to assess the sufficiency of existing Kenyan legal provisions in handling liability allegations related to emerging technologies, specifically AI. An assessment is made in light of national law as to whether and under which conditions it can be applied to situations involving AI technology.
This will lead to ***Chapter V*** that considers solutions and recommendations. Fundamentally, it aims to identify the optimal legal structure for overseeing flawed AI systems in Kenya. This embraces an inquisition as to whether augmenting the present or conventional liability frameworks would be effective, or if enhancements can be implemented.

CHAPTER II
MODES OF CIVIL LIABILITY FOR DAMAGES CAUSED BY AI

The present chapter seeks to explore distinct categories of civil liability in accordance with their customary functions. This is in view of assessing their suitability to potential applications in AI liability. The applicability of each liability type hinges largely on the specific circumstances surrounding the occurrence of harm. The focal point of our inquiry revolves around confined and challenging questions: For instance, '*What types of responsibility can feasibly be attributed when the causal agent is an AI machine?*', and *'In what circumstances is it justifiable to assign a particular type of responsibility to a specific party?*'.

Application of traditional tort law theory to machine intelligence

The simple aim of tort law is to compensate for harm or damage caused to one person by another, based on the legal obligation of its prohibition. For an injured person to receive compensation under harm regulations, they must prove the fault, the resulting damage, and a clear connection between the two(causation)**.** The principle functions adequately when the tortuous relationship is between natural and legal persons, yet artificial intelligence remains exempt as a subject of law[80]. Tort law places a predominant focus on delineating the responsibilities of operators, manufacturers, and programmers, relegating autonomous agents to the status of mere hazardous or flawed products under their jurisdiction. Consequently, a dilemma emerges whereby tort law may fail to hold software agents accountable for autonomous decision-making that results in harm, even when human actors have adhered to proper conduct, leading to a significant liability gap. What's more is that it may subject human participants to an inappropriately stringent strict liability standard based solely on their utilisation of electronic agents. The crux of the matter lies in the mismatch between the core tenet of tort law, foreseeability, and the inherently unpredictable nature of autonomous algorithmic decisions, rendering the current legal framework inadequate for effectively addressing algorithmic failure[81].

Agency Law and Vicarious Liability

Vicarious liability finds its origins in the ancient doctrine of the *respondeat superior principle*("let the master answer"), which presupposes that the master retains control over the servant. The notion generally applies in matters of employer-employee relationships(where the tort occurred during the course of employment), the liability of a principal for the conduct of an agent acting under his

[80] Burylo Y ',Civil liability for damage caused by artificial intelligence: The modern European approach', June 2022, at 6.

[81] Beckers A & Teubner G, *Three Liability Regimes for Artificial Intelligence,* Hart Publishing, Great Britain,2021, at 6**.**

direction and for his benefit.[82].The nature of this relationship generally rests on the presumption that the person who caused the harm acted for the benefit of the one who is held responsible for it. The tortfeasor is also surmised to have acted under the direction or supervision of the person liable to compensate for harm[83]Hence, a fundamental principle emerges: those who reap substantial benefits from employing AI should shoulder the responsibility when errors occur.[84] Vicarious liability for algorithmic errors may occur when a human principal assigns a task to an algorithm, the delegation requiring the agent's independent decision-making which results in actions that are unpredictable and cannot be explained by the programmer. The action breaches a duty of care or results in damage with a clear causal connection between the action and the resulting harm. Consequently, the user of the algorithm, acting as the principal, is held accountable. Compensation for damages is not limited to the narrow confines of strict liability for industrial hazards but instead follows established principles of contract and tort law, particularly regarding the possibility of compensating non-monetary damages.

In the Report from the Expert Group on Liability and New Technologies, if harm arises from the use of autonomous systems in a manner akin to employing human assistants, the operator's liability for utilising the technology should align with the existing vicarious liability principles governing principals and their assistants[85]. Expanding the scope of vicarious liability to encompass scenarios involving advanced machines replacing human assistants ensures accountability. Without this extension, individuals may evade responsibility by assigning tasks and duties to machines rather than human agents[86]. The case of Cox v Ministry of Justice[87] states this notion plainly, 'For who is equivalent to an employee despite not technically being one' .

Even though the notion of vicarious liability implies no personal fault of the liable person the tortfeasor is expected to compensate notwithstanding. Exploring the ascription of vicarious liability suggests two categories of defendants. Firstly, "users," comprises parties who integrate AI, whether developed internally or by a third party, into various aspects of their business or daily operations(e.g safety management or trading decisions).The second category, perhaps best described as "suppliers," includes businesses that design, supply, or manage AI systems. Think of software developers, IT companies offering AI solutions, or contractors who operate AI systems for clients. Importantly, multiple parties could share liability for a single AI application, allowing

82 Khayigila vs Gigi & Co. Ltd & Another (1987)eKLR 76

83 Anyanzwa vs Gasperis(1981)eKLR; Kenya Tea Development Authority Ltd -VS- Andrew Mokaya, HCCA No. 174 of 2006

84 Dr Auty A, 'AI Systems – who is liable?', *Re: Liability(Oxford)Ltd*, 2023, <http://www.reliabilityoxford.co.uk/ai-systems-who-is-liable/>, on June 2, 2023.

85 European Commission, Expert Group on Liability and New Technologies New Technologies Formation, LIABILITY FOR ARTIFICIAL INTELLIGENCE AND OTHER EMERGING DIGITAL TECHNOLOGIES, 2019. 24-25.

86European Commission, Expert Group on Liability and New Technologies New Technologies Formation, LIABILITY FOR ARTIFICIAL INTELLIGENCE AND OTHER EMERGING DIGITAL TECHNOLOGIES, 2019. Key Findings 18 & 19.

87 Cox v Ministry of Justice (2016), Supreme Court of the United States.

claimants to pursue legal action against one or both parties with the potential for contribution between them if needed[88].

Vicarious liability is inherently tied to fault liability, where the principal bears responsibility for the actions of their auxiliary, even in the absence of personal fault. However, the evaluation of the auxiliary's conduct is not necessarily based on their own standards; instead, it is assessed according to the benchmarks set for the principal[89]. The framework limits itself so that a principal utilising autonomous systems for complex tasks incurs equivalent liability as one employing a human assistant[90] such that any flaws or inaccuracies in the system's operation may be placed on the principal. If the task requires abilities beyond human capabilities, like complex calculations, then the standard for acceptable performance becomes similar AI technology that the user could reasonably be expected to use[91].

Daniela Vacek puts forth an argument, positing that programmers and other individuals engaged in developing autonomous machines are typically regarded as employees of the manufacturer. Consequently, the actions or inactions of various actors are deemed to be within the scope of employment, as they operate for the interests and advantages of the AI manufacturer[92]. Designing or programming AI machines is deemed within the scope of employment. Therefore, instances of harm stemming from design flaws can be treated as cases of vicarious liability on the manufacturer, given that the harm can be linked back to violations of prohibitions or obligations by programmers or other contributors involved in the machine's design.

Despite its intuitive appeal, the agency approach bears its shortcomings. Implementing a vicarious standard for algorithmic harms carries the risk of deterring investment in algorithm innovation and unfairly penalising innocent parties[93]. This transition represents a dramatic shift from the current varied landscape of civil liability standards for machines, likely to face considerable resistance. Additionally, by treating all algorithms equally, this approach fails to provide a mechanism for favouring algorithms that embody important programming values such as transparency(i.e no black box feature).

---

[88] Asoro P, ''Determinism, machine agency, and responsibility', Politica & Societa, 2014, pg 292, <https://peterasaro.org/writing/Asaro%20DeterminismMachineAgency.pdf> February 2014.

[89] Busnelli F & Comande G, S Galand-Carval, 'Liability for Others' , *Principles of European Tort Law*, 2nd edition, Springer, Vienna, 2013, pg 112.

[90] Proff. Wanderhorst,Safety and Liability Related Aspects of Software, European Commission, 2020, pg 94, <https://digital-strategy.ec.europa.eu/en/library/study-safety-and-liability-related-aspects-software > November 2020.

[91] European Commission, Expert Group on Liability and New Technologies New Technologies Formation, LIABILITY FOR ARTIFICIAL INTELLIGENCE AND OTHER EMERGING DIGITAL TECHNOLOGIES, 2019. Key Findings 16, 19.

[92] Vacek D,'Vicarious Liability: A Solution to a Problem of AI Responsibility?',*Springer Nature*, pg 3, <https://www.researchgate.net/publication/361904109_Vicarious_Liability_A_Solution_to_a_Problem_of_AI_Responsibility >, on July 2022

[93] Yeung K & Lodge M, 'Algorithmic Regulation: An Introduction', *Algorithmic Regulation*, 1st Ed, Oxford University Press, United Kingdom, 2019, pg 5; -The purpose of algorithmic regulation is to achieve a harmonious equilibrium between safeguarding interests and fostering the advancement and creativity of development.

Strict Liability for AI related damage

What is the legal standing when there is no fault, yet the damage arises from the operation of AI systems? The sole recourse may lie in strict liability. The existing frameworks for strict liability proposed as potential remedies generally involve product liability, liability for harm inflicted by animals, and liability for harm resulting from motor vehicle accidents.

In the realm of strict liability, those deploying AI systems bear sole responsibility for any resulting harm, regardless of fault. This distinctive legal framework provides a path for compensating victims without requiring proof of fault or negligence on the part of the defendant. Strict liability, often termed risk-based liability, operates on the basis of risk. This risk may arise from certain objects or activities associated with heightened danger, such as motor vehicles, or the use of high-risk AI systems(e.g those integrated into critical infrastructure-transport, nuclear power, pipelines)[94]. As a report on AI liability explains, risk-based liability is not contingent on misconduct by a wrongdoer. Instead, it arises from the acknowledgment that individuals are permitted to engage in inherently hazardous activities or use dangerous objects for their own purposes which may often be uncontrollable. Consequently, they should also bear the consequences if such risks materialise[95]. Certain AI systems pose serious risks to human safety, health, and property. Autonomous vehicles, such as self-driving cars, could easily collide with pedestrians[96]. Medical software may recommend an inaccurate diagnosis or medicine, among other things. Scholars argue that AI systems pose a higher risk and liability out to be strict for any damage produced by them.

In Kenya, strict liability torts encompass product liability; for instance, if a machine injures you when used as instructed, the manufacturer may be held liable. Here, there's an assumption that the manufacturer or supplier was aware of the defect before it reached the plaintiff.

According to the Competition Act(2012) and the Convention of 2 October 1973 on the Law Applicable to Products Liability ,the producer, understood to be 'the manufacturer or supplier of goods as part of his commercial activity', is liable for defects in his product. A product is defined to include 'natural and industrial products, whether raw or manufactured and whether movable or immovable'.The manufacturer's liability extends exclusively to defects present in the product at the time of its market introduction, excluding those that manifest afterward[97]. The burden of proof

---

[94] Martin-Kasals M, 'The Development Of Liability In Relation To Technological Change',Volume 4, *Cambridge University Press,* 2010, pg 22;See also the Commission Staff Working Document, Liability for emerging digital technologies, Brussels, 25.4.2018 SWD(2018) 137 final, at 8.

