Thesis

Ayushman Kheterpal

261079267

Choosing the appropriate standard of liability for Artificial Intelligence-related harms.

Faculty of Law, McGill University, Montreal

August 2024

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of Master of Laws (LL.M)

Table of Contents

Land Acknowledgement

McGill University is on land which has long served as a site of meeting and exchange amongst Indigenous peoples, including the Haudenosaunee and Anishinabeg nations. I acknowledge and thank the diverse Indigenous peoples whose presence marks this territory on which peoples of the world now gather.

Abstract

With recent advancements in generative Artificial Intelligence technology, AI has become a buzzword that captivates everyone's attention whenever it is mentioned during any social or professional conversation. The importance of AI is so significant that it is being termed the steam engine of the fourth industrial revolution. However, the adoption of AI is outpacing the legal developments. While some progress has been made with respect to overall AI regulation, particularly in the EU, lawmakers have been sluggish in clarifying the contours of jurisprudence surrounding AI liability laws.

What, then, should be the standard of liability for AI to foster its adoption and innovation while balancing it with protection for those adversely impacted by AI-related harms? To begin with, this thesis discusses the concept of a standard of liability and argues for a deliberate approach by policymakers towards adopting a standard of liability. This thesis employs doctrinal analysis across various legal domains, including banking, insurance, health, and environmental law, to evaluate different existing standards of liability. This thesis extracts and delineates the existing standards of liability, which range from good faith-based due diligence to common law-based negligence, to utmost good faith-based due diligence, and to no-fault standard of liability. Thereafter, this thesis deploys analogical reasoning to doctrinally compare the suitability of each of these standards of liability for AI-related harms. Therefore, this thesis presents different options for policymakers when selecting an appropriate standard of liability.

After presenting different options for adopting a standard of liability, this thesis discusses factors that should be considered while selecting an appropriate standard of liability. It discusses how traditional liability systems, including the strict standard of liability, consider risk as an important consideration when the standard of liability is being chosen for a specific use case and how a risk-based approach has also been adopted by the EU to regulate AI through its overarching regulation. Thus, when selecting the appropriate standard of liability for a particular economic activity, the risk factor is an important consideration; however, this thesis discusses other crucial factors that should also be considered while making this selection. This thesis also explores how the distribution of liability among different stakeholders can be guided by environmental law doctrines such as the "polluter pays principle", "deep pocket theory", and assessment of the degree of control.

Resumé

Avec les récents progrès de la technologie de l'intelligence artificielle générative, l'IA est devenue un mot à la mode qui captive l'attention de tous chaque fois qu'elle est évoquée lors d'une conversation sociale ou professionnelle. L'importance de l'IA est telle qu'elle est qualifiée de machine à vapeur de la quatrième révolution industrielle. Cependant, l'adoption de l'IA dépasse les évolutions juridiques. Bien que certains progrès aient été réalisés en ce qui concerne la réglementation globale de l'IA, en particulier au sein de l'UE, les législateurs ont mis du temps à clarifier les contours de la jurisprudence concernant les lois sur la responsabilité en matière d'IA.

Quelle devrait donc être la norme de responsabilité de l'IA pour favoriser son adoption et son innovation tout en la mettant en équilibre avec la protection des personnes affectées par les dommages liés à l'IA ? Pour commencer, cette thèse aborde le concept de norme de responsabilité et plaide en faveur d'une approche délibérée des décideurs politiques en vue d'adopter une norme de responsabilité. Cette thèse utilise une analyse doctrinale dans divers domaines juridiques, notamment le droit bancaire, des assurances, de la santé et de l'environnement, pour évaluer les différentes normes de responsabilité existantes. Cette thèse extrait et délimite les normes de responsabilité existantes, qui vont de la diligence raisonnable fondée sur la bonne foi à la négligence fondée sur la common law, en passant par la diligence raisonnable fondée sur la bonne foi et la norme de responsabilité sans faute. Par la suite, cette thèse déploie un raisonnement analogique pour comparer doctrinalement la pertinence de chacune de ces normes de responsabilité pour les dommages liés à l'IA. Par conséquent, cette thèse présente différentes options aux décideurs politiques lors du choix d'une norme de responsabilité appropriée.

Après avoir présenté différentes options pour l'adoption d'une norme de responsabilité, cette thèse discute des facteurs qui devraient être pris en compte lors de la sélection d'une norme de responsabilité appropriée. Il explique comment les systèmes de responsabilité traditionnels, y compris la norme de responsabilité stricte, considèrent le risque comme une considération importante lorsque la norme de responsabilité est choisie pour un cas d'utilisation spécifique et comment une approche basée sur le risque a également été adoptée par l'UE pour réglementer l'IA. à travers sa réglementation globale. Ainsi, lors de la sélection du niveau de responsabilité approprié pour une activité économique particulière, le facteur de risque est une considération importante; cependant, cette thèse aborde d'autres facteurs cruciaux qui devraient également

être pris en compte lors de cette sélection. Cette thèse explore également comment la répartition de la responsabilité entre les différentes parties prenantes peut être guidée par des doctrines du droit de l'environnement telles que le « principe du pollueur-payeur », la « théorie des poches profondes » et l'évaluation du degré de contrôle.

Acknowledgements

Firstly, I would like to thank my friends, family, and classmates, who have been a source of wisdom and emotional strength during the past two years.

I am grateful to my supervisor, Professor Omar Farahat, who expediently provided me with valuable insights despite challenging circumstances, which have been immensely helpful in improving this thesis. I am also very thankful to Professor Mark Antaki and Professor Richard Gold, whose lectures significantly helped me understand the process of academic writing and critical thinking. I am grateful to Professor Gold and Professor Lara Khoury for their invaluable guidance in helping me define the correct scope for the present research during the proposal drafting stage.

I am also very grateful to the Graduate Programs Office of McGill's Faculty of Law, particularly Bianca Bourgeois, for critical support and help.

List of figures and tables

Abbreviations

AI: Artificial Intelligence

AILD: Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)

EO 14110: Executive Order 14110 on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, issued by the Biden administration on October 30, 2023.

EU: The European Union

EU AI Act: Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828.

FLOPs: floating-point operations

FTC: Federal Trade Commission

GPAI: General-Purpose Artificial Intelligence

IEEE: Institute of Electrical and Electronics Engineers

IP: Intellectual Property

ISO: International Organization for Standardization

NIST: National Institute of Standards and Technology

PLD: Product Liability Directive - Proposal for a Directive of the European Parliament and of the Council on liability for defective products COM/2022/495 final

The US: The United States of America

RMF: Risk Management Framework

SCI: Supreme Court of India

SDO: Standards Development Organization

Introduction

The human world today is becoming more and more data-driven with the help of Artificial Intelligence (AI). AI helps improve economic efficiency and provides excellent tools to solve problems quickly, including those which are not humanly possible to achieve. Owing to these advantages, AI is already a significant part of various industries, such as finance, marketing, manufacturing, healthcare, weather forecasting, and digital services (including web search engines, online booking, and social media). However, since AI is partially a self-training tool, its functionality is not always fully understood, and there are risks associated with its usage. Against this backdrop, the question arises as to how the liability should be determined in a case when any damage, bodily[1] or material, is caused to an individual, entity, or the public in connection with the application of AI (hereinafter referred to as "AI-related harms").

Shortcomings of the existing liability systems when applied to AI-related harms - a discussion based on the literature review

The need for customising standard of liability beyond torts law for AI-related harms

While tort law and its guiding principles, essential elements, and procedures are well-established, its application in AI-related harms faces numerous challenges. One such main challenge is that AI is unpredictable and inscrutable[2], making the legal analysis of essential elements like causation and duty of care under tort law, specifically under negligence assessment, difficult.[3] Owing to its unpredictability and inscrutability, AI can also result in novel types of harm, for example, harm to the personal integrity of the individuals and harm to the social, cultural or political fabric of society due to deep fakes and discriminative AI,

[1] Including, but not limited to, physical harm and mental distress for an individual.
[2] Difficult for humans to interpret or understand.
[3] Kristen Thomasen, "AI and Tort Law" in Florian Martin-Bariteau & Teresa Scassa, eds, *Artificial Intelligence and the Law in Canada* (New York: Rochester, 2021).

respectively.[4] It is difficult to provide a remedy under existing tort law for such harms since even the existence of these harms is challenging to ascertain. Traditional legal systems provide only narrow and limited remedies to tackle such society-wide distress, usually through criminal law provisions such as public nuisance, which are limited by the constitutional guarantees relating to basic human rights of freedom of speech and freedom to conduct business activities.

In this context, Thomasen raises a pertinent question: whether new rights must be recognised to tackle such AI-related harms.[5] If such rights are formally recognised, then there might be an option to use a standard of liability which is customised for particular use cases, for example, implementing a strict standard of liability with minimal burden of proof on victims in case of severe harm. Another significant challenge that the courts need to deal with while implementing the existing tort law with respect to AI-related harms is to balance policy considerations, especially the perceived social benefits of innovation through AI.[6] Due to these variations in policy, perceptions regarding the benefits of AI, and technical factors of unpredictability and inscrutability, the courts will struggle to uniformly apply the existing tort law in a uniform and predictable manner. This uncertainty erodes the trust of developers, operators and end users of AI systems.

This also results in uncertainty regarding the extent of liability of various stakeholders, such as the end user, hardware supplier, software developer(s)[7], data provider[8], operators and, if used by public institutions, the government. Such uncertainty leads to hesitation in improving the adoption of AI in various economic activities and, hence, will deter innovation and

[4] *Ibid.*
[5] *Ibid.*
[6] *Ibid.*
[7] There could be multiple software developers, for example different software developer at the architectural level, and different software developer at the implementation (core program) level, and still another developer at the interface level (Graphics or User Interface), all of which may contribute to a potential undesired consequence.
[8] The data provider may not always be same as the entity engaged in software development and optimization. Erroneous, unorganized, or incomplete datasets may lead to undesired consequences.

investment, including in critical sectors such as healthcare[9], which is against the interests of society.[10] On the other hand, clarity regarding liability can encourage entities to develop and integrate risk-mitigation measures, improve product design, and also build public trust and adoption.[11]

And while AI is booming, lawmakers are still playing catch-up in a race to regulate AI-related harms.[12] For example, in the case of self-driving cars, Tesla's Autopilot has been involved in a significant number of accidents and even fatalities, which in itself are indicative of defects in the technology and possibly hasty rollouts.[13] Tesla claims that its autopilot is safer than that of human drivers; however, statistics are not clear in this regard.[14] One of the reasons for these fatalities is also because of the excessive trust in technology.[15] If left unchecked and unregulated, society may lose trust not only in technology but also in public institutions' ability to compensate for technology-related harms. As such, in the absence of a well-drawn and customised standard of liability, the AI developers of self-driving technology currently have no real incentive to implement guardrails[16] or other safety protocols within AI to make self-driving technology safer, and their design is largely driven by business and marketing decisions to attract more consumers. Currently, in the absence of a customised standard of liability, the defence strategy has evolved organically, wherein the attempt has been to shift the liability to

[9] Mélanie Forcier, Lara Khoury & Nathalie Vézina, "Liability issues for the use of artificial intelligence in health care in Canada: AI and medical decision-making" (2020) 46 Dalhousie Medical J at 7.
[10] Miriam Buiten, Alexandre de Streel & Martin Peitz, "The law and economics of AI liability" (2023) 48 Computer L & Security Rev 105794 at 9.
[11] *Ibid*.
[12] Adam Satariano & Cecilia Kang, "How Nations Are Losing a Global Race to Tackle A.I.'s Harms", *The New York Times* (6 December 2023), online: <https://www.nytimes.com/2023/12/06/technology/ai-regulation-policies.html>.
[13] "17 fatalities, 736 crashes: The shocking toll of Tesla's Autopilot", (10 June 2023), online: *Washington Post* <https://www.washingtonpost.com/technology/2023/06/10/tesla-autopilot-crashes-elon-musk/>.
[14] *Ibid*.
[15] *Ibid*.
[16] Guardrails is a term commonly used in the context of AI. It typically refers to algorithmic mechanisms or AI training practices that make AI systems safer and more ethically sound.

the end user (car driver), claiming that ultimate control rests with the end users.[17] Creating a customised and well-defined standard of liability could create the right incentives for AI developers to make the technology safer; for example, if a self-driving car manufacturer knows that there is a safe harbour that protects them from liability if they apply appropriate safety protocols within their AI systems, then focus of defence would shift towards making the technology safer. Safe harbours have been tested and proven to provide a boost to innovation. For example, in the case of intermediary liabilities, safe harbours are well recognised to have been major enablers for the existence of the internet, an extremely crucial technology on which the world relies today.[18]

And, although there have been efforts from various countries to come up with AI regulations and policies[19], there is no clarity regarding standards of liability and the extent of sharing of liabilities amongst various stakeholders. For example, some of the literature suggests careful drafting of contracts, clear and true upfront disclaimers, and using the precautionary principle[20] to counter the uncertainties in the law and also raises interesting questions about the liability of not using AI.[21] Hence, the existing literature acknowledges the wide lacuna in existing liability frameworks and primarily suggests non-technical measures to overcome this shortcoming in the law. While the literature does suggest adopting additional measures proactively as a precautionary measure for a possible defence, there is no guarantee that such a defence would be accepted in the courts. This clearly demonstrates that existing laws need to

[17] "Lawsuits test Tesla claim that drivers are solely responsible for crashes", (28 April 2024), online: *Washington Post* <https://www.washingtonpost.com/technology/2024/04/28/tesla-trial-autopilot-lawsuit/>.
[18] Ernest Lim & Phillip Morgan, eds, *The Cambridge Handbook of Private Law and Artificial Intelligence* (Cambridge University Press, 2024) at 385–386.
[19] For example, AI Bill of Rights (2022) in the USA, and White Paper on Artificial Intelligence - A European approach to excellence and trust (2020) of the EU based on the risk-based approach.
[20] The precautionary principle implies adopting additional measures proactively and being guided through non-binding international and renowned guidelines for good AI design practices to prevent potential harm in the absence of any guidelines issued by the state.
[21] Forcier, Khoury & Vézina, "Liability issues for the use of artificial intelligence in health care in Canada", *supra* note 9 at 10.

be modified and that there needs to be deliberation on inducting safe harbours for the developers or operators of AI to encourage them to make AI safer, and integrating these safe harbours with the standard of liability can be one potent tool for lawmakers and policymakers to encourage safe development of AI.

Hence, customising the standard of liability for AI is the need of the hour. Any choice of implementing a more customised standard of liability, which differs from the existing tort law, will require an active deliberation of the state, either executive, legislature, or judiciary. The actions of policymakers and lawmakers will be more crucial in this regard since the judiciary can only make exceptions to general jurisprudence when exceptional facts prevail in a particular case. Policymakers and lawmakers have an opportunity to act proactively and create an environment where AI can be designed for the good of society and is closely aligned with the state's policy objectives. Further, policymakers and lawmakers can deploy their vast pool of resources and even interact with stakeholders freely before customising the standard of liability in an optimal and balanced manner.

However, customising and selecting an appropriate standard of liability is complex. Policymakers can apply a broad framework based on the economic activity's risk levels, importance, degree of control exercised by stakeholders, geopolitical circumstances, and financial, social, and political capacity of the entities involved. The first step in this exercise would still be to ascertain the available options for the standard of liability.

Distribution of liability

Whenever there is a discussion regarding AI-related harms, there is a presumption that AI is acting autonomously; however, most of the applications of AI, including in medicine, finance, and car driving assistance, are where human users team up with AI systems and human

users take the final decision.[22] Hence, a person or entity simply cannot get rid of their liability merely because of their reliance on AI systems. This is similar to the jurisprudence regarding situations where the car owner is unable to get rid of their liability when the owner has consented to someone else driving the owner's car while the owner is a passenger, owing to the presence of the direct connection of the substitute driver to the owner of the car.[23] However, depending upon the severity and extent of the harm, the end user may not always be in the best position to compensate the victims, and in the absence of protections for the end user, the end user, such as medical practitioners, would be hesitant to adopt new technologies. In other literature, it was noted that since various countries do not recognise AI as a legal entity, the liability shall have to be borne by the other stakeholders and discusses vicarious or strict standards of liability.[24] Hence, there is also a need for developing/adopting legal principles that would assist courts in determining the distribution of liability in the most optimal manner.

Since many parties are involved in the development and operation of AI-based systems, products, and services, each of these parties must take care to minimise AI-related harms.[25] When the harm does occur, it is difficult for the courts to determine the exact source of the failure.[26] This may make the judiciary more inclined to impose full liability on a strict standard basis on just the entity with more degree of control or more financial resources. However, Buiten et al. argue that both developers and operators should be liable for AI to reach a socially efficient level of care, even if operators have less degree of control over the AI system.[27] While existing literature does recognise the importance of the locus/degree of control in determination

[22] Andrew D Selbst, "Negligence and AI's Human Users" (2019) 100 BUL Rev 1315.
[23] "Torts. Negligence. Liability of Owner of Automobile for Negligent Act of Driver. Section 282-e of New York Highway Law Interpreted" (1927) 27:7 Colum L Rev 886–887.
[24] Paulius Cerka, Jurgita Grigienė & Gintarė Sirbikytė, "Liability for damages caused by artificial intelligence" (2015) 31 Computer Law & Security Review.
[25] Buiten, de Streel & Peitz, *supra* note 10 at 12.
[26] *Ibid.*
[27] *Ibid.*

of distribution of liability, there is a lack of discussion regarding what exact legal principles can be used to determine exact distribution.

