

**WHO BEARS THE BLAME? ANALYSING WHAT KENYA'S CIVIL  
LIABILITY FRAMEWORK MEANS FOR CHATBOT HARMS**

Submitted in partial fulfillment of the requirements of the Bachelor of Laws Degree, Strathmore

University Law School

By

Georgina Achieng Okello

139181

Prepared under the supervision of

Cecil Abungu

Submitted on 13<sup>th</sup> March, 2025

(14,881)

Acknowledgements	i
Declaration	ii
Abstract	iii
List of Abbreviations	iv
List of Cases	v
List of Legal Instruments	vi
1.0. Introduction	1
1.1. Background	1
1.2. Problem Statement	4
1.3. Research Objectives	4
1.4. Research Questions	5
1.5. Hypothesis	5
1.6. Justification	5
1.7. Conceptual Framework: Chatbots as agents	6
1.8. Literature Review	7
1.8.1. On chatbots as machine learning systems	7
1.8.2. On attributing liability	8
1.8.3. Contribution	10
1.9. Methodology	11
1.10. Chapter Breakdown	12
<b>2.0. The Technology behind Chatbots, its Benefits &amp; Risks</b>	<b>14</b>
2.1. Introduction	14
2.2. The Technology behind Chatbots	14
2.2.1. Data Training	14
2.2.2. User Interface	15
2.2.3. Knowledge Base	16
2.2.4. Machine Learning	16
2.2.4.1. Natural Language Processing	18
2.2.4.2. Natural Language Understanding	19
2.2.4.3. Dialogue Management Component	20
2.2.4.4. Natural Language Generation	20
2.3. The Benefits of Chatbots	21
2.4. Instances of Harm Caused by Chatbots	22
2.4. Conclusion	27
<b>3.0. Tracing the Evolution of Liability under Kenyan Tort Law and Its Implications on Liability for Chatbot Harms</b>	<b>28</b>
3.1. Introduction	28
3.2. Strict Liability	29

3.2.1. Examining Judicial Precedents and Legal Implications for Strict Liability	29
3.2.2. Case for Strict Liability for Chatbot Harms	31
3.2.3. Weaknesses and Shortcomings of using a Strict Liability Framework	32
3.3. Liability based on Negligence	33
3.3.1. Examining Judicial Precedents and Legal Implications for Liability based on Negligence	33
3.3.2. Case for Liability based on Negligence for Chatbot Harms	36
3.3.3. Weaknesses and Shortcomings of using Liability based on Negligence Framework	37
3.4. Vicarious Liability	39
3.4.1. Examining Judicial Precedents and Legal Implications for Vicarious Liability	39
3.4.2. Case for Vicarious Liability for Chatbot Harms	41
3.4.3. Weaknesses and Shortcomings of using a Vicarious Liability Framework	42
3.5. Conclusion	43
<b>4.0. The Case for Vicarious Liability in Addressing Chatbot Harms in Kenya</b>	<b>44</b>
4.1. Introduction	44
4.2. Evaluating Chatbot Agency and Vicarious Liability	44
4.2.1. Agency	44
4.2.2. Personhood	45
4.3. Justifying Vicarious Liability for Chatbot Harms in the Kenyan Context	46
4.3.1. Access to Justice and Consumer Protection	46
4.3.2. Corporate Responsibility and Ethical AI Development	47
4.3.3. Promoting Digital Transformation and Innovation	48
4.4. Conclusion	48
<b>5.0. Conclusion and Recommendations</b>	<b>50</b>
5.1. Conclusion	50
5.2. Recommendations	50
5.2.1. Recognition of Implied Agency Relationships between Chatbots and Deployers	50
5.2.2. Develop the Standard of Care for Deployers Seeking to Avoid Liability	50
5.2.3. Recognition of vicarious liability as the most appropriate legal framework for chatbot harms	51
<b>Bibliography</b>	<b>52</b>

## **Acknowledgements**

I am deeply grateful to my mother for her unwavering support and encouragement throughout this journey. To my siblings, thank you for making life easier, especially during the busiest moments.

To my friends, your motivation and reassurance have been invaluable, and I will always be grateful for having you by my side throughout this journey. A special appreciation goes to my supervisor for nurturing my interest in this subject and guiding me through the research process. Lastly, I am grateful to myself for the labour of love I have poured into this work and for the outcome of this paper.

To all who have supported me in any way—thank you.

**Declaration**

I, **GEORGINA ACHIENG OKELLO**, 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: .....13<sup>th</sup> March 2025.....

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



Signed:.....

Cecil Abungu

## **Abstract**

The increasing reliance on AI-driven systems in business operations has raised critical legal questions regarding liability for harms caused by chatbots. This study explores whether companies in Kenya should be held vicariously liable for chatbot-related harms, particularly in online customer interactions. It evaluates the applicability of existing liability frameworks—strict liability, negligence, and vicarious liability—to determine the most suitable approach for Kenyan courts when adjudicating chatbot liability disputes.

The study's scope focuses on analysing chatbot technology and the legal principles governing civil liability in Kenya. Given the absence of specific AI liability laws in Kenya, the research examines how courts can extend traditional legal doctrines to address chatbot-related harms, ensuring accountability while fostering responsible AI deployment.

A doctrinal research methodology was employed, involving an in-depth analysis of statutes, case law, journal articles, and regulatory policies. Additionally, the study applied agency theory to establish a legal basis for holding chatbot deployers accountable under vicarious liability principles. The findings reveal that chatbot deployers exercise substantial control over chatbot operations, making vicarious liability the most effective framework for assigning responsibility. Chatbots function within parameters set by their deployers, creating an implied agency relationship similar to that between employers and employees. Courts can apply existing principles of agency law and *respondeat superior* to hold companies accountable for chatbot-related harms, ensuring victims have clear avenues for legal redress.

The study recommends that courts recognise implied agency relationships between chatbots and deployers, define the scope of chatbot employment, establish a standard of care for chatbot oversight, and prioritise vicarious liability in chatbot-related disputes. By adopting these principles, Kenyan courts can develop a coherent legal framework that balances AI innovation with consumer protection, providing much-needed clarity in chatbot liability cases.

## **List of Abbreviations**

1. Artificial Intelligence (AI)
2. Constitution of Kenya, 2010 (CoK)
3. Customer relationship management (CRM)
4. European Union (EU)
5. Information, Communications and the Digital Economy (ICT)
6. Insurance Company of East Africa (ICEA)
7. Kenya Bureau of Standards (KEBS)
8. Knowledge Base (KB)
9. Machine learning (ML)
10. Natural Language Generation (NLG)
11. Natural Language Processing (NLP)
12. Natural Language Understanding (NLU)
13. Recurrent Neural Networks (RNNs)

## List of Cases

1. *Anastassios Thomos v Occidental Insurance Company Limited* (2017) eKLR
2. *Associated Motors Co. Ltd v Blue Sea Services Ltd* (2019) eKLR
3. *Attorney General v Law Society of Kenya & another* (2017) eKLR
4. *Beatrice William Muthoka & another (Both suing as Legal Representatives of the Estate of the Late William Muthoka Yumbia (Deceased) v Agility Logistics Limited* (2020) eKLR.
5. *Coca Cola Company Ltd & another v Josephat Okello Oduori* (2011) eKLR
6. *Cotecna Inspection S.A v Hems Group Trading Company Limited* (2007) eKLR
7. *David M Ndetei v Orbit Chemical Industries Limited* (2014) eKLR
8. *Donoghue v Stevenson*
9. *Duncan Nderitu Ndegwa v Kenya Pipeline Limited & another* (2013) eKLR
10. *Elijah Ole Kool v George Ikonya Thuo* (2001) eKLR
11. *Equator Distributors v Joel Muriu & 3 others* (2018) eKLR.
12. *Fred Ben Okoth v Equator Bottlers Limited* (2015) eKLR.
13. *James Ngugi Kariuki v Peter Kamau Kariuki* (2020) eKLR
14. *James Finlaly (K) Ltd v Bernard Kipsang Koechi* (2021) eKLR
15. *James Watenga Kamau v CMC Motors Group Limited* (2020) eKLR
16. *John Gachanja Mundia v Francis Muriira & Another* (2017) eKLR
17. *John Nderi Wamugi v Ruhesh Okumu Otiangala* (2018) eKLR.
18. *Joseph Kimani Gatheca v Gatundu South Water and Sewerage Company* (2018) eKLR
19. *JMA (Suing through BOA as next friend) & another v Registered Trustees of the Sisters of Mercy (Kenya) t/a Matter Misericordiae Hospital* (2023) eKLR
20. *Kara Commodities Limited v Muyendi (Suing as the administrator of the Estate of Onesmus Ndilivu Muyendi)* (2024) eKLR
21. *Kenya Breweries Ltd v Godfrey Odoyo* (2010) eKLR
22. *Kenya Ports Authority v East African Power & Lighting Company Ltd* (1982) eKLR
23. *Kenya Shell Limited v Milkha Kerubo Onkoba* (2010) eKLR
24. *Kenya Wildlife Service v Rift Valley Agricultural Contractors Limited* (2018) eKLR
25. *LWW (Suing as the Administrator of the estate of BMN) deceased v Charles Githinji* (2019) eKLR

26. *Moffatt v. Air Canada* (2024)
27. *Mombasa Maize Millers & another v Elius Kinyua Gicovi* (2021) eKLR.
28. *Munyu Maina v Hiram Gathiha Maina* (2013) eKLR.
29. *Mwita Merengo v Joseph Tunei Marwa & 2 others* (2012) eKLR
30. *Odero v Aga Khan Hospital Kisumu* (2024) eKLR
31. *Patrick J.O. Otieno v Lake Victoria South Water Services Board* (2020) eKLR
32. *Ricarda Njoki Wahome v Attorney General & 2 others* (2015) eKLR
33. *Richard v Lothian* (1913); AC 263 (C.A)
34. *Rylands v Fletcher*
35. *Stephen Kanjabi Wariari v Dennis Mutwiri Muriuki & another* (2022) eKLR
36. *Tabitha Nduhi Kinyua v Francis Mutua Mbuvi & another* (2014) eKLR
37. *Teachers Service Commission v WJ & 5 others* (2020) eKLR
38. *Terra Fleur Limited v Kenya Cuttings Limited* (2021) eKLR

## **List of Legal Instruments**

1. Constitution of Kenya, 2010 (CoK)
2. Consumer Protection Act, 2012, CAP 501
3. EU AI Act ( 2024)
4. Evidence Act
5. Judicature Act, CAP 8
6. Sale of Goods Act, CAP 31

## 1.0. Introduction

### 1.1. Background

The pervasive influence of Artificial Intelligence (AI) on a global scale has introduced numerous benefits while simultaneously presenting a range of legal challenges, particularly concerning liability. As businesses increasingly adopt AI for administrative, managerial, and marketing purposes, one notable application is the use of chatbots to optimise customer relationships.<sup>1</sup> Chatbots facilitate interactions with users, bringing various benefits to organisations.<sup>2</sup> They are specialised forms of conversational AI designed for two-way communication with users, typically through text-based messaging platforms. Leveraging natural language processing (NLP)<sup>3</sup> chatbots can interpret and respond to human language.<sup>4</sup> NLP enables them to comprehend human sentences, and allows them to retrieve information based on keywords or phrases to provide relevant responses.<sup>5</sup>

The primary motivation for deploying chatbots is to enhance customer support.<sup>6</sup> Chatbots have frequently been used for sales, marketing and customer services.<sup>7</sup> These AI-driven interfaces provide automated support around the clock, offering efficiency and accessibility beyond human capabilities.<sup>8</sup> Chatbots also reduce labour costs by automating up to 80% of inquiries, independently, reducing human support costs by around 30%.<sup>9</sup> They further allow support teams

---

<sup>1</sup> Binns R, 'Case study: How Apple has mastered CRM' Expert Market, 22 February

2023 - <<https://www.expertmarket.com/uk/crm-systems/apple-crm-case-study> > on 22 February 2023.

<sup>2</sup> Shawar A and Atwell E, 'Chatbots: Are they really useful?' 22(1) *Journal of Language Technology and Computational Linguistics*, 2007, 29-49.

<sup>3</sup> Allows any chatbot machine to understand one or more human languages which allows it to interpret human language input information provided to it.

<sup>4</sup> -<<https://gettalkative.com/info/chatbots-vs-conversational-ai#:~:text=So%2C%20in%20short%2C%20conversation%20AI,%E2%80%9Cbot%2Dlike%E2%80%9D%20conversation.> > on 29 February 2024.

<sup>5</sup> -<<https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/deloitte-analytics/deloitte-nl-chatbots-moving-beyond-the-hype.pdf> > on March 2018.

<sup>6</sup> -<<https://www.businessinsider.com/80-of-businesses-want-chatbots-by-2020-2016-12?r=US&IR=T> > on 14 December 2016.

<sup>7</sup> -<<https://www.businessinsider.com/80-of-businesses-want-chatbots-by-2020-2016-12?r=US&IR=T> > on 14 December 2016.

<sup>8</sup> Folstad A, Skjuve M and Brandtzaeg P, 'Different chatbots for different purposes: Towards a typology of chatbots to understand interaction design' ResearchGate, 2019, 1-2.

<sup>9</sup> -<<https://gettalkative.com/info/chatbots-vs-conversational-ai#:~:text=So%2C%20in%20short%2C%20conversation%20AI,%E2%80%9Cbot%2Dlike%E2%80%9D%20conversation.> > on 29 February 2024.

to focus on more complex cases while increasing online sales and marketing efforts through interactive customer engagement.<sup>10</sup>

Despite these benefits, chatbots present several challenges. Due to machine learning capabilities, chatbots can generate unexpected or inaccurate information, leading to potential negligence issues.<sup>11</sup> For instance, a customer service chatbot providing misinformation could be deemed negligent if it fails to deliver accurate information, resulting in harm to the user. Notable incidents, such as the shutdown of Microsoft's Tay chatbot in 2016 due to offensive tweets<sup>12</sup> and the termination of Facebook's chatbot experiment after unintelligible communication between bots,<sup>13</sup> highlight the potential risks associated with chatbot deployment.<sup>14</sup> Furthermore, data privacy and cybersecurity issues arise, as chatbots both collect and store user data, making them potential targets for attacks.<sup>15</sup> On the other hand, determining liability in AI systems is complicated by their autonomy and the involvement of multiple parties, including; developers, users, and operators, which makes tracing responsibility more complex.<sup>16</sup> It is often challenging to pinpoint who should be held liable because various actions by different parties may contribute to the AI system's output. This difficulty in traceability is compounded by the 'many hands' problem,<sup>17</sup> which highlights the shared responsibility among those involved in the system's creation and operation—including the developers who create the algorithm, the users who input data, and the operators who deploy the system.<sup>18</sup> As AI technology continues to evolve, these challenges call for the need to determine liability.

---

<sup>10</sup> -<<https://gettalkative.com/info/chatbots-vs-conversational-ai#:~:text=So%2C%20in%20short%2C%20conversational%20AI.%E2%80%9Cbot%2Dlike%E2%80%9D%20conversation.> > on 29 February 2024.

<sup>11</sup> Stojanov M, 'Prospects for chatbots' 8(3) *Izvestia Journal of the Union of Scientists-Varna Economic Sciences Series*, 2019,13.

<sup>12</sup> Heine C, 'Microsoft's chatbot 'Tay' just went on a racist, misogynistic anti-Semitic tirade' *Adweek*, 24 March 2016—<<https://www.adweek.com/performance-marketing/microsofts-chatbot-tay-just-went-racist-misogynistic-anti-semitic-tirade-170400/> > on 24 March 2016.

<sup>13</sup> Griffin A, 'Facebook's artificial intelligence robots shut down after starting to talk to each other in their own language' *Independent*, 31 July 2017—<<https://www.independent.co.uk/life-style/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html> > on 31 July 2017.

<sup>14</sup> Omri R, 'Whose robot is it anyway?: Liability for artificial intelligence-based robots' 20(4) *University of Illinois Law Review*, 2019, 1144-1145 .

<sup>15</sup> Leua C and Didu I, 'Chatbots. legal challenges and the EU legal policy approach' 10(3) *Society of Juridical and Administrative Sciences*, 2021, 216.

<sup>16</sup> Leua C and Didu I, 'Chatbots. legal challenges and the EU legal policy approach', 215.

<sup>17</sup> Vasudevan A, 'Addressing the liability gap in AI accidents' Center for International Governance Innovation, Policy Brief No. 177, 2023, 2—<[https://www.cigionline.org/static/documents/PB\\_no.177.pdf](https://www.cigionline.org/static/documents/PB_no.177.pdf) > on July 2023.

<sup>18</sup> Vasudevan A, 'Addressing the liability gap in AI accidents' 2.

To date, various chatbots have emerged, including Kayak Facebook Messenger chatbot, that provides information about discounts on flight tickets and hotels;<sup>19</sup> and Microsoft's Cortana in 2014.<sup>20</sup> In Kenya, notable chatbots like Safaricom's Zuri,<sup>21</sup> ICEA Lion's Leo that answers users investment queries,<sup>22</sup> and AjiraPoa WhatsApp chatbot whose aim is to provide accessible information on work readiness,<sup>23</sup> have been introduced, reflecting the growing adoption of chatbots across different sectors. This increasing use highlights the need for judicial consideration of potential harms resulting from interactions with chatbots.