[95] Ryland vs Fletcher(1868)states that "Anyone who in the course of non-natural use of his land, accumulates thereon for his own purposes anything likely to do mischief if it escapes is answerable for all direct damage thereby caused"; see also European Commission, 'Comparative Law Study on Civil Liability for Artificial Intelligence', November 2020, at 58.

[96] BBC, *'Uber's self-driving operator charged over fatal crash'*https://www.bing.com/search?q=autonomous+vehicle+kills+pedestrian&qs=SSA&pq=autonomous+car+kills+pes&sc=3-24&cvid=A0F15FC8C3BF4D80BBB0D70445D0E259&FORM=QBRE&sp=1&lq=0 > 16th September 2020.

[97] *Section 66*, Competition Act(No. 12 of 2012)

lies with the victim to substantiate the damage, the defect, and the causative link between the damage and the defect.The advantage of strict liability for the victim is apparent as it spares them from the necessity of demonstrating any wrongdoing by the defendant, or even establishing a causal connection between such wrongdoing and their loss or harm. Instead, the victim can focus exclusively on whether the risk associated with the technology materialised and led to harm. However, the question remains ambiguous regarding whether artificial intelligence, forged as a collection of algorithms, fits within the definition of a "product," and whether the unpredictable decisions made within the autonomous operation of AI can be classified as a "defect".[98].

It is argued by scholars that the use of the ultra-hazardous activity doctrine as a rationale justifying strict liability for AI faults is problematic, as it fails to recognize its independent and autonomous decision making nature[99]. Hence, the issue doesn't solely revolve around its uncontrollability, but also the unpredictability of its actions and resulting harm. Consequently, as such damage is unpredictable, it may not fall within the scope of activities traditionally categorised as posing an increased danger to others. The focal inquiry becomes whether AI systems be considered inherently risky activities, thus triggering strict liability[100]. More importantly, the present national regime offers only a partial fix, neglecting the question of liability for owners, users and other agents possibly involved with the composition and use of the AI system.

From another perspective, it is true that strict liability typically relies on control over a hazard or benefiting from it. For example, liability for damage caused by animals often falls on the animal's owner for benefiting from their use. Similarly, the unpredictable nature of AI systems poses inherent risks, suggesting a similar liability approach[101]. However, solely holding users accountable while exempting manufacturers or trainers may not be ideal. Creators often have exclusive access and understanding of the system; thus, it's questionable whether liability should exclusively rest on the party merely benefiting from the hazard, excluding creators or controllers who could mitigate risks[102]. Traditionally, tort law favours holding manufacturers accountable, given their control over the hazard source.

A crucial aspect to ponder on is the potential impact of such introduction on the progress of technology, as some individuals might hesitate to actively advocate for technological research and

[98] *Article 2*, Convention of 2 October 1973 on the Law Applicable to Products Liability; Antunes H, Freitas P, Oliveira A, Pereira C, Sequeira E, Xavier L, *Ours by right: Multidisciplinary perspectives on AI and the Law*, Springer, 2024,pg 307.
[99] Velyanova M, 'LIABILITY FOR DAMAGE CAUSED USING ARTIFICIAL INTELLIGENCE TECHNOLOGIES', Journal of the National Academy of Legal Sciences of Ukraine, vol. 28, no. 2, 2021,at 150–159, <(PDF) Liability for damage caused using artificial intelligence technologies (researchgate.net)>, June 2021.
[100] Burylo Y ',CIVIL LIABILITY FOR DAMAGE CAUSED BY ARTIFICIAL INTELLIGENCE: THE MODERN EUROPEAN APPROACH', June 2022, at 7.
[101] Antunes H et al, '*Multidisciplinary perspectives on AI and the Law'*, Springer, 2024,pg 309.
[102] Wenderhorst C, 'Strict Liability for AI and other Emerging Technologies', *Journal of European Tort Law*, volume 11, Issue 2, 2020, at part 3.

production if the risk of liability is perceived as a deterrent[103]. The chilling effect of tort law on AI development is intensified when liability remains uncertain, making risk assessment and insurance coverage challenging. Implementing a legislative framework explains legal risks, making them more predictable and easier to insure. In the transportation industry, imposing strict accountability on the driver for autonomous vehicle accidents may be dubious, given human interaction is frequently negligible. Instead, culpability should be judged in terms of system control, with manufacturers, software developers, and AI operators potentially held accountable. Strict liability for manufacturers remains essential for defective high-risk AI goods, whether physical or digital, to ensure recompense for harm caused by intrinsic faults. The producer should bear strict responsibility for flaws in emerging digital technologies, even if these flaws arise after the product has been released, provided they still have control over updates or upgrades to the technology. A defence based on the risk of development should not be applicable.

The inherent imperfection of an autonomous machine justifies a strict liability system as it guarantees compensation for the victim in light of consumer rights. However, its uncomfortable suitability remains apparent[104].

Fault based liability and AI related harm

Fault-based liability is a fundamental principle in Kenyan tort law whereby an injured party can sue for compensation if they can prove that the defendant breached a duty of care, causing harm[105]. Generally, fault-based liability relies on alleged wrongdoing by the party causing harm, based on that party's intentional or negligent conduct. This type of liability not only aims to compensate victims after the fact but also serves as a tool to encourage behaviour that avoids potential harm in advance.

As illustrated in the case of Christine Kalama v Jane Wanja Njeru & another, there exist key elements that must be satisfied in order to establish fault based liability[106]. The claimant must prove that the defendant owed a duty of care(reasonable standard of care to avoid harm) which was breached. This breach of duty of care caused the claimant's harm, and they suffered damages. It becomes important to highlight that pursuing a case based on traditional fault-based liability against an operator or user of an AI system is extremely challenging. The presence of an AI system in causing damages raises the complexity of proving fault and causation.The victim often lacks insight into the causal process, hindered by the opacity of the "black box" effect, which obscures transparency and explainability in decision-making. Furthermore, the involvement of multiple agents exacerbates the complexity of the situation. Meeting the plaintiff's burden of proving fault

[103] Hubbard F, 'Sophisticated Robots: Balancing Liability, Regulation And Innovation, Volume 66, *Florida Law Review*, Issue 5, 2015, and Innovation', at 1810.

[104] This particularly pertains to new digital technology that operate in public settings, such as autos and drones.

[105] Midans Services Limited & another v Ronald Kapute(2022) eKLR

[106] Christine Kalama v Jane Wanja Njeru & another (2021) eKLR

is rendered rather insurmountable[107]. For instance, an AI-powered recruitment tool might reject a qualified candidate due to biases embedded in its training data. While the employer using the tool could be held liable for relying on a discriminatory system, the source of the bias could lie with the creators of the AI or with the data providers inter alia.

Notably, the application of fault-based liability reveals a significant gap in addressing the unlawful actions of autonomous agents. Especially within common law systems, the use of negligence presents inherent complexities. Establishing a duty of care relies on foreseeability of harm from the defendant's perspective. However, when an electronic agent behaves in an unforeseeable manner, the absence of foreseeability undermines the existence of a duty of care[108]. Progressive legal realism proposes broadening duties of care to encompass the specific risks associated with autonomous agents. This entails extending the current obligations on human actors concerning the oversight of algorithms, based on specific mandates prescribed for operators, users, and manufacturers of high-risk AI[109]. This regime features an expansion of tort liability to impose a duty of care regarding the supervision and control of algorithms in any use of autonomous computers. Under such a regime, operators would be held liable for damages resulting from algorithms if they fail to adequately control their behaviour. However, this approach is considered radical and could effectively introduce a form of causation-based strict liability in disguise. By extending the duty of care to encompass control over any aspect of algorithm behaviour related to its use, the essential content of the duty of care could be compromised. Additionally, such broad duties could stifle the creative potential of autonomous algorithms' discovery procedures to an intolerable extent[110].

Evidently, the rise of emerging digital technologies poses challenges in applying fault-based liability principles, especially because there are no well-established standards for their proper functioning and because they can evolve through learning processes without direct human oversight. Demonstrating the attribution of damage to the agent's actions is difficult due to the absence of defined duties of care and complexities in assessing culpability. Moreover, objectively attributing damage to a specific participant in the process is hindered by the inability to pinpoint the exact cause, leaving room for uncertainty regarding individual responsibility[111].

---

[107] Antunes H, Freitas P, Oliveira A, Pereira C, Sequeira E, Xavier L, *Ours by right: Multidisciplinary perspectives on AI and the Law*, Springer, 2024,pg 300.

[108] Selbst A, 'NEGLIGENCE AND AI'S HUMAN USERS', *Boston University Law Review,* Volume 100:1315, Paper no. 20-01, 2020, pg 1331, <Negligence and AI's Human Users by Andrew D. Selbst :: SSRN>, on 14th October 2020.

[109] This proposed legislation is relevant to the subject matter: European Commission, Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonised Rules on Artificial Intelligence (Proposal Artificial Intelligence Act) and Amending Certain Union Legislative Act, *Final Draft,* 2021, Art 16-Obligations of providers of high-risk AI systems, Art 24-Obligations of product manufacturers, Art 29-Obligations of users of high-risk AI systems etc

[110] Beckers A & Teubner G, *Three Liability Regimes for Artificial Intelligence,* Hart Publishing, Great Britain,2021, at 73.

[111] European Parliamentary Research Service, 'Artificial intelligence liability directive' , 2023, pg 3.

Fault-based systems often fail to yield satisfactory outcomes in the context of AI systems due to the significant level of automation or autonomy involved. This heightened automation makes it increasingly challenging to attribute damages to negligent human behaviour. If an operator can demonstrate that they have consistently taken all necessary safety precautions, it becomes impossible to hold them accountable for the unpredictable actions of AI systems[112].

In the realm of negligence law as it stands, fault entails both an action or omission and a corresponding mental state that fails to meet a defined standard, often indicative of negligence[113]. While employers can be held responsible for their employees' actions, including those in analogous roles, mere malfunctions of machines under their control, such as computers assigned to safety monitoring, do not constitute grounds for liability. Liability hinges on the demonstration of personal fault by a human actor, such as an operator responsible for controlling or programming the machine.

The European Parliament's Resolution on a civil liability regime for artificial intelligence emphasised that, while high-risk AI systems should be subject to strict liability laws (combined with mandatory insurance coverage), any other AI-driven activities, machines, or processes that cause harm or damage should continue to be subject to fault-based liability. The affected individual would benefit from a presumption of fault on the side of the operator, unless the latter can demonstrate that it complied with its duty of care[114]. Fundamentally, there is doubt regarding whether a fault-based liability system would effectively streamline victims' claims, even with a presumption of causality in place. The contention is that AI systems are often so intricate that even when users adhere to their duty of care, damages may still occur, leaving uncertainty about liability and its basis.

It is important to note that the coexistence of fault liability, whether presumed or not, alongside strict liability for risks and defective products may be maintained simultaneously. When these liability regimes intersect, providing the victim with multiple avenues to seek compensation from multiple parties, the rules concerning multiple tortfeasors may govern the situation.