Innovation concerns

Asimov's three laws of robotics form the foundation of discussion for any autonomous system/ robot whose purpose is to avoid causing serious bodily injury. These laws require the robot to prioritise not to injure a human being either by action or inaction, obey human commands as a second priority, and protect its own existence as a last priority.[28] However, Asimov's laws limit the usefulness of autonomous systems due to their lack of autonomy.[29] How much of a free hand should, then, be given to these AI-based systems? As per the current legal system, safety is addressed through contract and tort law.[30] Hubbard argues that with the rise in the sophistication of robots, the tort system will have to adapt and will require human users to use a higher level of operation and maintenance skill in a fashion similar to operators of heavy machinery, such as additional licensing mechanisms for operators to establish that they were not at fault.[31]

Marchant and Lindor are wary that the increasing autonomous nature of the vehicles can lead to an increase in exposure to liability-based claims against the manufacturer and may shift the blame from the driver to the vehicle manufacturer.[32] This liability exposure can deter the adoption of a socially beneficial new technology.[33] They believe if the shift in liability exposure is serious enough, there would be a need for better defences for the car manufacturers, including the assumption of risk defence (user/car owner assumes the risk and remains liable)

[28] F Patrick Hubbard, "'Sophisticated Robots': Balancing Liability, Regulation, and Innovation" (2015) 66 Fla L Rev at 1808.
[29] *Ibid* at 1809.
[30] *Ibid* at 1812–1820.
[31] *Ibid* at 1858–1862.
[32] Gary Marchant & Rachel Lindor, "The Coming Collision Between Autonomous Vehicles and the Liability System" (2012) 52:4 Santa Clara L Rev 1321 at 1339–1340.
[33] *Ibid*.

and legislative safe harbours. [34] At the same time, these defences can lead to complacency in car manufacturers to innovate and make incremental improvements with respect to safety measures. [35] Hence, the right balance is required between excess exposures of innovators to liability and excess protection from the same to incentivise the safe development of AI systems.

The argument in the previous section was that the existing liability framework, such as tort law, is too broad and lacks specificity to incentivise entities with sufficient locus of control to take meaningful measures and implement guardrails in the AI systems. On the other hand, the liability framework, if customised, will put restraints on the hands of the developers or operators of AI, which can limit innovation. What if, in the development cycle of AI, to get to a safer and more efficient version of the technology, we need to go through a phase of unhindered innovation? Will the encouragement of implementing safety measures not be a concern for innovation? Should we, as a society, begin with less stringent standards of liability to encourage innovation and make it more stringent as the technology matures? Regardless of the answer to these questions, it is crucial to identify the available spectrum of the standard of liability, which will offer policymakers and lawmakers an option to fine-tune the liability framework. Thus, this thesis aims to identify the available options for the standard of liability in the context of AI-related harms.

AI-specific practical concerns

Hubbard acknowledges that expert testimony would become important for determining negligence before the courts.[36] AI is highly technical and may not be straightforward for assessment by the courts comprising of generalist judges. However, similar practical issues are also faced in patent disputes, and there are various additional mechanisms that can be imported

[34] *Ibid*.
[35] *Ibid*.
[36] Hubbard, *supra* note 28 at 1858–1862.

from patent litigation practices and can be used to assist the court in tackling complex technologies. Nevertheless, it is of utmost importance that the state, and especially the executive arm of the government, take necessary measures to build institutional capacities so that the technology does not race away fast beyond the understanding of the state machinery. A lack of institutional capacity of the state may put the state in a situation where they have to rely on the potential wrongdoers themselves for making technical assessments and even policymaking.

Hubbard opposes the proposals for the introduction of no-fault insurance schemes to safeguard victims and argues for maintaining the current liability system to ensure fair compensation and incentives to innovate in a safe manner.[37] Hubbard is wary that using no-fault insurance schemes can result in the creation of a pay-to-be-unsafe model.[38] If there are any alternate pathways, whether legal, monetary or financial, giant corporations are likely to adopt the method which generates maximum profits for the shareholders. However, such an approach to AI development may not always be the best for society.

Hubbard argues that the current legal system is already fair, creates the right balance between innovation and safety, and does not require any structural changes.[39] However, the current liability system is uncertain owing to technical complexities in AI, which helps neither the AI-related stakeholders nor the victims.

Legal Research Questions

With AI becoming an omnipresent and necessary part of our daily lives, it is essential to define the mechanisms for regulation of liability to balance innovation in AI and safeguard

[37] *Ibid* at 1866–1867.
[38] *Ibid* at 1866–1867.
[39] *Ibid* at 1872.

the interests of individuals and society. The primary question of the presently proposed research is: What should be the standard of liability for damages for AI-related harms, and would that liability vary based on the industry or economic activity in which it is being implemented? A further question that needs to be investigated is the mechanisms through which liability can be distributed among various stakeholders.

Methodology

The proposed research will carry doctrinal research across different spheres/domains of law and jurisdictions[40], and explore various mechanisms available for the regulation of liability. Doctrinal research will study different standards in the fields of banking law, insurance law, health law and environmental law, wherein the standards of liability will vary from no-liability if due diligence[41] has been carried out (for example, in banking law), to the utmost due diligence standard, and to the no-fault liability standard of environment law (which can be further categorised into a strict liability[42] and an absolute liability[43]). The doctrinal research will rely on analogical reasoning[44], which will include a three-step approach[45] wherein, firstly, a "base point"[46] for comparison will be identified, and secondly, similarities and differences between the "base point" and the instantaneous case point[47] will be analysed, and finally, it will be determined whether the two points can be treated in a same manner based on this analysis.

[40] Primarily based on the common law system.
[41] With its own varying standards based on the area of law and jurisdiction.
[42] Liability standard of Indian environmental law, even when there is no fault or negligence and despite carrying out due diligence.
[43] Exemplary damages standard of Indian environmental law, for inherently dangerous activities.
[44] Cass R Sunstein, *Analogical Reasoning* (New York: Rochester, 2021).
[45] "The Bridge: How Reasoning By Analogy Works in Law", online: <https://cyber.harvard.edu/bridge/Analogy/analogy3.htm>.
[46] In our case, the liability regime of other spheres of law where liability regimes are well established.
[47] The possible adoption of such a regime to the AI.

To ascertain the distribution of liability among various stakeholders, the doctrinal research will also try to draw parallels between the “polluter pays” principle and the “deep pocket theory” of environmental law. The research will look at the perspective of Intellectual Property (IP) Law to derive some takeaways for the distribution of liability.

Chapter 1: Current legal framework on AI, including its liability.

The world has been ushered into a new era where the adoption of AI is exponentially expanding in diverse sectors, in both manufacturing and service sectors, including medicine and healthcare, transportation, energy, education, tourism, culture, primary industries such as mining and agriculture, and government-related services such as social safety, defence and welfare.[48] The current state of technology with respect to AI is still semi-autonomous with substantial human control, and hence, the liability regime is still applied as if being applied to such humans, i.e., principals.[49] But if we want to explore the full potential of AI, we will need to use fully automated AI tools. However, policymakers and lawmakers are still playing catch up to the significant advancements in AI, specifically with respect to education, financial markets, and labour markets,[50] and are even more underprepared to tackle AI-related harms.[51]

1.1 What is AI?

It is crucial to delve into the definition of AI first to understand the context of why AI is being regulated and enable a deeper discussion regarding the appropriate manner of regulation. The process of defining AI is not only a legislative challenge but also a technical one. Definitions of AI can be very different based on the purpose of constructing them, which could be philosophical, technical/scientific, or regulatory/legal. These definitions of AI have also evolved over time with the progression of the technology and its prevalence.

[48] Jee-Sun Oh, Moon-Koo Kim & Duk Hee Lee, "A study on the selection of future AI+X promising fields and the direction to strengthen competitiveness" (2021) 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) 371–374.
[49] David C Vladeck, "Machines without Principals: Liability Rules and Artificial Intelligence Essay" (2014) 89:1 Wash L Rev 117–150.
[50] Mark Fenwick, Wulf A Kaal & Erik P M Vermeulen, "Regulation Tomorrow: What Happens When Technology Is Faster than the Law" (2016) 6:3 Am U Bus L Rev 561–594 at 565.
[51] Satariano & Kang, *supra* note 12.

1.1.1 Technical Definitions of AI

Alan Turing, widely considered to be the father of modern computer science, in 1950 gave his theory of 'imitation game'[52]. Alan Turing opposed asking abstract questions as a metric to determine whether a machine is intelligent or not; for example, he argued that the question "Can machines think?" was not objective and was merely a public opinion poll[53]. He instead proposed his theory of imitation game wherein, in an experiment, the machine is able to successfully pass off as a human, and humans perceive that the output of the machine is coming from a human and not a machine. Therefore, the imitation game theory lays emphasis on the assessment of the output of the machine instead of its process for generating such output to classify it as intelligent or not.

Some of the recent discussions partially critique Turing's way of understanding and defining artificial intelligence. The critique mainly lies against comparing outputs with that of humans and being human-like as being the test of intelligence by machines. For example, Russell and Norvig argue that comparing the intelligence of aircraft against flying birds wouldn't be relevant to assessing the intelligence of flying machines[54]. However, this may not be an appropriate analogy since the goal of aeronautical engineering is not to imitate birds but to learn and improve beyond the capability of birds. What about anthropomorphic robots in that case? Some suggestions have been made to improve the definition of intelligence by considering the imitation of motor abilities as an additional component of the Turing test since the human brain also has a component that stems from interaction with a physical medium[55]. Assuming such expansion is allowed, we might then have to look to expand the definitions to

[52] A M Turing, "I.—COMPUTING MACHINERY AND INTELLIGENCE" (1950) LIX:236 Mind 433–460.
[53] *Ibid*.
[54] Stuart Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed (USA: Prentice Hall Press, 2009) at 3.
[55] Ruth Stock-Homburg et al, *Evaluation of the Handshake Turing Test for anthropomorphic Robots* (2020) arXiv:2001.10464 [cs].

consider other senses of humans as well, including the five senses of smell, touch, sight, taste, and hearing. But, perhaps, we should even consider proprioception and immune sense[56]. While AI may never be truly human, it can and has surpassed human capabilities in some of the tasks, and at the same time, it could struggle in some of the other tasks.

Other computer scientists argue that comparing the intelligence of AI systems with human-like or other biological beings-like intelligence is not rational. John McCarthy, in his later definitions, defines intelligence simply as "…the science and engineering of making intelligent machines…" and focuses on the ability to solve real-world problems and says that AI is not bound by the biological methods of intelligence[57]. The term "Artificial Intelligence" was first coined together by McCarthy, Minsky, Rochester and Shannon in their 1955 research proposal, wherein they defined AI in a manner where they compared it to human intelligence and stated that "*For the present purpose the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving*".[58] McCarthy later was of the view that while comparison with human-like intelligence provides us with some understanding of the capabilities of AI, comparing AI with humans is not fully appropriate because human abilities and mechanisms that humans use to solve problems have not been fully understood.[59] Geoffrey Hinton, considered to be one of the godfathers, commented that AI is extremely adept at replicating its learnings, as opposed to humans, who are biological beings.[60] Hence, later definitions observe that there are differences between AI and biological beings and go beyond Turing's definition to define AI. Google adopts a broader definition and defines AI as "reason, learn, and act" in a human-like manner,

[56] Jonathan Kipnis, "Immune system: The 'seventh sense'" (2018) 215:2 J Exp Med 397–398.
[57] "What is AI? / Basic Questions", online: <http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html>.
[58] J McCarthy et al, "A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE", (31 August 1955), online: <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>.
[59] *Ibid.*
[60] "Watch: Geoffrey Hinton tells BBC of AI dangers", *BBC News*, online: <https://www.bbc.com/news/av/world-us-canada-65453192>.

but Google's AI definition also includes the machines' ability to analyse data beyond human capacity[61].

The common theme among the definitions given on a technical basis is that they all agree with the main aspect of the definition given by Alan Turing, wherein instead of asking philosophical questions, the emphasis is on assessing outputs to ascertain whether a system or machine is intelligent. Most of the definitions try to gauge machine intelligence by comparing it to human intelligence. Importantly, since the definitions have been postulated as well as assessed by humans from a human standpoint, all these definitions invariably have a 'human factor' associated with them.[62]

Turing's approach focused mainly on the outputs of the machines to ascertain whether they were intelligent by assessing whether they were human-like, whereas later definitions, including McCarthy's, focused on the underlying capabilities that constitute intelligent or somewhat intelligent behaviour. Initially, the definition of AI was more from a conceptual standpoint when AI was nascent, and technology's focus was on applications such as gaming and natural language processing only. Currently, we have a more diverse type of AI with applications in almost all industries and services. Hence, a shift in the definition of AI is expected with the improvement of technology and progress in the discourse regarding AI. This progress can be seen from the shifting views of technologists such as McCarthy themselves. Today, AI "*statistically learn[s]*" in a continuous manner owing to the ongoing processing of data and is referred to as a shift from "*deterministic to probabilistic computing*".[63]

---

[61] "What Is Artificial Intelligence (AI)?", online: *Google Cloud* <https://cloud.google.com/learn/what-is-artificial-intelligence>.
[62] *Supra* note 57.
[63] Iria Giuffrida, "Liability for AI Decision-Making: Some Legal and Ethical Considerations" (2019) Fordham Law Review, online: <https://www.semanticscholar.org/paper/Liability-for-AI-Decision-Making%3A-Some-Legal-and-Giuffrida/7b0e4436bffa0a627e3efc33c1f3f2a3480d74a3> at 441.

In line with later McCarthy's views, I believe that the essence of intelligence, whether originating in the brains of carbon-based organisms or silicon-based neural networks, is that intelligence is a form of computational capacity to achieve objectives in physical or virtual world problems. Thus, the debate should pivot from comparing biological and artificial intelligence towards understanding the mechanisms through which any form of intelligence navigates and manipulates its environment to achieve desired outcomes, be it human, animal, or machine.

This approach to understanding the mechanisms of AI to define it is echoed in the document published by the European Commission's High-Level Expert Group on Artificial Intelligence, wherein it recommends updating the definition of AI by incorporating the scientific perspective and describes the current approach as a "*simple abstract description*".[64] The key mechanisms, notions and/or terminologies discussed in the said document include perception, reasoning and decision-making, knowledge and its representation and reasoning, planning, scheduling activities, searching, optimising, learning (machine learning, neural networks, deep learning and decision trees), speech and language understanding, computer vision, behaviour prediction, supervised learning, unsupervised learning, reinforcement learning, robotics, narrow AI, general AI, data bias, black-box AI, explainability, and goal-directed AI.[65] The said document suggests that this discussion be read together with the following updated definition for a more accurate definition of AI[66]:

> "Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans3 that, given a complex goal, act in the physical or digital

[64] "A definition of Artificial Intelligence: main capabilities and scientific disciplines | Shaping Europe's digital future", (18 December 2018), online: <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>.
[65] *Ibid.*
[66] *Ibid.*

dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.

As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (whi ch includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber -physical systems)."

The technical definitions of AI become important when they are imported into regulatory laws or policies, thereby impacting economic activities. Hence, the technical foundational understanding also helps in accurately determining liability aspects surrounding AI. Since the objectives of computer scientists and regulators in defining AI are different, we should not expect similarities between the definitions. For computer scientists and technologists, the goal in defining AI is accuracy and precision, aiming to capture the essence of AI's capabilities and limitations. However, the regulators are primarily concerned with meeting their policy objectives, which are aligned with societal values and political priorities. Political priorities can also vary significantly based on the jurisdiction. A jurisdiction where many AI developers exist could develop a tendency to foster innovation by providing developers with more freedom and minimal control and choosing to adopt a narrow definition of AI while formulating legal frameworks. In other jurisdictions where citizen rights are more

important and don't have a concentration of AI developers; however, they choose to adopt a broader definition to improve consumer protection. This variation in objectives underscores the complexity of creating a universally accepted definition of AI, highlighting the need for a nuanced understanding that considers both technological intricacies and the socio-political context. Nevertheless, the regulators must keep their definitions closer to reality, and these technical definitions can then serve as a starting point for the assessment of these definitions.

1.1.2 Legal Definitions of AI

The definitions used by government institutions for regulatory purposes differ significantly from those definitions where the objective is to define AI accurately and scientifically. The reason behind this is that the main objective of the government is not to provide a precise definition of AI but to meet the policy objectives in the most effective manner possible.[67]

This analysis of definitions will focus on the legal frameworks with respect to AI for the jurisdictions of the US and Europe owing to their leadership in technology, early regulatory efforts, economic impact and market size, and global influence, including through trade and diplomacy. The standards adopted in a significant market such as Europe led to the creation of de facto global standards and have even led to the coining of terms such as "Brussels effect" and "California effect" and show the ability of these jurisdictions to influence global regulations beyond its borders. The US is home to major AI developers, including Alphabet, Amazon, Adobe, IBM, Meta, Microsoft, OpenAI and Anthropic, which highlights the massive economic stake of the US in the AI industry. Europe has always been a trendsetter in setting regulatory precedents for the tech industry, notably with its early adoption of comprehensive

[67] The North American jurisdictions are about staying competitive in the AI industry and hence are preferring to stay agnostic in their frameworks, more discussion in the next chapter.

laws such as the General Data Protection Regulation (2016), Digital Markets Act Regulation (2022), The Digital Services Act Regulation (2022), and, especially given the context of this research, the Artificial Intelligence Act (2024).