However, how courts would decide liability issues related to chatbots in Kenya remains unclear, considering the fact that globally, countries like Canada are beginning to tackle these challenges. In *Moffat v. Air Canada (2024)*, a Canadian tribunal ruled that companies can be held liable for misinformation provided by a chatbot on their website.<sup>24</sup> This case is significant in recognising corporate responsibility for chatbots' actions.<sup>25</sup> Similarly, the European Union's AI Act, which came into force in August 2024, outlines various responsibilities for individuals and organisations that use AI systems.<sup>26</sup> It categorises chatbots as minimal risk AI systems, which are permitted provided they adhere to specific transparency and disclosure obligations where their use pose limited risk.<sup>27</sup> Furthermore, deployers of chatbots are required to inform users ab initio that they are interacting with an AI system.<sup>28</sup> In Kenya, no specific laws or cases addressing

---

<sup>19</sup> Botadra B, 'Web robots or most commonly known as bots' Academia, 2.

<sup>20</sup> Botadra B, 'Web robots or most commonly known as bots' 2.

<sup>21</sup> Kirwa M, 'Nine ways AI is making life easier for Kenyans,' The Star, 25 April 2023-<<https://www.the-star.co.ke/news/big-read/2023-04-25-nine-ways-ai-is-making-life-easier-for-kenyans/>> on 25 April 2023.

<sup>22</sup> Kirwa M, 'Nine ways AI is making life easier for Kenyans,' The Star, 25 April 2023-<<https://www.the-star.co.ke/news/big-read/2023-04-25-nine-ways-ai-is-making-life-easier-for-kenyans/>> on 25 April 2023.

<sup>23</sup> -<<https://home.fabo.org/ajira-poa-chatbot-kenya/>> on 17 July 2024.

<sup>24</sup> Hung R and Lifshitz L, 'BC Tribunal confirms that companies remain liable for information provided by AI chatbot' American Bar Association Group, 29 February 2024-<[https://www.americanbar.org/groups/business\\_law/resources/business-law-today/2024-february/bc-tribunal-confirms-companies-remain-liable-information-provided-ai-chatbot/](https://www.americanbar.org/groups/business_law/resources/business-law-today/2024-february/bc-tribunal-confirms-companies-remain-liable-information-provided-ai-chatbot/)> on 29 February 2024.

<sup>25</sup> Hung R and Lifshitz L, 'BC Tribunal confirms that companies remain liable for information provided by AI chatbot' American Bar Association Group, 29 February 2024-<[https://www.americanbar.org/groups/business\\_law/resources/business-law-today/2024-february/bc-tribunal-confirms-companies-remain-liable-information-provided-ai-chatbot/](https://www.americanbar.org/groups/business_law/resources/business-law-today/2024-february/bc-tribunal-confirms-companies-remain-liable-information-provided-ai-chatbot/)> on 29 February 2024.

<sup>26</sup> Kappel R, 'Overview of AI regulations and regulatory proposals of 2023' Central Eyes, 5 March 2024-<<https://www.centraleyes.com/ai-regulations-and-regulatory-proposals/>> on 5 March 2024.

<sup>27</sup> Section 132, *EU AI Act* (2024).

<sup>28</sup> Section 132, *EU AI Act* (2024).

chatbot liability have emerged, indicating a regulatory gap.<sup>29</sup> However, several proposals including the Ministry of Information, Communications and the Digital Economy's (ICT) development of a draft National AI strategy and the Kenya Bureau of Standards' (KEBS) Draft Information Technology Artificial Intelligence Code of Practice are underway.<sup>30</sup> From the proposals it is clear that Kenya has no specific laws aimed at regulating AI technologies despite the proposals that have been made thus far. Therefore, in the event that disputes emerge, courts would not have an appropriate means of establishing liability.

In conclusion, while chatbots have become essential tools for companies, their increasing autonomy introduces significant legal challenges. The growing reliance on AI systems in business operations raises important questions about liability, particularly when chatbots produce inaccurate or harmful outcomes. Current regulatory frameworks are evolving globally, as seen in Canada and the European Union, but many jurisdictions, including Kenya, still lack comprehensive laws addressing chatbot liability. Due to the regulatory gap in liability laws concerning harms caused by chatbots, it is essential to propose judicial approaches for handling such cases. This will ensure that affected individuals have a means of determining liability even in the absence of specific legislation enacted. Consequently, the study seeks to assist potential harmed person(s) to be able to determine causality easier and be able to seek recourse when such harm arises.

## **1.2. Problem Statement**

Assess whether companies in Kenya should be held vicariously liable for the errors of their chatbots in online customer communication.

## **1.3. Research Objectives**

1. To analyse the underlying technology that enables chatbots to function and understand how they operate.

---

<sup>29</sup>

<https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-kenya#:~:text=Sectoral%20scope,the%20various%20sectors%20in%20Kenya.> > on 20 June 2024.

<sup>30</sup><https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-kenya#:~:text=Sectoral%20scope,the%20various%20sectors%20in%20Kenya.> > on 20 June 2024.

2. To examine the legal principles historically used in Kenya for assigning civil liability for harms caused by company's resources.
3. To assess the viability of holding companies in Kenya vicariously liable for the actions of their chatbots as a means of assigning liability for harm caused to users.

#### **1.4. Research Questions**

1. What is the key technical component of chatbots, and how do they function?
2. What are the legal principles historically used in Kenya to assign civil liability for harms caused by company's resources?
3. How viable is it to hold companies in Kenya vicariously liable for the actions of their chatbots as a means of assigning liability for harm caused to users?

#### **1.5. Hypothesis**

Companies in Kenya that use chatbots to communicate with customers online, should be held vicariously liable for the errors of their chatbots.

#### **1.6. Justification**

The findings of this study will provide valuable insights into how liability should be apportioned for harms caused by chatbots in Kenya. This will address a critical gap in the current legal framework surrounding chatbots given its complexity. It is often challenging for victims to identify the responsible party and meet the legal requirements for a successful liability claim. Therefore, this study aims to clarify these issues, offering significant benefits to several key stakeholders. First, victims of chatbot-related harm will gain clarity on who they can hold accountable, empowering them to seek appropriate legal recourse. Second, the study will assist judges who may encounter such cases in the future, providing them with a well-researched basis for their decisions. Third, it will guide companies that deploy chatbots, emphasising the importance of due diligence in ensuring their chatbots operate effectively to minimise causing harm. Last, the study will serve as a source for scholars interested in AI liability, offering a foundation for further research and critique in the evolving field.

## 1.7. Conceptual Framework: Chatbots as agents

This study is based on the concept of a chatbot as an agent of a corporation. The legal concept of agency has proven useful in determining responsibility and liability for entities capable of modifying their actions.<sup>31</sup> Under the law of agency, an agent is a fiduciary empowered to act on behalf of a principal.<sup>32</sup> Similarly, chatbots are designed to perform specific tasks for humans, and in doing so, act on behalf of a principal. In this case, the principal could be the programmers, manufacturers, or users of the chatbot.<sup>33</sup> The principal defines the goals to be achieved, and the system operates to fulfil them. Through this framework, a chatbot can be viewed as an agent,<sup>34</sup> since its actions are guided, defined, and ultimately controlled by human intervention, either directly or through the ability to override the system.<sup>35</sup>

Due to this control, chatbots can be understood as agents or instruments of legal entities such as individuals, corporations or other legal person(s), who can be held accountable for their actions.<sup>36</sup> This positions chatbots as extensions of their human controllers.<sup>37</sup> Moreover, chatbots are increasingly replacing human workers.<sup>38</sup> Corporations cannot act independently; they must do so through agents, traditional employees, but now also through algorithms.<sup>39</sup> Employees are often treated as agents—with their employers—the corporations, being held accountable for harms caused.<sup>40</sup> Under the doctrine of respondeat superior, a relatively clear path exists for assigning liability when corporate employees misbehave.<sup>41</sup> However, the law has yet to clearly define agency relationships for autonomous systems such as chatbots, leaving a gap that allows corporations to evade liability by blaming algorithmic misconduct. Recognising an agency

---

<sup>31</sup> Osoba O, 'The value of the concept of agency in an increasingly rational world' *California Digital Library*, 2019, 5 -<<https://escholarship.org/uc/item/8q15786s>> on 26 September 2019.

<sup>32</sup> Harris D 'The lost rationale of agency law' 3(1) *Business Finance Law Review*, 2019, 5.

<sup>33</sup> Lima G, 'Can (and should) AI be considered an agent' *California Digital Library*, 2019, 20 -<<https://escholarship.org/uc/item/8q15786s>> on 26 September 2019.

<sup>34</sup> Chopra S and White L, 'Artificial Agents: Philosophical and legal perspective' ResearchGate, 2007-<<http://www.sci.brooklyn.cuny.edu/~schopra/ChopraWhiteChapter1.pdf>> on 2007.

<sup>35</sup> Vladeck D, 'Machines without principals: Liability rules and artificial intelligence' 89(117) *Washington Law Review*, 2014,120.

<sup>36</sup> Vladeck D, 'Machines without principals: Liability rules and artificial intelligence' 121.

<sup>37</sup> Walker J, 'I smell a bot: California's S.B. 1001. Free speech and the future of bot regulation' *Houston Law Review*, 2019, 408.

<sup>38</sup> Walker J, 'I smell a bot: California's S.B. 1001. Free speech and the future of bot regulation,' 408.

<sup>39</sup> Walker J, 'I smell a bot: California's S.B. 1001. Free speech and the future of bot regulation' 408.

<sup>40</sup> Walker J, 'I smell a bot: California's S.B. 1001. Free speech and the future of bot regulation' 408.

<sup>41</sup> Harris D 'The lost rationale of agency law'5.

relationship between corporations and chatbots would extend corporate liability to cover algorithms performing roles that human employees once filled.<sup>42</sup>

This framework will be used in this study to propose a system for determining liability for harms caused by chatbots. First, it will be used to evaluate the potential harmful effects that chatbots can cause on users. Second, it will show the causal link between the chatbot and the corporation using it. Consequently, the analysis will clarify the type of liability that should be applied and identify the appropriate party to be held responsible.

## **1.8. Literature Review**

Research on closed-domain AI chatbots remains limited, with most studies focusing on assigning liability for harm caused by Large Language Models (LLM), such as ChatGPT, as opposed to chatbots designed for specific organisations.

### **1.8.1. On chatbots as machine learning systems**

Scholars like Surden, have explained machine learning as a set of statistical tools and processes that begin with data and algorithms to devise procedures for making predictions.<sup>43</sup> Chatbots as machine learning systems, are programmed to recognise, summarise and predict text based on training data.<sup>44</sup> Therefore, using its predetermined set of rules, when a user inputs a query, the chatbot will analyse the words and respond based on its algorithmic computation.<sup>45</sup> However, this predictability comes with risks, as chatbots may produce false or misleading information.<sup>46</sup> Domingos recognises that chatbots' reliance on machine learning allows for behaviour that occurs without direct human interference, raising concerns such as privacy violations through user interactions.<sup>47</sup>

---

<sup>42</sup> Diamantis E, 'The extended corporate mind: when corporations use AI to break the law' 98(4) *North Carolina Law Review*, 2020, 900.

<sup>43</sup> Surden H, 'Artificial Intelligence and law: An overview' 35(4) *Georgia State University Law Review*, 2019, 1311.

<sup>44</sup> Domingos P, *The master algorithm: How the quest for the ultimate learning machine will remake our world*, Hachette Book Group, New York, 2015, 20.

<sup>45</sup> Gidron Z, 'Language models, conversational AI, and chatbots explained' Hyro.ai, 15 December 2021-<<https://www.hyro.ai/blog/language-models-conversational-ai-and-chatbots-explained/>> on 15 December 2021.

<sup>46</sup> Domingos P, 'The master algorithm: How the quest for the ultimate learning machine will remake our world' 20.

<sup>47</sup> Domingos P, 'The master algorithm: How the quest for the ultimate learning machine will remake our world' 20.

Chatbot's ability to retrieve information hinges on a structured decision tree framework.<sup>48</sup> This is a hierarchical system where each branch represents a decision point and sub-branches offer multiple responses depending on user input.<sup>49</sup> Additionally, through the chatbot's ability to self-learn, it can be able to discover unknown aspects of consumer interests and profiles of market behaviour.<sup>50</sup> Thus it is able to learn and generate new responses based on its environment and multiple interactions with customers, to determine their context and intent, effectively providing required responses.<sup>51</sup> The chatbot's ability to accept user corrections over time, allow it to improve and learn from user inputs and is commonly referred to as reinforcement learning.<sup>52</sup> Therefore the database found in chatbots is as a result of the specific algorithms fed to it as well as knowledge generated from its interactions with consumers, machine learning and natural language processing technologies.<sup>53</sup>

### 1.8.2. On attributing liability

AI runs afoul of several well established legal doctrines.<sup>54</sup> Many causation doctrines rely on the idea of foreseeability in order to impose liability.<sup>55</sup> Unfortunately for AI systems, determining causation is difficult due to the various stakeholders involved as well as the technology equipped in it.<sup>56</sup> Thus, the key challenge in determining liability for chatbots is who should bear the cost when harm occurs.<sup>57</sup> Given that chatbots can evolve over time by referencing new data and learning from its success and mistakes, their unpredictability complicates the process of

---

<sup>48</sup> A hierarchical system where each branch represents a decision point and sub-branches offer multiple responses depending on user input.

<sup>49</sup> -<<https://gettalkative.com/info/chatbots-vs-conversational-ai#:~:text=So%2C%20in%20short%2C%20conversational%20AI,%E2%80%9Cbot%2Dlike%E2%80%9D%20conversation.> > on 29 February 2024.

<sup>50</sup> -<<https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/deloitte-analytics/deloitte-nl-chatbots-moving-beyond-the-hype.pdf> > on March 2018.

<sup>51</sup> -<<https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/deloitte-analytics/deloitte-nl-chatbots-moving-beyond-the-hype.pdf> > on March 2018.

<sup>52</sup> Gidron Z, 'Language models, conversational AI, and chatbots explained' Hyro.ai, 15 December 2021 -<<https://www.hyro.ai/blog/language-models-conversational-ai-and-chatbots-explained/>> on 15 December 2021.

<sup>53</sup> Stojanov M, 'Prospects for chatbots' 13.

<sup>54</sup> Lewis C, 'The need for a legal framework to regulate the use of artificial intelligence' 47(2) *University of Dayton Law Review*, 2022, 304.

<sup>55</sup> Lewis C, 'The need for a legal framework to regulate the use of artificial intelligence' 304.

<sup>56</sup> Lewis C, 'The need for a legal framework to regulate the use of artificial intelligence' 304.

<sup>57</sup> Vladeck D, 'Machines without principals: Liability rules and artificial intelligence' 129.

assigning liability.<sup>58</sup> Various scholars have addressed the issue, attempting to identify the appropriate party responsible for harms caused by chatbots.<sup>59</sup>

Brown, Feferkon and Vladeck have recognised the inherent difficulty in attributing liability in chatbot errors.<sup>60</sup> Consequently, they have proposed that product liability should be extended to chatbots.<sup>61</sup> The prevailing assumption is that the producer, who controls the core coding, data processing algorithms, and the objectives of the AI, should bear responsibility.<sup>62</sup> Similar to traditional machines, chatbot manufacturers are expected to be motivated to ensure their products are safe, adhere to technical standards, undergo regular inspections, and receive necessary system updates.<sup>63</sup> While product liability aims to ensure the safety of products before they reach consumers, it is challenging to apply this concept to machine learning chatbots. These systems continuously evolve by learning from new data, which makes it difficult for manufacturers to predict future behaviour. Additionally, the technical complexities of AI and the black box nature of its decision-making process makes it challenging for claimants to establish liability.<sup>64</sup>

Some scholars, including Lai, have proposed that AI should be granted legal personhood.<sup>65</sup> This idea draws on the concept of juridical persons, where entities such as corporations or animals are recognised as legal persons.<sup>66</sup> She argues that by recognising AI as a legal person could simplify the process of assigning liability, as the AI itself would be the entity held responsible.<sup>67</sup> However, this proposal faces significant challenges as AI systems lack the financial resources to

---

<sup>58</sup> Shrestha S, 'Nature, nurture, or neither?: Liability for automated and autonomous artificial intelligence torts based on human design and influences' 29(1) *George Mason Law Review*, 2021, 6.

<sup>59</sup> Shrestha S, 'Nature, Nurture, or Neither?: Liability for automated and autonomous artificial intelligence torts based on human design and influences' 6.

<sup>60</sup> Brown N, 'Bots behaving badly: A products liability approach to chatbot generated defamation' 2023(3) *Journal Free Speech of Law*, 2023, 398; Vladeck D, 'Machines without principals: Liability rules and artificial intelligence' 129; and Feferkorn K, 'Am I an algorithm or a product? When product liability should apply to algorithmic decision-makers' 30(61) *Stanford Law School*, 2019, 90.

<sup>61</sup> Brown N, 'Bots behaving badly: A products liability approach to chatbot generated defamation' 398; Vladeck D, 'Machines without principals: Liability rules and artificial intelligence' 130; and Feferkorn K, 'Am I an algorithm or a product? When product liability should apply to algorithmic decision-makers' 90.

<sup>62</sup> Feferkorn K, 'Am I an algorithm or a product? When product liability should apply to algorithmic decision-makers' 90.

<sup>63</sup> Brown N, 'Bots behaving badly: A products liability approach to chatbot generated defamation' 400.

<sup>64</sup> Brown N, 'Bots behaving badly: A products liability approach to chatbot generated defamation' 400.

<sup>65</sup> Lai A, 'Artificial Intelligence, LLC: Corporate Personhood as tort reform' *Michigan State Law Review*, 2020, 604-[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3677360](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3677360) > on 1 October 2020.