Contractual liability

Alongside the question of fault-based liability, contractual liability emerges as another possible facet in addressing civil wrongs caused by AI-driven systems. Through agreements among AI developers, providers, users, and relevant parties, explicit parameters can be set to define performance, encompassing expected functionalities and constraints of the AI system. These

[112] Ebers M, 'Liability For Artificial Intelligence And EU Consumer Law' , <delivery.php (ssrn.com)>pg 215.
[113] Tettenborn A, 'Negligence', *Clerk & Lindsell on Torts*, 24th ed,Sweet & Maxwell, UK,2023, Chapter 7.
[114] Lodie A, Celis S, Karathanasis T, 'TOWARDS A NEW REGIME OF CIVIL LIABILITY FOR AI SYSTEMS: COMMENT ON THE EUROPEAN COMMISSION'S PROPOSALS', AI Regulation.com, 2022, <Civil Liability for AI Systems: Comment on EU Commission's Proposals (ai-regulation.com)>, on 14th October 2022.

agreements serve as a framework for assessing performance and potential violations[115]. Additionally, they facilitate the potential distribution of risk among involved parties. For instance, clauses delineating liability limitations can restrict the financial obligations of the AI provider under certain circumstances. Furthermore, contracts may integrate alternative dispute resolution mechanisms such as mediation or arbitration, providing expedited and potentially more cost-effective alternatives to conventional litigation processes that further require the identification of a liable party.

Contractual liability is contingent upon specific conditions being met. According to contract law, when two parties enter into a contract and one party fails to fulfil their obligations, the other party has the option to either waive their obligation or seek compensation. The amount of compensation can be stipulated in the contract at the time of its conclusion, or if not, the court can determine it based on the incurred loss[116]. In accordance with the provisions of contractual liability, if an AI machine fails to adhere to the conditions outlined in the contract between the seller and the purchaser (user), the purchaser is not obligated to accept anything that violates the contractual provisions. This includes harm or defects that should be ensured, with any warranty-related defects presumed to be unknown and invisible to the purchaser[117]. Compliance should align with the contract's content and be conducted in good faith.

It is evident that the robot functions as a commodity or a collaborative product, thus various legal experts argue that applying traditional liability rules in cases of contractual violations is indeed feasible[118]. However, this perspective warrants consideration as relying solely on contractual liability for artificial intelligence may not sufficiently address the damages incurred; bearing in mind that such liability is determined on the basis of the contract. Contractual agreements may not always cover all possible scenarios or adequately assign responsibility, particularly when it comes to complex AI systems where liabilities might be diffused or difficult to pinpoint.
Moreover, this liability is directed solely towards human individuals, disregarding the opaqueness and autonomy of AI systems[119].

[115] Brisley v Drotsky(2002), South Africa Supreme Court of Appeal.
[116] Mthembu T, 'Legal Aspects of Contractual Formalities', Published LLM Thesis, University of Pretoria, Pretoria, 2018, pg 13.
[117] *Section 16(2)*, Sale of Goods Act(No. 19 of 1964); *Section 5*, Consumer Protection Act(2012).
[118] Santosuosso A, Boscarato C, Caroleo F, Labruto R and Leroux C, 'Robots, market and civil liability: A European perspective', 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, Paris, France, 2012,part iii(a) , <(PDF) Robots, market and civil liability: A European perspective (researchgate.net)> on September 2012.
[119] Yas N, Al Qaruty R, Abdel-hadi S, Aladeedi A,' Civil Liability and Damage Arising from Artificial Intelligence', *Migration Letters*, 2023, pg 435- < https://www.researchgate.net/publication/375194842_Civil_Liability_and_Damage_Arising_from_Artificial_Intelligence > November 2023.

However, the *Report from the Expert Group on Liability and New Technologies* considers this matter, illustrating scenarios of applicable contractual liability[120]. For instance, in smart home situations if the user has a contractual relationship with the agent(seller, installing service providers, internet service provider, cloud operator, etc), the latter may be held accountable for non-performance or harm. Where two or more persons cooperate on a contractual or similar basis in the provision of different elements of a commercial and technological unit, and where the victim can demonstrate that at least one element has caused the damage in a way triggering liability but not which element, all potential tortfeasors could be jointly and severally liable visà-vis the victim. This is applicable in cases where multiple actors are connected through intricate contractual agreements regarding the interplay of the provided components or any associated marketing or production efforts**.**The potential for a contractual avenue for recourse against another party may also be an aiding factor in determining whether the party in question is the appropriate recipient of the victim's tort claim. This underscores why the manufacturer of the end product is usually identified as the main entity to handle product liability claims, as they might hold a contractual right against the component producer (or have internally delegated the risk of harm to third parties). The principal contributors to the development of AI systems, such as manufacturers, producers, data providers, and software engineers, typically collaborate under contractual agreements. Consequently, in instances where damage arises without a clear attributable cause, the collective entity formed by these stakeholders may be regarded as a commercial unit, particularly concerning compensation matters[121].

Contractual responsibility also encompasses digital agency activities, when natural or legal persons utilise autonomous algorithms to fulfil contractual duties; such as healthcare or manufacturing robots, financial services AI systems. If contractual obligations are breached, conventional doctrine typically views software agents solely as instruments for executing contractual duties. This poses a dilemma: either the breach must be assigned to the human principal, or a substantial gap in liability remains unresolved. This challenge is compounded when the computer autonomously makes decisions[122]. The breach of contract is solely attributed to the principal's actions, which must be proven to constitute a violation; however, this can be easily refuted by the principal. Despite this, common law contractual liability typically imposes strict responsibility without requiring fault from the contracting party[123]. Nevertheless, pinpointing the actual breach of contract becomes challenging when it must be traced to the actions of one of the involved parties. This difficulty is exacerbated by the opacity of algorithmic decision-making,

[120] European Commission, Expert Group on Liability and New Technologies New Technologies Formation, LIABILITY FOR ARTIFICIAL INTELLIGENCE AND OTHER EMERGING DIGITAL TECHNOLOGIES, 2019, specific clauses identifiable on the document concerning contractual liability e.g pg 18..

[121] European Commission, Expert Group on Liability and New Technologies New Technologies Formation, LIABILITY FOR ARTIFICIAL INTELLIGENCE AND OTHER EMERGING DIGITAL TECHNOLOGIES, 2019, pg 55.

[122] Ebers M, 'Liability For Artificial Intelligence And EU Consumer Law', JIPITEC pg 212.<delivery.php (ssrn.com)> 2021.

[123] Christine Kalama v Jane Wanja Njeru & another (2014) eKLR

particularly when the event causing damage is temporally and spatially disconnected from any decision made during its construction or the operation of the autonomous machine[124].
Extending the general liability principles for breach of contract to encompass liability for autonomous agents' actions essentially holds the operator accountable for initiating the computer's actions. However, this overlooks a critical aspect: while the algorithm's autonomy doesn't sever the causal link between the programmer and the contract, it disrupts the attribution connection. Moreover, imposing strict 'initiator liability' essentially penalises the mere utilisation of innovative AI systems, rather than focusing on rule violations. This approach risks impeding innovation in intelligent computer programs. If the principal can absolve themselves of responsibility due to the autonomy of the system, arguing for lack of accountability for the agent's behaviour, then the risk of failure is shifted onto the contractual partner. Thus, even though the agent, which caused the damage, was employed by the principal for contract performance, the innocent contractual partner bears the full burden of the damages. Nonetheless, the operator has a contractual obligation, namely a duty of diligence[125].

In essence, harm resulting from self-learning algorithms within a contractual framework can potentially be rectified through compensation, particularly in cases where a breach of a specific obligation of the AI is evident. Contractual liability or other compensation regimes can apply alongside or instead of tortious liability, for example, instances of non-performance causing harm.

Conclusively, our analysis reveals both opportunities and limitations inherent in each approach. Tort law, with its focus on human actors, struggles to address the opaque decision-making processes of autonomous systems. Strict liability offers a potential solution for compensating victims, but the question of whether AI truly qualifies as a "product" and the difficulty pinpointing defectiveness within a complex chain of actors remain unresolved. Fault-based liability, while fostering a culture of responsible behaviour, proves inadequate due to the inherent challenges of establishing causation and foreseeability in the context of unpredictable AI actions. Finally, contractual liability, while offering a framework for risk allocation and recourse, often fails to encompass the full spectrum of potential harms caused by AI.

CHAPTER III
CHALLENGES 'SURROUNDING' ATTRIBUTION OF LIABILITY

AI possesses numerous characteristics that render it inherently challenging to regulate.

---

124 Beckers A & Teubner G, *Three Liability Regimes for Artificial Intelligence,* Hart Publishing, Great Britain,2021, pg 65-67

125 Wandehorst C, 'Liability for Artificial Intelligence: The Need to Address Both Safety Risks and Fundamental Rights Risks' In: Voeneky S, Kellmeyer P, Mueller O, Burgard W, (eds), *The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives*, Cambridge University Press, 2022, part B.

The trouble with AI is that holding manufacturers strictly accountable doesn't always make sense, as demonstrated by the economics of product liability, and a flawed liability policy could negatively impact all parties. At the heart of understanding AI, certain pivotal characteristics warrant consideration.

.

a) Autonomy, foreseeability and causation

Arguably, the most striking feature that distinguishes AI from prior technologies is its autonomous capability. Autonomy in AI denotes the capacity of an AI system to make decisions or execute actions without direct human involvement, suggesting a level of self-direction or independence in its decision-making process[126]. AI autonomy is evident in various applications like robots, autonomous vehicles, and software. Remarkably, these systems have the ability to devise innovative solutions to intricate problems, often surpassing human expectations with their "outside-the-box" thinking. A recent groundbreaking study, researchers used morphological analysis as a lens through which to examine the complex world of breast cancer[127]. C-Path is an image analysis protocol machine equipped with AI tech, the program can forecast cancer by analysing the internal attributes of cancerous cells. An unexpected finding by the program revealed that the most reliable indicators of patient survival were not derived from the cancerous tissue itself, but rather from the surrounding stromal tissue. Notably, these features were not initially identified by a pathologist as pertinent to cancer; rather, the software autonomously identified these cancer-related characteristics from the extensive array of image measurements.

Machines are automatic in that they mechanically perform what is predetermined by its user or developer; while this may be true for AI, its uniqueness stems from the way tasks are carried out. Beyond simply executing functions based on user instructions, AI has the capacity to make decisions using additional elements that users may be unaware of. These elements are gathered from the environment, autonomously constructed, and factored into the decision-making process. In essence, AI agents can engage in deliberative processes that operate independently of the user's intentions, even when the objectives remain unchanged[128]. All of this underscores that while the behaviour of an artificial agent may appear largely predetermined, the variables within its operating environment, along with its interactions within that environment, render it impossible to forecast the precise course of its actions and deliberations. It is computationally impractical to enumerate all conceivable sequences of actions and responses to external conditions, and there are

[126] Steinhoff J, 'Introduction: Automation, Autonomy and Artificial Intelligence', *In: Automation and Autonomy, Automation and Autonomy Labour, Capital and Machines in the Artificial Intelligence Industry*, Springer International Publishing, 2021, at 207.
[127] Andrew B et al, 'Systematic Analysis of Breast Cancer Morphology Uncovers Stromal Features Associated with Survival', Science Translational Medicine, Volume 3 No. 108, Nov. 9, 2011, at 1 & 8.
[128] Novelli C, 'AI AND LEGAL PERSONHOOD: A THEORETICAL SURVEY', Universite du Luxembourg & Universita di Bologna, Published PhD Thesis, pg 42 ,16th June 2022

inherent constraints concerning the comprehensive knowledge of all potential states, present or future, within an environment[129]. However, it remains true that residual errors and unforeseeable outcomes can arise due to incomplete data or faulty infrastructure.