The United States

In the United States, AI regulation is not exclusively for federal or state (provincial) jurisdiction. Hence, the initial tranche of legal frameworks specifically for AI, in addition to the application of existing laws, has been observed from both levels of government.[68] However, the federal legal framework is more important since the most important competencies, including interstate commerce and national security and defence, are within the federal jurisdiction in the US.[69]

The National Artificial Intelligence Initiative Act of 2020 in (15 U.S.C. 9401(3))[70] defines and expands the definition of AI in the following manner:

> "*The term "artificial intelligence" means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments.*" (for ease of analysis, referred to as "the first part") "*Artificial intelligence systems use machine and human-based inputs to— (A) perceive real and virtual environments; (B) abstract such perceptions into models through analysis in an automated manner; and (C) use model inference to formulate options for information or action.*" (for ease of analysis, referred to as "the second part")

[68] "The United States' Approach to AI Regulation: Key Considerations for Companies", online: <https://www.morganlewis.com/pubs/2023/05/the-united-states-approach-to-ai-regulation-key-considerations-for-companies>.
[69] In accordance with the interpretation of "enumerated powers" in Article I, Section 8 of the Constitution of the United States.
[70] *National Artificial Intelligence Initiative Act of 2020*, H.R.6216 (12 March 2020).

The definition, *prima facie*, seems very broad since as long as any type of machine is working on objectives set by humans to tackle real or virtual environments, it falls into this category. The second part of the definition elaborates a bit more about the functionality of AI systems and does not narrow down the definition in the first line. It is to be noted here that the objective of the National Artificial Intelligence Initiative Act is to, *inter alia*, bolster US leadership in AI research and development, including by improving private, defence and intelligence partnerships, to advance trustworthy AI, to promote AI engagement with allies of the US and to prepare the US workforce for integration with AI.[71]

The definition of AI in 15 U.S.C. 9401(3) is important because it is also being further relied on by various institutions of the government of the United States. Firstly, the Federal Trade Commission (FTC)[72], while creating measures to enhance oversight of evolving AI landscape, has adopted this definition and has stated that "*AI includes, but is not limited to, machine-based systems that can, for a set of defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments*".[73] Pertinently, while FTC borrows the key phrases of definition in 15 U.S.C. 9401(3), it adds the phrase "…*includes, but is not limited to…*" which keep the definition of AI open-ended and can virtually end up regulating any software, and herein I would argue that technical definitions of AI provide some guidance, especially if any dispute goes to the next level, such as an appeal before an adjudicatory body. FTC also uses the wide phrase "*...for a set of defined objectives…*" and omits the term "*human-defined*", which shows the intent of keeping the greater discretion within itself in the quickly evolving AI landscape.

[71] Dr Lynne Parker, *National Artificial Intelligence Initiative* (2022).
[72] The Federal Trade Commission (FTC) is an important independent agency in the US established statutorily through the Federal Trade Commission Act (1914) and is tasked with ensuring compliance with antitrust law and protecting consumers.
[73] "FTC Authorizes Compulsory Process for AI-related Products and Services", (21 November 2023), online: *Federal Trade Commission* <https://www.ftc.gov/news-events/news/press-releases/2023/11/ftc-authorizes-compulsory-process-ai-related-products-services>.

Secondly, this definition was also relied upon in Executive Order 14110 ("EO 14110") issued by the Biden administration on October 30, 2023, wherein the objective was to use AI for good while recognising the substantial risks, such as safety, security, and trust, and taking measures to minimise such risks.[74] One of the very interesting aspects of EO 14110 is the new reporting requirement, which is triggered when a certain threshold of floating-point operations ("FLOPs" [75]) is exceeded. The requirement of EO 14110 is that if the AI developer develops (or intends to) dual-use AI model training which exceeds the computing power of 10^26 FLOPs[76], the entity must report certain information to the US government, such as model training, testing, and data ownership. The reporting requirements are also in place for computing clusters, which refers mainly to the combination of numerous computing hardware to optimise computing power. The justification for this reporting requirement is that such AI models could have potential national security implications owing to their dual-use capabilities, i.e., civilian and military applications.[77]

Now, while the definition of AI in EO 14110 still relies on the broad definition mentioned in 15 U.S.C. 9401(3), which is much wider than and certainly includes the AI models trained well below 10^26 FLOPs; however, the reporting requirements do not apply to those models. Therefore, having a broad definition in a legal framework does not imply that a legal framework would have measures to regulate all AI. Interestingly, OpenAI's GPT-4 has been estimated to have been trained with around five times less than this threshold.[78] Hence, this requirement appears to be very narrow and doesn't place any serious burden on those

[74] The White House, *Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (2023).
[75] This is Not to be confused with floating-point operations per second, which is the rate of computing resources used/available per unit of time, whereas floating-point operations refer to the quantity of computing resources.
[76] Or 10^23 FLOPs when using primarily biological sequence data.
[77] Laurie Harris & Chris Jaikaran, "Highlights of the 2023 Executive Order on Artificial Intelligence for Congress (R47843)", (17 November 2023), online: *Congressional Research Service (CRS)* <https://crsreports.congress.gov/product/pdf/R/R47843/3>.
[78] Will Henshall, "Why Biden's AI Executive Order Only Goes So Far", (1 November 2023), online: *TIME* <https://time.com/6330652/biden-ai-order/>.

crossing this threshold since only disclosure to the US government has been mandated. It is to be noted, however, that the computing resources required to train AI models will not remain the same. It has been observed that, in the last decade, the computing resources required have doubled every six months.[79] It is hard to say if the rate of increase of computing resource requirement for training AI would remain the same; however, it is understandable to witness an increase in the same owing to the race between developers to build better AI models. Regardless, it could be observed that the definition would still remain narrow since the disclosure requirement applies only to dual-use AI. On the other hand, it might be tough to draw an exact line between merely civilian AI and dual-use AI. For an AI to not have any connection to military use requires building rather expensive technical safety measures (commonly referred to in the tech industry as "guardrails"), which would require extensive testing in various products and services, but the government may argue that if a company has the resources to spend resources to perform 10^26 FLOPs, it might also have the resources to build these guardrails.

The approach by the US legislature and executive, based on the analysis of the definitions above, is largely to define AI in the broadest manner possible, which could keep the door open for future regulatory measures. The definition of 15 U.S.C. 9401(3) does not evaluate AI by the Turing test standard of similarity with humans and only requires that to be classified as AI, it should be able to "…*make predictions, recommendations or decisions influencing real or virtual environments*…" "*… for a given set of human-defined objectives*". This definition is more in line with McCarthy's later views. Hence, the definition of AI in the US remains largely broad and flexible.

[79] *Ibid.*

I believe that the measures applicable to a much narrower subset of broadly defined AI, such as the reporting requirement of EO 14110, are not significantly cumbersome in affecting the dynamics of the innovation and adoption of AI. This is because there is no requirement to disclose the reporting to the public as well, and it is being monitored by the government for defence purposes, which is clear since EO 14110 derives significant authority from the Defense Production Act.[80] However, the definition adopted by the FTC is broad and could provide them with greater discretion and flexibility to promote competitiveness and also protect consumer interests.

The US seems to be adopting broader definitions and a flexible regulatory framework, which, based on how discretion is exercised, has the possibility to create an environment that encourages R&D in AI. Interestingly, AI-related lobbying surged by 185% in 2023, with new entities joining the lobbying process, including major AI developers.[81] It can be argued that the laissez-faire approach in narrowing the reporting requirement to only a subset of broadly defined AI is because of the massive presence of the Tech industry in the US, which results in lobbying for minimal regulation. However, this concentration of tech companies also encourages the laissez-faire approach since it aligns with the economic interest of the US, and hence, public policy favours it as well.

Europe

After extensive deliberations and negotiations, the EU legislature recently passed a comprehensive Artificial Intelligence Act on March 13, 2024, which the European

[80] House, *supra* note 74.
[81] Hayden Field, "AI lobbying spikes 185% as calls for regulation surge", (2 February 2024), online: *CNBC* <https://www.cnbc.com/2024/02/02/ai-lobbying-spikes-nearly-200percent-as-calls-for-regulation-surge.html>.

Commission had proposed on April 21, 2021 ("the EU AI Act").[82] The objective of this law is to ensure that AI systems marketed and/or used in the EU jurisdiction are "*safe and respect fundamental rights and EU values*", are fostered by investment and innovation, and that the EU can attain leadership by setting a global benchmark for AI regulation similar to its impact of GDPR.[83] While the proposed AI Liability Directive has been stalled in the EU, Article 2 (1) of the AI Liability Directive relies on the definition of AI as adopted in the EU AI Act.[84]

Hence, it is important to consider the latest version of the definition of AI in the EU AI Act, as adopted on March 13, 2024, in Article 3 (1), which defines AI as follows[85]:

"(1) 'AI system' is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments;"

This definition has a substantial resemblance to the previously discussed definition of AI affirmed by the US executive in the EO 14110; however, there are certain differences as well. The comparison of these two definitions has been done in the **Table 1** below:

Relevant portion of the US definition mentioned in 15 U.S.C. 9401(3) and relied upon in EO 14110	Relevant portion of the EU adopted in the EU AI Act	Comments

[82] *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS*, (2021).
[83] "Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world", online: <https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>.
[84] *Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)*, (2022).
[85] "Texts adopted - Artificial Intelligence Act - Wednesday, 13 March 2024", online: <https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html>.

a machine-based system that can … make predictions, recommendations or decisions …	*'AI system' is a machine-based system … that, for … objectives, infers, … how to generate outputs such as predictions, content, recommendations, or decisions …*	Both definitions describe AI in terms of its main functionality to receive inputs and generate outputs with very similar terminology.
…for a given set of human-defined objectives… *… use machine and human-based inputs to…*	*… for explicit or implicit objectives,* *…from the input it receives, …*	The US definition explicitly mentions that the objectives are set by humans, but the EU definition is silent in this regard making EU's definition broader than the US one. However, it is possible that the objective could also be categorized as just an input of the AI system, and since the US definition does mention machine-based inputs, it could make the scope of both the definition similar in practice during implementation. Inputs in both definitions seem to include both human and machine inputs, owing to silence in the EU's definition regarding the same.
…influencing real or virtual environments.	*…that can influence physical or virtual environments…*	Both the US definition and the EU definition mention influence on real /physical environment and virtual environment with an "or" between them, thereby keeping the doors of regulation more open.

One could argue that the automated analysis portion mentioned in the subhead (B) narrows the US definition as opposed to the "*varying levels of autonomy*" recognised in the EU definition. However, it can also be argued that the second part of the US definition is a mere expansion of AI's definition for explanation purposes and doesn't limit the scope of the definition set out in the first part and that the second part only delineates the general functioning of AI. Nevertheless, I believe the second part is still silent about the extent of automation, and

hence, the executive, judiciary, or legislature could benefit from this flexibility when required to make such a determination.

The EU definition incorporates the phrase *"may exhibit adaptiveness after deployment*", whereas the US definition is silent about this aspect. Adaptive AI, in general, is the ability of AI to adapt its outputs based on new data. Regardless, the use of the term "may" in the EU definition keeps the scope of both definitions the same.

Both definitions, i.e., the US and the EU, have significant similarities, as shown above. The approach of the EU legislature, like that of its US counterparts, also defines AI in a broad manner, which provides them regulatory flexibility. The EU definition also does not evaluate AI by the Turing test standard of similarity with humans and only requires that to be classified as AI, it should be able to "…*from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions…*".

Similar to the US approach, where the regulation was then carried out on a specific subset of AI[86], the EU AI Act also regulates AI based on its five subsets, wherein the subsets are primarily defined on their perceived potential risk levels, i.e., unacceptable risk, high risk, limited risk and minimal risk, and have a special category for the General-Purpose Artificial Intelligence ("GPAI")[87]. The overview of the regulation has been depicted in **Table 2** in a simplified manner as follows[88] [89] [90]:

[86] The disclosure requirement discussed above in the context of the US definition.
[87] Mohamad M Nasr-Azadani & Jean-Luc Chatelain, *The Journey to Trustworthy AI- Part 1: Pursuit of Pragmatic Frameworks* (arXiv, 2024) arXiv:2403.15457 [cs].
[88] *Ibid.*
[89] "High-level summary of the AI Act | EU Artificial Intelligence Act", online: <https://artificialintelligenceact.eu/high-level-summary/>.
[90] "Commission welcomes political agreement on AI Act", online: *European Commission - European Commission* <https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473>.

Risk Level	Minimal	Limited	GPAI	High	Unacceptable
Classification basis	Minimal risk	AI creates or manipulates content for humans.	AI built using 'Foundational Models' and may have inherent risks	AI can have negative impact on safety or fundamental human rights.	AI is hazardous to individuals.
Examples	Video games and spam filters (may change with generative AI)	Chatbots, deepfakes, biometric categorization, and emotion recognition systems.	Large Language Models. GPAI classified as having 'systemic' risk when training's cumulative computing is greater than 10^25 FLOPS.	Products otherwise subject to safety laws (toys, medical devices), non-banned biometrics, critical infrastructure, education, employment, essential services, law enforcement, and other public services.	Social scoring systems, biometric categorisation systems inferring sensitive attributes, and manipulative AI
Regulation	Unregulated.	Light transparency obligations: End-users must be informed that their interaction is with AI (and not a human). Labelling of AI generated data. End users should be made aware when biometric	All GPAI: a) disclose technical documentation, b) usage instructions, c) summary of training data used and d) comply with copyright law. Free and open licence GPAI (no systemic risk): only c) and d) above.	AI providers are required to: a) create a risk management system and a quality management system, b) data governance (error free and relevant data used), c) provide authorities with technical documentation to showcase compliance,	Prohibited

		categorisation or emotion recognition systems are deployed.	GPAI with systemic risk – in addition to a), b), c) and d), also required to: e) conduct model evaluation, f) conduct adversarial testing, g) track and report serious incidents and h) ensure cybersecurity protections.	d) design AI with record-keeping feature and to allow human oversight, e) usage instructions provided to downstream deployers so that they can comply as well, f) ensure accuracy, robustness, and cybersecurity.	

The EU's strategy of defining AI broadly in the EU AI Act and then further defining subsets of AI based on different risk levels to have tiered regulation is a rather interesting approach. On the one hand, it gives both consumers and AI developers more clarity regarding what to expect, which builds trust in the regulatory environment. On the other hand, the product or service based on AI that has been classified as Minimal Risk enjoys minimum scrutiny, and such classification can even encourage developers to colour their product or services to those lying in a different category than their main intended purpose. For example, a chatbot could be bundled with a video game to avoid scrutiny. Further, AI classified as High-Risk face cumbersome due diligence measures, even if they pose virtually no risk because of the strict classification. For example, it is possible that simple medical devices, such as those used for diagnosis at home by patients, could now have to undergo additional due diligence measures, which increases the cost of the devices and reduces accessibility. Interestingly, in such cases, the manufacturer could argue that the algorithms used therein are not AI; however, the

definition of AI is very broad and could include virtually any software if the regulators deem them to be. However, these issues appear whenever a new legal framework is adopted, and these possible anomalies can get ironed out in due course of time.

The US and the EU share the common theme of adopting a broader legal definition of AI as a starting point and then regulating only a subset of that AI based on their varying policy objectives. Hence, the policymakers are conscious that if their legal frameworks are too restrictive, their choices in defining AI could deter innovation in their jurisdiction, propel the AI developers to flee towards easier regulatory jurisdictions and even have a long-term economic impact. Making broad definitions along with narrow subset definitions allows policymakers some discretion to steer AI innovation in line with public policy in a more flexible manner.

1.1.3 Why create separate legal frameworks for AI?

AI is garnering increasing attention from policymakers worldwide. So, what is so special about it that makes it distinctive from traditional IT systems that policymakers are giving it such special attention, even making standalone statutes and creating new regulatory institutions to deal with AI? Why not just supplement the existing laws with a few additional provisions to deal with AI?

Firstly, with improvements in AI, it is now becoming highly autonomous. This increasing autonomy and the unsolved legal issue of attributing responsibility to AI, given its lack of intent, raises significant legal and ethical questions.[91] This is further complicated when we try to compare AI's actions to human actions.[92] Secondly, algorithm complexity is another

[91] Enas Mohammed Alqodsi & Dmitry Gura, "High tech and legal challenges: Artificial intelligence-caused damage regulation" (2023) 9:2 Cogent Social Sciences 2270751 at 8–9.
[92] *Ibid.*

serious impediment to the application of existing laws owing to their complexity and opacity, which makes it troublesome for legal institutions.[93] This is termed a "black box" issue wherein, owing to its complexity, it is difficult to ascertain how a decision was made by AI, and hence, certain regulators are pushing to improve explainability and transparency in AI.[94]

Thirdly, AI is also prone to bias and discrimination, especially if the training data is not free from bias and existing legal frameworks are not adequate to deal with such issues.[95] Fourthly, AI, and especially generative AI, has an extraordinary ability to automate tasks, at least the tasks that are repetitive and require less creativity, which has the potential to transform the labour market quickly and have a socio-economic impact on society. [96] Fifthly, since AI uses large datasets for its training, it poses unique challenges and risks to privacy rights to the extent that the policymakers now treat privacy law as a field of its own.

In view of these novel and unique challenges posed by AI, policymakers and regulators have two choices. They can either supplement the existing regimes by adding minor clauses to existing laws or they can create an overarching framework for AI itself. I would argue that the first option of supplementing the existing law would be a mere patchwork and would reduce clarity for both the public and the AI developers. A single unified, broad, and flexible regulatory framework could set basic minimum standards to ensure that AI is developed and deployed safely, ethically, transparently, and in an accountable manner. Pursuant to formulating such a legal framework, the existing laws can then be supplemented or amended to avoid overlaps and to fine-tune the larger public policy narrative.

---

[93] *Ibid.*
[94] "TechDispatch #2/2023 - Explainable Artificial Intelligence | European Data Protection Supervisor", (2 April 2024), online: <https://www.edps.europa.eu/data-protection/our-work/publications/techdispatch/2023-11-16-techdispatch-22023-explainable-artificial-intelligence>.
[95] Alqodsi & Gura, "High tech and legal challenges", *supra* note 91.
[96] Valerio Capraro et al, *The impact of generative artificial intelligence on socioeconomic inequalities and policy making* (arXiv, 2023) arXiv:2401.05377 [cs].