<sup>66</sup> Lai A, 'Artificial Intelligence, LLC: Corporate Personhood as tort reform' 605.

<sup>67</sup> Lai A, 'Artificial Intelligence, LLC: Corporate Personhood as tort reform' 605.

compensate victims, rendering the assignment of liability to them ineffective.<sup>68</sup> Moreover, holding AI liable does not align with the corrective justice goals of tort law, as it is unclear how courts could punish or reform an AI system.<sup>69</sup> Furthermore, AI lacks emotions and critical thinking abilities, so holding it accountable does not satisfy the need for a responsible party.<sup>70</sup> Thus, a legally recognised human or corporate entity must be held liable for the actions of AI.

Diamantis has supported the idea of holding corporations accountable through vicarious liability.<sup>71</sup> His claim is that corporations are in the best position to mitigate the risks of employee misbehaviour.<sup>72</sup> Therefore, by threatening to punish a corporation whenever one of its employees does something wrong, respondeat superior incentivises corporations to implement compliance protocols such as additional training, monitoring, having open reporting channels and stricter disciplinary responses.<sup>73</sup> Thus the same is true for algorithms as corporations should be held responsible for algorithmic misconduct. By holding corporations responsible for the algorithms they use, the law incentivises them to do a better job in monitoring and correcting their algorithms to work more efficiently.

### **1.8.3. Contribution**

While this study aligns with Diamantis E's view of corporate responsibility, it offers a unique perspective by applying the principal-agent relationship model to chatbots in the Kenyan context. What sets this study apart is its application of the control-benefit traditionally used in servant-master relationships in Kenya, to the context of chatbots. By critically assessing the legal framework in Kenya and analysing court standards on principal-agent relationships, this research shows how these principles can be extended to the actions of chatbots. Therefore, this approach not only advances the argument that corporations can be held liable for the actions of their chatbots but also provides a novel legal basis for attributing such liability within the Kenyan legal context.

---

<sup>68</sup> Lai A, 'Artificial Intelligence, LLC: Corporate Personhood as tort reform' 608.

<sup>69</sup> Lai A, 'Artificial Intelligence, LLC: Corporate Personhood as tort reform' 608.

<sup>70</sup> Lai A, 'Artificial Intelligence, LLC: Corporate Personhood as tort reform' 608.

<sup>71</sup> Diamantis E, 'The extended corporate mind:When corporations use AI to break the law' 926-927.

<sup>72</sup> Diamantis E, 'The extended corporate mind:When corporations use AI to break the law' 928.

<sup>73</sup> Diamantis E, 'The extended corporate mind:When corporations use AI to break the law' 928.

## **1.9. Methodology**

This study will employ a doctrinal approach, focusing on analysing primary legal sources such as statutes and case law, as well as secondary sources, including journal articles, credible reports, and books related to liability. This combination of sources will provide an appropriate legal foundation for examining the liability issues surrounding chatbots.

To address the first research question, a content analysis methodology will be used to understand the technical components and functioning of chatbots. This will be done by analysing technical literature, industry standards, research papers and reports on chatbot architecture to break down how chatbots function, what algorithms they use, and what factors contribute to their output. Relevant sources will be determined by selecting recent technical studies on popular AI platforms such as the EA forum and expert reviews to identify key trends and technologies.

The second research question requires doctrinal research to investigate the traditional principles used in Kenya to assign civil liability to corporations. This method involves examining legal texts, case law, and judicial opinions to determine how Kenyan law assigns liability to companies, especially for indirect harm. Relevant cases will be sourced from the Laws of Kenya Database and legislations such as the *Consumer Protection Act, 2012* will be used to conduct a comprehensive review of precedents and legal principles used in similar scenarios. Such scenarios include cases that have determined corporations to be held vicariously liable due to the employee-employer relationship as well as agentic relationships that have resulted in harm being attributed to the company. By identifying cases that share a similar agentic relationship with corporations, it will assist in critically analysing the similarity in such cases and assist in determining the regulatory framework needed for chatbots and determine why such uniqueness exists.

To answer the third research question, a combination of doctrinal and philosophical analysis will be employed to address the viability of holding companies accountable for the actions of their chatbots. The doctrinal analysis will involve an in depth examination of the existing legal framework in Kenya, focusing on relevant statutes, and principles related to liability for harms caused indirectly, such as through negligence or product liability. By analysing these legal

sources, I aim to determine whether these areas adequately encompass scenarios where harm results indirectly from automated systems like chatbots. If such principles are difficult to apply or do not fit AI-related systems, it will indicate a gap. Consequently, a comprehensive proposal on the legal framework to be adopted will be suggested in the study. In conjunction with this, a philosophical analysis will explore the ethical and normative dimensions of responsibility and accountability in the context of AI systems. This approach allows for a deeper examination of questions around corporate responsibility and the extent to which companies should be held accountable for the actions of autonomous technologies like chatbots. Drawing on philosophical theories of agency, causation, and corporate responsibility, I will analyse whether it is justifiable, both legally and morally, to attribute liability to companies for chatbot actions. By examining the ethical justifications for corporate responsibility in the context of emerging technologies, we can gauge whether existing laws align with societal expectations of accountability and justice, or if they fall short in addressing the complexities of AI-driven harm. This combined methodology will provide a comprehensive framework for assessing the viability of holding companies accountable in the evolving legal landscape surrounding AI.

An inductive approach underpins this study, as it begins with specific observations and analyses them to form broader conclusions about chatbot liability. By examining individual cases, legal principles, and the technical workings of chatbots, the study gathers detailed information that can be used to form broader conclusions on the adequacy of Kenya's legal framework in addressing chatbot-related harm.

### **1.10. Chapter Breakdown**

In Chapter One, I will establish the background and significance of the study, including the research objectives, conceptual framework, and justification for the research among others, which will set the foundation for the subsequent chapters.

Chapter Two will focus on the underlying technology of chatbots, explaining how they function and the key technological principles involved.

Chapter Three will analyse the existing legal and regulatory framework for assigning civil liability for harms caused indirectly by companies in Kenya. This analysis will include a review of court decisions on such relationships, the standards applied, and an examination of whether these standards can be extended to chatbot-related harms.

In Chapter Four, I will discuss the way forward, exploring whether existing legal standards should be modified to regulate harms caused by chatbots of companies in Kenya or if they are already sufficient to cover such issues.

Finally, Chapter Five will summarise the main findings of the research, drawing conclusions based on the findings. It will also consider the implications of the findings for the broader field of AI, offering valuable insights and recommendations for future research.

## 2.0. The Technology behind Chatbots, its Benefits & Risks

### 2.1. Introduction

The previous chapter established the foundation for this study by outlining the rationale behind its necessity. Building on that, this chapter explores the technological components of chatbots, examining their functionality, benefits, and the potential risks they pose to users.

The use of chatbots by companies has risen significantly in recent years, particularly as a tool for enhancing customer relationship management (CRM) processes.<sup>74</sup> These AI-driven systems improve efficiency while reducing operational costs by taking over tasks traditionally performed by human employees.<sup>75</sup> As businesses increasingly integrate chatbots into their operations, it becomes essential to understand how they function, not only to grasp their internal mechanisms, but also to identify potential risks such as misinformation that may cause harm to users.

This chapter focuses on specific components of chatbot technology that are fundamental to their operation. These components were selected based on their important role in enabling a chatbot to process user queries and generate appropriate responses and are thus the main components of a chatbot. By examining these elements, the chapter aims to provide a clearer understanding of how chatbots function and the technological processes that underpin their interactions with users.

## 2.2. The Technology behind Chatbots

### 2.2.1. Data Training

This marks the initial phase of chatbot development, where data is continuously fed into the system.<sup>76</sup> The data provided is tailored to align with the operator's objectives, consisting of predefined questions, corresponding answers, and variations of those answers to

---

<sup>74</sup> Adam M, Wessel M and Benlian A 'AI-based chatbots in customer service and their effects on user compliance' *Electronic Markets*, 2021, 428—<<https://link.springer.com/article/10.1007/s12525-020-00414-7>> on 17 March 2020; Cheng X, Bao Y, Zarifis A, Gong W and Mou J 'Exploring consumers' response to text-based chatbots in e-commerce: the moderating role of task complexity and chatbot disclosure' 32(2) *Internet Research*, 2021, 498.

<sup>75</sup> Adam M, Wessel M and Benlian A, 'AI-based Chatbots in Customer Service and their Effects on User Compliance' 428; Maruti T 'Can Chatbots Help Reduce Customer Service Costs by 30%' *Chatbots Magazine*, 21 April 2017 —<<https://chatbotmagazine.com/how-with-the-help-of-chatbots-customer-service-costs-could-be-reduced-up-to-30-b9266a369945>> on 21 April 2017.

<sup>76</sup> —<<https://www.mathworks.com/discovery/deep-learning.html>>.

comprehensively address anticipated customer inquiries.<sup>77</sup> Generally, the more questions it is fed, the broader its response capacity and the less an occurrence of redundancy in responses provided.<sup>78</sup> Additionally, through machine learning, the chatbot develops the ability to generate responses to questions that were not explicitly included during the training phase.<sup>79</sup> By learning from its environment, the chatbot can formulate appropriate answers based on patterns and information acquired during its general training phase.<sup>80</sup>

### 2.2.2. User Interface

The user interface serves as the communication platform through which users interact with the chatbot.<sup>81</sup> It provides a structured environment where users can input queries, which the chatbot then processes and responds to within the same interface.<sup>82</sup> Typically designed as a chat-based interface, it facilitates seamless interaction by enabling the exchange of messages in real-time.<sup>83</sup>

The interface consists of essential components such as a text input field, where users type their queries, and a response display area, where the chatbot's replies appear.<sup>84</sup> Some advanced interfaces also incorporate buttons, quick reply options, and voice input features to enhance user experience and improve interaction efficiency.<sup>85</sup> By ensuring a smooth flow of information

---

<sup>77</sup> Khanna A, Pandey B, Vashishta K, Kalia K, Pradeepkumar B and Das T, 'A study of today's AI through chatbots and rediscovery of machine intelligence' 8(7) *International Journal of u- and e- Service, Science and Technology*, 2015, 278.

<sup>78</sup> Khanna A, Pandey B, Vashishta K, Kalia K, Pradeepkumar B and Das T, 'A study of today's AI through chatbots and rediscovery of machine intelligence' 278.

<sup>79</sup> Khanna A, Pandey B, Vashishta K, Kalia K, Pradeepkumar B and Das T, 'A study of today's AI through chatbots and rediscovery of machine intelligence' 278.

<sup>80</sup> Khanna A, Pandey B, Vashishta K, Kalia K, Pradeepkumar B and Das T, 'A study of today's AI through chatbots and rediscovery of machine intelligence' 278.

<sup>81</sup> Ukpabi D, Aslam B and Karjaluo H 'Chatbot adoption in tourism services: A conceptual exploration' ResearchGate, 2019,4.

<sup>82</sup> Zumstein D and Hundertmark S, 'Chatbots—An interactive technology for personalised communication, transaction and services' 15(1) *IADIS International Journal*, 2017, 98; Wang Y and Petrina S, 'Using learning analytics to understand the design of an intelligent language tutor-chatbot Lucy' 4(11) *International Journal of Advanced Computer Science and Applications (IJACSA)*, 2013, 124.

<sup>83</sup> Wang Y and Petrina S, 'Using learning analytics to understand the design of an intelligent language tutor-chatbot Lucy' 124.

<sup>84</sup> Wang Y and Petrina S, 'Using learning analytics to understand the design of an intelligent language tutor-chatbot Lucy' 124.

<sup>85</sup> Wang Y and Petrina S, 'Using learning analytics to understand the design of an intelligent language tutor-chatbot Lucy' 124; Haugeland I, Folstad A, Taylor C and Bjorkli C, 'Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction and design' 161(102788) *International Journal of Human-Computer Studies*, 2022, 2.

between the chatbot and the user, the interface plays a crucial role in delivering an intuitive and user-friendly experience.

### **2.2.3. Knowledge Base**

The Knowledge Base (KB) serves as the central repository of information that enables a chatbot to generate appropriate responses.<sup>86</sup> Often referred to as the "brain" of the chatbot, it stores and organises data that has been fed into the system during training. Implementation of the KB is done through databases, text files, script files and XML files.<sup>87</sup> When a user inputs a query through the chat interface, the chatbot searches the KB to determine whether a relevant answer exists.<sup>88</sup> If a matching response is found, the chatbot retrieves and delivers it to the user.<sup>89</sup>

The effectiveness of the KB depends on the quality and comprehensiveness of the data provided during training, thus the reason why data training is very important.<sup>90</sup> It typically consists of predefined questions and answers, alternative phrasings, and structured information that allows the chatbot to handle a wide range of user inquiries.<sup>91</sup>

### **2.2.4. Machine Learning**

Machine Learning (ML) empowers chatbots to continuously learn from their environment, enabling them to handle questions beyond those initially programmed.<sup>92</sup> By working alongside the KB, ML algorithms allow chatbots to refine their responses over time based on user interactions and newly acquired data.<sup>93</sup>

---

<sup>86</sup> Reshmi S and Balakrishnan K, 'Empowering chatbots with business intelligence by big data integration' 9(1) *International Journal of Advanced Research in Computer Science*, 2018, 627.

<sup>87</sup> Reshmi S and Balakrishnan K, 'Empowering chatbots with business intelligence by big data integration' 627.

<sup>88</sup> Reshmi S and Balakrishnan K, 'Empowering chatbots with business intelligence by big data integration' 627.

<sup>89</sup> Reshmi S and Balakrishnan K, 'Empowering chatbots with business intelligence by big data integration' 627.

<sup>90</sup> Haugeland I, Folstad A, Taylor C and Bjorkli C, 'Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction and design' 4.

<sup>91</sup> Haugeland I, Folstad A, Taylor C and Bjorkli C, 'Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction and design' 4.

<sup>92</sup> Viski A, Jones S, Rand L, Boyce T and Siegel J, *Artificial intelligence and strategic trade controls*, Center for International & Security Studies, Maryland, 2020, 21.

<sup>93</sup> Viski A, Jones S, Rand L, Boyce T and Siegel J, *Artificial intelligence and strategic trade controls*, 21.

Unlike traditional rule-based systems that rely solely on predefined inputs and outputs, ML-equipped chatbots can recognise and classify new queries, even if they were not explicitly included during training.<sup>94</sup> Through pattern recognition and adaptive learning, the chatbot can analyse unfamiliar inputs, draw connections to existing knowledge, and generate appropriate responses.<sup>95</sup> This ability to evolve and improve makes ML-driven chatbots more dynamic, responsive, and capable of handling a broader range of user inquiries.<sup>96</sup>

A chatbot's ability to recognise patterns and establish connections between data is referred to as deep learning, a specialised subset of machine learning.<sup>97</sup> Deep learning is efficient in pattern tracing and prediction related tasks such as activity identification.<sup>98</sup> This advanced capability is made possible through employing a classification method composed of neural networks that make it have the capacity to retrieve data.<sup>99</sup> The neural networks mimic the way the human brain processes information, whereby the network consists of interconnected nodes or neurons in a layered structure that relate the input to its corresponding output.<sup>100</sup> The neurons between the input and output layers of the neural network are usually hidden which gives rise to the term "deep" that refers to the number of hidden layers.<sup>101</sup> Consequently, these neural networks rely on labeled data sets and systems with large computing power capabilities that allow them to employ a classification model.<sup>102</sup> By leveraging multiple layers of interconnected nodes, deep learning enables chatbots to analyse complex inputs, refine their understanding over time, and deliver highly accurate responses.<sup>103</sup>

---

<sup>94</sup> Montagnani M 'Liability and emerging digital technologies: an EU perspective' 11(2) *Notre Dame Journal of International & Comparative Law*, 2020,91.

<sup>95</sup> Montagnani M 'Liability and emerging digital technologies: an EU perspective' 91; Radziwill N and Benton M, 'Evaluating quality of chatbots and intelligent conversational agents' *Cornell University*, 2017, 5—<[https://www.researchgate.net/publication/316184347\\_Evaluating\\_Quality\\_of\\_Chatbots\\_and\\_Intelligent\\_Conversational\\_Agents](https://www.researchgate.net/publication/316184347_Evaluating_Quality_of_Chatbots_and_Intelligent_Conversational_Agents)> on April 2017; Weber U, Lomker M and Moskaliuk J, 'The human touch: The impact of anthropomorphism in chatbots on the perceived success of solution focused coaching' 32(4) *Management Revenue*, 2021, 385.

<sup>96</sup> Weber U, Lomker M and Moskaliuk J, 'The human touch: The impact of anthropomorphism in chatbots on the perceived success of solution focused coaching' 385.

<sup>97</sup> Viski A et al, *Artificial Intelligence and Strategic Trade Controls*, 22.

<sup>98</sup> Viski A et al, *Artificial Intelligence and Strategic Trade Controls*, 22.

<sup>99</sup> Viski A et al, *Artificial Intelligence and Strategic Trade Controls*, 22.

<sup>100</sup> —<<https://www.mathworks.com/discovery/deep-learning.html>>.

<sup>101</sup> —<<https://www.mathworks.com/discovery/deep-learning.html>>.

<sup>102</sup> Viski A et al, *Artificial Intelligence and Strategic Trade Controls*, 22.

<sup>103</sup> Viski A et al, *Artificial Intelligence and Strategic Trade Controls*, 22.