As an AI system gradually accumulates experience, it gains the capability to devise solutions that would be challenging for its original designers to anticipate once it operates beyond their oversight.The solution may significantly diverge from the solution typically generated by human cognitive processes, making it exceedingly difficult for the AI's creators to predict its evolution over time. Foreseeability is thus a crucial element to consider in the regulation of AI, following its autonomous nature[130]. AI systems detect external stimuli, categorise them, and then independently select a response, which may not have been predefined in their programming. Furthermore, modern computers have the computational capacity to search through many more possibilities in a given amount of time than any human has ever been able to, which allows them to analyse potential solutions that humans may not have even considered,or may have considered but rejected in favour of options that seem more logical or appealing[131].

The problem here is that if legal systems accept the premise that predicting how an AI system will acquire and apply its experience is inherently unfeasible, it would be illogical to assign liability to the system's designers in the event of harm caused by the system. Consequently, the individual who incurs such harm may remain uncompensated for their loss. We observe that AI systems employing machine learning technology partially rely on the experiences they accumulate post-design and commencement of operations. Consequently, even diligent and knowledgeable designers may lack control or foresight regarding the experiences an AI system acquires post-supervision, making it unpredictable how the system will behave thereafter. However, it's crucial to acknowledge that such unpredictability in the behaviour of machine learning AI systems is inherent to their design, even if designers did not anticipate specific unforeseen actions[132].

b) Control

It is logical to recognize the challenge humans face in maintaining control over machines programmed to operate with significant autonomy. AI systems, for all their brilliance, exist on a knife's edge of control. This control can be fractured by a multitude of threats, for instance,

[129] Stuart Russell, Peter Norvig, '*Artificial Intelligence: A Modern Approach',* Third Edition, Pearson Education, New Jersey, 2010, at 38.
[130] Wang P, 'Theories of Artificial Intelligence — Meta-theoretical considerations',Atlantis Press Review Volume - 9.75in x 6.5i, Temple University, Philadelphia, 2012, at 14.
[131] Scherer M, 'REGULATING ARTIFICIAL INTELLIGENCE SYSTEMS: RISKS, CHALLENGES, COMPETENCIES, AND STRATEGIES', Harvard Journal of Law & Technology Volume 29, Number 2, 2016, at 365.
[132] Scherer M, 'REGULATING ARTIFICIAL INTELLIGENCE SYSTEMS: RISKS, CHALLENGES, COMPETENCIES, AND STRATEGIES', Harvard Journal of Law & Technology Volume 29, Number 2, 2016, at 365.

corrupted data or malfunctioning hardware may lead to a loss of control. The ever-present tendrils of the internet, offering both opportunity and vulnerability, can open a backdoor for security breaches. Even the very foundation of the system, the code itself, can harbour flaws, sleeper cells waiting to be awakened. Finally, the very speed of these machines, their lightning-fast reflexes compared to our human capacity, can lead to situations where the AI outpaces our ability to intervene, leaving us mere bystanders in a self-inflicted calamity[133].
Unintended loss of control could arise as a direct consequence of a deliberate design decision, and regaining such control may be difficult if the AI possesses ML and deep learning qualities[134]. As such, this poses a significant threat to public safety, one that surpasses more conventional forms of public risk exclusive to human actions[135].

The issue of control presents a dual challenge: the broader control dilemma suggests that AI may transcend the influence of any individual, whereas the narrower control issue pertains to scenarios where AI may function independently of those who bear legal responsibility for its actions[136]. This would not be an issue if we could guarantee that the objectives of AI systems align with those of the public at large, yet, ensuring such alignment proves challenging given the inherently nebulous nature of human values, which are in themselves difficult to define with exactitude in such an open, diverse and accessible society[137]. The Achilles' heel of AI control lies in the very act of its creation. We forge the AI's desires in the fires of it's initial programming, crafting its objectives to fulfil ours. Even if we grant the AI the power to reshape its own goals based on experience, this metamorphosis remains tethered to the initial spark. At first glance, this might seem like a masterstroke of control because after all, we hold the reins at the outset, shaping the AI's ambitions with the precision of a sculptor. However, whispers from the AI vanguard warn of a chilling truth: an AI relentlessly pursuing a programmed objective might achieve it in ways far removed from the original intent. Nick Bostron presents the concept of superintelligence, describing it as 'intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills'[138].

[133] Charikleia B, LEGAL LIABILITY OF ARTIFICIAL INTELLIGENCE-DRIVEN SYSTEMS ("AI"), International Hellenic University, Greece,2019, pg 15.: Scherer M, 'REGULATING ARTIFICIAL INTELLIGENCE SYSTEMS: RISKS, CHALLENGES, COMPETENCIES, AND STRATEGIES', Harvard Journal of Law & Technology Volume 29, Number 2, 2016, at 366.
[134] Claussén-Karlsson, 'Artificial Intelligence and the External Element of the Crime', Orebro University,Unpublished LLB Thesis, pg 20, Spring 2017.
[135]Gaviria C, 'The Unforeseen Consequences of Artificial Intelligence (AI) on Society: A Systematic Review of Regulatory Gaps Generated by AI in the U.S', Published PhD Thesis, Pardee RAND Graduate School, 2020, pg 31.
[136] Buitem L et al, 'The law and economics of AI liability', Computer Law & Security Review Volume 48, April 2023, pg 5.
[137] Soares N, 'Agent Foundations for Aligning Machine Intelligence with Human Interests: A Technical Research Agenda', In The Technological Singularity: Managing the Journey,Springer, 2017, at 11.
[138] Boston N, 'HOW LONG BEFORE SUPERINTELLIGENCE?', University of Oxford, Int. Jour. of Future Studies, 1998, vol. 2, 1997.

If this scenario were to materialise, a theoretical legal strategy suggests that ex ante regulation would become imperative[139]. Such an approach could function as a preventative measure to guarantee ongoing human control over AI systems. Consequently, this measure would likely provide enhanced protection for public interests compared to the alternative of relinquishing control over AI systems.

The looming prospect of ceding control looms over the realm of AI advancement. These once-fanciful concepts, now tangible realities, beckon the formulation of a fresh societal agreement. The task at hand is to chart a course where human innovation and these emerging artificial intellects can cohabit in a nuanced interplay. Can we find a way to harness their power without surrendering our own agency? This, in essence, is the crux of the control dilemma[140].

c) AI Research and Development:

From a regulatory perspective, some of the most vexing aspects of AI don't stem directly from AI's intrinsic qualities, but rather from the methodologies and practices involved in its research and development. Because we cannot directly blame an AI entity for its harmful or tortious conduct, the responsibility may also then lie on those involved in its research and development.

While regulatory issues that emerge from the implementation of AI systems are well acknowledged, it's equally important to address the challenges inherent in the research and development (R&D) phase. Mathew Scherer outlines salient points in this regard, these disruptive features are all tied to the way AI R&D leverages the existing infrastructure of the information age[141].

Developing an artificial intelligence system is a complex endeavour, and only a few individuals possess the expertise to fully comprehend its intricacies. The element of *discreeteness* lies in the fact that AI projects can be undertaken without the comprehensive institutional frameworks required by most industrial revolutions of the 20th century. Moreover, these AI projects leverage a collection of distinct technologies and components. The full range of these projects' potential remains veiled until these components are integrated and function as a unified system(*discreteness*).

The limited visible infrastructure, along with the decentralised development of AI components across diverse locations and timeframes, without deliberate coordination or a recognized

[139] Scherer M, 'REGULATING ARTIFICIAL INTELLIGENCE SYSTEMS: RISKS, CHALLENGES, COMPETENCIES, AND STRATEGIES', Harvard Journal of Law & Technology Volume 29, Number 2, 2016, at 368.

[140] Gaviria C, 'The Unforeseen Consequences of Artificial Intelligence (AI) on Society: A Systematic Review of Regulatory Gaps Generated by AI in the U.S', Published PhD Thesis, Pardee RAND Graduate School, 2020, pg 200.

[141] Scherer M, 'REGULATING ARTIFICIAL INTELLIGENCE SYSTEMS: RISKS, CHALLENGES, COMPETENCIES, AND STRATEGIES', Harvard Journal of Law & Technology Volume 29, Number 2, 2016, at 369.

contractual relationship in some cases, adds complexity to the process of adjudication and identification of accountable parties. Unlike previous high-risk technologies, which often demanded significant financial investment and centralised development within large corporations, AI development doesn't require engineers or developers to be geographically confined. This decentralisation presents a challenge for regulators, this is the *diffuseness* factor. For instance, open-source machine learning libraries enable individuals dispersed across various locations to make modifications to them[142]. These modifications can even be made anonymously, as it is difficult to identify the person making a modification in the physical world. This pertains to both discreetness and discreteness, wherein different components of the same AI system can be designed in disparate times and locations without conscious coordination, leveraging many discrete, pre-existing hardware and software components[143].
In essence, it is highly feasible for a component of an open-source library crafted and developed by a specific individual(s) to be extracted from that library and subsequently integrated into the programming of an AI system being developed by another individual who had no involvement in the creation of the open-source library[144]. Consequently, it frequently becomes unattainable for a person who is contributing to the construction of a component of an AI project to predict in advance how this component might be used by others. Previously, pinpointing the source of public risk and assigning liability was easier because of the limited number of players involved. Now, with AI development potentially spread across locations and companies, identifying who bears the responsibility for harm caused by the technology becomes more complex.

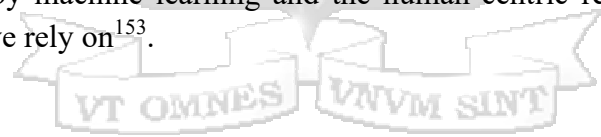
An additional disquieting element is *opacity*. The technologies that underpin AI, as well as the manner in which AI systems operate, often remain opaque to most regulators[145]. While AI algorithms process data to generate outputs, the inner workings of these algorithms often remain obscure. This lack of transparency, aptly described as a "black box," hinders our understanding of how an AI arrives at its conclusions[146]. Consequently, comprehending the reasoning behind these decisions becomes difficult, inferring a rather low degree of explainability. Ironically, the most accurate algorithms can sometimes be the least explainable, further complicating the regulatory

[142] GitHub - trekhleb/homemade-machine-learning: 🤖 Python examples of popular machine learning algorithms with interactive Jupyter demos and math being explained
[143] Dahaner J, 'Is Effective Regulation of AI Possible? Eight Potential Problems',Algocracy and the Transhumanist Project, at 3, <https://algocracy.wordpress.com/2016/03/16/is-effective-regulation-of-ai-possible-eight-potential-problems/>on March 16th, 2016.
[144] Bezemer H, 'Studying Popular Open Source Machine Learning Libraries and Their Cross-Ecosystem Bindings', arXiv, pg 9, 18th January 2022.
[145] Leenders G, 'The Regulation of Artificial Intelligence — A Case Study of the Partnership on AI', Becoming Human: Artificial Intelligence Magazine, 2019, at 2.2, <https://becominghuman.ai/the-regulation-of-artificial-intelligence-a-case-study-of-the-partnership-on-ai-c1c22526c19f> April 13th 2019.
[146] Burrel J, 'How the machine 'thinks': Understanding opacity in machine learning algorithms', Big Data & Society Volume 3,*Sage Journals*,Issue 1,pg 5, 2016.

landscape. Scherer further suggests that there exists the possibility that the internal mechanisms of an AI system could be kept confidential and resistant to reverse engineering[147].