1.2 Current liability regime

For this thesis, the focus will be on civil liability. Criminal liability is a jurisdiction-specific issue, and since it is a public law where the wrong is against the society and not just one individual, it would also involve various questions related to criminal law and constitutional law, which are different from the research questions postulated for this thesis.

1.2.1 Key terms in tort law.

Let us discuss some of the key terms of tort law to enable the discussion; however, the detailed discussion regarding the standards of liability, including with respect to tort law and its relevance to AI-related harms, will be in the next chapter. Tort law is the branch of civil liability in which a person who is damaged by another's mistake (intentional or unintentional) can seek restitution. Such liability is extracontractual and, hence, this type of legal framework is extracontractual responsibility in jurisdictions such as France.

A "tort" has the following elements: first, an act or omission of an act; second, such an act or omission of an act results in an injury or harm to another; and third, it amounts to a civil wrong.[97] Injury and harm are similar terms; however, injury refers to *de jure* detriment to an individual's rights, and harm refers to the *de facto* detriment.[98] There are many types of torts; however, they are primarily classified into three categories, i.e., intentional, unintentional and strict liability type.[99]

With respect to AI-related harms, I believe that the most important tort is negligence. Negligence occurs if the four elements exist; as per the existing principles of the common law of torts, to make a party liable, the following conditions are required to be satisfied: firstly,

[97] "tort", online: *LII / Legal Information Institute* <https://www.law.cornell.edu/wex/tort>.
[98] *Ibid.*
[99] *Ibid.*

there should be a relationship/duty of care; secondly, there should be a breach of this duty of care; thirdly, the aggrieved must suffer an injury; and fourthly the said injury must have been cause by thus breach and this connection must have been reasonably foreseeable. [100] [101] The civil law regime draws liability from the principles of contract law; however, the proof of burden is similar in both civil law and common law regimes.[102] However, the premise for this thesis is that even after such existing torts principles are applied, there is no clarity regarding the standards of liability that will apply to the domain of AI.

The term "standard" in tort law context is often used to discuss the concepts of 'standard of care' or 'standard of conduct' of the person or entity who may have committed the tort and against whom the action could be brought (hereinafter the "alleged tortfeasor" or simply the "tortfeasor"). While referring to the 'standard of liability' in this thesis, I am referring to the threshold of law, which determines if the alleged tortfeasor (or also referred to as "tortfeasor" hereinafter for the sake of simplicity in the discussion) is liable to compensate the injured party or not, and, if liable, to what extent. The argument here is that this threshold of law is not standardised and is based on many factors, including expectations regarding 'standard of care' and 'standard of conduct' in accordance with the law, precedents, cultural and socio-economic nuances, extent of regulatory and statutory compliance by the tortfeasor, and policy and political climate. Pertinently, different standards of liability place different importance on these factors. For example, strict liability, also known as no-fault liability, is a well-defined standard of liability wherein inherently dangerous action (or inaction) is sufficient to make the tortfeasor liable without establishing any fault of the tortfeasor.

[100] *Ibid.*
[101] Forcier, Khoury & Vézina, "Liability issues for the use of artificial intelligence in health care in Canada", *supra* note 9.
[102] *Ibid.*

We can preliminarily observe that the main issue with existing tort law is that there is no uniform obligation on adjudicatory authorities to consider the extent of compliance with AI safety legal frameworks to determine the first step of relationship/duty of care in the assessment of torts. While one could argue that there is a broad obligation to consider all the relevant factors pertaining to facts and law in each case, there is no uniform legal framework regarding which factors should be weighed in more to reach the final determination regarding liability.

1.2.2 The European Union

The jurisdiction of the European Union (EU) always grabs attention when significant regulatory shifts are made, owing to the previously mentioned "Brussels effect", wherein, owing to its large market power and size, the EU regulation has a global effect on products and services. The Brussels effect puts an impetus on corporations to align their products and services in compliance with the strictest version of the regulation. This is so because the EU is one of the largest and most important markets and has a stringent regulatory tendency, which, coupled with the globalised nature of the world, makes the importance of EU regulation very significant. It is expected that the effects of Brussels effect on AI will also be significant. When applying the law in a member state, both member state-specific law and the EU-wide legal framework apply simultaneously. For the purpose of this research, we are going to focus only on legal frameworks designed by the institutions of the EU. Usually, the legal frameworks of the EU are broad and flexible, which allows enough wiggle room for member states to adapt the EU's legal framework according to their own needs and public policy while still providing a larger vision.

<u>The European Union – the general regime</u>

The general product-related frameworks may seem like obsolete legal frameworks when it comes to new technologies like AI; however, they are also applicable to AI, either

directly or indirectly. The current general product regulatory regime in the European Union (EU) can be divided into Product Safety laws and Product Liability laws.[103] The objective of both Product Safety and Product Liability laws is to ensure that consumers' interests are protected, to encourage businesses to follow the best possible practices to ensure the safety of society, and to improve trust in the marketplace to promote economic activities. The approach of product safety laws is to act as a safeguard that prevents unsafe products from entering the market. On the other hand, product liability laws help consumers obtain compensation from the erring business entity by defining the procedure, standard, and extent of liability of such business.

Product Safety Laws, such as the General Product Safety Regulation (GPSR), which will replace the previous General Product Safety Directive[104], have the objective of ensuring that only safe products are sold in the EU market.[105] The product safety laws apply to a product as a whole, wherein the AI is just one component out of many components within the product. GPSR applies to products where no other regulation exists. Where area-specific regulations exist, for example, toys, electrical and electronic goods, cosmetics, chemicals complements, etc.[106], GPSR complements those specific laws or regulations. However, in accordance with provisions (9) and (10) of GPSR[107], GPSR is not applicable to medicinal products, food, feed, and related products, which have their own niche regulatory framework. As such, before

[103] *AI liability in the EU and the US: stifling or securing innovation?* (2023).
[104] GPSR came into force on June 12, 2023, and will be applicable from December 13, 2024. The objective of the change was to modernize the law and ensure that consumers' interests were protected regardless of whether the mode of transaction was online or through traditional brick-and-mortar shops.
[105] "Consumer product safety", online: <https://commission.europa.eu/business-economy-euro/product-safety-and-requirements/product-safety/consumer-product-safety_en>.
[106] "Product liability and safety in the EU: overview", online: *Practical Law* <http://uk.practicallaw.thomsonreuters.com/w-013-0379?transitionType=Default&contextData=%28sc.Default%29>.
[107] *Regulation (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023 on general product safety, amending Regulation (EU) No 1025/2012 of the European Parliament and of the Council and Directive (EU) 2020/1828 of the European Parliament and the Council, and repealing Directive 2001/95/EC of the European Parliament and of the Council and Council Directive 87/357/EEC (Text with EEA relevance)*, OJ L 2023Legislative Body: CONSIL, EP.

launching the product into the market, the businesses need to comply with all the applicable regulations. However, in case a tort arises that involves AI, it is likely that the adjudicatory authorities will also consider the facts surrounding the extent to which the AI developers comply with these product safety laws as a factor to determine the existence of liability and/or the extent of restitution to be awarded in accordance with the local laws of that member state.

The product liability regime in the European Union is twofold in nature, i.e., it is comprised of member state-specific law and the EU-wide Product Liability Directive[108]. The Product Liability Directive ("PLD") was adopted in 1985 "*to harmonise the fragmented legal protection on damage caused by defective products*"[109]. The main feature of the PLD regime is that it adopts a strict standard for liability known as the no-fault liability system, wherein the producers are liable regardless of their fault or negligence, as long as the damage is caused by defective products and a causal link exists between the damage and the defective product.[110] Currently, PLD allows consumers to claim compensation for death, personal injury, or material damage caused by defective products above a certain threshold.

The proposed revisions in PLD aim to widen the definition of 'product', and it also widens the scope of parties that can be made liable than the old PLD.[111] According to the proposed amendments, PLD will cover both tangible and intangible products.[112] Hence, when amendments in PLD go through, PLD will also be applicable to products comprising digital manufacturing files, digital services and software, including AI systems.[113] Revised PLD

---

[108] *Supra note 103.*
[109] "New Product Liability Directive", online: <https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739341/EPRS_BRI%282023%29739341_EN.pdf>.
[110] *Ibid*.
[111] *Ibid* at 4.
[112] "All change to the EU's strict product liability regime", (25 November 2022), online: <https://www.taylorwessing.com/en/insights-and-events/insights/2022/10/all-change-to-the-eus-strict-product-liability-regime>.
[113] *Supra* note 109 at 4.

would also make 'development risk defence' uniformly available in all member states, wherein the 'economic operators'[114] would be exempted from liability if it is shown that either the product wasn't circulated by them or that the product wasn't defective when it was first circulated or that the defect couldn't have been detected at the time when the product was placed in the market owing to technical limitations at that time.[115] Article 8 of the revised PLD will, subject to trade secrets and confidentiality concerns, compel manufacturers to disclose necessary and proportionate information when an injured party presents sufficient facts and evidence to support the "*plausibility of the claim for compensation*".[116] This requirement could be onerous for AI developers since it will cause legal costs in an attempt to reduce the scope of necessary and proportionate information. Further, AI developers also have to realign their internal mechanisms and retask some of the staff, at least temporarily, to be able to pull out the requisite information. Pulling out information from AI might also require the AI developers to redesign the AI itself, which is not always the most efficient way, and the resulting AI is not as potent or cost-effective.

In addition, Article 9 of the revised PLD reduces the burden of proof of the injured party by including 'presumptions of defectiveness and causal link' into the law.[117] In accordance with Article 9 (2) of the revised PLD, the defectiveness would be presumed against the producer if the producer either fails to disclose information if necessitated under Article 8 (1) of the revised PLD, or if the product doesn't meet safety standards prescribed by the law ("*Union or national law*"), or if the damage results from a malfunction which can be deemed to be obvious.[118] In accordance with Article 9 (3) of the revised PLD, the causal link will be

[114] Economic operator is a much broader term used in the EU's PLD. For ease, the much narrower term "AI developers" has been used interchangeably in the context of this paper's research questions.
[115] *Supra note 109 at 7.*
[116] *Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on liability for defective products*, (2022).
[117] *Ibid.*
[118] *Ibid.*

presumed if either the damage aligns with the suspected defect or if complex technical or scientific issues make it difficult to prove the liability, for example, in 'black box' AI systems.[119] The manufacturers do have the ability to challenge these presumptions according to Article 9 (4) of the revised PLD.[120] Interestingly, Article 4 (6) (a) of the revised PLD regime also recognises "*medically recognised harm to psychological health*" as personal injury and, hence, damage. [121] Article 14 of the revised PLD proposes a very relaxed 10-year limitation period after the defective product was circulated in the market or a 3-year limitation period after the injured person became or should have become aware, which applies to the initiating of proceedings for claiming compensation for damage falling within the scope of this Directive.[122]

While the disclosure requirement under Article 8 does not make the standard of liability more stringent, at least in legal terms; however, it can change the notion of legal terms themselves. This is so because AI developers are now required to comply not only with Product Safety laws, but also with procedural laws that would elevate the 'standard of care' and 'standard of conduct' as higher requirements many years after the tort has occurred when more knowledge regarding safety has been gained. The AI developers can argue that the limitation period is so long that the only way to backtrack and find information years later after the product was circulated in the market is to change the processes, which makes the compliance beyond what is required under safety law, in anticipation of a lawsuit.

The revised PLD has also removed the minimum damage threshold of 500 euros, making it much easier for injured consumers to seek restitution.[123] This reversal of the burden

[119] *Ibid.*
[120] *Ibid.*
[121] *Ibid.*
[122] *Ibid.*
[123] "Defective products: revamped rules to better protect consumers from damages | News | European Parliament", (12 March 2024), online: <https://www.europarl.europa.eu/news/en/press-room/20240308IPR18990/defective-products-revamped-rules-to-better-protect-consumers-from-damages>.

of proof, upon proving plausibility, also significantly lowers the barriers for the injured consumers to raise liability claims; however, it can potentially increase the legal and operational costs for the AI developers. The next chapter provides a detailed discussion of how notional changes in key terms are also redefining standards of liability.

The European Union – the proposed AI-specific regime

In addition to the existing national and EU-wide laws governing general product safety and liability, the EU has proposed specific legislative measures regarding AI. The EU AI Act has been attracting a lot of attention in the technology sector for some time, and in a recent development, it has also been passed. The main requirements of the EU AI Act have been discussed in Table 2 above and are not repeated herein for the sake of brevity. The EU AI Act is largely a safety law that regulates the way AI-based technologies are implemented by classifying them into different groups based on their risk assessment, as mentioned previously. As such, based on the risk category classification, AI developers ('providers' is the term used in the EU legal framework) are required to design AI in a manner that ensures compliance and takes steps to showcase such compliance. Most of the risk requirements are in the category of high-risk AI.

The European Commission, in its report to the European Parliament, raised concerns that the national tort laws may not be adequate since they may be expensive and complex when the consumer is required to prove its liability claims, thereby leading to inadequate compensation[124]. Hence, the proposed AI Liability Directive ("AILD") was framed to overcome these concerns. The proposed AILD has the objective of providing victims of AI system-related harms with a safety net in a manner similar to that of other technologies. The

[124] *REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL AND THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics*, (2020).

AILD has links to the EU AI Act; however, as of now, the PLD regime is moving faster, and AILD is stalled. AILD mainly creates certain procedural measures to ensure that the injured party is not unnecessarily burdened with proving the tort when it doesn't have the capability to do so. Article 4 of the AILD creates a presumption of causal link in the case of fault, which can be done in many ways. One important way to show this causal link is in case the AI provided does not comply with a court order for disclosure or preservation of evidence under Article 3(5), which can be raised under Article 3 (2) by the claimant/ injured party if the injured party made all possible attempts up to a proportionate level and was unsuccessful in gathering this evidence from the defendant. The presumption can also be established in case of non-compliance regarding the duty of care of the alleged tortfeasor, as required as per the law at the EU level, especially the EU AI Act. This requirement has some similarities to the previously discussed Article 9 requirements of the PLD.

The proposed changes in PLD and above AILD provisions attempt to reduce the burden of proof on the end consumers. The European Commission, in its 2020 report to the European Parliament, noted that AI systems are complex, making it excessively difficult for the injured person to identify the liable entity or person and to prove all the necessary ingredients required to establish the fault, including difficulty in establishing the causal link between that fault/defect and the damage suffered.[125] This was considered in the briefing of the EU parliament in respect of AILD, which noted that a uniform liability regime is required to avoid legal fragmentation and to avoid scenarios where the judiciary is forced to interpret the general liability regime, which wasn't from the perspective of complex technologies such as AI.[126] This causes difficulty not only to the injured person but also to the businesses of AI developers, who

[125] *Ibid*.
[126] "Artificial intelligence liability directive", online: <https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI%282023%29739342_EN.pdf>.

face legal uncertainty.[127] This briefing also discusses the possibility where a duty of care is complied with in detail, and damage may still occur, in which case it might be difficult to determine who is at fault.[128] This briefing also notes a possible negative impact of safety laws and liability laws on innovation.[129] As such, AILD proposes to harmonise a fault-based liability regime with an additional invokable measure of presumption of causality, whereas PLD revisions harmonise a strict (no-fault) liability framework.[130]

[127] *Ibid*.
[128] *Ibid*.
[129] *Ibid*.
[130] *Ibid*.

Chapter 2: Standards of Liability and their application in respect of AI.

As discussed in the previous chapter, AI is usually characterised by its ability to, at least partially, mimic human cognition and make autonomous decisions. This ability allows entities to deploy AI and automate processes in manufacturing and services sectors and remove humans from decision-making processes, which raises novel challenges in applying the traditional liability frameworks uniformly in cases where injuries occur due to the application of AI. AI is a different type of intelligence than human intelligence and could be more accurate and efficient than humans in, say, repetitive tasks such as arithmetic computations, and at the same time, AI might fail in tasks that one-year-old humans can easily do.[131] Hence, even well-tested and well-trained AI could still be prone to errors, which would have otherwise been uncommon if human decision-making had been implemented instead. Specifically, Natural Language Processing AI is prone to hallucinations, biases, and ethical violations, and as such, concerns have been raised regarding their application in the medical field, wherein erroneous outputs can induce incorrect decisions which can be detrimental to patient safety.[132] Additionally, as mentioned in the previous chapter, AI may have characteristics such as the autonomous nature and opacity of algorithms, lack of intent, inclination to imitate the bias in its training data, the potential to impact the labour market and potential of privacy risks, creates challenges for the traditional liability framework. This, in turn, results in unpredictability regarding liability for both AI developers and consumers. A clear and well-defined legal framework of liability regarding AI would tackle these challenges and provide certainty, which would enable innovation and build trust among consumers.

[131] J E (Hans) Korteling et al, "Human- versus Artificial Intelligence" (2021) 4 Front Artif Intell 622364.
[132] Reza Khanmohammadi et al, *An Introduction to Natural Language Processing Techniques and Framework for Clinical Implementation in Radiation Oncology* (arXiv, 2023) arXiv:2311.02205 [cs].

In light of the challenges faced by traditional liability frameworks, this chapter looks at the existing liability law principles that have evolved in various domains[133] of law and compares them through analogical reasoning to assess the possibility of adopting such legal frameworks into AI. Each domain of law deals with certain social, economic, or socio-economic activities, which usually have their own peculiarities and history, based on which the governments regulate them. These varying peculiarities have resulted in the evolution of different liability-related frameworks that supplement traditional tort law and together form a specific framework of law for a certain domain of law. By examining these existing standards in different domains of law, this chapter would then assess if, in a fashion similar to EU safety law (the AI Act), these varying standards of liability can be implemented based on the risk assessment/criticality of a particular industrial sector.