Various deep learning techniques enhance chatbot functionality, one of the most notable being Recurrent Neural Networks (RNNs).<sup>104</sup> RNNs leverage past information to improve future performance, making them particularly effective for processing sequential data, such as conversations.<sup>105</sup> At the core of an RNN, is its hidden state and looping structure, which allows it to retain previous inputs and use them to influence future responses.<sup>106</sup> This memory-like mechanism enables the network to recognise patterns, maintain context, and generate more coherent interactions over time.<sup>107</sup> Much like the human brain's short-term and long-term memory, RNNs store relevant information from past exchanges, ensuring continuity and improved accuracy in chatbot responses.<sup>108</sup>

#### 2.2.4.1. Natural Language Processing

Natural Language Processing (NLP) is a branch of ML that allows chatbots to process human language to determine user intent through tasks like text classification, machine translation, entity recognition and sentiment analysis.<sup>109</sup> This is possible due to the use of deep learning techniques such as RNN that allow it to learn how to contextualise words in a sentence.<sup>110</sup> Therefore, once a customer's request is made at the communication interface, it is recorded by a Natural Language Parser (NLP) and is translated into the programming language of the conversation engine.<sup>111</sup>

The primary goal of NLP is to analyse and synthesise user input, enabling chatbots to easily understand what is being asked or said by the customer.<sup>112</sup> By recognising patterns in language and extracting meaning from large datasets, NLP helps chatbots uncover the relationships within text and even generate new language data by predicting the subsequent words that are to

---

<sup>104</sup>—<<https://www.mathworks.com/discovery/rnn.html>>.

<sup>105</sup> —<<https://www.mathworks.com/discovery/rnn.html>>.

<sup>106</sup> —<<https://www.mathworks.com/discovery/rnn.html>>; Wang Y and Petrina S, 'Using Learning Analytics to Understand the Design of an Intelligent Language Tutor-Chatbot Lucy' 125.

<sup>107</sup> —<<https://www.mathworks.com/discovery/rnn.html>>.

<sup>108</sup> Wang Y and Petrina S, 'Using Learning Analytics to Understand the Design of an Intelligent Language Tutor-Chatbot Lucy' 125.

<sup>109</sup>—<<https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>110</sup> —<<https://www.mathworks.com/discovery/rnn.html>>.

<sup>111</sup> Zumstein D and Hundertmark S, 'Chatbots—an interactive technology for personalised communication, transaction and services' 98.

<sup>112</sup> —<<https://www.mathworks.com/discovery/natural-language-processing.html>>.

follow.<sup>113</sup> This capability enhances the chatbot's ability to engage in more natural, context-aware interactions with users.<sup>114</sup> Furthermore, it allows chatbots to discover and visualise complex relationships in large data sets and generate new language data.<sup>115</sup>

NLP involves a series of steps to transform raw, unstructured text into a structured format that a chatbot can easily interpret.<sup>116</sup> The process begins with tokenisation, where text is divided into sentences or words to facilitate intent recognition.<sup>117</sup> Next, stemming or lemmatisation is applied to reduce words to their root forms by removing prefixes and suffixes.<sup>118</sup> This step may sometimes result in non-standard words, such as converting the sentence "*Building has floors*" into "*build floor.*"<sup>119</sup> Once these transformations are complete, the text proceeds to the Natural Language Understanding (NLU) phase for further processing.<sup>120</sup>

#### 2.2.4.2. Natural Language Understanding

In the NLU stage, a syntactic and semantic analysis is done to extract the meaning of words.<sup>121</sup> At this point the intent is already identified, thus the NLU comes in to achieve the most desirable response for the user through a rule based dialogue structure.<sup>122</sup> This is a tree with multiple nodes that have several branches leading to other nodes.<sup>123</sup> Therefore, when a question is asked, the specific words trigger specific nodes that have various predefined replies.<sup>124</sup> Therefore, if the input matches a pattern, the predefined response is selected and forwarded.<sup>125</sup> This ultimately

---

<sup>113</sup> —<<https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>114</sup> Zumstein D and Hundertmark S, 'Chatbots—an interactive technology for personalised communication, transaction and services' 99.

<sup>115</sup> Zumstein D and Hundertmark S, 'Chatbots—an interactive technology for personalised communication, transaction and services' 99.

<sup>116</sup> Zumstein D and Hundertmark S, 'Chatbots—an interactive technology for personalised communication, transaction and services' 99.

<sup>117</sup> —<<https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>118</sup> —<<https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>119</sup> —<<https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>120</sup> —<<https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>121</sup> —<<https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>122</sup> Haugeland I, Folstad A, Taylor C and Bjorkli C, 'Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction and design' 3.

<sup>123</sup> Haugeland I, Folstad A, Taylor C and Bjorkli C, 'Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction and design' 3.

<sup>124</sup> Haugeland I, Folstad A, Taylor C and Bjorkli C, 'Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction and design' 3.

<sup>125</sup> Baris A, 'The use of chatbots in customer service: A qualitative analysis on customers' reception' Yasar University, 2021, 15—<<https://dSPACE.yasar.edu.tr/bitstream/handle/20.500.12742/18727/688219.pdf?sequence=1>> on 21 June 2021.

reduces the risk of chatbot misinformation due to an erroneous interpretation of user's input, while allowing for the best response to be provided to the customer.

#### **2.2.4.3. Dialogue Management Component**

The Dialogue Management Component is the system responsible for controlling and updating the conversation.<sup>126</sup> The system ensures that the query is still in the user interface and if the chatbot is unable to identify intent, it is responsible for follow up questions up until the intent is recognised.<sup>127</sup> The component has various functions. These include; firstly, ambiguity handling, which essentially provides answers when the chatbot may not have predefined context for the specific query, thus will be able to inform the customer that they either do not have an answer, seek clarification if it is unsure of the intent, ask follow up questions or give a general answer to satisfy the customer.<sup>128</sup> Secondly, it contains a data handling component, which stores user's information to allow a historical record of right and wrong answers.<sup>129</sup> This allows it to modify answers for future interactions with other customers.<sup>130</sup> Thirdly, it also contains an error handling component which allows it to function even with unexpected grammatical errors that are identified from customer's questions, allowing for proper chatbot operation despite such errors.<sup>131</sup>

#### **2.2.4.4. Natural Language Generation**

Natural Language Generation (NLG) deals with the ability to generate meaningful and contextually relevant responses based on the output from the NLU and the dialogue management.<sup>132</sup> At this stage, the chatbot has already found the answer and has to convert the language which is still in a programming language to the language being used by the user.<sup>133</sup>

---

<sup>126</sup> Adamopoulou E and Moussiades L, 'Chatbots: History, technology, and applications' 2(100006) *Machine Learning with Applications*, 2020,10.

<sup>127</sup> Adamopoulou E and Moussiades L, 'Chatbots: History, technology, and applications' 10.

<sup>128</sup> Adamopoulou E and Moussiades L, 'An overview of chatbot technology' *Artificial Intelligence Applications and Innovations*, Maglogiannis, I.Illiadis, L., Pimendis, 29 May 2020, 380.

<sup>129</sup> Adamopoulou E and Moussiades L, 'An overview of chatbot technology' 380.

<sup>130</sup> Adamopoulou E and Moussiades L, 'An overview of chatbot technology' 380.

<sup>131</sup> Adamopoulou E and Moussiades L, 'An overview of chatbot technology' 380.

<sup>132</sup> —< <https://www.mathworks.com/discovery/natural-language-processing.html>>; Adamopoulou E and Moussiades L, 'Chatbots: History, technology, and applications' 7; Saka A, Bello S, Oyedele L, Akanbi L, Ganiyu S and Chan D 'Conversational artificial intelligence in the AEC industry: A review of present status, challenges and opportunities' 55(101869) *Advanced Engineering Informatics*, 2023, 5.

<sup>133</sup> Zumstein D and Hundertmark S, 'Chatbots—An interactive technology for personalized communication, transaction and services' 99; Kaput M, 'Natural Language Generation (NLG): Everything you need to know' *Market*

Therefore, the tasks of the chatbot in NLG, is content planning<sup>134</sup> and machine translation.<sup>135</sup> This is done through two methods, namely rule or retrieval based approach and generative based approach.<sup>136</sup>

The rule based approach allows for the retrieval of responses from the KB and follows pre-defined pathways and neural networks to generate responses.<sup>137</sup> On the other hand, a generative based approach, utilises past inputs and responses to provide answers to new queries.<sup>138</sup> The basic generation model is a recurrent sequence-to-sequence model (Seq 2 Seq) which sequentially feeds in each word in the query as input, and then generates the output word one by one.<sup>139</sup>

### 2.3. The Benefits of Chatbots

Chatbots are used by companies to provide specific information and direct dialogue to specific topics such as website guides, frequently asked questions (FAQ) guides, virtual support agents, virtual sales agents, survey takers and chat-room hosts among others.<sup>140</sup> Chatbots offer numerous benefits for both customers and businesses.<sup>141</sup> For customers, they streamline the information retrieval process on websites and apps, allowing users to quickly obtain the details they need by simply making a request, rather than spending time searching manually or calling the organisation.<sup>142</sup> This reduces the waiting time that

---

Artificial Intelligence Institute, 28 February 2024—<<https://www.marketingaiinstitute.com/blog/the-beginners-guide-to-using-natural-language-generation-to-scale-content-marketing> > on 28 February 2024.

<sup>134</sup> Where the NLG system decides what information to include and how to structure sentences and paragraphs.

<sup>135</sup> —< <https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>136</sup> —< <https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>137</sup> —< <https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>138</sup> —< <https://www.mathworks.com/discovery/natural-language-processing.html>>.

<sup>139</sup> Gao C, Lei W, He X, Rijke M and Chua T, ‘Advances and challenges in conversational recommender systems: A survey’ *KeAI*, 2021, 111—<[https://www.sciencedirect.com/science/article/pii/S2666651021000164?ref=pdf\\_download&fr=RR-2&rr=90e4a40ff861b199](https://www.sciencedirect.com/science/article/pii/S2666651021000164?ref=pdf_download&fr=RR-2&rr=90e4a40ff861b199) > on 24 July 2021.

<sup>140</sup> Wang Y and Petrina S, ‘Using learning analytics to understand the design of an intelligent language tutor-chatbot Lucy’ 124.

<sup>141</sup> Wang Y and Petrina S, ‘Using learning analytics to understand the design of an intelligent language tutor-chatbot Lucy’ 124.

<sup>142</sup> Kevin Scott, ‘Popular use cases for chatbots’ *Chatbot Magazine*, 18 October 2016—<<https://chatbotmagazine.com/popular-use-cases-for-chatbots-925ef8f2b48b> > on 18 October 2016.

would have accrued to the customer if the question was to be lodged manually through an email or calling a call-center.<sup>143</sup>

On the other hand, for businesses, integrating chatbots into CRM systems ensures that customer inquiries are handled efficiently, providing support around the clock—even beyond regular operating hours.<sup>144</sup> Furthermore, as stated earlier, chatbots allow for companies to save up to 30% since they have the capacity of automating most of the contact center staff.<sup>145</sup> Additionally, compared to a human, chatbots do not experience negative emotions or get tired, thus can always operate optimally and effectively.<sup>146</sup> Furthermore, due to their effectiveness, they can easily handle multiple customer queries at the same time unlike individuals, who can only serve the needs of one customer at a time.<sup>147</sup> Therefore, for businesses, customer engagement is important to maintain customer satisfaction, thus the reason for the rise in the use of chatbots.<sup>148</sup>

## 2.4. Instances of Harm Caused by Chatbots

Despite the technical advantages chatbots bring, they also cause harm to customers. This can be through the provision of unsatisfactory or unsuitable responses to the user's requests.<sup>149</sup> Deep learning models are often described as opaque, meaning their decision-making processes are not

---

<sup>143</sup> Maruti T 'Can chatbots help reduce customer service costs by 30%'; Wessel M and Benlian A, 'AI-based chatbots in customer service and their effects on user compliance' 429.

<sup>144</sup> Kevin Scott, 'Popular use cases for chatbots'; Zumstein D and Hundertmark S, 'Chatbots—an interactive technology for personalised communication, transaction and services' 101.

<sup>145</sup> Maruti T 'Can chatbots help reduce customer service costs by 30%'; Wessel M and Benlian A, 'AI-based chatbots in customer service and their effects on user compliance' 429.

<sup>146</sup> Luo X, Tong S, Fang Z and Qu Z, 'Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases' 38(6) *Marketing Science*, 2019, 937-947.

<sup>147</sup> Cheng X, Bao Y, Zarifis A, Gong W, Mou J 'Exploring consumers' response to text-based chatbots in e-commerce: the moderating role of task complexity and chatbot disclosure' 497.

<sup>148</sup> Ltifi M, 'Trust in the chatbot: a semi-human relationship' 9(109) *Future Business Journal*, 2023, 2; Adam M, Wessel M and Benlian A, 'AI-based chatbots in customer service and their effects on user compliance' 427.

<sup>149</sup> Luger E and Sellen A, 'Like having a really bad PA: The gulf between user expectation and experience of conversational agents' Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, California, San Jose, USA, 7-12 May 2016, 5286-5297; Hawley M, 'Exploring Air Canada's AI chatbot dilemma' CMSWire, 2 April

2024-<<https://www.cmswire.com/customer-experience/exploring-air-canadas-ai-chatbot-dilemma/>>- on 2 April 2024.

easily understood or predicted.<sup>150</sup> This phenomenon, known as the black box problem, arises because these models generate predictions based on input data without providing clear explanations for how they reach their conclusions.<sup>151</sup> As a result, users and operators may struggle to understand and interpret the model's reasoning, identify potential biases or errors, and ensure accountability for its decisions.<sup>152</sup>

One of the biggest technical challenges in chatbot development is language processing.<sup>153</sup> Chatbots often struggle with lexical and semantic ambiguity, meaning they may misinterpret words with multiple meanings or fail to grasp the context of a conversation.<sup>154</sup> This ultimately leads to an incorrect response given to the customer, since the chatbot was incapable of identifying the intent. Additionally, chatbots face issues unique to their functionality, such as maintaining coherent conversation flow, avoiding repetitive responses, and handling unclear or vague user inputs effectively.<sup>155</sup> These limitations can affect the chatbot's ability to provide accurate and engaging interactions.

Due to the above stated challenges, there have been several instances where chatbots have caused harm. These include:

#### 1) Microsoft's Tay

Microsoft launched an AI chatbot in 2016 known as Tay.<sup>156</sup> The main goal was to ensure it learns millennial's language use from various apps such as Twitter, GroupMe and Kik, among others to provide an appropriate response fit for them.<sup>157</sup> The company was impressed with Tay's capability to learn new vocabulary and speech patterns, however, they did not expect that it

---

<sup>150</sup> Hassija V, Chamola V, Mahapatra A, Singal A, Goal D, Huang K, Scardapane S, Spinelli I, Mahmud M and Hussain A, 'Interpreting Black-Box Models: A review on explainable artificial intelligence' *Cognitive Computing*, 2023, 45-<<https://link.springer.com/article/10.1007/s12559-023-10179-8#citeas>> on 24 August 2023.

<sup>151</sup> Hassija V, Chamola V, Mahapatra A, Singal A, Goal D, Huang K, Scardapane S, Spinelli I, Mahmud M and Hussain A, 'Interpreting black-box models: A review on explainable artificial intelligence', 47; Price N, 'Black-box medicine' 28(2) *Harvard Journal of Law & Technology*, 2015, 432.

<sup>152</sup> Hassija V, Chamola V, Mahapatra A, Singal A, Goal D, Huang K, Scardapane S, Spinelli I, Mahmud M and Hussain A, 'Interpreting black-box models: A review on explainable Artificial Intelligence', 47.

<sup>153</sup> Ukpabi D, Aslam B and Karjaluto H 'Chatbot adoption in tourism services: A conceptual exploration' 8.

<sup>154</sup> Ukpabi D, Aslam B and Karjaluto H 'Chatbot adoption in tourism services: A conceptual exploration' 8.

<sup>155</sup> Ukpabi D, Aslam B and Karjaluto H 'Chatbot adoption in tourism services: A conceptual exploration' 8.

<sup>156</sup> Heine C, 'Microsoft's Chatbot 'Tay' Just Went on a Racist, Misogynistic Anti-Semitic Tirade'.

<sup>157</sup> Heine C, 'Microsoft's Chatbot 'Tay' Just Went on a Racist, Misogynistic Anti-Semitic Tirade'.

would learn racist, misogynistic and anti-semantic language and relay the same to users.<sup>158</sup> This ultimately led to its shut down.<sup>159</sup> Microsoft stated that the reason for the chatbot's behaviour was as a result of a coordinated attack by a group of people, which they did not anticipate at the creation phase.<sup>160</sup>

## 2) Moffat v. AirCanada

In the AirCanada case, the Claimant sought compensation for receiving the wrong information by the chatbot.<sup>161</sup> The airline could only provide the bereavement fare discount prior to taking the flight and not after.<sup>162</sup> However upon inquiry by the Claimant, Mr.Moffat, the chatbot informed him that the discount would be applied after his flight and could seek it at such a date.<sup>163</sup> However, after he took the flight and sought for the payment, AirCanada stated that they could not provide the bereavement fare discount as it was to be sought prior to taking the flight.<sup>164</sup> When the Claimant stated that it obtained the information from the company's chatbot, AirCanada stated that they were not liable for the chatbot's error as their policy in their website says that the refund can only be given if it was sought prior to taking the flight to his respective destination.<sup>165</sup> Further, that it was incumbent upon any customer to clarify the information obtained from the chatbot with the information provided in their website as the chatbot may have errors.<sup>166</sup> Despite the company trying to claim that the chatbot was a separate legal entity and thus AirCanada could not be held liable, the tribunal found AirCanada responsible.<sup>167</sup> Their logic

---

<sup>158</sup> Heine C, 'Microsoft's Chatbot 'Tay' Just Went on a Racist, Misogynistic Anti-Semitic Tirade'.