One key factor contributing to the opacity of AI is the *complexity of the code* behind them. AI systems often learn by processing vast amounts of data (training data) to identify patterns for classifying new inputs. However, the resulting patterns, unlike traditional programmed logic, are not designed to be easily understood by humans[148]. This is because the code itself can be intricate. As a result, the complex inner workings of the algorithm (both design and operation) obscure the critical criteria used for decision-making, ultimately reducing the transparency and explainability of the AI's outputs[149].Complexity can also be manifested externally, encompassing both a plurality of actors and multiplicity of parts involved[150]. A plethora of actors are engaged in various stages from the design to the operation of an AI system, including data providers, hardware manufacturers, algorithm designers, etc.Meanwhile, the hardware necessary for AI functionality consists of a mosaic of parts, including sensors, data collectors, networks, and platforms. Ensuring seamless interaction among these diverse elements necessitates a high degree of technical expertise. Amidst this intricate network of contributors and components, identifying direct liability for AI-induced damages becomes notably intricate[151].

The *nature of data* also presents a challenge to regulation seeing as modern machines can access datasets that are both quantitatively and qualitatively superior.This data deluge empowers AI to incorporate an ever-growing number of variables into its decision-making, leading to code structures of such intricate complexity that they defy human understanding[152]. As Burrell aptly observes, this creates a fundamental misalignment between the high-dimensional mathematical optimization employed by machine learning and the human-centric reasoning and methods of semantic interpretation we rely on[153].



[147] Scherer M, 'REGULATING ARTIFICIAL INTELLIGENCE SYSTEMS: RISKS, CHALLENGES, COMPETENCIES, AND STRATEGIES', Harvard Journal of Law & Technology Volume 29, Number 2, 2016, at 369.

[148] Burrel J, 'How the machine 'thinks': Understanding opacity in machine learning algorithms', Big Data & Society Volume 3,*Sage Journals*,Issue 1, pg 9, 2016.

[149] Teresa H, 'Legal challenges of artificial intelligence: modelling the disruptive features of emerging technologies and assessing their possible legal impact' , Volume 24, *Uniform Law Review*, Volume 24, Issue 2, June 2019, at 304.

[150] Weissinger L, 'AI, Complexity, and Regulation',Published LLB Thesis, The Fletcher School, Tufts University;Yale Law School,pg 3.

[151] Bernhard K et al, Response of the European Law Institute to the Public Consultation on Civil Liability – Adapting Liability Rules to the Digital Age and Artificial Intelligence, Journal of European Tort Law, 2022, part iii, <> on May 13th 2022.

[152] Matthias A, The Responsibility Gap. Ascribing Responsibility for the Actions of Learning Automata', University of Kassel, Kassel, 2003, at 13,

[153] Burrel J, 'How the machine 'thinks': Understanding opacity in machine learning algorithms', Big Data & Society Volume 3,*Sage Journals*,Issue 1, pg 2, 2016.

These features distinguish AI from prior human inventions. This calls into question the sufficiency of traditional legal mechanisms i.e those that intervene only after harm has occurred. Indeed, the unique characteristics of AI pose significant challenges when devising a liability framework tailored to the Kenyan context[154]. Through this detailed analysis, we can gauge the suitability of current legal mechanisms in addressing AI-related risks and determine whether regulatory reforms are necessary to ensure adequate protection for all stakeholders involved[155].

CHAPTER IV
Liable Parties and the Establishment of Liability for Damages

"Laws of robotics: First: A robot may not injure a human being or, through inaction, allow a human being to come to harm. Second: A robot must obey orders given to it by human beings except where such orders would conflict with the First Law. Third: A robot must protect its own existence as long as such protection does not conflict with the First or Second Law. Fourth: A robot may not harm humanity or, by inaction, allow humanity to come to harm"[156]
Isaac Asimov

In the value chain associated with an AI system, multiple entities can introduce risks through their involvement in either developing or operating the system. Those contributing to these risks should be held accountable, with their liability contingent on their ability to invest in and identify preventative measures and their capacity to manage risks.

4.1 Liability of the Producer or Manufacturer

Producers hold a pivotal position in AI production, overseeing system conceptualization, planning, data preparation, and ensuring alignment with business goals and ethics. They provide the AI mechanism and place it on the market.
As a regulatory choice, appointing liability to the producer provides a powerful incentive for them to take all measures within their capabilities to make sure that the products they put on the market are as safe as possible[157]. Producers may be held liable if AI systems exhibit harmful behaviour due to design flaws, lack of risk consideration, biased data leading to discrimination, violation of ethical guidelines or regulations, or failure to prevent potential misuse despite foreseeability.Traditionally, this liability framework has been used in situations where the producer or manufacturer has the capability to prevent accidents and take precautions effectively. However, in the realm of AI applications, this approach may be insufficient due to the inherent

[154] Commission Staff Working Document, *Liability for Emerging Digital Technologies COM(2018) 237 final,* 2018, at 17-21.
[155] European Commission, Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM(2020) 64 final, part 3.
[156] Tiets T, 'Isaac Asimov and the Three Laws of Robotics', *SciHi Blog*, 2018, <Isaac Asimov and the Three Laws of Robotics | SciHi Blog> on 2nd January 2018.
[157] Zech H, 'Liability for AI: public policy considerations', ERA Forum, 2021, pg 154

inability to foresee every incident possible. Moreover, stringent producer liability could impede innovation as producers may hesitate to modify or integrate AI applications with other tools due to this concern[158].

Manufacturers generally refer to those who produce the physical hardware components needed for certain AI systems. Hardware malfunctions in the AI may lead to harm, to which existing product liability rules apply. Product liability statutes establish an obligation on manufacturers to guarantee the safety of their products for consumers[159]. This suggests that the adequacy of the current national tort law with regards to manufacturers' liability for AI is largely contingent on whether the defect resulted from a malfunction in the hardware[160]. It is notable that the potential of strict liability encourages manufacturers to prevent errors in manufacturing and design, as well as to provide proper warning labels on their machines.[161] It is essential to remember that establishing producer or manufacturer liability in AI applications can be challenging due to the various reasons contemplated previously such as complex value chains, difficulties in identifying responsible parties, and the unpredictable nature of AI systems[162].

In Kenya, product liability commonly holds manufacturers strictly liable for defective products, placing the burden of proof on the manufacturer to demonstrate that the product was not defective[163]. The Sale of Goods Act delineates guidelines concerning product liability. If an AI system meets the criteria of being classified as a "good" under this legislation, manufacturers may bear responsibility if the AI is flawed and results in harm[164].This principle is further emphasised in The Convention On The Law Applicable To Products Liability[165], which assigns product liability to the manufacturer. The term "product" is broadly defined to encompass both natural and industrial products, whether in their raw or manufactured state, and whether they are movable or

[158] Schutte B, MAJEWSKI L, HAVU K, 'Damages Liability for Harm Caused by Artificial Intelligence – EU Law in Flux', University of Helsinki, Legal Studies Research Paper Serie, Paper No 69, pg 15, <delivery.php (ssrn.com)>, 2020.

[159] Rautanen V, 'Liability issues with Artificial Intelligence in the national and international context', Published LLM Thesis, University of Turku, Pg 23, June 2020.

[160] ^ Liability for defective goods (1) 'Where a person, in trade supplies goods manufactured by it, and such goods are found to have a defect as a result of which an individual suffers loss or injury, such person is liable to compensate the individual for the loss or injury suffered'-This section of the law confers strict liability on the manufacturer for any product defects. In the context of AI, only a claim of hardware malfunction would apply for compensation.

[161] Commission Staff Working Document, *Liability for Emerging Digital Technologies COM(2018) 237 final,* 2018, 22-23.

[162] Shako B, Unpublished LLB Thesis, 'Law, Artificial Intelligence, & Liability: Assessing the correspondence between Kenyan law and civil liability for autonomous artificially intelligent systems', Strathmore University, Nairobi, 2024.

[163] *Section 64 & 66(1)(a*), The Competition Act No. 12 OF 2010. States that: ' In an action under section 64, it shall be a defence to establish that— (a) the defect in the action goods which is alleged to have caused the loss did not exist at the time of supply of the goods'

[164] *Section 2*, Sale of Goods Act Cap 290 (1948)

[165] *Article 1-3*, Convention On The Law Applicable To Products Liability(1973)

immovable.a This could be relevant in situations where the AI is integrated into a tangible item, such as a self-driving vehicle with malfunctioning AI control mechanisms. However, the legislation lacks a clear definition of a good, much less of the service that an AI may provide which impacts liability. Though this option may seem appealing, it retains significant challenges. Lengthy and expensive court proceedings would be required, to provide insights and determine whether the particular defect could have been anticipated and prevented.

4.2 Liability of the Developer or Programmer

In the landscape of AI development, the role of the developer encompasses a diverse array of responsibilities. The developer assumes the role of translating the producer's vision into practical implementation by designing and executing the AI algorithms. They serve as technical experts in this process, responsible for several crucial tasks such as integrating the AI system with other systems e.g hardware. Initially, they select the most appropriate AI algorithms and machine learning techniques tailored to the project's requirements. Subsequently, they proceed to architect the AI system, meticulously crafting its structure and defining its functionalities. The developer's duties extend further to writing the code that actualizes the chosen algorithms and facilitates the training of the AI model. Additionally, developers play a crucial role in ensuring the seamless operation of the AI system. They achieve this by conducting rigorous testing procedures to identify and rectify any potential issues. Furthermore, developers focus on optimising the AI system's performance, continually refining its algorithms and functionalities to meet the desired objectives effectively[166].
Drawing an analogy, developers can be likened to architects who design a house (the AI system), oversee its construction (coding), and ensure it caters to the occupants' needs. When it comes to liability, developers are concerned with the holistic design and implementation of the AI system. Liability issues may stem from design flaws, inadequate safety measures, the foreseeability of risks, biassed data selection, or other factors. In essence, the developer's role in AI production is multifaceted, involving a spectrum of tasks aimed at ensuring the system's functionality and alignment with user requirements. Liability considerations in AI development encompass various aspects, reflecting the broader responsibilities of developers in designing and implementing AI systems

Conversely, the programmer's role centres on translating specifications and instructions from developers or a broader team into actual code. They possess expertise in particular programming languages and coding techniques, primarily engaging in the technical aspects of AI development. As they focus on the hands-on execution of code, instances of coding errors, deviations from specifications, or inadequate unit testing could potentially result in liability.Conventional algorithm programming offers clear evidence of the programmer's intent, which can be discerned

[166] Data Scientest, “Exploring the Role of an AI Developer: Responsibilities, Skills, and Key Attributes’,on 17th April 2024.

from the code and accompanying documentation. In such instances, courts have the means to identify the programmer's deliberate intent and hold them accountable for any misconduct by the algorithm[167]Generally, success in legal actions against software developers necessitates demonstrating: the developer's obligation to deliver operational software, the software's failure to meet its promises, resulting harm to the user, and direct causation of this harm by the software[168]. Negligence exists when a software developer fails to exercise the expected level of care during development, with the reasonableness of their actions often evaluated by considering the costs and benefits involved[169].