It is crucial to understand that not all applications of AI are fully autonomous and that decision-assistance tools make up the bulk of applications of AI, especially in services which make critical decisions, such as those related to medicine, financial advice, and assisted driving (as opposed to fully autonomous driving).[134] In cases where the decision-making has not been transferred to AI, can the existing liability systems adapt as they have in the past and apply either directly or vicariously to relevant stakeholders, wherein AI is just another tool made by technological advancement? Or are there any specific challenges these existing liability systems face when applied to AI-related harms? What standard of liability would then be able to address these challenges?

[133] The laws are either general or formulated to regulate a specific (or a group of) economic activity that applies to specific economic, social, or natural sectors. For ease, we will refer to this as "domain" in this thesis.
[134] Selbst, *supra* note 22.

2.1 General liability systems.

Before we define and discuss the standard of liability, it is essential to discuss the evolution of the concept of civil liability and related terminology, including the standard of care and negligence. Civil liability arises for "wrongs" committed by a party against another party, which can be classified into different categories, i.e., claims in contract, claims in quasi-contract, breach of trust, and torts.[135] Issues related to claims in contract law arise when a legally binding contract has been breached, and compensation is being pursued in accordance with the stipulations of the contract.[136] On the other hand, quasi-contracts primarily concern obligations that emerge from the legal implications, such as a statute.[137] The breach of trust deals with the wrongs associated with vesting ownership under trust rules.[138] "Tort" is usually defined in a "negative" sense, and any wrong that is residuary after the said wrong could not be classified in any of the previous three categories.[139] Claims in contracts, claims in quasi-contracts, and torts have been largely developed under the common law courts.[140]

Types of liability laws and their relevance to AI

Which of these four categories of wrongs is more suitable for the wrongs being caused by AI? The answer, of course, would be that it depends on the jurisdiction and specific case. The potential wrongs would largely fall under the category of torts, along with some overlap with the quasi-contractual wrongs. This is because, firstly, contracts are likely to be drafted skillfully by the AI developers to ensure that contracts are drafted with blank indemnification clauses, which protect them against most liabilities since tech companies are likely to be in a

[135] J S Colyer, *A Modern View of the Law of Torts* (Elsevier Ltd, 1966) at 1–2.
[136] *Ibid.*
[137] *Ibid.*
[138] *Ibid.*
[139] *Ibid.*
[140] *Ibid.*

far stronger position.[141] Secondly, the wrongs based on trust law are irrelevant because of the lack of vesting in the case of implementation of AI, which eliminates the breach of trust category of wrongs.

Thirdly, the quasi-contractual wrong category will be highly jurisdictional and case-specific in nature. Historically, laws related to product liability in the US have been adapted and applied to the new age technologies, and hence, the product liability laws should also be applicable to AI-related harms covered under relevant laws.[142] As discussed in the previous chapter, there has been a surge in AI-related lobbying in the US, with new entities joining the lobbying process, including major AI developers, which will also be a factor in how the product liability law shapes up in the US. On the other hand, in addition to formulating AI-specific laws and amending the product liability laws (as discussed in the previous chapter), the EU has also recently formulated the Digital Markets Act Regulation 2022[143] and the Digital Services Act Regulation 2022[144] to rein in the tech giants. In addition to the EU AI Act and product liability laws, these two regulations will likely impose some quasi-contractual obligations on AI developers.

The remaining fourth category of civil wrong, i.e., torts, casts a much wider net and is crucial in respect of wrongs committed during the implementation of AI. The term "tort" originates from the archaic language of Law French, which was used in the Courts of medieval England, and simply means "wrong".[145] Yet, the term "tort" had been derived from the Latin

---

[141] "What Happens When You Click 'Agree'?", *The New York Times* (23 January 2021), online: <https://www.nytimes.com/2021/01/23/opinion/sunday/online-terms-of-service.html>.
[142] "Products liability law as a way to address AI harms", online: *Brookings* <https://www.brookings.edu/articles/products-liability-law-as-a-way-to-address-ai-harms/>.
[143] *Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) (Text with EEA relevance)*, OJ L 2022Legislative Body: CONSIL, EP.
[144] *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance)*, OJ L 2022Legislative Body: EP, CONSIL.
[145] J L R Davis & Rosalie P Balkin, *Law of torts*, 5th ed (LexisNexis Butterworths, 2013) at 3.

word “torquere”, which means twisted.[146] Hence, a tort is a wrongful act (or omission thereof) and depicts a figurative and literal “twisting” that injures another and legally entitles the victim to seek a remedy before the courts of law to “set things straight”. [147] While tort law casts a wide net owing to its negative definition, it still requires the victim to establish various essential elements to succeed in their tort claim.[148] This requirement varies based on the kind of tort action. Torts are broadly classified as falling under the categories of intentional, negligence, and strict liability, wherein negligence is the most common.[149]

Intentional torts

AI systems lack human-like “intent”; hence, the claims of intentional liability will struggle to succeed in the context of AI.[150] AI systems are designed in a manner so that they can be functional even in unpredictable situations, and to do that, they are trained to achieve objectives on a finite data set, which can result in unpredictable outcomes.[151] Therefore, AI does not always act with the same ‘intention’ or in a manner that was intended by the developer. Hence, holding the developer or operator liable for an error under intentional torts is difficult.[152] For example, when injuries result from AI decisions, such as an autonomous vehicle misinterpreting environmental data and causing a collision, it is generally accepted that the harm was not intended by the manufacturer or operator.[153] It does, however, lead to the question of whether AI is inherently risky and if strict liability is the most relevant category of tort in the case of applications of AI.

[146] John C P Goldberg et al, *Tort Law: Responsibilities and Redress* (Aspen Publishing, 2021).
[147] *Ibid.*
[148] This is usually the case under common law but subject to variations in specific jurisdictions. Civil law jurisdictions largely have the same ingredients, too, but with some variations.
[149] Cross & Frank B, *Business law : negligence and torts*, Great courses (Teaching Company, 2015).
[150] Thomasen, *supra* note 3.
[151] *Ibid.*
[152] *Ibid.*
[153] *Ibid.*

Nevertheless, intentional torts have a special place in the context of AI. Intentional torts provide remedies in cases of very particular categories of damages, and in some cases, intentional torts are the most relevant form of torts. For example, AI systems can be used intentionally to harm individuals by creating defamatory generative material or to collect and use private data maliciously.[154] The most common types of intentional torts include battery, assault, false imprisonment, trespass (to either land or chattels), and intentional infliction of emotional distress.[155] Each of these intentional torts has varying sets of elements of requirements required to establish such torts. However, they all have a common requirement in respect of intent. AI systems, especially black-box AI, make decisions based on their complex algorithms, and it may not be technically feasible to ascertain the source of a particular decision.[156] The remedy under intentional torts can also be useful since it requires a different set of elements than negligence and, also, has different procedural aspects in the courts. For example, in certain jurisdictions such as Canada, intentional torts have a lower bar of the requirement and proof of damages is not required to be established to obtain injunctions.[157]

Negligence

Negligence is a much broader remedy since, as opposed to the intentional torts wherein intention is required to be established to prove negligence, the duty and causation elements are required, which have a lower threshold than the mental capacity required to establish intention. To prove negligence, a victim must prove four ingredients: that the victim suffered an injury (the injury element)[158], the tortfeasor had a duty to exercise reasonable care not to cause the kind of injury that occurred to a class of persons including the victim (the duty element),

[154] *Ibid* at 4-5.
[155] "intentional tort", online: *LII / Legal Information Institute* <https://www.law.cornell.edu/wex/intentional_tort>.
[156] Yavar Bathaee, "The Artificial Intelligence Black Box and the Failure of Intent and Causation" 31:2 : Spring 2018 Harv JL & Tech.
[157] Thomasen, *supra* note 3.
[158] Injury suffered by the victim can be either related to property, bodily harm or even emotional distress.

tortfeasor breached such duty of care despite the reasonable foreseeability of possibility of such harm (the breach element), and such breach was an actual and proximate cause of the injury suffered by the victim (causation element).[159] [160] These elements are very important since they are usually assessed in most civil liability claims outside of negligence as well.

The flexibility of the fault-based negligence framework makes it a potential remedy for AI-related harms.[161] Since negligence law is applied by the courts under the common law system on a case-by-case basis, the courts will have the opportunity to use flexibility and interpret the elements of injury, duty, breach, and causation and develop the jurisprudence in a fair and effective manner. The duty of care arising from the negligence framework can be imposed on developers, operators, designers, and users, which would include selecting appropriate AI, monitoring, and maintenance.[162] The extent of the duty of the care industry required could depend on the industry practices, opinions of experts in AI, and soft law such as technical standards and government policy or guidelines.[163] This duty of care can create an incentive for rigorous testing and responsible usage. Establishing the causation element will be a complicated task since the harm will be required to undergo the tracing back process, which will be further complicated owing to the multiplicity of stakeholders involved in an AI system.[164] 'Reasonable foreseeability' is a key factor in the assessment of negligence in the modern practice of law.[165] However, in a bid to make AI systems more capable, they are made unpredictable by their very design, and their actions can be unforeseeable and could lead to unforeseeable harm.[166] These complications would make litigation, at least in the initial stages,

[159] Goldberg et al, *supra* note 146 c 2.
[160] It is to be noted that foreseeability analysis is done from the perspective of a reasonable person who is acting with moderate care, as will be discussed further later in this chapter. The foreseeability analysis provides insights into both duty and causation elements.
[161] Lim & Morgan, *supra* note 18 at 143–147.
[162] *Ibid*.
[163] *Ibid*.
[164] *Ibid*.
[165] Goldberg et al, *supra* note 146.
[166] Lim & Morgan, *supra* note 18.

in negligence matters related to AI-related harms likely to be associated with high costs owing to the requirement for specialised expertise.[167] The question that needs to be asked here is whether a liability framework designed by the legislative or executive would be a more cost-effective strategy at the cost of some loss of fairness.

It is prudent to discuss the duty element in negligence further since it will be key in our discussion related to the standard of liability aspect. Negligence is a flexible remedy that focuses on reasonableness and especially the standard of a reasonable person, especially regarding the duty requirement. The simple principle of negligence is that everyone has a duty of ordinary care to others, and if a person fails such duty, then the breach element is established.[168] However, there has been criticism regarding a shift towards risk levels[169] rather than staying with the original principle of duty of ordinary care[170] with a focus on moderate care and mutuality towards the needs of others.[171] In assessing what level of care should be considered moderate, the focus is on the conduct of the person or entity while keeping in mind the kind of person or entity it is.[172] A moderate level of care does not mean an optimal or extraordinary level of care; rather, it requires a person or entity to moderate risky activities and act reasonably carefully.[173] Mutuality is the principle of being fair-minded and sensitive to the real needs of others that can be anticipated by a reasonably prudent person.[174]

The principles of moderate care and mutuality keep the duty of care grounded to a person or entity who can likely make a difference and prevent damage, thereby achieving the objective of tort law to minimise harm and compensate fairly when harms occur. From the

---

[167] *Ibid*.
[168] Benjamin C Zipursky, "Reasonableness In and Out of Negligence Law" 163 U Pa L Rev at 2169.
[169] Relevant terminology being given priority, according to Zipursky, is "reasonable risk" or "unreasonable risk".
[170] According to Zipursky, the relevant terminology should have been "reasonably prudent person," "reasonably careful person," "reasonable person," or "unreasonable person."
[171] Zipursky, *supra* note 168.
[172] *Ibid* at 2157.
[173] *Ibid* at 2169.
[174] *Ibid* at 2160.

perspective of a person or entity, does adherence to moderate care and mutuality mean 'reasonably' ensuring that others are not adversely impacted by their action or inaction? Wouldn't the type of activity in question and its associated levels of risk be an important factor in determining the extent of levels of care required by the person or entity to avoid the possibility of adverse impact reasonably? Isn't there a strong link between the risk of an activity and the duty of ordinary care? Despite this link, Zipursky's concerns are valid because shifting to risk-based assessment takes away the individual's (person or entity) responsibility to evaluate and implement the necessary level of care to prevent harm.

The risk-based approach is already being adopted in the context of AI-related harms[175], and as such, the risk will also seep into the AI-related negligence regime. This is especially true since the judges, who are primarily responsible for laying down the jurisprudence under the case of tort law under the common law system, are not immune to politics and developments in policy and are even influenced by academic articles.[176] A risk-based negligence assessment for AI-related harms can reduce the flexibility of the negligence framework since such an assessment will tend more towards a one-size-fits-all approach than a case-by-case approach. This risk-based negligence assessment can make the assessment more mechanical and reduce the importance of the context and real-world dynamics between parties; however, it can provide quick remedies and make the framework more predictable, which would be conducive to the innovation and adoption of AI systems. However, assuming that the courts in any jurisdiction would entirely switch to risk-based negligence assessment without any legislative mandate would be a stretch. And, in case of a legislative mandate, the negligence-based remedy simply wouldn't exist, and the remedies would then be available under the quasi-contractual wrongs.

[175] In the EU through the legislative measures, and in the US through Executive Orders, as discussed in the previous chapter.
[176] Dan Priel, "The Indirect Influence of Politics on Tort Liability of Public Authorities in English Law" (2013) 47:1 Law & Soc'y Rev 169–198 at 30.

AI is a complex technology, and flexibility in liability is essential for ensuring the fair administration of justice; hence, negligence is a critical remedy for AI-related harms, especially in the absence of other legal frameworks. It is clear, however, that technical guidelines and industry practices are going to dictate the standard required for duty of care. Further, expert testimonies will be critical in determining liabilities for AI-related harms, specifically when assessing 'reasonable foreseeability' in AI, which will likely raise litigation costs. Overall, negligence provides a more flexible and fair way to address liability-related issues for AI-related harms; however, it may not be cost-effective and could make the overall framework unpredictable.

Strict Liability

Under strict liability, the tortfeasor is held liable for the loss of an injured party from an activity, regardless of fault or preventability of harm.[177] Unlike intentional tort and negligence, strict liability is a type of no-fault tort wherein fault is not required to be established.[178] In particular, strict liability is oblivious of the intent of the tortfeasor and the reasonable person standard and, hence, suspends the duty of care requirement.[179] Under strict liability, even if a person or entity takes appropriate care, they would still be liable in certain circumstances, including ultrahazardous activities.[180] The most critical factor in ascertaining liability under the strict liability regime is only the nature of the alleged conduct of the tortfeasor.[181] The three types of conduct that could incur strict liability are possession of animals known to be harmful, abnormally dangerous activities, and product liability.[182] For the

[177] Colyer, *supra* note 135 at 35.
[178] *Ibid* at 202.
[179] "Understanding the Interplay Between Strict Liability and Product Liability", online: <https://www.lexisnexis.com/community/insights/legal/b/thought-leadership/posts/understanding-the-interplay-between-strict-liability-and-products-liability, https://www.lexisnexis.com/community/insights/legal/b/thought-leadership/posts/understanding-the-interplay-between-strict-liability-and-products-liability>.
[180] Cross & B, *supra* note 149 at 4.
[181] *Supra* note 179.
[182] *Ibid.*

purposes of AI-related harms, the most relevant category of strict liability would be product liability; however, there still would be plenty of applications for the abnormally dangerous activities category. To understand the strict liability, let's first discuss its modern origins.

The landmark ruling in the case of *Rylands v. Fletcher* (1868)[183] was a pivotal point in the modern history of law, formally establishing the rule of strict liability in tort law. In this case, the tortfeasor had constructed a water reservoir on their land, which overflowed and flooded the neighbour's mine. The House of Lords ruled against the tortfeasors and held them to be liable despite the absence of negligence since such use of the land was categorised as a "non-natural use".[184] To establish liability in accordance with *Rylands v. Fletcher*, the essential requirements were that the use of land by the tortfeasor was non-natural, there was an accumulation of dangerous items like water or explosives, there was an escape of these items from the tortfeasor's land/control, and injured party's land suffered damage.[185]

The House of Lords, in its ruling in Rylands v. Fletcher, also laid out some possible defences to this strict liability, including an act of God and cases where the harm resulted from the actions of the injured party itself.[186] Further defences available in the strict liability framework are available when the unforeseeable act of a stranger causes harm, the act of tortfeasor was not non-natural owing to approval of the state, the injured party had consented to the non-natural use, and the hostile action of the enemy of the state.[187]

Strict liability removes the requirement to prove fault, which could be helpful, especially in the case of black box AI systems, wherein proving 'reasonable foreseeability' is not easy. Since it is the tortfeasor who is using a resource (land otherwise) in a "non-natural"

[183] *Rylands v Fletcher* (1868) LR 3 HL 330 (HL (Eng)).
[184] Colyer, *supra* note 135 at 202.
[185] *Ibid* at 202-203.
[186] *Ibid* at 202.
[187] *Ibid* at 203-205.

which makes it inherently risky, strict liability tilts the scales of justice away from the tortfeasor and towards the injured party, which creates incentives for the tortfeasor to take a more cautious attitude and raise their standards of care. Strict liability-based litigation proceedings also have lower financial costs and are quicker than negligence proceedings since the fault is not required to be proven. This is similar to how summary trials, which are based on some assumptions, are quicker and more cost-effective remedies in complex patent trials, which are being adjudicated on a techno-legal basis and require technical expertise. These assumptions, of course, come with a risk of resulting in unjust initial rulings. Thus, the strict liability framework faces criticism for being too harsh and sometimes even unjust. The primary reason for such criticism is that strict liability does not permit tortfeasors to avoid liability by proving faultlessness.[188]

However, it is crucial to note that there are varying degrees of no-fault liability, and the availability of these defences also varies according to the type of no-fault liability being adopted. Generally, no-fault liability can be classified into three categories: "strict tortfeasor liability", where the tortfeasor is liable, and the defence of victim fault is available to the tortfeasor; "absolute tortfeasor liability", where the tortfeasor is liable and where the defence of victim misconduct is not available; and "absolute victim liability", where the injured party is liable without any defence relating to tortfeasors wrongfulness.[189] The main difference between the strict liability framework and absolute liability, as noted in the landmark Oleum gas leak case by the Supreme Court of India, is the lack of defences in the absolute liability framework.[190] Further, absolute liability is associated with exemplary damages so as to have an additional deterrent effect when the environment is damaged by ultrahazardous activity.[191] The damages awarded in absolute liability have the purpose of not only compensating the

[188] Jules Coleman, "The Morality of Strict Tort Liability" (1976) 18:2 Wm & Mary L Rev 259 at 273.
[189] *Ibid* at 274.
[190] *M.C. Mehta and Another v Union of India and Others*, AIR 1987 SC 1086 (1986) (SC India).
[191] *Ibid.*

victims but also restoring the degraded environment.[192] Currently, based on the discussions of the previous chapter, we can see that the EU's product liability approach with respect to AI-related harms is heading towards strict tortfeasor liability wherein the presumption of liability is in favour of the victim/injured party, with developers having an opportunity to rebut such presumption, wherein the policymakers intend to balance out the disadvantage of inscrutability that victims face.