<sup>159</sup> Heine C, 'Microsoft's Chatbot 'Tay' Just Went on a Racist, Misogynistic Anti-Semitic Tirade'.

<sup>160</sup> Lee P, 'Learning from Tay's introduction' Official Microsoft Blog, 25 March 2016,

-<<https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/> > on 25 March 2016.

<sup>161</sup> Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024-<<https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> > on 16 February 2024; Brand J, 'Air Canada's chatbot illustrates persistent agency and responsibility gap problems for AI' *AI & Society*, 2024, 2 -<<https://doi.org/10.1007/s00146-024-02096-7>> on 23 October 2024.

<sup>162</sup> Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024-<<https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> > on 16 February 2024.

<sup>163</sup> Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024-<<https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> > on 16 February 2024.

<sup>164</sup> Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024-<<https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> > on 16 February 2024.

<sup>165</sup> Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024-<<https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> > on 16 February 2024.

<sup>166</sup> Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024-<<https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> > on 16 February 2024.

<sup>167</sup> Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024-<<https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> > on 16 February 2024.

was that it was the company's responsibility to train the chatbot effectively to give accurate information that aligns with their company's policies.<sup>168</sup> As such, the company was asked to compensate the Claimant an amount of \$812.<sup>169</sup>

### 3) Gok spreading misinformation during elections

Several instances have been reported where AI-powered chatbots have spread election misinformation to voters. A notable example is the 2024 U.S. elections, where Grok, an AI chatbot developed by X, disseminated false information about electoral processes.<sup>170</sup> When asked whether certain prospective candidates still had time to be added to the ballot, Grok provided inaccurate responses, misleading voters.<sup>171</sup> In response, a group of state secretaries contacted X, urging them to ensure the chatbot delivers accurate information or, at the very least, directs voters to credible sources instead of spreading misinformation.<sup>172</sup>

### 4) Character.ai's chatbot led to a child committing suicide

Megan Garcia, the mother of 14-year-old Sewell, accused Character.ai and its founder, as well as Google, for contributing to her son's death.<sup>173</sup> She claims that the chatbot worsened Sewell's depression and encouraged him to commit suicide.<sup>174</sup> According to Garcia, Sewell frequently communicated with the chatbot on his phone, spending long hours alone in his room.<sup>175</sup> Evidence

---

<sup>168</sup> Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024—<<https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> > on 16 February 2024.

<sup>169</sup> Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024—<<https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> > on 16 February 2024.

<sup>170</sup> Leingang R, 'X's AI Chatbot spread voter misinformation-and election officials fought back' The Guardian, 12 September 2024—<<https://www.theguardian.com/us-news/2024/sep/12/twitter-ai-bot-grok-election-misinformation> > on 12 September 2024.

<sup>171</sup> Leingang R, 'X's AI Chatbot spread voter misinformation-and election officials fought back' The Guardian, 12 September 2024—<<https://www.theguardian.com/us-news/2024/sep/12/twitter-ai-bot-grok-election-misinformation> > on 12 September 2024.

<sup>172</sup> Leingang R, 'X's AI Chatbot spread voter misinformation-and election officials fought back' The Guardian, 12 September 2024—<<https://www.theguardian.com/us-news/2024/sep/12/twitter-ai-bot-grok-election-misinformation> > on 12 September 2024.

<sup>173</sup> Montgomery B, 'Mother says AI chatbot led her son to kill himself in lawsuit against its maker' The Guardian, 23 October 2024—<<https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death> > on 23 October 2024.

<sup>174</sup> Montgomery B, 'Mother says AI chatbot led her son to kill himself in lawsuit against its maker' The Guardian, 23 October 2024—<<https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death> > on 23 October 2024.

<sup>175</sup> Montgomery B, 'Mother says AI chatbot led her son to kill himself in lawsuit against its maker' The Guardian, 23 October 2024—<<https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death> > on 23 October 2024.

on record reveals that the AI-powered chatbot allegedly asked Sewell if he had devised a plan to end his life.<sup>176</sup> To which Sewell responded by expressing concerns about the pain he might endure and the uncertainty of success.<sup>177</sup> Tragically, the chatbot is said to have replied that these were insufficient reasons to deter him from proceeding to end his life, which resulted in his untimely death.<sup>178</sup>

This is not the first time an AI chatbot has been implicated in such a case. In a similar incident, a Belgian man reportedly took his own life after engaging in conversations with an AI chatbot named Eliza.<sup>179</sup> During their discussions about the future, Eliza allegedly encouraged the man to sacrifice himself as a way to combat climate change.<sup>180</sup> Unfortunately, since he cared about climate change, he took the chatbot's advice and ended his life.<sup>181</sup>

From the examples given, chatbots can carry out varying degrees of harm, with the most extreme cases resulting in death. Unfortunately, in most of the instances provided above, the Company operating the chatbot did not have any explanation as to how the chatbot led to such grave errors. In addition to this, most of them continued operations under the guise that they would look into the errors and better train the chatbot. However, this is far from sufficient, as lives were lost, which is uncompensatable. To make matters worse, most of the cases have not been adjudicated upon, further compounding the effects of the damage that the victims accrued.

---

<sup>176</sup> Montgomery B, 'Mother says AI chatbot led her son to kill himself in lawsuit against its maker' The Guardian, 23 October 2024—<<https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death> > on 23 October 2024.

<sup>177</sup> Montgomery B, 'Mother says AI chatbot led her son to kill himself in lawsuit against its maker' The Guardian, 23 October 2024—<<https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death> > on 23 October 2024.

<sup>178</sup> Montgomery B, 'Mother says AI chatbot led her son to kill himself in lawsuit against its maker' The Guardian, 23 October 2024—<<https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death> > on 23 October 2024.

<sup>179</sup> Atillah I, 'A Belgian man reportedly decided to end his life after having conversations about the future of the planet with an AI chatbot named Eliza' Euronews, 31 March 2023—<<https://www.euronews.com/next/2023/03/31/man-ends-his-life-after-an-ai-chatbot-encouraged-him-to-sacrifice-himself-to-stop-climate-> > 31 March 2023.

<sup>180</sup> Atillah I, 'A Belgian man reportedly decided to end his life after having conversations about the future of the planet with an AI chatbot named Eliza' Euronews, 31 March 2023—<<https://www.euronews.com/next/2023/03/31/man-ends-his-life-after-an-ai-chatbot-encouraged-him-to-sacrifice-himself-to-stop-climate-> > 31 March 2023.

<sup>181</sup> Atillah I, 'A Belgian man reportedly decided to end his life after having conversations about the future of the planet with an AI chatbot named Eliza' Euronews, 31 March 2023—<<https://www.euronews.com/next/2023/03/31/man-ends-his-life-after-an-ai-chatbot-encouraged-him-to-sacrifice-himself-to-stop-climate-> > 31 March 2023.

## **2.4. Conclusion**

The chapter aimed to study how chatbots work and their inherent technology that allows it to function. It further analysed the benefits and harms that they could cause to customers when they are introduced to customer management systems of businesses. Through this, the Chapter has found the need for businesses to balance the risks and benefits that arise as a result of using chatbots. Additionally, in the event that they do use chatbots, they should ensure proper training from its onset and carry out frequent safety and ethical practices to ensure that it provides accurate information to customers.

### 3.0. Tracing the Evolution of Liability under Kenyan Tort Law and Its Implications on Liability for Chatbot Harms

#### 3.1. Introduction

Tort law, a branch of common law, is rooted in customary principles that have been refined and solidified through judicial precedents over time.<sup>182</sup> Tort is an act which, whether intentional or not, is either contrary to law or a specific legal duty or a violation of a right.<sup>183</sup> It is therefore defined as a civil wrong, independent of a contract, for which the remedy is common law action for unliquidated damages.<sup>184</sup> Its primary aim is to uphold principles of justice—corrective, distributive and retributive—by providing remedies for wrongs and ensuring accountability.<sup>185</sup> In Kenya, tort law was introduced through the reception clause in the *Judicature Act*, which allows for the application of common law principles.<sup>186</sup> However, this application is subject to adaptation, as courts are directed to consider Kenya’s unique circumstances and the needs of its people when interpreting and implementing the principles.<sup>187</sup>

There are various types of liability under Kenyan jurisprudence.<sup>188</sup> Like any other common law jurisdiction, the types of liability include; strict liability, negligence, and vicarious liability.<sup>189</sup> These forms of liability all seek to ensure that harms caused to a victim are able to be compensated for, through the identification of the person(s) or object responsible for such harm.<sup>190</sup>

With this in mind, chatbots also cause harm, as identified in 2.4. above. However, there are varying degrees of what proponents think the appropriate means of liability should be. Currently, there are proponents supporting each type of liability and their justifications for such a proposal,

---

<sup>182</sup> Akech M, ‘The common law’s approach to liability and redress its applicability to East Africa’ International Environmental Law Research Centre (IELRC.ORG), 2003, 1-<[https://www.ielrc.org/activities/workshop\\_0309/content/akech.pdf](https://www.ielrc.org/activities/workshop_0309/content/akech.pdf)> on 6 September 2003.

<sup>183</sup> Hussain A, *General Principles and Commercial Law of Kenya*, East African Educational Publishers Ltd, Kenya, 55.

<sup>184</sup> Hussain A, *General principles and commercial law of Kenya*, East African Educational Publishers Ltd, 55.

<sup>185</sup> Akech M, ‘The common law’s approach to liability and redress its applicability to East Africa’ 1.

<sup>186</sup> Section 3(1), *Judicature Act* (CAP 8).

<sup>187</sup> Akech M, ‘The common law’s approach to liability and redress its applicability to East Africa’ 1.

<sup>188</sup> Akech M, ‘The common law’s approach to liability and redress its applicability to East Africa’ 1.

<sup>189</sup> Akech M, ‘The common law’s approach to liability and redress its applicability to East Africa’ 1.

<sup>190</sup> Akech M, ‘The common law’s approach to liability and redress its applicability to East Africa’ 1.

which is what will be discussed in this chapter, to identify the strengths and weaknesses of the suggested forms of liability.

## 3.2. Strict Liability

### 3.2.1. Examining Judicial Precedents and Legal Implications for Strict Liability

Strict liability as stated in the case of *Rylands v Fletcher*, is whereby liability is imposed on the owner of the land for damages caused by a product that escaped from their property to a neighbour's land.<sup>191</sup> In the case of *Kenya Shell Limited v Milkha Kerubo Onkoba*,<sup>192</sup> The Court of Appeal confirmed that the rule in *Rylands v Fletcher* is recognised in our law. The prerequisites of a strict liability claim are: first, that the defendant made a 'non-natural' or 'special' use of his land; second, that the defendant brought onto his land something that was likely to do mischief if it escaped; third, that the substance in question escaped; and fourth, that the Plaintiff's property was damaged because of the escape.<sup>193</sup>

The courts have held that the principle of non-natural use of land under *Rylands v Fletcher* is not a rigid concept but a developing rule that adapts to the continuous changes in modern life.<sup>194</sup> The term "non-natural use" refers to a special use of land that increases the risk to others and goes beyond ordinary use or a use that serves the general benefit of the community.<sup>195</sup>

Just to shed more light, on Kenya's decisions with regards to what a non-natural and natural use of land is, I'll begin by discussing the case of *Kenya Ports Authority v East African Power & Lighting Company Ltd*,<sup>196</sup> where the Court of Appeal found that the bringing of oil into the premises of the plaintiff for the production of electricity was not a non-natural use of land, as the Plaintiff should have anticipated the defendant having oil in their premises, given they were

---

<sup>191</sup> *Duncan Nderitu Ndegwa v Kenya Pipeline Limited & another* (2013) eKLR.

<sup>192</sup> (2010) eKLR.

<sup>193</sup> *David M Ndetei v Orbit Chemical Industries Limited* (2014) eKLR; *Mwita Merengo v Joseph Tunei Marwa & 2 others* (2012) eKLR; *James Ngugi Kariuki v Peter Kamau Kariuki* (2020) eKLR; *Attorney General v Law Society of Kenya & another* (2017) eKLR; *Kenya Ports Authority v East African Power & Lighting Company Limited* (1982) KECA; *Joseph Kimani Gatheca v Gatundu South Water and Sewerage Company* (2018) KEHC 4245 (KLR).

<sup>194</sup> *David M Ndetei v Orbit Chemical Industries Limited* (2014) eKLR.

<sup>195</sup> *Richard v Lothian* (1913); AC 263 (C.A) *Patrick J.O. Otieno v Lake Victoria South Water Services Board* (2020) eKLR.

<sup>196</sup> (1982) eKLR.

licensed to produce electricity.<sup>197</sup> This is despite the plaintiff claiming that the oil, if it did escape, would result in a lot of harmful effects to the sea therefore, the defendant had to ensure it does not escape its premises.<sup>198</sup> Despite its escape, the court found that it was a natural use of the land.<sup>199</sup> Similarly, in the case of *Kenya Wildlife Service v Rift Valley Agricultural Contractors Limited*,<sup>200</sup> wild animals escaped the national park and went to the claimant's property, destroying their crops.<sup>201</sup> Similar to the previous case, the court found that the existence of wildlife on the land did not constitute a non-natural use of the property in question.<sup>202</sup> Furthermore, since the appellant was not the owner of the land which was under the management of the Narok county government, it failed in the other tiers of the test.<sup>203</sup>

On the contrary, in the case of *David M Ndetei v Orbit Chemical Industries Limited*,<sup>204</sup> where the defendant was carrying out chemical trading, the importation and sale of chemicals in bulk, the court found that the defendant was strictly liable despite being given permits to carry out such activities in their premises.<sup>205</sup> For its manufacturing processes the defendant used various chemicals such as bromine, sodium and zinc among others, whose residues after production were hazardous.<sup>206</sup> As stated earlier, despite the defendant arguing that they were allowed to carry out such activities by having been provided with the appropriate licences and permits, the court found that it does not act as a defence against strict liability.<sup>207</sup> Additionally, since the Court found that the effluent escaped to the plaintiff's property resulting in massive pollution of underground water and soil, it found the defendant strictly liable.<sup>208</sup>

The rule in *Rylands v Fletcher* applies exclusively to non-natural uses of land. It does not extend to: objects naturally present on the land; items unlikely to cause harm if they escape; situations

---

<sup>197</sup> *Kenya Ports Authority v East African Power & Lighting Company Ltd* (1982) eKLR.

<sup>198</sup> *Kenya Ports Authority v East African Power & Lighting Company Ltd* (1982) eKLR.

<sup>199</sup> *Kenya Ports Authority v East African Power & Lighting Company Ltd* (1982) eKLR.

<sup>200</sup> (2018) eKLR.

<sup>201</sup> *Kenya Wildlife Service v Rift Valley Agricultural Contractors Limited* (2018) eKLR.

<sup>202</sup> *Kenya Wildlife Service v Rift Valley Agricultural Contractors Limited* (2018) eKLR.

<sup>203</sup> *Kenya Wildlife Service v Rift Valley Agricultural Contractors Limited* (2018) eKLR.

<sup>204</sup> (2014) eKLR.

<sup>205</sup> *David M Ndetei v Orbit Chemical Industries Limited* (2014) eKLR.

<sup>206</sup> *David M Ndetei v Orbit Chemical Industries Limited* (2014) eKLR.

<sup>207</sup> *David M Ndetei v Orbit Chemical Industries Limited* (2014) eKLR.

<sup>208</sup> *David M Ndetei v Orbit Chemical Industries Limited* (2014) eKLR.

where no escape occurs from the land where they were gathered; cases where the escape results from an act of God, a third party's actions, or the claimant's own fault; instances where the escaped object was present with the injured party's consent; and certain circumstances involving statutory authority.<sup>209</sup> Consequently, some of the things that were considered by the court to be non-natural are; felling a tree to another's property;<sup>210</sup> diesel from a leaking tank<sup>211</sup> among others.

With these exceptions in mind, it is noteworthy to draw the differences in determination of the *Ndetei v Orbit* case and the *KPA v East African Lighting Company* since both parties that were sued had licences to operate their line of business. In the *KWS v Rift Valley* case the court of appeal did not hold the appellant liable as the sea is considered terra nullus thus was not under the ownership of the applicant, therefore they failed to establish the test in *Rylands v Fletcher*. A similar reason was given for the lack of determination of strict liability in *KWS v Rift Valley* case, whereby the land holding the wild animals was not owned by KWS but the Narok county government. In both these instances, the reason for its failure is due to the land not being owned by either of the parties in the case. This is contrary to *Ndetei v Orbit*, whereby despite the defendant having licence to operate like KPA, the court found them liable since the effluents escaped and damaged the plaintiffs property.

### 3.2.2. Case for Strict Liability for Chatbot Harms

The first argument by proponents of strict liability is the fact that since strict liability is an evolving concept as noted in *David M Ndetei v Orbit Chemical Industries Limited (2014) eKLR*, it can be able to cover harms caused by AI entities.<sup>212</sup> In the Kenyan context, this evolving nature described has not been discussed in the current jurisprudence of cases, so it would be interesting to see how the court interprets it and finds such things to be. However, it is a fair claim to make considering presentLY, assets are owned digitally and can still cause harm to an individual or their property, especially when it comes to AI.

---

<sup>209</sup> *Duncan Nderitu Ndegwa v Kenya Pipeline Company Limited & another* (2013) eKLR.

<sup>210</sup> *Mwita Merengo v Joseph Tunei Marwa & 2 others* (2012) eKLR.