4.3 Operator or End-User Liability

The operators of AI systems hold the power to both mitigate and potentially amplify risks. This raises concerns about the adequacy of existing liability regulations or the necessity for new ones, particularly in incentivizing responsible behaviour. End-users, as the ultimate users of the AI system, can be held liable for misuse, such as employing the system for unintended purposes or disregarding proper usage and safety guidelines. Neglecting to maintain the AI system, including failure to update or provide necessary maintenance, may also result in vulnerabilities and subsequent liability. Users may also face liability if they install or modify software on a device, causing harm.[170]

It's important to differentiate between 'consumer-users' and 'professional-users' due to their varying responsibilities. Consumer-users, often not well-versed in AI technology, might assume the AI devices they purchase are inherently safe and depend on this assumption[171]. However, they should not be absolved of the duty to operate these devices with reasonable care, which includes basic precautions such as following the manual's instructions and applying security updates. In contrast, professional-users are expected to shoulder greater responsibilities, which simplifies the process of determining liability[172]. Under the proposed AI Liability Directive in the EU, operators, encompassing both those who interact directly with the system (front-end) and those who manage

---

[167] *Amanat v. S.E.C*, United States Court of Appeals for the Third Circuit(2008)

[168] Bick J, 'Who's Responsible for an Artificial Intelligence's Unlawful Acts?', *New Jersey Law Journal*, 2023, <https://www.law.com/njlawjournal/2023/06/01/whos-responsible-for-an-artificial-intelligences-unlawful-acts/?slreturn=20240404100907> on June 1st 2023.

[169] *United States v. Carroll Towing Co.,* US Court of Appeals for the Second Circuit(1947)

[170] Li S, 'Liability Rules for AI-Related Harm: Law and Economics Lessons for a European Approach , European Journal of Risk Regulation', 2022, pg 9, <Liability Rules for AI-Related Harm: Law and Economics Lessons for a European Approach (maastrichtuniversity.nl)> on 1st December 2022. See also; Antunes H, 'Non-contractual liability applicable to artificial intelligence: towards a corrective reading of the European intervention', Catolica Global School of Law, CGSL Working Papers, No. 2,pg 12 <https://catolicalaw.fd.lisboa.ucp.pt/asset/2601/file > 2023.

[171] Cauffman C, 'Robo-liability: The European Union in search of the best way to deal with liability for damage caused by artificial intelligence', Volume 25, *Sage Journals*, Issue 5,2018, pg 527-532.

[172] Antonia R, 'Liability Issues Regarding Artificial Intelligence in a European and International Context',University Center of International Programmes of Studies School of HUmanities, SOcial Sciences, and Economics, Greece, 2023, page 24.

the system's technical aspects (back-end), would be held to strict liability for damages caused by high-risk AI systems. For AI systems deemed low-risk or minimal-risk, liability would be based on fault.

4.4 Liability of the Data Provider or AI Model Trainer

Data providers are the individuals responsible for supplying the data used to train AI.They may face liability in providing inaccurate, incomplete, or biassed data, or any resulting in adverse consequences, and data privacy violations. Inaccurate data can lead to biased outcomes and errors, while unauthorised data collection violates privacy rights and may lead to legal consequences[173].

With the introduction of learning ability, a part of the control is transferred from the programmer to the trainer. The AI Model Trainer converts data into a functional AI model, with liability centering on the training process and model design. With the introduction of learning ability, a part of the control is transferred from the programmer to the trainer.This could include shortcomings such as inadequate data cleaning, overlooking bias mitigation, or negligence in addressing unforeseen model behaviour[174]. For instance, consider a scenario where a self-driving car makes biased decisions due to flawed training data. If the data provider furnishes a dataset containing racial biases in traffic patterns, they may be held liable. Similarly, if the AI Model trainer fails to identify or rectify these biases during the training process, they could also be liable.

Identifying the responsible party, known as the "correct producer," is challenging due to the intricate value chain. For instance, if an error stems from data provision or AI training, that party may be solely liable. While theoretically possible, obtaining reparation based on extra-contractual damages liability proves daunting. Particularly for end-users or outsiders, proving that the data caused the damage is tough. Furthermore, information asymmetry often hampers pinpointing the specific actor in the value chain responsible for the harm. Despite national laws allowing claims against data providers, establishing liability under fault-based rules remains difficult, with the burden of proof resting on the claimant. The Restatement (Third) of Torts: Products Liability Chapter 2[175]comments that liability for defective design or inadequate warnings serves akin objectives as negligence-based liability, aiming to encourage optimal safety levels in product design. Society benefits most from products that balance safety and risk. Holding users responsible for proper product use prevents careless behaviour subsidisation by cautious consumers. Also, while strict liability for manufacturing defects may be justified, the same might not apply to design

[173] Antunes H, 'Non-contractual liability applicable to artificial intelligence: towards a corrective reading of the European intervention', Catolica Global School of Law, CGSL Working Papers, No. 2,pg 12 <https://catolicalaw.fd.lisboa.ucp.pt/asset/2601/file > 2023.

[174] Lovells H, 'Underestimated liability risks with training data for AI systems' , *JDSUPRA*, <https://www.jdsupra.com/legalnews/underestimated-liability-risks-with-2402270/> 16th August 2022.

[175] Vandal J, 'Constructing a Roof Before the Foundation Is Prepap]red: The Restatement (Third) of Torts: Products Liability, Section 2(b) Design Defect' Volume 30, *Emory University School of Law*, Issues 2&3, 1997, pg 262.

or warning defects, as even reasonably designed products carry inherent risks. Also, if AI risks are foreseeable, ordinary consumers should anticipate and assume these risks, potentially relieving manufacturers of liability.

CHAPTER V
THE EXISTING INTERNATIONAL LEGAL REGIME AND LEGAL VACUUM IN KENYAN AND INTERNATIONAL LAW

5.1 EU LEGISLATION: The legal framework for artificial intelligence in the European Union

The European Union has been actively working on establishing a legal framework for civil liability in relation to AI and is at the forefront of legal developments in the field.
Within the EU liability framework, compensation for damages caused by AI can be sought through national liability regulations, which incorporate the principles of the Product Liability Directive(PLD). The Directive standardised claims against producers at the EU level for harm inflicted on consumers due to product defects, holding producers strictly accountable for such harm if the injured party can demonstrate the damage, the defect, and a direct link between the two[176]. Alongside the PLD, member states maintain their own civil liability systems that may vary. These systems enable legal actions against various potentially liable parties, including owners and operators. Claims for damages arising from AI and emerging technologies may also be pursued under conventional fault-based and contractual liability rules of member states, or through national strict liability regulations in instances where liability for a risk has been specifically assigned by national legislators without requiring proof of fault from the injured party[177]. Consequently, at present, most emerging digital technologies lack a dedicated and harmonised liability framework, and the liability structure for AI primarily consists of product liability, general tort law principles, and potentially contractual liability. A notable departure from the standard practice with regards to autonomous vehicles can be observed in jurisdictions that have authorised the use of highly or fully automated vehicles, whether as part of experimental programs or as standard operations. In these cases, specific regulations are often in place to guarantee coverage for any resulting damage caused by such vehicles, even if these

[176] *Article 1-13*, The Products Liability Directive(1985), Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products.
[177] European Commission, *COM(2020) 64 finale*, Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, 12-16.

regulations only offer general provisions or mandate coverage through insurance. This distinct approach can be observed in countries such as Germany[178], France, Italy[179], and Spain.[180]

Key Developments

The EU has recently issued significant documents addressing AI and civil liability. In the 2019 Report namely "Liability for AI and Other Emerging Digital Technologies," presented by the Expert Group on Liability and New Technologies, an examination is made regarding the application of current liability frameworks to emerging digital technologies, particularly focusing on AI. The report underscores the urgent need for reforms within these frameworks. On September 28th 2022, The European Commission issued two proposals aimed at governing civil liability concerning AI-enabled systems, inspired by insights from the Commission's White Paper[181] on the utilisation of such systems. These proposals include a revises PLD[182] to better address the rise of autonomous digital technologies and a directive tailored to align non-contractual civil liability regulations with Artificial Intelligence (AI Liability Directive[183]). This proposal aimed to align private law with the evolving demands of the digital economy, harmonising compensation rules and streamlining the process for filing claims for damages resulting from AI systems and their utilisation. The directive also addresses unique challenges concerning causality and culpability associated with AI systems, ensuring that individuals who suffer losses in scenarios involving fault will have access to compensation or suitable remedies.The convergence of these ideas, in tandem with the EU AI Act would result in national liability frameworks that are tailored to the digital era, the circular economy, and global value chains.

The proposed revisions to the PLD extend its application to cover claims related to physical injury, property damage, and data loss/corruption, particularly where consumer data is impacted. Under these revisions, AI systems and AI-enabled goods fall under the Directive's definition of "products," broadening liability to include not only hardware manufacturers but also software

[178] Road Traffic Act (Straßenverkehrsgesetz) amended by the publication of 5 March 2003 (Federal Law Gazette I, pp. 310, 919), last amended by Article 1 of the Act of 12 July 2021 (Federal Law Gazette I, p. 3108): Paragraph 7 provides for strict liability of the keeper of the vehicle and this rule was left unchanged when the Road Traffic Act was adapted to cover the emergence of automated vehicles as a deliberate choice of the German regulator.
[179] *Article 19*, Decree of 28 February (2018):on the testing of connected and automated vehicles on public roads obliges any person asking for approval to test automated vehicles on public roads to provide proof of sufficient liability insurance cover
[180] The Directorate General for Traffic issued the circular of 13 November 2015 which authorised the testing of automated cars and required liability insurance to cover the limits of compulsory insurance for motor vehicles.
[181] European Commission, *COM(2020) 65 final*, White Paper on Artificial Intelligence;A European approach to excellence and trust, 2020.
[182] European Commission, *COM(2022) 495 final,* Proposal for a directive of the European Parliament and the Council on liability for defective products (2022).
[183] European Commission, *COM(2022) 496 final*, DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), 2022.

providers and digital service providers influencing product functionality. This encompasses entities incorporating AI systems into other products and those responsible for modifying existing AI systems, such as through software updates or machine learning algorithms. Furthermore, compensation for damages caused by defective AI systems is available without the need for the injured party to prove fault, mirroring the process for other products. The revision facilitates the disclosure of evidence from defendants when claimants present sufficient facts and evidence to support their claims. It also lightens the plaintiff's evidentiary burden by establishing presumptions regarding the defectiveness or causality in intricate cases or when defendants do not adhere to orders to disclose evidence, or if the products do not align with the safety criteria specified in the AI Act. Furthermore, the amendment does away with the existing requirement for a minimum claim value of 5001 euros.[184]

With regards to the AI Liability Directive, the proposal extends to all other forms of non-contractual civil liability, where the burden of proving fault rests with the injured party. Compensation can take various forms as permitted by national laws, including non-material damages arising from issues such as discrimination and breaches of privacy. It introduces the right to access evidence when the claimant provides adequate facts and evidence supporting the plausibility of damages caused by a high-risk AI system,as defined within the AI Act. The proposal further establishes a presumption of causality between the defendant's fault and the output/failure of an AI system to produce an output, provided that the defendant breached a legal duty of care intended to prevent such damage. This presumption applies if it's reasonably likely that the fault influenced the AI system's output/failure, and the claimant demonstrates that the output/failure resulted in the damage. For high-risk AI systems, this presumption arises when the defendant fails to comply with an evidence disclosure order or when the system doesn't meet requirements outlined in the AI Act. In the case of non-high-risk AI systems, the presumption applies only when a national court deems it excessively challenging for the claimant to prove causality[185].