Having examined various traditional systems of torts available and their applicability to AI-related harms, we will now define various standards of liability based on our knowledge of liability law and existing regimes in different areas of law. Before we do that, however, it is necessary to understand the legal test of analogical reasoning to assess whether existing frameworks in other domains of law could be suitable for AI-related harms.

2.2 Analogical reasoning

Analogical reasoning is a comparative method where two items are compared by identifying their common shared traits and also identifying uncommon additional traits to assess if any new hypothesis can be discovered by this reasoning process about either of the items.[193] It is crucial that the additional trait(s) should appeal to intuition as being normally coexisting with the common traits or that the presence of additional trait(s) does not make two items dissimilar if the analogous condition has to be established so that the same conclusions can be drawn about the two items.[194] The weight of the traits is very crucial.[195] Hence, if the high-weight traits are common traits, it is easier to establish analogous conditions and a similar

[192] Aruna Venkat, *"Polluter Pays" Principle: A Policy Principle* (New York: Rochester, 2012) at 3–5.
[193] *Supra* note 45.
[194] *Ibid*.
[195] *Ibid*.

conclusion for both items.[196] On the other hand, if there are additional traits in both items and these traits are opposite of each other, a disanalogous condition can be established, and it sometimes allows us to draw opposite conclusions for the two items.[197] If it is difficult to ascertain whether additional traits exist in harmony or in conflict, the test of analogical reasoning will fail to provide any insights.

Usually, analogical reasoning is used in Court opinions, for example, to assess the relevancy of precedence in a particular case. However, analogical reasoning is also a fruitful mechanism when comparing a better-known existing legal framework with a yet-to-be-determined legal framework, wherein the existing legal framework would become the fixed point, and the yet-to-be-determined legal framework would be the premise of legal analogy.[198] Analogical reasoning includes a three-step approach.[199] In the first step, a "base point" for comparison is identified, which, for the current paper's purposes, would entail the identification of relevant existing liability approaches in other domains of law as the base point. Second, similarities and differences between the traits of the "base point" and the "instantaneous point" are analysed, where the instantaneous point would be AI-related harms; and finally, it is determined whether the two points/items can be treated in the same manner by balancing the weightage of similarities and differences, and then a conclusion could be drawn if the identified existing liability approach is useful to address AI-related harms.

Analogical reasoning has many advantages in the context of current research. It can help generate suggestions to fill gaps and lacunae in the legal framework by relying on existing legal frameworks in other domains.[200] Analogical reasoning is a creative process that allows

[196] Maciej Koszowski, *Analogical Reasoning in Law*, 1st ed (Newcastle upon Tyne: Cambridge Scholars Publishing, 2019) at 34.
[197] *Ibid* at 127.
[198] *Ibid* at 180.
[199] *Supra* note 45.
[200] Maciej Koszowski, "Analogical Reasoning in Statutory Law" (2017) 8:2 J of Forensic Research 372 at 372.

us to work from a familiar and clear position.[201] Sunstein argues that analogical reasoning makes conditions for consensus among people with different theoretical tendencies.[202] Analogical reasoning can adapt to new situations and include a variety of perspectives, potentially leading to more holistic and informed judicial decisions.[203] However, the ambiguities and fallacies persisting in the existing framework can also seep into our understanding developed during analogical reasoning.[204] However, one has to be careful while exercising intuition and weighing the traits to avoid making incorrect conclusions since it is a subjective test and is prone to errors.

2.3 What is a standard of liability?

Standard of liability should not be confused with the term standard of care. As discussed previously, in fault-based tort law and specifically under negligence, the term 'standard of care' is a critical element that determines the duty owed by individuals to prevent harm to others based on the 'reasonable person' standard and helps determine the extent of the duty of care required under negligence. While varying requirement levels of standard of care could indeed form a critical barometer in ascertaining the standard of liability, there is much more to the standard of liability than only standard of care.

In academic discussions, the standard of liability is not defined *per se* and has not been defined even in the famous Black's Law Dictionary (9th ed.).[205] The term standard of liability is usually discussed only through examples, particularly in terms of Intentional Liability, Negligence, and Strict Liability.[206] An article from Harvard acknowledges that there are

---

[201] "Criticisms & Defenses of Reasoning by Analogy in Law", online: <https://cyber.harvard.edu/bridge/Analogy/analogy4.htm>.
[202] *Ibid*.
[203] *Ibid*.
[204] *Ibid*.
[205] Bryan A Garner, ed, *Black's Law Dictionary*, 9th ed (St. Paul, MN: West Group, 2009) at 1535.
[206] "CED: An Overview of the Law - Torts: Principles of Liability | Westlaw Canada Portal", online: <https://www.westlawcanada.com/blog/insider/ced-an-overview-of-the-law-torts-principles-of-liability-808/>.

multiple possible standards of liability from which lawmakers can choose.[207] The article recognises prominent standards of liability as being negligence and strict liability and acknowledges one of the available variants as a strict liability with contributory negligence, but does not define what exactly is a "standard of liability'.[208] Even in journal publications, the term standard of liability is rarely explored and only referred to through examples of the same. For example, in their paper titled "In Pursuit of the Appropriate Standard of Liability for Defective Product Designs", Moylan only discusses whether negligence or strict liability framework is the appropriate standard of liability.[209] Similarly, Sachs also explored the tussle between negligence and strict liability frameworks by referring to these two as standards of liability.[210] The discussion on the liability standard is constrained to the frameworks within tort law. However, I argue that understanding the standard of liability has to evolve beyond that, especially since contractual and quasi-contractual norms will play a significant role in the case of new-age technologies, including AI.

As discussed before, liability primarily arises from these four sources: claims in contract, claims in quasi-contract, breach of trust, and torts. While the examples given in academic settings usually refer to tort law-based standards, we should not overlook other forms of liability when assessing the standard of liability. To define the term, it would be helpful first to consider the definitions of 'standard' and 'liability'. One of the two definitions of standard in Black's Law Dictionary defines a standard as "*A criterion for measuring acceptability, quality, or accuracy.*"[211] The term liability has been defined in Black's Law Dictionary as "*The*

[207] "The Bridge: Economic Analysis of Alternative Standards of Liability in Accident Law", online: <https://cyber.harvard.edu/bridge/LawEconomics/neg-liab.htm>.
[208] *Ibid*.
[209] Christina M Moylan, "In Pursuit of the Appropriate Standard of Liability for Defective Product Designs" (1990) 42 Me L Rev 453.
[210] Reynold M Sachs, "Negligence or Strict Product Liability: Is There Really A Difference in Law or Economics?" 8 Ga J Intl & Comp L 259.
[211] Garner, *supra* note 205 at 1535.

quality or state of being legally obligated or accountable; legal responsibility to another or to society, enforceable by civil remedy or criminal punishment".[212]

Since the definition of 'standard' in Black's Law Dictionary is general in nature, it refers to measuring "*acceptability, quality, or accuracy*", which are more relevant in a general context. Hence, we can replace these terms in the context of civil liability and define the term standard of liability as a set of criteria that can be used to measure if a party is legally accountable to another party. Therefore, the standard of liability is a set of criteria which assesses if the tortfeasor surpassed a threshold by their actions (or lack thereof), making them liable. We will explore different criteria as we assess different relevant standards of liability further in this chapter.

2.4 Good faith-based due diligence standard of liability.

The good faith-based due diligence standard of liability (or simply "due diligence standard") rides on the negligence framework along with some quasi-contractual elements. One of the best examples of this standard comes from banking law in India. Section 131 of the Negotiable Instruments Act, 1881 in India shields a banker from liability while receiving payment of the cheque.[213] According to this, when a banker receives a cheque payment and acts in good faith and without negligence, the banker does not incur any liability in case the title to the cheque proves defective.[214] At first glance, it might seem that this is a case of straightforward negligence standard or perhaps even imposition of a stricter standard than moderate care required under the negligence regime. However, this is not true, and reading the explanation and surrounding jurisprudence provides us with more insight into how safe harbour modifies the standard of liability.

---

[212] *Ibid* at 997.
[213] The Negotiable Instruments Act 1881 (India), s 131.
[214] *Ibid*.

The requirement under Explanation II of Section 131 was held to be a *sine qua non* by the Supreme Court of India (hereinafter referred to as "SCI") in *Pradeep Kumar and Anr. v Post Master General and Ors*.[215] Explanation II of Section 131 requires the banker to check the electronic image of a truncated cheque to verify "*prima facie*" genuineness and to check for any fraud, forgery or tampering "apparent on the face" of the instrument based on "due diligence" and "ordinary care". The usage of the terms "prima facie", "apparent on the face", "due diligence", and, above all, "ordinary care" clearly shows that the legislature intends to keep the standard of care at "ordinary levels", unlike the common law requirement of keeping it at moderate levels as discussed previously in Zipursky's assessment.

In Pradeep Kumar (supra), the SCI held that an assessment of whether the bank followed the rules or instructions was necessary; however, this assessment alone was inconclusive, and the general practice of the bankers would be crucial to determine whether there was any negligence.[216] It was also held that the bank bears the burden of proof that it acted in good faith and without negligence to avail the statutory defence.[217] The court noted the definition of 'good faith' as an act done honestly, as defined in Section 3(22) of the General Clauses Act, 1897.[218] Negligence will be established in circumstances when the banker ignores suspicion regarding the cheque and its ownership, for example, when the amount is very large or the credibility and identity of the customer are unclear.[219] Negligence is also established when the banker acts in contrast to the characteristics and mandate of the instrument itself.[220] The court also mentioned other relevant factors, such as contractual relationships once the bank

[215] *Pradeep Kumar and Another v Post Master General and Others*, 2022 SCC Online SC 154, (2022) (SC India).
[216] *Ibid*.
[217] *Ibid*.
[218] *Ibid*.
[219] *Ibid*.
[220] *Ibid*.

opens a bank account for a customer and accompanying quasi-contractual obligations arising from the law. [221]

The court noted that the standard of care required to avoid negligence has to be "*realistic and pragmatic*", and the statutory defence should not be watered down since that would be disadvantageous to the expansion of the banking business.[222] Therefore, the safe harbours modify the standard of liability wherein the banker needs to do the bare minimum and is only required not to ignore the obvious, which is well below the requirement of the moderate standard of care under the negligence regime. The court held that a microscopic examination was not ordinarily necessary unless facts raised a reasonable suspicion and that the banking officers were not expected to act like amateur detectives.[223] The court cautioned against making liability stricter than the statutory regime.[224] Further, the court noted that banking services have penetrated and are now widespread; therefore, the standard of care required is constantly evolving based on the changes in the general practice of the bankers, and even case laws have become obsolete with the evolution in the banking industry practice.[225] Hence, the usage of the phrase "negligence" in the statutory provision does not mean that the standard of care required is stricter than the negligence standard established on a precedential basis, but rather that negligence herein is being assessed not on a common law or precedential basis but on the basis of industry-specific practices.

Similar provisions exist in other jurisdictions as well. In the UK, Section 4 of the Cheques Act 1957 protects the banker from any liability if the banker acted in good faith and without negligence and processed the payment of an instrument despite a defect in the title of

[221] *Ibid*.
[222] *Ibid*.
[223] *Ibid*.
[224] *Ibid*.
[225] *Ibid*.

such instrument.[226] Section 175 of the Bills of Exchange Act in Canada also uses very similar language and provides a safe harbour on the basis of good faith and lack of negligence.[227] In New Zealand, the law protects bankers who acted in good faith and without negligence in a similar fashion.[228] However, this safe harbour is based on a good faith standard, is very specific to the banking industry and has been created to promote the banking system and make day-to-day banking seamless. It is to be noted that such a safe harbour is likely not to exist when liability is related to the field with very high stakes, such as in the case of medical negligence. For example, in *Helling v Carey*[229], the Supreme Court of Washington did not accept that mere abidance to widely endorsed clinical guidelines could exonerate the tortfeasor.[230] [231] However, this raised concerns in the medical community as to the standards of care required or if this will result in the strict standard of liability. Hence, pursuant to the *Helling v Carey* ruling, various states in the US proceeded to define their own standard of care, and as such, an interesting safe harbour was made in the state of Washington wherein, to succeed in their claim, the victims will be required to prove that the tortfeasor "failed to exercise a degree of skill, care and learning possessed by other persons in the same profession".[232] Nevertheless, in the medical profession, the courts are likely to take a tougher stance since it results in bodily harm.

[226] Cheques Act of 1957 (UK), s 4.
[227] Bills of Exchange Act, RSC 1985, c B-4, s 175.
[228] Cheques Act 1960 (NZ), 1960/17, s 5.
[229] *Helling v Carey*, 83 Wn. 2d 514, 519 P.2d 981 (Wash 1974)
[230] Ash Samanta, Jo Samanta & Michael Gunn, "Legal considerations of clinical guidelines: will NICE make a difference?" (2003) 96:3 J of the Royal Society of Medicine 133–138.
[231] In *Helling v Carey*, the victim was suffering from glaucoma, which is a medical condition wherein fluids do not discharge from the eye, and abnormal pressure builds on. This condition can be detected by a pressure test. However, the existing medical standard prescribed such a test only for patients who were older than 40 years. Since the victim was less than this age, the test was not administered for many years, and the doctors thought it was a case of irritation because of the contact lens. Glaucoma was finally detected; however, permanent damage had been caused by that stage. Tortfeasors were held liable, and the contention of tortfeasors that adhering to the medical standards was enough to establish that the requisite standard of care was applied.
[232] WestJEM, "The Standard of Care: Legal History and Definitions: the Bad and Good News", (24 February 2011), online: *The Western Journal of Emergency Medicine* <https://westjem.com/articles/the-standard-of-care-legal-history-and-definitions-the-bad-and-good-news.html>.

We can hence observe that the relevance of the guidelines in determining negligence varies depending on the activity, and exceptions have been created by the legislatures only with respect to activities legislatures deem worthy of promoting in favour of the larger public interest. Keeping this in context, we will need to apply analogical reasoning to see how fruitful such a regime could be for AI-related harms.

Base Point in this good faith-based due diligence regime such a case is a liability regime, wherein the traits are that the banker is not liable if, firstly, the banker acts in good faith, and secondly, the banker is not negligent according to current general banking practices and guidelines. This liability regime provides statutory safe harbours if the person or entity is abiding by the guidelines, thereby altering the negligence regime and lowering the overall standard of liability compared to the negligence regime since the standard of care required is lower. While the legislature provides a safe harbour to the bankers if the bankers carry out acts in good faith and are not negligent, in accordance with the practices in the banking sector, this safe harbour is significantly influenced by the acceptable industry practices, which the court can ascertain on the expert evidence basis and on different guidelines or due diligence norms.

Due diligence norms wound, in the case of banks, stem from guidelines of the relevant central bank of the jurisdiction, the guidelines made by the Basel Committee on Banking Supervision (Basel III norms is the minimum standard for internationally active banks currently[233]), the nature of the instrument itself, the contractual relation between the parties, including the bank, the bank's internal guidelines, and the instructions or manuals issued by the technology partner of the bank.

[233] "Basel III: international regulatory framework for banks" (2017), online: <https://www.bis.org/bcbs/basel3.htm>.

Table 3: Trait comparison chart of good faith-based due diligence liability framework and comparison with potential implementation as a liability regime for AI-related harms.