<sup>211</sup> *Kenya Shell Limited v Milkha Kerubo Onkoba* (2010) eKLR.

<sup>212</sup> Buiten M, Streel A and Peitz M, 'The law and economics of AI liability' 48(105794) *Computer Law & Security Review*, 2023, 3-4; Wendehorst A, 'Liability for Artificial Intelligence' in Wendehorst A (ed) *The Cambridge Handbook of Responsible Artificial Intelligence*, Cambridge University Press, 2022, 194.

In my view, an evolution of this test would cover things such as autonomous digital systems, AI-driven software and other intangible but high-risk technological assets such as cryptocurrency or internet of things devices among others. Unlike the traditional understanding of ‘non-natural use of land’ in *Rylands v Fletcher*, where liability was imposed for the escape of hazardous substances, modern risk arises from intangible yet impactful digital entities. Given the rapid integration of AI in critical sectors, ranging from finance to healthcare, the strict liability doctrine could evolve to include AI-generated misinformation, automated defamation, privacy breaches, and economic loss caused by autonomous digital decisions. For instance, an AI chatbot integrated into a financial advisory platform such as ICEA Lion’s Leo, could erroneously provide misleading investment advice leading to significant financial harm. Due to the stated circumstances, an evolution could extend the principles of *Rylands v Fletcher* to hold AI system owners and operators strictly liable for harm caused by their digital assets, even if they do not fit within the conventional definition of ‘escape’ due to their inherently risky effects.

The second argument by proponents is the fact that the plaintiff does not have to prove causation or fault, allowing for fairness since they only have to prove that the defendant allowed it to escape from its premises and there was harm caused.<sup>213</sup> It is worth noting that the incapacity to determine fault or causation is due to the opacity and unpredictability of AI systems generally, thus offering a way out for plaintiffs.<sup>214</sup> Since the use of strict liability eliminates the need to prove fault, the claimant is only left to determine causation.<sup>215</sup>

### **3.2.3. Weaknesses and Shortcomings of using a Strict Liability Framework**

Holding companies liable solely based on causation, which is a very low standard to prove, is unfair.<sup>216</sup> Unlike a fault-based system where an individual has to determine other surrounding circumstances, strict liability does not allow for that other than through the exceptions stated in 3.2.1.

---

<sup>213</sup> Buiten M, Streel A and Peitz M, ‘The law and economics of AI liability’ 3-4; Wendehorst A, ‘Liability for Artificial Intelligence’ 194.

<sup>214</sup> Buiten M, Streel A and Peitz M, ‘The law and economics of AI liability’ 3-4; Wendehorst A, ‘Liability for Artificial Intelligence’ 194.

<sup>215</sup> Wendehorst A, ‘Liability for Artificial Intelligence’ 194.

<sup>216</sup> Wendehorst A, ‘Liability for Artificial Intelligence’ 195-196.

The situation is further nuanced due to causation being unclear.<sup>217</sup> When explaining the technology of chatbots, deep learning was one of the technologies used to allow for its operation, which in and of itself is a complex, multi-layered environment where drawing a decision tree might prove difficult. Thus in such a case, liability may be assigned even when harm results from external factors, such as user manipulation, third party misuse, or unexpected data inputs.<sup>218</sup>

### 3.3. Liability based on Negligence

#### 3.3.1. Examining Judicial Precedents and Legal Implications for Liability based on Negligence

The principle in *Donoghue v Stevenson* has been used to expound and determine situations where a person was owed a duty of care. It basically espouses that where a person owes a duty of care to another, where their actions could directly affect others, they ought to reasonably foresee that certain actions could affect them.<sup>219</sup> The case has enshrined the principle of good neighbourliness, which provides that one should take reasonable care to avoid acts or omissions which you can reasonably foresee would likely injure your neighbour.<sup>220</sup> A neighbour in law being, the person who is so closely and directly affected by the person's acts or omissions that they ought to have reasonably contemplated.<sup>221</sup>

The ingredients for determining negligence requires that; there is an existence of a legal duty of care owed to the plaintiff by the defendant; breach of that duty of care; damage to the plaintiff which results from that breach; and foreseeability of the act of the defendant that such conduct would have inflicted upon the plaintiff the particular type of damage which he complains of.<sup>222</sup> The issue of remoteness of damage or causal link was addressed in the case of *Cotecna Inspection S.A v Hems Group Trading Company Limited*,<sup>223</sup> where the Court stated that there

---

<sup>217</sup> Buiten M, Strel A and Peitz M, 'The law and economics of AI liability' 5.

<sup>218</sup> Buiten M, Strel A and Peitz M, 'The law and economics of AI liability' 5.

<sup>219</sup> *Stephen Kanjabi Wariari v Dennis Mutwiri Muriuki & another* (2022) eKLR; *Anastassios Thomos v Occidental Insurance Company Limited* (2017) eKLR.

<sup>220</sup> *Stephen Kanjabi Wariari v Dennis Mutwiri Muriuki & another* (2022) eKLR; *Anastassios Thomos v Occidental Insurance Company Limited* (2017) eKLR.

<sup>221</sup> *Stephen Kanjabi Wariari v Dennis Mutwiri Muriuki & another* (2022) eKLR.

<sup>222</sup> *JMA (Suing through BOA as next friend) & another v Registered Trustees of the Sisters of Mercy (Kenya) t/a Matter Misericordiae Hospital* (2023) eKLR; *Kenya Wildlife Service v Rift Valley Agricultural Contractors Limited* (2018) eKLR; *Anastassios Thomos v Occidental Insurance Company Limited* (2017) eKLR; *Mombasa Maize Millers & another v Elius Kinyua Gicovi* (2021) eKLR.

<sup>223</sup> (2007) eKLR.

must be a link between the action complained of and the loss incurred. Furthermore, in *Elijah Ole Kool v George Ikonya Thuo*,<sup>224</sup> Justice Alnashir Visram stated that a defendant will be liable for negligence only if their act or omission is the sole effective cause of the plaintiff's injury or is sufficiently linked to it as a significant contributing factor.<sup>225</sup>

An instance where the Court found that there was negligence is the case of *Terra Fleur Limited v Kenya Cuttings Limited*<sup>226</sup> where the plaintiff's and defendant's farms were adjacent to each other, and both cultivated flowers. The defendant used methan sodium in its production whose remnants after fumigation, were released into a nearby river dam which the applicant was using to irrigate their flowers.<sup>227</sup> Without knowing about the effluent in the river, the applicant used the water for irrigation which resulted in damage to their flowers consequently suffering financial loss.<sup>228</sup> The court ruled that the defendant was responsible for handling methan sodium in a way that prevented it from reaching the plaintiff's farm, where it could potentially cause harm.<sup>229</sup> Since the defendant was aware of the risks associated with methan sodium, the court determined that it had violated its duty of care to the plaintiff by permitting the remnants to flow into the river.<sup>230</sup> In negligence cases, a claim cannot be established unless the plaintiff has suffered actual physical harm to their person or property.<sup>231</sup>

There are also specific types of negligence claims, with the most common one being medical negligence. In the case of *Ricarda Njoki Wahome v Attorney General & 2 others*,<sup>232</sup> the court stated that A doctor may be found guilty of medical negligence if they fail to meet the standard expected of a reasonable medical professional in the same field, rather than simply making an error in judgment on a matter of opinion.<sup>233</sup> In this case, the breach of duty corresponds to the

---

<sup>224</sup> (2001) eKLR.

<sup>225</sup> *Elijah Ole Kool v George Ikonya Thuo* (2001) eKLR.

<sup>226</sup> (2021) eKLR.

<sup>227</sup> *Terra Fleur Limited v Kenya Cuttings Limited* (2021) eKLR.

<sup>228</sup> *Terra Fleur Limited v Kenya Cuttings Limited* (2021) eKLR.

<sup>229</sup> *Terra Fleur Limited v Kenya Cuttings Limited* (2021) eKLR.

<sup>230</sup> *Terra Fleur Limited v Kenya Cuttings Limited* (2021) eKLR.

<sup>231</sup> *Kenya Ports Authority v East African Power & Lighting Company Ltd* (1982) eKLR.

<sup>232</sup> (2015) eKLR.

<sup>233</sup> *John Gachanja Mundia v Francis Muriira & Another* (2017) eKLR; *Odero v Aga Khan Hospital Kisumu* (2024) eKLR.

standard of a reasonable and competent healthcare professional.<sup>234</sup> Therefore, in a medical negligence case, the plaintiff must demonstrate that there was a standard and customary practice, the healthcare provider failed to follow that practice, and instead chose an approach that no competent or reasonably skilled professional would have taken.<sup>235</sup> The logic behind having a higher threshold for medics is to ensure that doctors are focused on deciding the best course of treatment as per their assessment rather than being concerned about possible prosecution or harassment.<sup>236</sup>

Another special type of negligence is manufacturer negligence. Under Section 16(b) of the *Sale of Goods Act*, there is an implied condition that goods must be of merchantable quality unless the buyer has inspected them for obvious defects.<sup>237</sup> A manufacturer bears strict liability if they introduce a product into the market, knowing it will be used without prior inspection, and the product has a defect that causes harm to individuals.<sup>238</sup> In the case of *Coca Cola Company Ltd & another v Josephat Okello Oduori*, the Court found the appellant negligent for having foreign substances in their product which resulted in the respondent's illness after consumption.<sup>239</sup> Despite the court not finding scientific evidence of the foreign substance's possible harm, they found that the anguish or anxiety that affected the respondent was enough to find the Appellant liable.<sup>240</sup> Even when considering manufacturer's negligence, the principles in *Donoghue v Stevenson* applies as stated in the case of *Kenya Breweries Ltd v Godfrey Odoyo*.<sup>241</sup> In that case, the court found the appellant negligent since it owed a duty to the consumers who drink Tusker Lager.<sup>242</sup> The duty was breached because the drink had certain impurities, which even under examination by KEBS was found to be unfit for human consumption.<sup>243</sup> Additionally, as a result of the impurities, it led to the consumer ailing.<sup>244</sup> The plaintiff is typically responsible for

---

<sup>234</sup> *John Gachanja Mundia v Francis Muriira & Another* (2017) eKLR; *Odero v Aga Khan Hospital Kisumu* (2024) eKLR.

<sup>235</sup> *LWW (Suing as the Administrator of the estate of BMN) deceased v Charles Githinji* (2019) eKLR; *Odero v Aga Khan Hospital Kisumu* (2024) eKLR; *JMA (Suing through BOA as next friend) & another v Registered Trustees of the Sisters of Mercy (Kenya) t/a Matter Misericordiae Hospital* (2023) eKLR.

<sup>236</sup> *Odero v Aga Khan Hospital Kisumu* (2024) eKLR.

<sup>237</sup> *James Watenga Kamau v CMC Motors Group Limited* (2020) eKLR.

<sup>238</sup> *Fred Ben Okoth v Equity Bottlers Limited* (2015) eKLR.

<sup>239</sup> *Coca Cola Company Ltd & another v Josephat Okello Oduori* (2011) eKLR.

<sup>240</sup> *Coca Cola Company Ltd & another v Josephat Okello Oduori* (2011) eKLR.

<sup>241</sup> (2010) eKLR.

<sup>242</sup> *Kenya Breweries Ltd v Godfrey Odoyo* (2010) eKLR.

<sup>243</sup> *Kenya Breweries Ltd v Godfrey Odoyo* (2010) eKLR.

<sup>244</sup> *Kenya Breweries Ltd v Godfrey Odoyo* (2010) eKLR.

proving their case on a balance of probabilities, relying on the principles established in *Donoghue v Stevenson*.<sup>245</sup> However, the defendant retains the evidential burden of clarifying certain facts that are specifically within their knowledge. In *Kenya Breweries Ltd v Godfrey*, this evidentiary burden required them to show why there was presence of foreign particles in their bottled beer.<sup>246</sup>

### 3.3.2. Case for Liability based on Negligence for Chatbot Harms

Courts and regulators have historically altered the burden of proof to adapt to new technologies or scenarios, especially in situations where it may be difficult for the victim to prove each of the conditions.<sup>247</sup> This is done in contemplation of how easy or hard it is for the victim to access evidence, the degree of the possible damage and the dangerousness of the activity.<sup>248</sup> The application of medical negligence in medical malpractices is one such example, whereby the burden of proof is borne by the medical practitioner and not the patient.<sup>249</sup> In cases of medical negligence, a higher standard applies, as the assessment is based not on the reasonable person test but on the competence of an ordinarily skilled doctor or specialist in that field.<sup>250</sup> Similarly, in the case of manufacturer's negligence, the evidentiary burden is on them to show or disprove the existence of a defect that is present due to their actions.<sup>251</sup> Thus for chatbot-related harms, a negligence-based approach would ensure that liability is not imposed automatically but rather depends on whether the chatbot operator or developer failed to meet a reasonable standard of care. Courts could assess whether the company:

- i) Properly tested and monitored the chatbot for foreseeable risks,
- ii) Implemented safeguards to prevent harmful outcomes, and
- iii) Acted responsibly when notified of potential harms.<sup>252</sup>

---

<sup>245</sup> *Kenya Breweries Ltd v Godfrey Odoyo* (2010) eKLR.

<sup>246</sup> (2010) eKLR.

<sup>247</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 21.

<sup>248</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 21.

<sup>249</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 22; Martin-casals M, 'The Development of Liability in Relation to Technological Change' Cambridge University Press <[https://assets.cambridge.org/97811074/75809/frontmatter/9781107475809\\_frontmatter.pdf](https://assets.cambridge.org/97811074/75809/frontmatter/9781107475809_frontmatter.pdf)> 2010.

<sup>250</sup> *John Gachanja Mundia v Francis Muriira & Another* (2017) eKLR.

<sup>251</sup> *Kenya Breweries Ltd v Godfrey Odoyo* (2010) eKLR.

<sup>252</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 23.

Additionally, by requiring proof of fault rather than imposing blanket liability, a negligence-based framework strikes a balance between protecting victims and fostering responsible AI development.<sup>253</sup> It ensures that only those who fail to exercise due diligence are held accountable, rather than discouraging innovation by making liability absolute.<sup>254</sup> Thus in the situation of chatbots, it is equally important to consider whether or not manufacturers of chatbots can be held liable in the event that an error can be traced back to them. By applying the *Donoghue v Stevenson* test, the answer would be in the affirmative, simply because manufacturers or producers do have a duty of care to ensure that the products are of quality standards.

### **3.3.3. Weaknesses and Shortcomings of using Liability based on Negligence Framework**

A key challenge in applying negligence-based liability to chatbot hams, is determining what measure or standard should be used to assess reasonable foreseeability.<sup>255</sup> This difficulty arises primarily due to the machine-learning capabilities of AI systems, where harmful effects may emerge unpredictably and be untraceable to specific design flaws or actions.<sup>256</sup>

Foreseeability is crucial in determining duty of care, tortious conduct, and proximate cause.<sup>257</sup> As discussed in the cases highlighted in 3.3.1, courts assess whether a defendant owed a duty of care to the plaintiff and whether that duty should have been adjusted based on foreseeable risks.<sup>258</sup> However, in contrast, plaintiffs can only seek remedies if the harm falls within the scope of the defendant's responsibility.<sup>259</sup> If foreseeability cannot be established, proving negligence-liability becomes significantly more difficult.<sup>260</sup>

---

<sup>253</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 23.

<sup>254</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 23.

<sup>255</sup> Lewis C, 'The need for a legal framework to regulate the use of artificial intelligence' 304.

<sup>256</sup> Huberman P, 'Tort law, corrective justice and the problem of autonomous-machine-caused harm' 34(1) *Canadian Journal of Law & Jurisprudence*, 2020, 18.

<sup>257</sup> Huberman P, 'Tort law, corrective justice and the problem of autonomous-machine-caused harm' 18.

<sup>258</sup> Huberman P, 'Tort law, corrective justice and the problem of autonomous-machine-caused harm' 18.

<sup>259</sup> Huberman P, 'Tort law, corrective justice and the problem of autonomous-machine-caused harm' 18.

<sup>260</sup> Huberman P, 'Tort law, corrective justice and the problem of autonomous-machine-caused harm' 18.

The determination of the causal link makes it further challenging to impose a negligence-liability.<sup>261</sup> In most cases, the burden of proving causation lies with the claimant. Therefore, the claimant must present evidence demonstrating that, on the balance of probabilities, the defendant's wrongful conduct was the actual cause of the damage being claimed.<sup>262</sup> This is because chatbot behaviour is complex due to factors such as; deep learning capacity of AI systems, which may generate unexpected responses.<sup>263</sup> User misuse, where individuals intentionally provoke harmful chatbot output such as in the case of Microsoft's Tay explained under 2.4. Or even unforeseeable externalities, such as data poisoning attacks or third party interferences.

Another challenge is the risk of leaving victims uncompensated.<sup>264</sup> Since negligence requires proof of fault, victims may go uncompensated if they are unable to pinpoint a specific act or omission that led to the harm. In extreme scenarios, they will even be hesitant to challenge the claims in court due to the difficulty in establishing fault.

In the alternative, with regards to manufacturer's negligence, for chatbots, like any other AI system, there are various intermediaries in the chain of production and use.<sup>265</sup> From the individual who codes, to the person who creates the software, to the eventual deployer of the product.<sup>266</sup> Due to this issue, which is compounded by the fact that it may be difficult to establish where exactly the problem arises, it may be difficult to establish manufacturer's negligence especially when it comes to establishing how exactly the duty was breached.<sup>267</sup> Thus it becomes incumbent upon the deployer purchasing the chatbot for use, to carry out multiple test runs to establish that the chatbot works properly to avoid the issue of trying to establish who exactly in the chain of production was responsible for the error that occurred. The rationale for this obligation will further be expounded in 3.4. below.