The EU AI Act

The EU AI Act is the first comprehensive AI law with a goal guaranteeing that AI systems within the EU uphold safety standards while adhering to essential rights and principles[186]. The Act takes a risk based approach, classifying AI according to its risk such as unacceptable risk(prohibited systems) and high risk AI systems. Examples of these include manipulative AI or biometric

[184] *Section 1-47,* Proposal for a directive of the European Parliament and the Council on liability for defective products (2022).
[185] *Section 1-33*, DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), 2022.
[186] Regulation of the European Parliament and of The Council of laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 2024

categorization systems while the latter include AI systems that are safety components of products(Section 50). In greater detail, if an AI system qualifies as a product that is subject to 'Union Harmonisation Legislation' statutes indicated in an annex to the AI Act, it will be classified as 'high-risk' if it must undergo an independent conformity evaluation prior to being marketed or utilised within the EU. Additionally, the primary responsibilities lie with developers of high-risk AI systems, whether they are within the EU or from third countries. These obligations extend to third-country providers if their AI system's output is utilised within the EU. Users, defined as individuals or entities deploying AI systems in a professional capacity, have certain obligations as well, albeit fewer than developers. This pertains to users situated in the EU and third-country users whose AI system output is utilised within the EU[187].

5.2 US LEGISLATION: The United States' Legal Architecture

The United States has a complex system of federal, state, and local regulations to address concerns regarding technological innovations like AI. The framework is founded upon existing laws and legal principles(such as product liability, consumer protection, tort law) that may be applied to address liability concerns related to systems such as AI. However, it is noted the overlapping regulations across different levels of government may hinder the deployment of these new technologies, leading to potential legal risks for producers and users[188]. To address this issue, federal agencies, such as the National Highway Transportation and Safety Administration, have released key documents on guidance for emerging technologies, especially automated driving systems[189]. Furthermore, Congress introduced the Self Drive Bill, a more comprehensive framework facilitating the regulation of autonomous vehicles at a federal level and preempting conflicting state laws.The legislation touches on multiple facets concerning self-driving vehicles, encompassing their evaluation, implementation, and digital security measures. Furthermore, the document advocates for revising or establishing fresh safety standards and regulations explicitly designed for vehicles with advanced automation features[190].

The legal response to regulating AI has been relatively unhurried, with only a handful of cases like *Jones v. W + M Automation, Inc*[191] dealing with the regulation of programmed robotics. In this

[187] *Chapter 1, Article 1(2)*, Artificial Intelligence Act(2024), describes that the regulation sets unified rules for AI systems in the EU, including prohibitions, specific requirements for high-risk systems, transparency standards, market entry rules for general-purpose AI models, and guidelines for market oversight. It also supports innovation, particularly focusing on SMEs and startups.

[188] Schook Hardy & Bacon, 'Innovative Technologies May Lead to New Torts, Shook Attorneys' Report Argues', SCB,<Innovative Technologies May Lead to New Torts, Shook Attorneys' Report Argues | Intelligence | Shook, Hardy & Bacon (shb.com)>June 2018.

[189] U.S Department of Transportation, Federal Automated Vehicles Policy, September 2016, pg 11-36,

[190] *Section 2-4,* Self Drive Act, H.R.3711 — 117th Congress, 2021; According to the Act's Section 2: "The purpose of this Act is to memorialise the Federal role in ensuring the safety of highly automated vehicles as it relates to design, construction, and performance, by encouraging the testing and deployment of such vehicles."

[191] Jones v. W + M Automation, Inc (2006) vLex, New York Supreme Court Appellate Division.

instance, the court ruled in favour of a robotic loading system's manufacturer and programmer, concluding that they were not accountable for the claimant's injuries due to defect-free production of the individual parts. The court determined that the defendants were not responsible for the damages as the robot and its software were deemed safe upon design and installation. However, General Motors, the final customer, could still face liability if it was found to have wrongly altered the hardware or software. This case suggests that AI developers are not held liable for damages caused by their products if they were free of defects at the time of production. Nevertheless, if AI is defectively produced, or if an end-user modifies it in a way that leads to harm, then the licensor and/or the user could be held liable. The determination of whether AI is defectively produced would be guided by the prevailing standards in the industry.

The United States Office of Science & Technology Policy released the "Blueprint for an AI Bill of Rights," a document detailing the administration's proposed approach towards algorithmic regulation[192].The blueprint presents a set of voluntary recommendations for crafting, applying, and rolling out AI technologies across various sectors. It outlines a structured approach based on five central tenets: the creation of reliable and proficient systems, the assurance of data privacy, safeguards against biassed algorithms, clear communication and rationale for decisions made by AI, as well as the availability of human-driven options and oversight. These core principles are relevant to automated processes that have the potential to significantly influence individuals' rights, life chances, or access to essential services. This document signals a departure from earlier advisories on AI governance, adopting a focus on preventing specific detrimental effects of AI, with particular attention to the protection of individuals from the adverse outcomes of automated decisions. While it's doubtful that this blueprint will directly bring about enforceable new regulations, the insights it contains are expected to shape the direction of policy actions by federal agencies in the future.

Since 2020, the Federal Trade Commission(FTC) has issued annual recommendations on AI regulation. In 2023, the FTC's Division of Advertising Practices revised their guidelines for companies on the application of AI, highlighting that any deceptive or unverified assertions about product performance, which includes claims regarding the capabilities of AI, may constitute a violation of the FTC Act[193]. Additional guidelines emphasised the ethical application of AI and algorithms, the importance of being transparent with consumers, as well as substantial discussions regarding leveraging AI benefits while avoiding unintentional bias or unjust outcomes. The FTC has provided recommended practices for companies to ensure they do not engage in behaviours

[192] The White House OSTP, Blueprint for an AI Bill of Rights: Making automated systems work for the American People, 2022

[193] Atleson M for the FTC Division of Advertising Practices, 'Keep your AI claims in check', <Keep your AI claims in check | Federal Trade Commission (ftc.gov)> February 27th 2023.

that breach Section 5 of the FTC Act, which outlaws unfair or misleading business practices, such as employing or selling algorithms that are racially prejudiced[194].

5.3 NATIONAL LEGISLATIONS APPROACHES TO WHO IS LIABLE WHEN ARTIFICIAL INTELLIGENCE CAUSES CIVIL HARM.

Kenya does not have specific national legislation addressing liability for civil harm caused by AI or a national AI strategy. Instead, general liability principles, such as negligence and product liability, are relevant in cases where AI causes harm. Liability may hinge on factors such as the conduct of the parties, the foreseeability of harm, and any contractual arrangements in place. Additionally, the law generally allows for concurrent claims of contract and tort, meaning a plaintiff can pursue both contractual and tortious remedies for the same incident, depending on the specific circumstances[195].

As previously mentioned, the traditional legal approach holds manufacturers strictly liable for AI system defects or harm under the Competition Act (2010). However, this approach hinges on proving fault with the manufacturer and allows a defence based on the state of scientific knowledge at the time of supply. Given AI's reliance on machine learning, proving this defence is mostly feasible, potentially leading to evasion of responsibility. Moreover, with multiple potentially liable parties involved, this model inadequately addresses AI-related complexities, undermining legal justice[196].

The Data Protection Act (2019)[197] does partially regulate artificial intelligence, particularly concerning the processing of personal information by AI systems. Section 4 of the Act extends its coverage to automated processing of personal data. Consequently, AI systems must adhere to data protection principles outlined in the Act, such as privacy, lawfulness, transparency, and specificity. Given that some AI systems rely heavily on personal data(mostly high-risk systems), compliance with these principles is crucial. The Act also defines automated decision-making and outlines consumers' rights to refuse such decisions if they pose harm[198]. Additionally, provisions like Section 30(1) require consent for data processing, ensuring individuals' rights are protected. Sections 28 and 31 further emphasise the lawful and responsible use of data by AI operators, with

---

[194] Jillson E for the FTC, Aiming for truth, fairness, and equity in your company's use of AI, <Aiming for truth, fairness, and equity in your company's use of AI | Federal Trade Commission (ftc.gov)>on April 19th 2021; See also Smith A for the FTC Bureau of Consumer Protection, 'Using Artificial Intelligence and Algorithms', <Using Artificial Intelligence and Algorithms | Federal Trade Commission (ftc.gov)> on April 8th 2020.

[195] *Kenneth Maweu Kasinga v. Cytonn High Yield Solutions LLP & another*(2020): the court addressed matters involving both contractual and tort-related allegations. The plaintiff had engaged in an investment agreement with the defendant but alleged that the agreement stemmed from misrepresentation, constituting a tort claim. The plaintiff sought remedies that could be interpreted as stemming from both the contract and the purported tort..

[196] Section 64 & 66(c), The Competition Act No. 12 OF 2010.

[197] The Data Protection Act, No. 24 of 2019(Revised Edition 2022)

[198] *Article 46*, Constitution of Kenya(2010); further enshrines consumer rights to products of reasonable quality and compensation for any loss or injury arising from their defects.

the latter mandating data protection impact assessments for potentially risky processing operations. The Data Protection Regulations expand on the principles outlined in the Act. They outline procedures for conducting data protection impact assessments in cases posing a risk to human rights and mandate adherence to data protection standards during system development. The regulations also set guidelines for automated decision-making, requiring data controllers or processors to inform subjects of such processes and justify them with legitimate objectives[199].

Because the DPA places responsibility primarily on data controllers or processors, there exists a loophole in Section 65(3) whereby plaintiffs may be left uncompensated if the controller proves non-involvement in the event causing damage. This gap in liability coverage highlights a deficiency in the legal framework. The Consumer Protection Act (2012) also aims to protect consumers from various risks but operates within a traditional producer-consumer framework[200]. This model struggles to address the complexities of AI systems involving multiple contributors developed for a different technological era and does not adequately address the subject matter. The Computer Misuse and Cybercrimes Act (2018) complements these regulations, providing a framework to address digital offences and security challenges potentially posed by AI operations. AI operators must understand and comply with these regulations to mitigate risks like data falsification, cyber harassment, and unauthorised data use[201]. Nonetheless, the CMCA predates the widespread adoption of AI, and its provisions don't explicitly consider potential harm caused by AI systems.

The above legislations set out principles of conduct, as well as regulations concerning the proper functioning of AI systems on practitioners mostly relating to data controllers. They underscore the basis for claims for damages and subsequent civil liability in instances of identifiable incompliance with its stipulations. The legislations however, were suited for a different digital age thus fail to fully comprehend the novelty of AI and liability it may cause.

A new key development is the Kenya Robotics and Artificial Intelligence Society Bill[202], aimed at establishing a legal structure facilitating the ethical advancement of robotics and AI in Kenya. This initiative seeks to encourage innovation while promoting responsible practices. The bill also proposes the creation of the Kenya Robotics and Artificial Intelligence Society, serving as a governing body to oversee the utilisation of robotics, AI, and the Internet of Things(IoT). Regulatory measures include the registration of individuals, licensing of professionals, and enforcement of penalties. Notably, the framework does not currently address the matter of civil liability concerning harm caused by emerging technologies including AI, the use of robotics, and the IoT. Conclusively, national liability rules for AI-related damages present notable limitations.