Standard of liability → *Trait ↓*	**Good faith-based due diligence (base point)**	**Hypothetical framework (instantaneous point) and discussion on suitability for AI-related harms**
Essential elements	Two essential elements are required to be complied by the tortfeasor to avail the safe harbour under this standard: 1. Good faith: The alleged tortfeasor acted in good faith (honestly). Black's law dictionary defines good faith (*bona fide*) as a state of mind, and includes honesty in belief or purpose, faithfulness to duty, reasonable fair dealing particular to a given commercial activity, and/or without an intention to seek unjust advantage.[234] Hence, this standard can also be termed as *bona fide* standard. Black's law dictionary defines bona fide as "*made in good faith; without fraud or deceit*".[235] Hence, the requirement of good faith pertains to the absence of negative intentions. This understanding will	1. Good faith: For an AI developer or an operator of AI system to comply with good faith requirement will primarily mean having honest state of mind while developing or deploying the AI systems. Acting without negative intentions would mean not creating or deploying AI systems that could harm others or unfairly disadvantage stakeholders. This requirement under the current standard would mean a focus on non-maleficence requirement under the ethical guidelines in respect of AI. Non-maleficence is a principle of "*doing no harm*", primarily related to medical ethics, and first used to be represented as *primum non nocere* (first, do no harm).[236] Hence, this good faith requirement, requires a very basic minimum threshold of not doing any harm and relates to not ignoring any obvious issues encountered from the perspective of the stakeholder such as developer of AI. In context of AI, non-maleficence requirement varies significantly from guideline to guideline

[234] Garner, *supra* note 205 at 762.
[235] *Ibid* at 199.
[236] "nonmaleficence", online: *Oxford Reference* <https://www-oxfordreference-com.proxy3.library.mcgill.ca/display/10.1093/oi/authority.20110803100237642>.

	form a crucial difference between the current good faith standard and the next standard, which shall be discussed in the later sections of this chapter. 2. Due diligence (ordinary care): The tortfeasor acted without negligence, in accordance with "ordinary care" standard as expected in the relevant industry (banking in this case). Rules and instructions provide guidance but are not conclusive. Microscopic examination not expected from the bankers unless facts raise a reasonable suspicion. Safe harbour is available (statutorily or otherwise) to the tortfeasor once it is established that they acted in good faith and exercised ordinary care in line with minimum due diligence required in a specific industry.	based on application and jurisdiction,[237] especially since "not doing any harm" is a very subjective phrase to begin with. Further, it is widely acknowledged that AI-related harms are unavoidable, and the best that can be done is to minimize the risk and establish a liability mechanism for cases when harm does occur.[238] Ethical guidelines for AI, while referring to non-maleficence, generally refer to improved safety and security measures with a focus that "*AI should never cause foreseeable or unintentional harm*".[239] The question, in case of adoption of this standard to AI, is that whether this foreseeability is a similar standard to that of negligence, or is it different, and, if yes, how different? The question then is where does base point standard mentioned in the left draw the line for us which can be imported to the AI-related harms. As per the base point standard, the terminology of "good faith" is relevant and as per its legal definition it is clear that this standard is defined in negative sense, i.e., absence of negative intentions and as such does not impose any positive burden. That is to say, the AI developer is not required to actively promote security and safety under the good faith requirement in this standard. Rather, a developer is only required to solve and not ignore any problems or issues that it may encounter during normal course of development and deployment.

[237] Anna Jobin, Marcello Ienca & Effy Vayena, "The global landscape of AI ethics guidelines" (2019) 1:9 Nature Machine Intelligence 389–399 at 394.
[238] *Ibid*.
[239] *Ibid*.

		Of course, acting unfairly with ill intentions towards any stakeholder would be outrightly unacceptable under this bare minimum standard as well. For example, intentional misuse via cyberwarfare and malicious hacking would be unacceptable.[240] 2. Due diligence (ordinary care): The tortfeasor acted without negligence, in accordance with "ordinary care" standard as expected in the AI industry. Hence, it would be expected that the AI developer and operators would, at minimum, comply with the laws and guidelines as issued by the state in the relevant jurisdiction as a quasi-contractual obligation. However, is there more that is required to be done to meet this due diligence standard to minimize AI related harms? Are optional guidelines part of the quasi-contractual obligations. Under the base point, the answer to these questions would be answered either in affirmative or negative based on whether this is being done in the industry/ service sector for which AI is being developed and deployed. The general AI model development cycle includes gathering vector data through sensors and other data through other sources such as databases, carrying out data conditioning, adopting and integrating relevant algorithms based on the objectives and available data, modeling through human-machine teaming to optimize solutions, evaluating the effectiveness

[240] *Ibid*.

		and accuracy of the proposed AI solution at every step of this whole process, and using this evaluation to improve every step of such process.[241] There are multiples standards that have been formulated for AI by the standards development organisations (SDOs) such as ISO and IEEE. For example, in the US, National Institute of Standards and Technology (NIST) has adopted a risk management framework (RMF) that is largely based on ISO/IEC 23894:2023 (Artificial intelligence — Guidance on risk management) and ISO 31000:2018 (Risk management — Guidelines).[242] [243] However, the frameworks being adopted by regulators are only discussing high-level principles like 'fairness', 'robustness', 'transparency' and 'right to recourse'.[244] For example, AI RMF 1.0 of NIST is a voluntary framework focusses on basic requirements of govern (cultivating a culture of risk management), map (recognizing context and identifying related risks), measure (assessing, analyzing, or tracking identified risks) and manage (prioritizing and acting to manage risks based on predicted impact).[245] However, SDO's are primarily led by the industry and have technical expertise and not necessarily social expertise and further there are

[241] Lukas Fischer et al, "AI System Engineering—Key Challenges and Lessons Learned" (2020) 3 Machine Learning and Knowledge Extraction 56–83 fig 2.
[242] "Crosswalk AI RMF (1.0) and ISO/IEC FDIS 23894 Information technology - Artificial intelligence Guidance on risk management", online: <https://www.nist.gov/system/files/documents/2023/01/26/crosswalk_AI_RMF_1_0_ISO_IEC_23894.pdf>.
[243] "What will the role of standards be in AI governance?", online: <https://www.adalovelaceinstitute.org/blog/role-of-standards-in-ai-governance/>.
[244] *Ibid*.
[245] *Supra* note 242.

		many challenges such as opacity and entry barriers such as membership fees which further limits ability of SDOs to form comprehensive socio-technically apt standards.[246] And hence, there is concern if SDOs are the best entities for creation of standards in respect of AI where ethics and technological components are inextricably connected.[247] Further, to get the technicalities correct, it would be required that standards are very specific in nature and cater to the specific type of AI and specific type of application of such AI. [248] Hence, AI RMF 1.0 deliberately remains agnostic in this respect and does not delve into minute technical details.[249] Further, political leaders such as Justin Trudeau are concerned about staying competitive in AI industry as a nation, are not inclined to formulate regulations with very specific requirements, and prefer to give the AI developers and deployers a free reign as to how to want to reach the broad objectives defined under the policy or legal frameworks.[250] For a standard to be considered to be accepted industry wide, it would also be required to be acceptable across different jurisdictions. However, geopolitics and local politics may not allow for cross-jurisdictional uniformity despite an attempt of collaboration between the US and the EU.[251]

[246] *Supra* note 243.
[247] *Ibid.*
[248] *Ibid.*
[249] *Ibid.*
[250] Hard Fork, "Justin Trudeau Wants A.I. For Good | EP 86" (07 June 2024), online (video): <www.youtube.com> [perma.cc/FQK8-485N].
[251] *Supra* note 243.

		In view of the above, it is fair to state that it is going to be extremely difficult to establish what are the best industry practices before the Court. While reliance on the standards can be done in a complementary manner with expert witnesses and regulatory requirements, as is done in the medical malpractice related matters [252]; however, in the base point of this standard there is a statutory safe harbour exception once the AI developers and deployers carry out the basic minimum required due diligence. Will it be difficult to draw the socio-technical threshold line regarding when exactly the required due diligence has been met to avail the safe harbour under this standard? Yes, however, from the perspective of AI developer or deployer, the "ordinary care" requirement under this standard only requires them to carry out the due diligence which, in addition to the requirements mandated under the law, are so widespread in the industry that they have an implied obligation to carry out the same. In absence of any widely accepted practice or standard in the industry, the good faith-based due diligence does not require AI developers and deployers to carry out any extra steps to avail safe harbour.
Operational mechanism	From the point of view of the banker, when the banker receives the instrument, the banker looks for any obvious concern on a *prima facie* basis	From the perspective of AI developer and deployers, they are simply required ensure that they act with an honest state of mind and do not create or deploy AI systems that they know could

[252] *Ibid.*

	and then runs through the checklist and prescribed procedure based on their training and guidelines. If no alarms are raised, the transaction is allowed. The banker trusts the established process and does not actively engage in microscopic fault-finding exercise.	harm others or unfairly disadvantage relevant stakeholders. In respect of due diligence requirement, the AI developers and deployers would carry out their tasks with ordinary care, comply with basic minimum legal requirements and will implement general development and deployment practices to minimize risk. They will not carry out any extra steps beyond the ordinary care exercised during data acquisition, data conditioning, adopting and integrating relevant algorithms, human-machine teaming, improving the AI model based on the evaluation of effectiveness and accuracy, and also implement a basic minimum risk management strategy to identify relevant risks in the specific context to minimize them based on priority. The basic minimum risk management strategy is, however, only required to be a part of normal evaluation which is done to technically improve and maintain quality standard of AI.
How different from negligence?	At first glance, this standard may look like a stricter standard for the tortfeasor than the negligence standard since both good faith and due diligence (lack of negligence itself), is required. However, when attention is paid regarding the requirement to only *prima facie* assess if there is any genuineness issue with the negotiable instrument and to proceed with "ordinary care" in absence of any suspicion. This is as opposed to moderate care in negligence standard.	In context of AI, the difference between "ordinary care" in absence of any suspicion on a prima facie basis practically means not taking extra steps beyond basic minimum requirements of risk management in a manner integrated with the regular development and monitoring mechanisms discussed above. The moderate care of negligence standard would require more active steps. Once the AI developer or deployer proves that they acted in good faith and carried out basic minimum due diligence are established in the

	The present standard of good faith-based due diligence has been established to promote banking system and for the bankers to not hesitate when processing instruments. Negligence on the other hand is more general and requires higher standard of moderate care. Negligence requires 4 elements to be established to establish liability, viz, damage, duty, breach and causation. On the other hand, once good faith and good faith-based due diligence are established in the current standard, the tortfeasors are already protected under the safe harbour and further elements are no longer required to be assessed.	current standard, assessment of duty, breach and causation would be rendered futile.
Variance factors and consistency	The courts use rules and instructions for guidance, but relies on expert evidence to assess current industry-specific practices, and then negligence is assessed on such basis. The industry specific practices can even trump the precedence of the court law as being obsolete since banking industry evolves dynamically alongside technology and acceptable practices also evolve concurrently.	The courts can rely on expert evidence in case of AI related harms under this standard as well; however, as discussed above it might be difficult to ascertain what acceptable practices are in very specific technical terms owing to various socio-technical challenges faced during development and adoption of standards by the SDO, high variance of technical processes based on specificities of type of AI and applications, and reluctance by the policymakers to delve into specifics to appear as more industry friendly. This ambiguity would rather lead in surprising effect of having lower requirements, in line with basic minimum mandated laws and RMFs.

Remedy	If the tortfeasor doesn't act in good faith or if the tortfeasor is negligent, the safe harbour is not applicable and the tortfeasor is left open to various challenges including liabilities based on negligence, contract, statute(s), breach of trust, criminal (if applicable) and other common law-based liabilities.	Similar to the base point analysis.
Advantages	The biggest advantage of this standard of liability is that it enables banking to be swift and adapt dynamically to evolving technologies. It also makes the standard of liability a dynamic one, and hence a banker would have to implement latest safety measures to avoid incurring liability. This strikes a balance between individual customer interest and larger economic goals according to public policy.	This standard can be very friendly to innovation and adoption of AI. It will create conditions wherein the developers would not hesitate while experimenting as long as they carry out their minimum due diligence while acting in good faith. The AI developer and deployers would be required to stay updated with common minimum RMFs which will bring additional dynamism, which would not only increase safety but will also make AI models more accurate, thereby raising the quality of products and services.
Disadvantages	Since the standard is dynamic, there is a degree of uncertainty regarding liability and the banks will be always required to plan additional funds, litigation resources, and extra documentation to safeguard its interests, which lowers the efficiency of the system.	Compliance with the laws and guidelines, as issued by the state in the relevant jurisdictions, could be very stringent and may make this standard very high to begin with. Further, the policymakers are only defining broad requirements in an agnostic manner without any specificity which could be broadly interpreted by the courts, as opposed to lack of due diligence

		requirement as argued above, which may render this standard vague.
Relevant domains	Banking sector, specifically during processing of a payments using negotiable instruments such as cheques.	As we can see in the base point, the standard is specifically being applied during the operational stage of banking, on a day-to-day operations where the harms suffered by victims are primarily monetary (property) in nature which are unlikely to immediately translate to bodily or mental harm. Hence, this standard would be very suitable for deploying. For development purposes, this standard would be suitable for low-risk activities. As a particular category (type and application) of AI is applied for a longer period of time, not only the risk would reduce but also it would make the due diligence requirement more and more clear, making this good faith-based due diligence standard highly suitable in such cases.
Public policy	Promote banking system and enable bankers to act efficiently to promote overall economy.	Promote innovation in AI, especially in low-risk AI where the balance shifts in favor of more freedom and basic minimum compliance.

2.5 The utmost good faith-based due diligence standard of liability.

The next level of standard that is more stringent than good faith (*bona fide*) is that of utmost good faith (*uberrima fide*) based due diligence standard (or simply "utmost due diligence standard"). The term and concept of utmost good faith originates in contract law with respect to insurance law.[253] In accordance with the utmost good faith concept, there is a duty

[253] Garner, *supra* note 205 at 368.

to disclose all the material facts. The reasoning behind such a requirement is that the knowledge regarding the relevant and material circumstances is primarily in the exclusive control of one party (usually the proposer/insured) and that it is impossible for the other party (usually the insurer) to obtain this knowledge on its own which is material to the contract (correct calculation of risk).[254]

To understand this standard further, let us understand the difference between misrepresentation and nondisclosure. Misrepresentation, in the context of insurance, is a situation where, in response to the questions posed to the proposer, the proposer gives incorrect response(s).[255] The variation in intention can create different types of misrepresentation, such as fraudulent, innocent, or negligent; however, they are all misrepresentations regardless.[256] On the other hand, non-disclosure is simply a lack of disclosure when no information has been volunteered by the proposer/applicant to the insurer, most likely because of the absence of a specific relevant question.[257] The intention behind non-disclosure of the information is an irrelevant factor; the relevant question is "*whether information is "material"*".[258] To decide the nature of the information, an inquiry needs to be made regarding whether a different premium has been decided if the information had been known to the insurer and whether the insurer couldn't have known the fact through surrounding circumstances.[259]

Failure to comply with this disclosure requirement under the *uberrima fide* standard makes the contract voidable.[260] The principle of utmost good faith or *uberrima fide* originated

[254] *Ibid*.
[255] Ray Hodgin, *Insurance Law: Text and Materials*, 2nd ed (London: Routledge-Cavendish, 2002) at 171.
[256] Garner, *supra* note 205 at 1091–1092.
[257] Hodgin, *supra* note 255 at 171.
[258] John F Dobbyn & Christopher C French, *Insurance Law in a Nutshell*, 5th ed (West Academic Publishing, 2015) at 471–474.
[259] *Ibid.*
[260] Francis Rose, *Marine Insurance: Law and Practice*, 2nd ed (London: Informa Law from Routledge, 2012) at 103.

in the eighteenth-century case of *Carter v Boehm*[261] in a bid to minimise fraud and promote fairness by introducing a narrow requirement of disclosure.[262] However, there are criticisms of the modern adaptations of this concept wherein the modern adaptations of utmost good faith go against the weaker party, i.e., the insured.[263] Mansfield noted that this requirement applies to both parties equally, although it has been rarely used reciprocally by the insured party.[264]

The utmost good faith requirement is different from the caveat emptor principle ('let the buyer beware') in general contract law, where the parties must merely act in good faith.[265] The utmost good faith imposes a higher standard of intentions. To extend the idea of *uberrima fide* to a standard of liability, the AI developers and deployers would be in a similar position to that of a proposer/applicant/insured, and the public and government would be the insurer. This is so because, in the case of AI-related harms, AI developers and deployers have an exponentially clearer understanding of their own AI systems than the general public, who, at best, have a rudimentary understanding of these systems. There is widespread recognition that lawmakers also do not understand what AI is.[266] In such a scenario, it is fair that AI developers and deployers bear the duty of utmost good faith.