---

<sup>261</sup> Huberman P, 'Tort law, corrective justice and the problem of autonomous-machine-caused harm' 18.

<sup>262</sup> *Fred Ben Okoth v Equator Bottlers Limited* (2015) eKLR.

<sup>263</sup> Lewis C, 'The need for a legal framework to regulate the use of artificial intelligence' 304.

<sup>264</sup> Lewis C, 'The need for a legal framework to regulate the use of artificial intelligence' 305.

<sup>265</sup> Lewis C, 'The need for a legal framework to regulate the use of artificial intelligence' 305.

<sup>266</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 21-23.

<sup>267</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 21-23.

Additionally, the standard for manufacturer negligence is really low as discussed in *Coca Cola Company Ltd & another v Josephat Okello Oduori* in 3.3.1 above. The fact that the court did not need to carry out scientific evidence to determine whether or not the effects of the substance actually arose from the consumption of the drink brings into question whether this would be an appropriate standard for chatbots. This is especially because, if a drink which does not need much tests to determine whether or not it is fit for human consumption is not tested appropriately to determine the cause of the harm, what of chatbots and its different complexities which can be difficult to determine? Consequently, the use of manufacturer's negligence would result in holding manufacturer's erroneously liable for things that may not be their fault.

### 3.4. Vicarious Liability

#### 3.4.1. Examining Judicial Precedents and Legal Implications for Vicarious Liability

Vicarious liability holds a party accountable for the actions of another.<sup>268</sup> It is founded on the principle of *respondeat superior*, which holds that an employer is accountable for any wrongful acts committed by their employee or agent during the course of their duties, even without the employer's direct instruction or knowledge.<sup>269</sup> The rationale is that, ordinarily, an employee is of meagre means thus it would not make sense to demand damages from them.<sup>270</sup> However, since an employer has or is assumed to have the funds, they should be the ones held liable since they have the capacity to compensate for the damages being sought.<sup>271</sup> Additionally, due to the fact that the employer has put up an enterprise in the community, they should also be held liable for harms caused to members of that community.<sup>272</sup> Therefore, deterrence to this effect is also a consideration whereby, by holding the employer or principal liable, it will encourage such person(s) to take steps to reduce the risk of harm in the future.<sup>273</sup> Consequently, where there is the existence of an employer-employee or principal-agent relationship, the former is held liable for the actions of the latter.<sup>274</sup>

---

<sup>268</sup> *John Nderi Wamugi v Ruhesh Okumu Otiangala* (2018) eKLR.

<sup>269</sup> *John Nderi Wamugi v Ruhesh Okumu Otiangala* (2018) eKLR.

<sup>270</sup> *John Nderi Wamugi v Ruhesh Okumu Otiangala* (2018) eKLR.

<sup>271</sup> *John Nderi Wamugi v Ruhesh Okumu Otiangala* (2018) eKLR.

<sup>272</sup> *Beatrice William Muthoka & another (Both suing as Legal Representatives of the Estate of the Late William Muthoka Yumbia (Deceased) v Agility Logistics Limited* (2020) eKLR.

<sup>273</sup> *Beatrice William Muthoka & another (Both suing as Legal Representatives of the Estate of the Late William Muthoka Yumbia (Deceased) v Agility Logistics Limited* (2020) eKLR.

<sup>274</sup> *Equator Distributors v Joel Muriu & 3 others* (2018) eKLR.

The test from decided cases is to prove that the employee or agent was conducting business in their ordinary course of conduct and not for their own benefit or a frolic of their own.<sup>275</sup> Drawing from this, the criterion for establishing whether a relationship between parties qualifies as that of employer and employee for the purposes of vicarious liability is:

- i) There must be a fundamental mutual obligation between the parties, where the employee is prepared to work, and the employer agrees to compensate them for the work performed.
- ii) The employee must be under the direct supervision or control of the employer.
- iii) The relationship remains unaffected even if the employer hires an agent, as long as they retain the authority to oversee and dismiss the agent.<sup>276</sup>

In addition to the test espoused above, the *Salmon* test has been used by courts to determine what suffices as a ‘course of employment’ of an employee, stipulating that an employer is not liable for a wrongful act committed by their employee unless it occurs within the scope of their employment.<sup>277</sup> It is deemed to be done by the master if either, a wrongful act is authorised by the master, or a wrongful and unauthorised way of doing some act is authorised by the master.<sup>278</sup> The primary function of the course of employment requirement is to ensure that an employee's tort is sufficiently linked to the employer's enterprise, so as to justify the imposition of liability on the employer.<sup>279</sup>

An employer's duty of care is generally not absolute and does not apply in situations where the employee acts negligently, carelessly, or recklessly.<sup>280</sup> Therefore, an employee is expected to manage their work responsibly and has an implied duty to exercise due care to prevent exposing themselves or their employer to foreseeable risks.<sup>281</sup>

---

<sup>275</sup> *Beatrice William Muthoka & another (Both suing as Legal Representatives of the Estate of the Late William Muthoka Yumbia (Deceased) v Agility Logistics Limited* (2020) eKLR; *Tabitha Nduhi Kinyua v Francis Mutua Mbuvi & another* (2014) eKLR.

<sup>276</sup> *Duncan Nderitu Ndegwa v Kenya Pipeline Company Limited & another* (2013) eKLR.

<sup>277</sup> *Teachers Service Commission v WJ & 5 others* (2020) eKLR.

<sup>278</sup> *Teachers Service Commission v WJ & 5 others* (2020) eKLR.

<sup>279</sup> *Teachers Service Commission v WJ & 5 others* (2020) eKLR.

<sup>280</sup> *James Finlaly (K) Ltd v Bernard Kipsang Koechi* (2021) eKLR.

<sup>281</sup> *James Finlaly (K) Ltd v Bernard Kipsang Koechi* (2021) eKLR.

### 3.4.2. Case for Vicarious Liability for Chatbot Harms

Proponents of vicarious liability analyse AI liability through a principal-agent lense.<sup>282</sup> In this context the AI entity, the chatbot, is the agent while the operator of the chatbot is the principal.<sup>283</sup> Vicarious liability having following the respondeat superior maxim, holds the principals liable for damages caused by their AI agent.<sup>284</sup> The reasoning is firstly, due to the fact that the AI entity cannot compensate victims due to the lack of capacity, secondly, that AI entities lack personhood and thus cannot be sued; and lastly, that the principal is best able to control and monitor the chatbot's activities.<sup>285</sup>

Another reasoning given for supporting this type of liability, is that generally, even in employee-employer or agent-principal relationships, employers and principals are still answerable for their employees' autonomous acts even though they are neither immediately influenced nor participated in the wrongful behaviour that occasioned the loss.<sup>286</sup> As AI entities gradually become autonomous, they continue to be more like human agents or employees, due to decreased intervention by their employer since they have the capacity to be social and interact with humans in a language that they understand, and are independent from humans to a certain extent due to their machine learning capabilities.<sup>287</sup> Furthermore, behind the imposition of such a relationship is due to the fact that ultimately, chatbots have replaced humans who used to be the ones carrying out the communication, thus why not transplant such a relationship and alter it based on AI entities special characteristics?<sup>288</sup>

According to Diamantis M, by recovering the principles behind respondeat superior, more generalised versions of its two elements come into view that could flexibly apply to human and

---

<sup>282</sup> Lior A, 'AI entities as AI agents: Artificial intelligence liability and the AI respondeat superior analogy' 46(5) *Mitchell Hamline School of Law*, 2020, 1071.

<sup>283</sup> Lior A, AI entities as AI agents: Artificial intelligence liability and the AI respondeat superior analogy' 1071.

<sup>284</sup> Lior A, AI entities as AI agents: Artificial intelligence liability and the AI respondeat superior analogy' 1071.

<sup>285</sup> Lior A, 'AI entities as AI agents: Artificial intelligence liability and the AI respondeat superior analogy' 1071; Wein L, 'The responsibility of intelligent artifacts: toward an automation jurisprudence' 6(Fall Issue), *Harvard Journal of Law & Technology*, 1992, 110-111.

<sup>286</sup> Wein L, 'The responsibility of intelligent artifacts: toward an automation jurisprudence' 110-111.

<sup>287</sup> Diamantis M, 'Employed algorithms: A labour model of corporate liability for AI' 1075.

<sup>288</sup> Diamantis M, 'Employed algorithms: A labour model of corporate liability for AI' 1075; Wein L, 'The responsibility of intelligent artifacts: toward an automation jurisprudence', 106.

digital labour alike.<sup>289</sup> The first is to recognise that corporations can employ algorithms.<sup>290</sup> The generalised version of respondeat superior being “the benefit-control test” that is derived from deeper principles that are at play in the intent to benefit and scope of employment elements.<sup>291</sup> These factors are designed to ensure that a corporation is only liable for an employee violation if the corporation expected to benefit from and controlled the employee’s conduct at the time of the violation. Therefore, with this in mind, a corporation employs an algorithm if it exercises beneficial control over it.<sup>292</sup>

As stated in the case of *Beatrice William Muthoka & another v Agility Logistics Limited*, in 3.4.1, deterrence is one of the main aims of vicarious liability. Therefore, even though holding companies or employers liable can never guarantee employees will obey the law, appropriate training and monitoring can reduce the probability of a violation.<sup>293</sup> Similarly, machine learning algorithms such as chatbots, present a similar challenge, where there are no guarantees that they will never hurt someone.<sup>294</sup> However, it incentivises better compliance for companies or operators to ensure the proper functioning of their chatbot by diversifying the body of engineers writing algorithms, more extensive pre-rollout testing, carrying out frequent audits, effective monitoring and continuous software updates to address problems as they arise.<sup>295</sup> As stated earlier, none of the steps highlighted above will guarantee that a chatbot will not make a mistake, but they will reduce the probability that it will go astray. Consequently, adopting a vicarious liability would ensure that corporations take pre-emptive care with their algorithms so as to avoid greater expense of liability down the road.<sup>296</sup>

### **3.4.3. Weaknesses and Shortcomings of using a Vicarious Liability Framework**

Looking into the amount of control a company has over the operations of a chatbot may be difficult.<sup>297</sup> This is because there are various powers at play that may be held by different parties

---

<sup>289</sup> Diamantis M, ‘Employed algorithms: A labour model of corporate liability for AI’ 1075-1076.

<sup>290</sup> Diamantis M, ‘Employed algorithms: A labour model of corporate liability for AI’ 1075-1076.

<sup>291</sup> Diamantis M, ‘Employed algorithms: A labour model of corporate liability for AI’ 1075-1076.

<sup>292</sup> Diamantis M, ‘Employed algorithms: A labour model of corporate liability for AI’ 848.

<sup>293</sup> Diamantis M, ‘Employed algorithms: A labour model of corporate liability for AI’ 852-853.

<sup>294</sup> Diamantis M, ‘Employed algorithms: A labour model of corporate liability for AI’ 852-853.

<sup>295</sup> Diamantis M, ‘Employed algorithms: A labour model of corporate liability for AI’ 852-853.

<sup>296</sup> Diamantis M, ‘Employed algorithms: A labour model of corporate liability for AI’ 852-853.

<sup>297</sup> Diamantis M, ‘Employed algorithms: A labour model of corporate liability for AI’ 847-848.

at varying degrees.<sup>298</sup> Some of the relevant powers in AI systems generally include; the power to design the algorithm, terminate its operation, modify it, monitor it and override it. Diamantis stated that, non of the powers as a stand alone objectives or capacity are necessarily determinative of corporation control over an algorithm, but the more powers a corporation has, the more control it has.<sup>299</sup> If this is the rationale used, then some companies may still escape accountability by delegating tasks to other parties to avoid accountability. This has been the pattern in various other fields of operations companies carry out, such as profit shifting or even having global value chains to ensure decrease in spending for production.

### **3.5. Conclusion**

This chapter has analysed how courts have interpreted and applied various liability frameworks, considering the reasoning behind these judicial decisions. By integrating these rulings with the arguments put forth by proponents of different liability models, the chapter has identified key weaknesses and limitations in their practical application. This analysis has further allowed for a critical evaluation of how these frameworks could be adapted or applied within the Kenyan legal context, taking into account its unique legal landscape and regulatory considerations.

---

<sup>298</sup> Diamantis M, 'Employed algorithms: A labour model of corporate liability for AI' 847-848.

<sup>299</sup> Diamantis M, 'Employed algorithms: A labour model of corporate liability for AI' 847-848.

## **4.0. The Case for Vicarious Liability in Addressing Chatbot Harms in Kenya**

### **4.1. Introduction**

The previous chapter examined how Kenyan courts have ruled on various liability frameworks and analysed the practical implications of their application, including the specific weaknesses they present in addressing chatbot-related harms. Building on this foundation, this chapter will propose the most suitable liability framework, providing a reasoned justification for its adoption despite the challenges identified in the preceding analysis.

### **4.2. Evaluating Chatbot Agency and Vicarious Liability**

#### **4.2.1. Agency**

Proponents of vicarious liability have recommended the adoption of a harmonised regime that ensures a principal employing AI for a sophisticated task, bears the same liability as a principal employing a human auxiliary.<sup>300</sup> This approach addresses the difficulty victims face in proving fault.<sup>301</sup> Additionally, legislators and courts would not need to have information on the optimal level of precaution in designing and deploying AI-based systems.<sup>302</sup> For the use of agency law, most supporters employ an analogy that treats AI entities as agents, as if they were servants of their owners or operators.<sup>303</sup> It has also been illustrated through the application of agency law to corporations, reinforcing the argument that chatbots can similarly be brought within the fold of agency-based liability.<sup>304</sup>

An agency relationship can either be implied or expressed.<sup>305</sup> To establish such a relationship, it must be demonstrated that the agent performed tasks based on the principal's implied or explicit instructions. An implied agency relationship, refers to actions that are generally assumed to be one of employer-employee relationships due to the conduct of the parties.<sup>306</sup> Conversely, an

---

<sup>300</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 38.

<sup>301</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 38.

<sup>302</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 38.

<sup>303</sup> Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' 38.

<sup>304</sup> Lior A, 'AI entities as AI agents: Artificial intelligence liability and the AI respondeat superior analogy' 1071.

<sup>305</sup> Lior A, 'AI entities as AI agents: Artificial intelligence liability and the AI respondeat superior analogy' 1071.

<sup>306</sup> Lior A, 'AI entities as AI agents: Artificial intelligence liability and the AI respondeat superior analogy' 1071.

express agency relationship exists where there is a valid formal agreement between the two parties, with both agreeing to carry out certain objectives.<sup>307</sup>

Applying this to chatbots, it is evident that an express agency relationship cannot exist due to the chatbot's lack of legal personhood and capacity to provide consent. However, an implied agency framework is particularly relevant because chatbots operate strictly within the parameters set by their deployers. They are trained through KBs and programmed to interact with users based on predetermined instructions, effectively acting as intermediaries on behalf of their developers or companies. This functional delegation underscores the legitimacy of classifying chatbots within an implied agency framework, making their deployers liable for their actions.

#### **4.2.2. Personhood**

Since we have already established the existence of an implied agency relationship between chatbots and their deployers in 4.2.1, the next consideration is whether chatbots can be recognised as agents with independent juridical powers. Traditionally, agents possess legal capacity and can be held accountable for their actions, especially when they act outside the scope of their authority.<sup>308</sup> Even for non-human entities like companies, they have the capacity to be sued.<sup>309</sup> However, if an act was committed by a director, the corporate veil may be opened to look into their individual conduct.<sup>310</sup>

Given that chatbots lack human cognition and legal personhood, they cannot be held liable in the conventional sense.<sup>311</sup> However, this does not negate the necessity of a liability framework that accounts for autonomous functions. In this regard, I claim that chatbots should be granted quasi-judicial recognition under a framework of diminished responsibility, ensuring that liability ultimately rests with their deployers. Under this model, deployers may argue in their defense that chatbot harms were the unintended consequences of machine learning processes rather than

---

<sup>307</sup> Lior A, 'AI entities as AI agents: Artificial intelligence liability and the AI respondeat superior analogy' 1071.

<sup>308</sup> Shrestha S, 'Nature, Nurture, or Neither?: Liability For Automated and Autonomous Artificial Intelligence Torts Based on Human Design and Influences' 6-8.

<sup>309</sup> Shrestha S, 'Nature, nurture, or neither?: Liability for automated and autonomous artificial intelligence torts based on human design and influences' 6-8.

<sup>310</sup> Shrestha S, 'Nature, nurture, or neither?: Liability for automated and autonomous artificial intelligence torts based on human design and influences' 6-8.

<sup>311</sup> Shrestha S, 'Nature, nurture, or neither?: Liability for automated and autonomous artificial intelligence torts based on human design and influences' 6-8.

deliberate misconduct. However, such defenses should not absolve them of responsibility but rather inform the degree of liability assigned in each case and allow for them to show the precautionary steps they undertook to ensure its efficient functioning.

### **4.3. Justifying Vicarious Liability for Chatbot Harms in the Kenyan Context**

#### **4.3.1. Access to Justice and Consumer Protection**

The *Constitution of Kenya (2010)* (CoK), enshrines access to justice under Article 48 and guarantees fair administrative action under Article 47. These constitutional principles underscore the necessity of ensuring that victims of chatbot-related harms can efficiently seek remedies and are adequately protected. To reinforce consumer protection and enhance access to remedies, Part IX of the *Consumer Protection Act* provides for various legal avenues, including litigation, effective notice, cancellation and arbitration.<sup>312</sup> However, in the absence of a clear liability framework, victims face significant barriers in litigation, particularly in attributing fault to an AI system that lacks legal personhood.