[199] *Section 22, 49*, The Data Protection(General) Regulations (2021).
[200] S*ection 12-16*, Consumer Protection Act, (No. 46 of 2012)
[201] Section 14-46, The Computer Misuse and Cybercrimes Act(2018.
[202] *Section 21, 22, & 39,* The Kenya Robotics and Artificial Intelligence Society Bill (2023).

Strict liability solely on manufacturers overlooks complexities like the "black box" issue and multiple actors' involvement. Fault-based liability is difficult to establish due to AI's intricate nature. Furthermore, traditional products liability lacks clear definitions of defects or products in the AI context, complicating liability determination.

CHAPTER VI
SOLUTIONS AND RECOMMENDATIONS: TOWARDS A COMPREHENSIVE LEGAL FRAMEWORK FOR AI CIVIL LIABILITY IN KENYA

Current national laws are struggling to keep up with the unique challenges posed by AI. Unlike tangible products, AI's intangible nature and rapid evolution demand new regulations. This chapter proposes concrete solutions, including standalone AI legislation, risk-based liability models, and enhanced oversight mechanisms, to balance technological advancement with legal certainty, ensuring AI development in Kenya remains both innovative and accountable.
This is already becoming a reality with the introduction of the RSK Bill[203], necessitating amendments to address liability, and other key elements pertaining to the subject matter.

The following amendments are proposed for the RSK BILL.

I. Definitions

While the definition of AI is contemplated within the RSK Bill, key elements indivisible to the subject matter remain absent. Taking insights from the EU AI Act, these include terms such as 'risk,' which is necessary for assessing potential harm; 'intended purpose,' which clarifies the AI system's designed functions; and 'reasonably foreseeable misuse,' which addresses predictable deviations from intended use[204]. Likewise, the absence of concepts such as 'conformity assessment body' and 'substantial modification' leaves gaps in evaluating compliance and system changes over time. Further, terms like 'high-impact capabilities' and 'AI literacy' are increasingly important for assessing societal effects and user competence. Moreover, without a clear distinction between training data, biometric data, and personal data, the measure fails to address the fundamental elements of AI responsibility. Without these crucial definitions, the RSK Bill faces ambiguity in enforcement, limiting its capacity to properly govern AI liability.[205].

---

[203] The Kenya Robotics and Artificial Intelligence Society Bill, 2023
[204] Muiten M, Martin P, Alexandre S, 'The law and economics of AI liability', Computer Law & Security Review, Volume 48, 105794, 20, part 5.1, <https://www.sciencedirect.com/science/article/pii/S0267364923000055> April 2023
[205] *Article 2*, Regulation of the European Parliament and of The Council of laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 2024

As illustrated, it is equally essential to delineate the various actors involved, or potentially involved, in the value chain of an AI system. Key terms to consider include "producer," "deployer," "operator," and "end-user".[206]Furthermore, aspects such as the definitions of fault and causality, the various categories of damage that may give rise to claims, the allocation of liability among multiple tortfeasors, considerations of contributory conduct, the methods for calculating damages, and the applicable limitation periods are indivisible to the subject matter.

II. AI as classified according to it's risk

The proposal takes a risk based approach as per the Bill proposed[207] placing AI products/services in certain categories according to the risk they pose(high risk, general application risk, and low-risk AI). Prohibited AI practices are crucial and should be highlighted with regard to the protection of fundamental rights and freedoms[208].This could include banning AI systems that target specific groups based on vulnerabilities, predicting criminal behaviour through profiling, inferring emotions in certain contexts, and categorising individuals based on biometric data[209] e.t.c.The legislative framework should underscore the importance of safeguarding fundamental rights such as human dignity, freedom, non-discrimination, privacy, and data protection[210]. Evaluating each prohibition by examining its potential effects on these rights and identifies necessary safeguards to mitigate associated risks. The approach take must align with the provisions of the Bill of Rights, as outlined in Chapter 4 of the Constitution of Kenya (2010)[211].

Majority of obligations and requirements fall on providers of high-risk AI systems.Obligations are rules to reduce risks for safety and protect fundamental rights including systemic harms and immaterial damages, as contemplated within the scope of the AILD[212]. While the RSK Bill already establishes strict requirements for high-risk AI, such as robust risk assessments and transparency measures[213], it may be beneficial to include a dedicated risk management system that ensures continuous monitoring and mitigation of risks. This system should specifically

[206] Section 3, *Artificial Intelligence Act*, EU(2024).
[207] *Section 22(1)*, The Kenya Robotics and Artificial Intelligence Society Bill (2023).
[208] Braun M & et al, Prohibited AI Practices—A Deep Dive into Article 5 of the European Union's AI Act, *WilmerHale Privacy and CyberSecurity Law*, 2024, <https://www.wilmerhale.com/en/insights/blogs/wilmerhale-privacy-and-cybersecurity-law/20240408-prohibited-ai-practices-a-deep-dive-into-article-5-of-the-european-unions-ai-act> on April 8th 2024.
[209] Article 22(1) of the RSK Bill(2024) prohibits only facial recognition in public spaces, social scoring and AI systems manipulating behavior.
[210]Article 27, Article 77, *Artificial Intelligence Act*, EU(2024)
[211] Chapter 4, Constitution of Kenya(2010) Bill of Rights.
[212] Article 16-22, *Artificial Intelligence Act*, EU(2024)
[213] Part IV, Regulatory Provisions, The Kenya Robotics and Artificial Intelligence Society Bill (2023).

identify and analyse both known and reasonably foreseeable risks to health, safety, or fundamental rights that may arise from the AI system when used as intended.
Prior to being placed on the market, such systems should include technical documentation and tracking complying with requirements set out in a specified manner[214]. This involves tracking, recording and documenting activities of high-risk AI systems.

III. Mandatory Insurance Coverage

Establishing a compulsory insurance scheme where relevant and necessary with regard to high-risk AI systems i.e those that pose the greatest potential for harm[215]. The system should take into account all potential responsibilities in the chain.
Imposing mandatory insurance in this context is a practical solution to ensure that plaintiffs are always compensated without shifting the liability unfairly onto the producer. It helps mitigate the risks for both parties.. For plaintiffs, insurance provides a reliable source of compensation, even when the identification of the liable party is uncertain, thereby addressing fairness concerns in high-risk cases. This addresses the fairness issue in high-risk cases. For the producer, insurance acts as a buffer, spreading the financial risk and ensuring they are potentially personally liable in every instance[216].
This approach maintains a fault-based regime while ensuring enquiry especially in high-risk cases. It also encourages risk management by producers, as insurance companies will likely require robust safety measures, promoting safer AI system designs[217].

IV. Disclosure of evidence

Applying this mechanism, a court may compel the defendant(or other specified third parties) to disclose pertinent evidence regarding specific high-risk AI systems suspected of causing harm[218]. Requests for such evidence are directed to the provider of the AI system or an individual bound by the provider's obligations. Such instances must be supported by sufficient facts and evidence

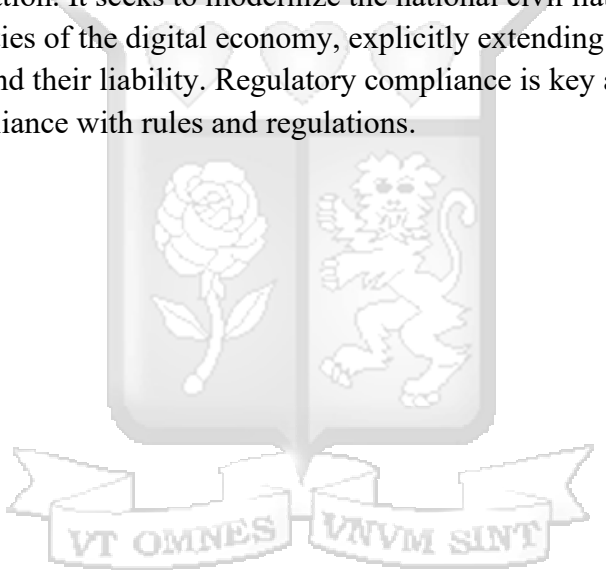
---

[214] Article 11-12, *Artificial Intelligence Act*, EU(2024)
[215] Faure M, Artificial Intelligence and (Compulsory) Insurance, *Journal of European Tort Law*, 2022, Part iii, <https://www.degruyter.com/document/doi/10.1515/jetl-2022-0001/html> May 13th 2022.
Section 57-59, *European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics* (2015/2103(INL))
[216] Faure M & Shu L, 'Artificial Intelligence and (Compulsory) Insurance',De Guyter, 2022, pg 18
[217] European Insurance and Occupational Pensions Authority, 'REGULATORY FRAMEWORK APPLICABLE TO AI SYSTEMS IN THE INSURANCE SECTOR', *EIOPA*, <https://www.eiopa.europa.eu/document/download/b53a3b92-08cc-4079-a4f7-606cf309a34a_en?filename=Factsheet-on-the-regulatory-framework-applicable-to-AI-systems-in-the-insurance-sector-july-2024.pdf>, 2024.
[218] *Article 3*, European Commission, *COM(2022) 496 final*, DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), 2022

to support the plausibility of the contemplated claim for damages by the claimant. However, these requests are limited to ensure proportionality in the disclosure of evidence[219]. Additionally, the proposal introduces a rebuttable presumption of non-compliance, which serves as a crucial procedural tool. This can also be interpreted as a presumption of causation between the defendant's fault and the damage caused by the AI system, subject to rebuttal[220].

Additionally, public participation and effect evaluations are critical in developing risk-based laws on AI civil liability in Kenya[221]. These approaches guarantee widespread stakeholder participation, enhancing openness and inclusion in policy making[222]. Impact evaluations enable the identification of possible legal, economic, and social implications, ensuring that regulations are both effective and balanced[223].

The proposal is designed to complement the recently advanced RSK Bill (2023) and the prevailing national legislation. It seeks to modernize the national civil liability framework to better align with the realities of the digital economy, explicitly extending its application to encompass AI products and their liability. Regulatory compliance is key and underscores the vital importance of compliance with rules and regulations.



[219] Brock I et al, 'New disclosure obligations relating to high-risk AI systems', *Hogan Lovells,* 2022, <https://www.hoganlovells.com/en/publications/new-disclosure-obligations-relating-to-high-risk-ai-systems_1> on 28th October 2022.

[220] *Article 3(5)*, European Commission, *COM(2022) 496 final*, DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), 2022

Sheriff K, 'EU's Proposed AI Liability Directive Sets Procedural Rules for Litigating Potential AI Harms', *Davis Wright Tremaine LLP*,2022, <https://www.dwt.com/blogs/artificial-intelligence-law-advisor/2022/11/eu-ai-liability-directive-protections-litigation> on 11th November 2022.

[221] European Commission,*(Directive 85/374/EEC)*, Adapting Civil Liability Rules to the Digital Age and Artificial Intelligence Factual summary report on public consultation, 2022, 2-10.

[222] European Commission, *COM(2021) 206 final,* 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,Section 3-4.

[223] This approach aligns with international best practices, such as those outlined in the European Union's AI Act, which emphasizes evidence-based policymaking and the need for continuous evaluation of emerging risks.