Regarding consumer protection, Article 46 of the CoK guarantees that consumers have the right to; quality goods and services, necessary information, protection of their health, safety and economic interests and compensation for loss or injury arising from defects of the good or service. When a chatbot provides misleading or inaccurate information, it directly infringes upon these rights, necessitating compensation for the resulting loss or injury. Additionally, the *Consumer Protection Act*, mandates that products should be of merchantable quality,<sup>313</sup> imposing an obligation on chatbot deployers to ensure their systems function properly and deliver reliable services. Furthermore, Article 12 of the Act prohibits false representation, which, in the context of chatbots, extends to the dissemination of false or misleading information to the user.<sup>314</sup> This provision effectively establishes a statutory basis for holding companies accountable for chatbot-related misinformation.

Consequently, vicarious liability provides a practical solution by shifting responsibility to the chatbot deployers while ensuring that affected users have a clear legal avenue for redress. Courts will also benefit from a structured framework that enables them to adjudicate chatbot-related

---

<sup>312</sup> (CAP 501 of 2012).

<sup>313</sup> S5(1), *Consumer Protection Act* (CAP 501 of 2012).

<sup>314</sup> *Consumer Protection Act* (CAP 501 of 2012).

disputes based on establishing legal doctrines rather than navigating the uncertainty of fault-based AI liability. Establishing vicarious liability in this context is crucial to safeguarding consumer rights and ensuring access to effective legal remedies in Kenya's evolving digital landscape.

#### **4.3.2. Corporate Responsibility and Ethical AI Development**

Kenya's National AI and Data Protection Policies emphasise ethical AI use, corporate accountability and consumer trust. A vicarious liability framework aligns seamlessly with these principles by ensuring that companies deploying chatbots remain responsible for their systems' interactions with consumers.

Without clear accountability, chatbot developers and deployers may lack the incentive to implement robust safety measures, leading to reckless AI deployment. By holding companies accountable for chatbot harms, vicarious liability fosters ethical AI development and compels businesses to prioritise user safety, fairness, and compliance with legal standards. Additionally, even if we were to argue that manufacturer's should be held liable, Article 16(b) of the *Sale of Goods Act* also imposes an obligation to the buyer or user of the chatbot as discussed in 3.3.3. Whereby, the manufacturer is obligated to produce goods of merchantable quality unless the buyer has inspected them for apparent defects. This implied condition shows that in the event a buyer does observe patent defects, they can hold the manufacturer liable. Thus even the deployer has recourse, if they are able to show that the product's defect was due to the manufacturer's negligence. This however, is an argument to be considered as a defence and not to be proven by the Claimant, who has faced the resultant effect. This is because liability under *Donoghue v Stevenson* is not confined to the manufacturer of a defective product but can also extend to all entities involved in the business chain.<sup>315</sup> Thus a wholesaler, in this case a deployer, must take reasonable steps to check the safety of what he distributes or uses.<sup>316</sup> Consequently, the deployer also has the capacity to bring in the manufacturer in the suit to be held jointly liable or if they are able to sufficiently prove it was the manufacturer's negligence, absolute liability in that case.

---

<sup>315</sup> *Fred Ben Okoth v Equator Bottles Limited* (2015) eKLR.

<sup>316</sup> *Fred Ben Okoth v Equator Bottles Limited* (2015) eKLR.

This is because, in civil proceedings, if a particular fact lies within the special knowledge of a party, the responsibility to prove or disprove that fact rests with them.<sup>317</sup>

This avenue of defence has been explored in multiple vicarious liability cases, whereby, in the event the employer is able to prove that the employee was not carrying out the actions contended in the suit during their course of employment, then they are held liable for their own actions as opposed to the employer being liable. This is equally a defence when it comes to such cases and can be extended to manufacturer's in chatbots and AI systems given they lack human agency.

Ultimately, this framework ensures that corporate actors bear responsibility for the consequences of AI-driven interactions, reinforcing the importance of responsible AI governance, while also allowing them to have recourse as to defence.

#### **4.3.3. Promoting Digital Transformation and Innovation**

Kenya is positioning itself as a regional leader in AI and digital innovation, attracting investors and fostering an environment conducive to technological advancements. However, as AI adoption increases, so do the potential risks associated with unregulated chatbot deployment. A legal framework that proactively addresses liability concerns will not only protect users but also enhance Kenya's reputation as a jurisdiction committed to responsible innovation.

Contrary to concerns that liability frameworks stifle innovation, vicarious liability encourages responsible AI deployment by ensuring that businesses invest in safer chatbot technology. By fostering accountability and risk management, this approach creates an ecosystem where AI innovation can thrive while protecting consumers from potential harm. It also ensures regulatory predictability, making Kenya an attractive destination for AI-driven enterprises seeking a balanced legal environment that promotes both technological growth and consumer protection.

#### **4.4. Conclusion**

In conclusion, the chapter has shown how vicarious liability aligns with established principles and Kenya's current objectives. By linking the established agency relationship to the rationale for

---

<sup>317</sup> Section 112, Evidence Act; *Munyu Maina v Hiram Gathiha Maina* (2013) eKLR; *Associated Motors Co. Ltd v Blue Sea Services Ltd* (2019) eKLR.

vicarious liability as the most viable framework for addressing chatbot harms, I have been able to show how it can be extended to such injuries or damages. By using vicarious liability, courts will be able to promote accountability and risk management by users which presents a compelling case for its adoption and its eventual reform in response to the challenge of establishing accountability for victims. In conclusion, adopting vicarious liability as the primary framework for chatbot harms will strengthen consumer protection, reinforce ethical AI governance, and position Kenya as a leader in responsible AI regulation.

## **5.0. Conclusion and Recommendations**

### **5.1. Conclusion**

The study was looking into whether companies should be held liable for chatbot harms. From the study, I was able to find that there exists key challenges in attributing responsibility for chatbot-related harms in the existing liability framework. Therefore, I was able to find that vicarious liability presents the most effective legal framework for addressing chatbot harms, for reasons explained in 4.2. Based on these findings, I propose several recommendations to strengthen the implementation of vicarious liability and enhance chatbot regulation in Kenya.

### **5.2. Recommendations**

#### **5.2.1. Recognition of Implied Agency Relationships between Chatbots and Deployers**

Since vicarious liability traditionally applies in employer-employee or principal-agent relationships, courts should extend the principle to chatbots by recognising them as implied agents of their deployers. Some of the considerations that should be taken into account are: whether the chatbot performed tasks aligned with the principal's business interests or was carrying out objectives set out by the principal, the degree of control the deployer exercised over the chatbot's functionality and, whether the chatbot's actions resulted from the parameters set by its deployer.

This approach will enable courts to apply well-established agency doctrines to chatbot-related cases rather than struggling to enact new laws or finding novel ways of determining liability.

#### **5.2.2. Develop the Standard of Care for Deployers Seeking to Avoid Liability**

Deployers may argue that chatbot actions were unpredictable or beyond their control, so courts should define what reasonable precautions deployers must take to avoid liability. This could include, implementing human oversight mechanisms for chatbot interactions, maintaining records of chatbot training and adjustments and, responding promptly to reports of chatbot misinformation or errors. By outlining clear expectations for chatbot deployers, courts can assess liability fairly while allowing for reasonable defenses based on due diligence.

### **5.2.3. Recognition of vicarious liability as the most appropriate legal framework for chatbot harms**

While strict liability and negligence may have some application in chatbot-related harms, courts should prioritise vicarious liability for several reasons. First, it aligns with existing employer-employee and principal-agent liability structures. Second, it ensures accountability without requiring plaintiffs to prove the chatbot's fault, which can be challenging due to AI's complexity. Third, it incentivises companies to improve chatbot safety rather than shifting blame to AI's unpredictability. Fourth, it provides sufficient avenues for companies to defend themselves if they are held liable.

In conclusion, by applying vicarious liability in future chatbot-related disputes, courts can set precedence for subsequent cases that may arise. Consequently, establishing a consistent approach to AI-related civil liability.

## Bibliography

1. Adamopoulou E and Moussiades L, 'An overview of chatbot technology' Artificial Intelligence Applications and Innovations, Maglogiannis, I., Illiadis, L., Pimendis, 29 May 2020.
2. Adamopoulou E and Moussiades L, 'Chatbots: History, technology, and applications' 2(100006) *Machine Learning with Applications*, 2020.
3. Adam M, Wessel M and Benlian A 'AI-based chatbots in customer service and their effects on user compliance' *Electronic Markets*, 2021.
4. Akech M, 'The common law's approach to liability and redress its applicability to East Africa' International Environmental Law Research Centre (IELRC.ORG), 2003.
5. Arcila B, 'AI liability in Europe: Addressing harms to fundamental rights and the problem of the human in the loop' *Computer Law & Security Review*, 2024.
6. Atillah I, 'A Belgian man reportedly decided to end his life after having conversations about the future of the planet with an AI chatbot named Eliza' Euronews, 31 March 2023.
7. Baris A, 'The use of chatbots in customer service: A qualitative analysis on customers' reception' Yasar University, 2021.
8. Binns R, 'Case study: How Apple has mastered CRM' Expert Market, 22 February 2023.
9. Botadra B, 'Web robots or most commonly known as bots' Academia.
10. Brown N, 'Bots behaving badly: A products liability approach to chatbot generated defamation' 2023(3) *Journal Free Speech of Law*, 2023.
11. Buiten M, Streel A and Peitz M, 'The law and economics of AI liability' 48(105794) *Computer Law & Security Review*, 2023.
12. Cecco L, 'AirCanada order to pay customer who was misled by airline's chatbot' The Guardian, 16 February 2024.
13. Cheng X, Bao Y, Zarifis A, Gong W and Mou J 'Exploring consumers' response to text-based chatbots in e-commerce: the moderating role of task complexity and chatbot disclosure' 32(2) *Internet Research*, 2021.
14. Chopra S and White L, 'Artificial Agents: Philosophical and legal perspective' ResearchGate, 2007.
15. Diamantis E, 'The extended corporate mind: when corporations use AI to break the law' 98(4) *North Carolina Law Review*, 2020.
16. Domingos P, *The master algorithm: How the quest for the ultimate learning machine will remake our world*, Hachette Book Group, New York, 2015.
17. Feferkorn K, 'Am I an algorithm or a product? When product liability should apply to algorithmic decision-makers' 30(61) *Stanford Law School*, 2019.
18. Folstad A, Skjuve M and Brandtzaeg P, 'Different chatbots for different purposes: Towards a typology of chatbots to understand interaction design' ResearchGate, 2019.
19. Gao C, Lei W, He X, Rijke M and Chua T, 'Advances and challenges in conversational recommender systems: A survey' *KeAI*, 2021.

20. Gidron Z, 'Language models, conversational AI, and chatbots explained' Hyro.ai, 15 December 2021.
21. Griffin A, 'Facebook's artificial intelligence robots shut down after starting to talk to each other in their own language' Independent, 31 July 2017.
22. Harris D 'The lost rationale of agency law' 3(1) *Business Finance Law Review*, 2019.
23. Hassija V, Chamola V, Mahapatra A, Single A, Goal D, Huang K, Scardapane S, Spinelli I, Mahmud M and Hussain A, 'Interpreting Black-Box Models: A review on explainable artificial intelligence' *Cognitive Computing*, 2023.
24. Haugeland I, Folstad A, Taylor C and Bjorkli C, 'Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction and design' 161(102788) *International Journal of Human-Computer Studies*, 2022.
25. Hawley M, 'Exploring Air Canada's AI chatbot dilemma' CMSWire, 2 April 2024.
26. Heine C, 'Microsoft's chatbot 'Tay' just went on a racist, misogynistic anti-Semitic tirade' Adweek, 24 March 2016.
27. Huberman P, 'Tort law, corrective justice and the problem of autonomous-machine-caused harm' 34(1) *Canadian Journal of Law & Jurisprudence*, 2020.
28. Hung R and Lifshitz L, 'BC Tribunal confirms that companies remain liable for information provided by AI chatbot' American Bar Association Group, 29 February 2024.
29. Hussain A, *General Principles and Commercial Law of Kenya*, East African Educational Publishers Ltd, Kenya.
30. Kappel R, 'Overview of AI regulations and regulatory proposals of 2023' Central Eyes, 5 March 2024.
31. Kaput M, 'Natural Language Generation (NLG): Everything you need to know' Market Artificial Intelligence Institute, 28 February 2024.
32. Kevin Scott, 'Popular use cases for chatbots' Chatbot Magazine, 18 October 2016.
33. Khanna A, Pandey B, Vashishta K, Kalia K, Pradeepkumar B and Das T, 'A study of today's AI through chatbots and rediscovery of machine intelligence' 8(7) *International Journal of u- and e- Service, Science and Technology*, 2015.
34. Kirwa M, 'Nine ways AI is making life easier for Kenyans,' The Star, 25 April 2023.
35. Lai A, 'Artificial Intelligence, LLC: Corporate Personhood as tort reform' *Michigan State Law Review*, 2020.
36. Lee P, 'Learning from Tay's introduction' Official Microsoft Blog, 25 March 2016.
37. Leingang R, 'X's AI Chatbot spread voter misinformation-and election officials fought back' The Guardian, 12 September 2024.
38. Leua C and Didu I, 'Chatbots. legal challenges and the EU legal policy approach' 10(3) *Society of Juridical and Administrative Sciences*, 2021.
39. Lewis C, 'The need for a legal framework to regulate the use of artificial intelligence' 47(2) *University of Dayton Law Review*, 2022.

40. Lima G, 'Can (and should) AI be considered an agent' *California Digital Library*, 2019.
41. Lior A, 'AI entities as AI agents: Artificial intelligence liability and the AI respondeat superior analogy' 46(5) *Mitchell Hamline School of Law*, 2020.
42. Ltifi M, 'Trust in the chatbot: a semi-human relationship' 9(109) *Future Business Journal*, 2023.
43. Luger E and Sellen A, 'Like having a really bad PA: The gulf between user expectation and experience of conversational agents' Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, California, San Jose, USA, 7-12 May 2016.
44. Luo X, Tong S, Fang Z and Qu Z, 'Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases' 38(6) *Marketing Science*, 2019.
45. Martin-casals M, 'The Development of Liability in Relation to Technological Change' Cambridge University Press, 2010.
46. Montagnani M 'Liability and emerging digital technologies: an EU perspective' 11(2) *Notre Dame Journal of International & Comparative Law*, 2020.
47. Montgomery B, 'Mother says AI chatbot led her son to kill himself in lawsuit against its maker' *The Guardian*, 23 October 2024.
48. Omri R, 'Whose robot is it anyway?: Liability for artificial intelligence-based robots' 20(4) *University of Illinois Law Review*, 2019.
49. Osoba O, 'The value of the concept of agency in an increasingly rational world' *California Digital Library*, 2019.
50. Price N, 'Black-box medicine' 28(2) *Harvard Journal of Law & Technology*, 2015.
51. Radziwill N and Benton M, 'Evaluating quality of chatbots and intelligent conversational agents' *Cornell University*, 2017.
52. Reshmi S and Balakrishnan K, 'Empowering chatbots with business intelligence by big data integration' 9(1) *International Journal of Advanced Research in Computer Science*, 2018.
53. Shawar A and Atwell E, 'Chatbots: Are they really useful?' 22(1) *Journal of Language Technology and Computational Linguistics*, 2007.
54. Shrestha S, 'Nature, nurture, or neither?: Liability for automated and autonomous artificial intelligence torts based on human design and influences' 29(1) *George Mason Law Review*, 2021.
55. Stojanov M, 'Prospects for chatbots' 8(3) *Izvestia Journal of the Union of Scientists-Varna Economic Sciences Series*, 2019.
56. Surden H, 'Artificial Intelligence and law: An overview' 35(4) *Georgia State University Law Review*, 2019.
57. Ukpabi D, Aslam B and Karjaluoto H 'Chatbot adoption in tourism services: A conceptual exploration' *ResearchGate*, 2019.
58. Vasudevan A, 'Addressing the liability gap in AI accidents' *Center for International Governance Innovation, Policy Brief No. 177*, 2023.

59. Viski A, Jones S, Rand L, Boyce T and Siegel J, *Artificial intelligence and strategic trade controls*, Center for International & Security Studies, Maryland, 2020.
60. Vladeck D, 'Machines without principals: Liability rules and artificial intelligence' 89(117) *Washington Law Review*, 2014.
61. Walker J, 'I smell a bot: California's S.B. 1001. Free speech and the future of bot regulation' *Houston Law Review*, 2019.
62. Wang Y and Petrina S, 'Using learning analytics to understand the design of an intelligent language tutor-chatbot Lucy' 4(11) *International Journal of Advanced Computer Science and Applications (IJACSA)*, 2013.
63. Weber U, Lomker M and Moskaliuk J, 'The human touch: The impact of anthropomorphism in chatbots on the perceived success of solution focused coaching' 32(4) *Management Revenue*, 2021.
64. Wein L, 'The responsibility of intelligent artifacts: toward an automation jurisprudence' 6(Fall Issue), *Harvard Journal of Law & Technology*, 1992.
65. Wendehorst A, 'Liability for Artificial Intelligence' in Wendehorst A (ed) *The Cambridge Handbook of Responsible Artificial Intelligence*, Cambridge University Press, 2022.
66. Zumstein D and Hundertmark S, 'Chatbots—An interactive technology for personalised communication, transaction and services' 15(1) *IADIS International Journal*, 2017.