It would be useful to understand the difference between the *uberrima fide* standard and the *bona fide* standard by comparing the legal terms of "actual knowledge" and "constructive knowledge". Actual knowledge is direct and clear knowledge where the person or entity is

[261] *Carter v Boehm* (1766), 3 Burr 1905 (King's Bench).
[262] R Hasson, "The Doctrine of Uberrima Fides in Insurance Law—A Critical Evaluation" (2011) 32 The Modern L Rev 615–637 at 633–634.
[263] *Ibid*.
[264] Hodgin, *supra* note 255 at 183.
[265] Jason David Strauss, *Uberrimae Fidei and Adverse Selection: the equitable legal judgment of Insurance Contracts* (2008).
[266] Cecilia Kang & Adam Satariano, "As A.I. Booms, Lawmakers Struggle to Understand the Technology", *The New York Times* (3 March 2023), online: <https://www.nytimes.com/2023/03/03/technology/artificial-intelligence-regulation-congress.html>.

expressly aware of the information of a fact.[267] [268] Constructive knowledge, on the other hand, is a legal inference wherein a person or entity should have known certain information by exercising reasonable care or diligence.[269] [270] Therefore, constructive knowledge does not require actual knowledge but merely implies an ability to acquire such knowledge. It is to be noted here that the concept of imputed knowledge implies a given party ought to know about another party's conduct owing to legal responsibility, a concept which is useful for vicarious liability.[271]

In the case of the good faith-based due diligence standard of liability, there was a reliance on a statutory safe harbour to shield the tortfeasor from the harshness of negligence in a limited manner, wherein the safe harbour was created by the lawmakers in the promotion of specific public policy. The good faith-based due diligence standard of liability, hence, evolves by lowering the standard of liability obligations (tortfeasor's perspective)[272] from the base level of the negligence regime through the safe harbour protection, for example, as discussed in the specific narrow case of payment processing of negotiable instruments in the banking industry. On the other hand, the utmost good faith standard of liability has the possibility of emerging in two manners. Firstly, the utmost good faith standard of liability is arrived at through certain case laws wherein the negligence tort law is applied in a manner that it has a higher/stricter standard of liability obligations (tortfeasor's perspective)[273] since the given specific circumstances require the tortfeasor to be more vigilant. Secondly, the utmost good faith

[267] Garner, *supra* note 205 at 950.
[268] "Intel Corp. Investment Policy Committee v. Sulyma", (26 November 2019), online: *LII / Legal Information Institute* <https://www.law.cornell.edu/supct/cert/18-1116>.
[269] Garner, *supra* note 205 at 950.
[270] *Supra* note 268.
[271] Garner, *supra* note 205 at 951.
[272] This makes the requirements to avoid liability less stringent for the tortfeasor. In other words, the threshold/standard of the liability has been raised, making it more difficult for the victim to obtain compensation from the tortfeasor.
[273] This makes the requirements to avoid liability more stringent for the tortfeasor. In other words, the threshold/standard of liability has been reduced, making it easier for the victim to obtain compensation from the tortfeasor.

standard of liability can be hypothetically arrived at in cases where normally strict liability would have applied (specifically in terms of environmental law); however, safe harbours lower the standard of liability obligations (tortfeasor's perspective). While this safe harbour is very rare, some exist in the form of shielding the strategic projects of the nation from excessive liability, particularly in the case of nuclear liability, where the projects would struggle to take off if the operators didn't have assurance from the state regarding this limited safe harbour. For example, the Civil Liability for Nuclear Damage Act of 2010 in India sets maximum monetary limits for the operators to be held liable.[274] Further, sovereign functions of the government would also be immune to the strict liability standard, for example, damages arising in case of war or nuclear testing. The source of the doctrine of sovereign immunity is based on the British common law system, and different jurisdictions have enacted laws to limit and define the contours of this overreaching immunity.[275]

In *Andrews v United Airlines*[276], the case was regarding a serious injury suffered by the victim when a briefcase fell from an overhead compartment of the aircraft cabin. In this case, the court held that since the airline was a common carrier, it had a duty of utmost care and required the airline to adopt "vigilance of a very cautious person towards its passengers". The airline had provided warnings in their arrival message and had claimed that this was sufficient. The airline further contended that additional safety measures would raise the cost and inconvenience the passengers. Without deciding whether the warnings were enough to safeguard the safety of passengers, the court reiterated that the airline had a heightened duty by virtue of being a common carrier and remanded the matter to the jury. This case further showcases that the factors of the relationship between the parties, the services being provided,

[274] The Civil Liability for Nuclear Damage Act 2010 (India).
[275] "sovereign immunity", online: *LII / Legal Information Institute* <https://www.law.cornell.edu/wex/sovereign_immunity>.
[276] *Andrews v United Airlines*, 24 F (3d) 39 (9th Cir 1994).

and the potentiality of the degree of harm are very important factors in deciding the standard of care and, hence, the standard of liability.

But above all, the most significant factor that modifies the standard of care and elevates it to the "utmost care" is the degree of control over the prevention of harm. This is the reason why the nineteenth-century cases set the standard of "utmost care" for railroads in the US, specifically in terms of roadbeds, cars, and machinery, wherein railways had complete agency, and passengers didn't have any control.[277] Hence, in cases where the passengers had some control over safety, the standard of care could not be set at "utmost care"; for example, lack of shovelling of the snow along the stairs leading to the train station.[278] It is interesting to note that the term "utmost" is also being used in insurance law under the *uberrima fide* standard, wherein the degree of control on disclosure of information is largely with the proposer. Hence, the law is doing a balancing act by requiring utmost care from the party who has gained a very high degree of control. Hence, this monopoly of one party over the degree of control forms the basis for the utmost good faith standard of liability and raises the standard of liability obligations (tortfeasor's perspective) above the negligence standard. Nevertheless, liability under the utmost good faith-based due diligence standard can still be avoided if the tortfeasor acted in an utmost good faith-based diligent manner (or simply "utmost diligent manner").

[277] Richard A Epstein, "Vicarious Liability of Health Plans for Medical Injuries" (2000) 34 Val U L Rev 581 at 585.
[278] *Ibid*.

Table 4: Trait chart of utmost good faith-based due diligence liability framework and comparison with potential implementation as a liability regime for AI-related harms.

Standard of liability → *Trait ↓*	**Utmost good faith-based due diligence (base point)**	**Hypothetical framework (instantaneous point) and discussion on suitability for AI-related harms**
Essential elements	The most critical factor to trigger this standard of liability is asymmetry in the degree of control to such extent that law tends to impose obligations on the potential tortfeasors beyond the moderate/reasonable care in negligence standard. If one party has a very high degree of control over the instrumentalities which can lead to harms, and the victim has very low control over such instrumentalities, the requisite standard of care rises to the standard of "utmost" to balance out this inherent unfairness. In this standard the tortfeasor is under constructive notice and cannot take refuge if tortfeasor lacked actual notice about the possibility of the harm arising from such instrumentality. However, the tortfeasor can still avoid liability by acting in an utmost diligent manner, wherein the tortfeasor carries out the minimum prescribed due diligence (good faith-based due diligence), then also addresses the issues the tortfeasor notices it	For an AI developer or deployer to comply with utmost good faith requirement will require firstly ensuring compliance with all the regulatory requirements specific to AI. While in case of good faith-based due diligence standard complying with local laws would have been enough, the utmost good faith would require that if the AI tool is being implemented in different jurisdictions, the AI system should have largely uniform guardrails in a manner that they comply with most stringent regulatory regime even in jurisdictions where stringent norms do not exist. This is simply because utmost standard would expect that the potential tortfeasors apply all the guardrails, they already have access to. The AI developers and deployers would also be expected to have constructive notice of relevant technical standards of SDOs such as ISO and IEEE (such as ISO/IEC 23894:2023 and ISO 31000:2018), and in addition would also be expected to be aware of voluntary guidelines such as AI RMF 1.0 of NIST, so that they can formulate best risk minimization strategies. Further, the AI developers and deployers will also have a constructive notice regarding ethical

	encounters in the normal course of its actions with moderate care (negligence standard), and then make extra steps to take all the precautions which are physically feasible, even if such measures would make less economic sense. Hence, had the medical practitioner conducted the pressure test *Helling v Carey* in abundant caution even though the same wasn't required under medical standards, and if the damage still would have occurred after such testing, the medical practitioner couldn't have been held liable.	guidelines. The AI developers and deployers would, hence, under this standard need to make their best efforts to promote transparency, explainability, justice, fairness, non-maleficence, responsibility and accountability, Privacy beneficence, freedom and autonomy, trustworthiness, sustainability, dignity and solidarity.[279] As discussed previously, the general AI model development cycle includes gathering vector data, carrying out data conditioning, adopting and integrating relevant algorithms, human-machine teaming, evaluating the effectiveness and accuracy and thus improving every step of such process.[280] The AI developer and deployers in this case would do their best efforts to ensure security and safety by also carrying additional research and development during the evaluation step to ensure that the requirements of technical standards of SDOs and ethical guidelines are being addressed. They would also not ignore any concern that any personnel might have and would create independent bodies such as ombudsman to ensure that engineers don't hesitate from being truthful and do not suffer any repercussions from highlighting any issues. The discovery of issues by their personnel can even be linked to performance-based incentive as a reward system for improving the safety of the AI system. This evaluation process should be kept alive even after the development of AI model is complete and

[279] Jobin, Ienca & Vayena, *supra* note 237.
[280] Fischer et al, *supra* note 241 fig 2.

		there should be systems in place where the knowledge base of developers can be transferred to deployers in case the development team has to shut down owing to cost constraints. While it may be extremely difficult to establish whether an AI developer or deployer fulfilled their utmost duty of care or not before the Court. However, having R&D teams for safety and ethics, and giving them full autonomy and also protection from corporate pressures could be indicative of compliance with utmost due diligence. Some of the tech companies already have such teams, for example, Google's DeepMind Ethics & Society[281] and Microsoft's AI and Ethics in Engineering and Research (AETHER) Committee [282] are already making strides in this regard. Microsoft has also disclosed some details regarding workings of this committee to the public and other AI developers can learn and improve upon their framework. Another crucial aspect of utmost due diligence would be establishing robust communication channels not only internally within the tortfeasor's organization but also between stakeholders, including processes for receiving feedback from the stakeholders including public, government and end users. Further, importantly, formulating end user manuals and guidelines with sufficient warnings

[281] "Why we launched DeepMind Ethics & Society", (3 October 2017), online: *Google DeepMind* <https://deepmind.google/discover/blog/why-we-launched-deepmind-ethics-society/>.
[282] "Responsible and trusted AI - Cloud Adoption Framework", (28 July 2023), online: <https://learn.microsoft.com/en-us/azure/cloud-adoption-framework/innovate/best-practices/trusted-ai>.

		are also crucial. It is critical to provide training by other mechanisms to different stakeholders when the technology or its interface is complex and manuals are not enough. Regardless, the onus of proving that the potential tortfeasor had acted in line with the requirements of the utmost good faith-based due diligence would rest on the potential tortfeasors themselves. Thus, above recommendations are only suggestive in nature and would vary on case-to-case basis depending on industry, economics, jurisdiction related political aspects, technical norms and socio-cultural norms.
Operational mechanism	If the potential tortfeasor finds itself in a position wherein it has asymmetrically high degree of control over the instrumentalities, it should make best efforts to first discover and understand all mechanisms which can result in harm to any stakeholder in the relevant economic activity, such that the actual knowledge does not have least possible disparity between constructive knowledge and actual knowledge. Next, the potential tortfeasor should act and establish mechanisms to minimize the identified risks in a vigilant and very cautious manner, even though such process may result in higher costs.	From the perspective of AI developer and deployers, they are required to make additional voluntary steps beyond what is required by the regulations and even after addressing the issues they observed in normal operations. Detailed discussion done above. Inspired by test of inventive step assessment adopted in India in the domain of patent law in *F. Hoffmann-La Roche Ltd. and Ors. v Cipla Ltd.*[283], I propose a following a five-step test to assess if the utmost due diligence requirements were met by the tortfeasor to avail the safe harbour under this standard, which is 1. Identify the characteristics of the domain and the skillset of technically, legally, and ethically adept persons relevant to current AI system.

[283] *F. Hoffmann-La Roche Ltd. and Ors. v Cipla Ltd.*, 2016(65) PTC 1 (Del), (2016) (HC India).

	If potential tortfeasor fails to do these steps, it would not be able to claim defense of utmost due diligence, in case of liability claims.	2. Identify the due diligence that was already being done based on the other less stringent standards of good faith-based due diligence and negligence. 3. Ask the question, "What more the AI developer or deployer could have done to avoid the kind of AI-related harm that occurred in the particular case". If the answer is nothing, then utmost due diligence requirements are met. Otherwise proceed to next step. 4. The query at this step is whether the tortfeasor had a constructive notice that the additional due diligence measures identified at Step 3 could have prevented the harm, even if it would have meant higher costs. If tortfeasor implemented these additional due diligence measures and harms occurred anyway, the tortfeasor is not liable. If tortfeasor didn't implement the additional due diligence measures, the test proceeds to the last step of the analysis. 5. The next and final question is if these additional steps were physically feasible to have been taken by the tortfeasor with the state of technology at the time when AI related harm occurred. This step is carried out to eliminate the hindsight perspective, since the technology may have evolved by the time the dispute enters the courts.

		This test makes the subjective test, as objective as possible and crucially also avoids the unfairness to the tortfeasor from the possibility of hindsight analysis.
How different from negligence?	This standard is a stricter standard for the tortfeasor than the negligence standard and less strict standard than strict liability standard as discussed previously.	In context of AI developers or deployers, this test poses much higher obligations if they wish to avoid liability.
Variance factors and consistency	The tortfeasors can use rules, instructions, guidelines (binding or otherwise), industry-specific expert evidence, and information regarding internal mechanisms in a bid to prove to the court that the best practices were implemented with utmost care. However, all the mechanisms in such discussion would only establish a base line. It would be subjective analysis every time to establish that extra steps could or couldn't have been done to prevent harm. For example, in *Helling v Carey*, carrying out the pressure test would have been enough. However, if the pressure in eye turned out to be normal and the disease was being caused by some other condition, and the test required to find that third possible out wasn't available in the same hospital, it could be argued that such a requirement	In case of AI, utmost due diligence would require compliance with binding regulations and laws, along with the adoption of technical standards of ISO and IEEE, ethical guidelines, and RMFs. This can be done preferably through independent and dedicated teams made to ensure safety and ethics. However, there is no one correct way to establish the utmost due diligence from the perspective of the tortfeasor and showing that proactive steps were taken beyond good faith based due diligence and negligence standard, would be helpful. These additional proactive steps would not include compliance with binding regulations and laws because the same would be expected in the less stringent standards already. The extra proactive steps would then be assessed based on subjective analysis.

	would go beyond utmost good faith standard. Even in such case, medical practioners could still recommend their peers when they are not able to fully diagnose a condition, as is the normal practice and would proactively satisfy the jurisprudence in Washington wherein, as discussed above, it would be statutorily enough for a tortfeasor to prove that it had exercised a degree of skill, care and learning possessed by other persons in the same profession, to seek the safe harbour. Hence, this analysis of defining when does is utmost due diligence requirement met, does remain very subjective unless the lawmakers or policymakers or a precedent step in and define the exact contours.	
Advantages	Provides recourse to victims where negligence standard would have denied them the compensation. In line with the principle *ubi jus ibi remedium* or "where there is a right there is a remedy".[284] The utmost due diligence standard creates incentives for the tortfeasor who has significant degree of control to	The increased trust regarding safety makes it easier to convince stakeholders to adopt new technology in areas where there is a hesitation. While this standard is stringent and can seem vague, it will create conditions encouraging to use AI where previously the only option was to apply strict liability. For example, product liability, and environmental law related liabilities. For example, AI would face less resistance to be

[284] "ubi jus ibi remedium", online: *Oxford Reference* <https://www-oxfordreference-com.proxy3.library.mcgill.ca/display/10.1093/oi/authority.20110803110448446>.

	exercise it in the interest of the society, and thus balances the scales of justice by imposing higher burden on the entity which has more degree of control. This standard of liability builds trust between contracting parties and also in general public, which would make them more open adoption of new technology. Ensures that the party with greater degree of control takes extra care to ensure safety of others.	applied in cases of dangerous activities such as handling ultrahazardous materials. The AI developer and deployers would not only stay updated with industry practices but would also have to constantly innovate since this standard requires taking extra steps, which would increase safety and productivity of society.
Disadvantages	This test is open to hindsight analysis in case when harms do occur. For example, in *Helling v Carey* the action of the medical practitioner may seem reasonable by moderate care standard because not only the doctor was following the guidelines but also the statistics favored the decision of medical practitioner, wherein only one in 25,000 people below the age of 40 years was known to be prone to glaucoma. However, it can be still argued that this standard is still lower than strict liability standard which could have otherwise applied since this standard is being applied in cases where the harms are severe and the	Compliance with this standard can be very difficult for small AI developers and deployers, specifically because this standard requires compliance with additional unknown due diligence requirements and the only way to comply is to invest more resources in terms of material, money, and talent, which makes it expensive and difficult for smaller developers. However, it is still better to have a diverse set of standards and perhaps the small developers can be promoted in other ways by the state. Adopting this standard over negligence standard can open the floodgates of litigation against tech companies whenever a new safety measure is developed, and it wasn't applied beforehand. However, such concerns can be addressed using the five-step test suggested above.

	The standard is dynamic, and the obligations of the potential tortfeasor will also be vague. It will require support from lawmakers to make this standard more well defined.	On the other hand, it can be argued that this standard of liability can open the doors for AI developers or deployers to avoid strict liability. This concern can be addressed by better and transparent decision-making by policymakers while deciding which domains should be selected for a particular standard of liability.
Relevant domains	Since the term "utmost good faith" already exists in scheme of insurance law, it is already applicable in therein. Further, as seen in *Helling v Carey* case, this doctrine is also applicable in case of medical malpractice. Similarly, it is also applicable for railroad and airline industry, for being common carriers, as discussed above in respect of *Andrews v United Airlines*. This standard can be further applied to domains of environmental law and product liability by lowering the standard of liability from strict liability to achieve a promote innovation, without sacrificing safety.	Since, in case of AI, there is likely to be a greater degree of control with AI developers, it is an apt standard for AI. However, that doesn't mean that this standard can be used for AI applications in all domains of applications, for example, the technologies recognized as minimal risk under the EU AI Act, such as video games, might benefit from a less stringent standard. The applications of AI in the industries mentioned in the left-hand side can benefit from this standard.
Public policy	This standard encourages entities with greater degree of control to take extra measures with abundant vigilance to protect public safety and build trust in the law. May also help encourage bring requisite changes in certain risky economic activities (primary,	Enables application of AI in high-risk categories of domains, rather than banning these applications outrightly.

	secondary and tertiary) at a more rapid pace which were previously conservative to the risky changes.	

2.6 No-fault standard of liability.

The strict liability regime has already been discussed in this chapter above since it is one of the most well-defined standards of liability. Even the usage term "standard of liability" in the academic literature is largely confined to distinguishing fault from no-fault standards of liability. In the context of AI, the strict standard of liability can still be useful in cases where the intentions are difficult to determine, for example, in proprietary black box AI systems, which are also being used in high-risk activities without human supervision. The application of this standard would ensure that the decision-making is not transferred to AI completely, especially in cases where the risk of bodily harm to individuals is high, such as in healthcare or in critical infrastructure where it would cause massive economic and social repercussions if these systems fail, such as energy, water, food, transport, and telecommunications. Further, extremely high-risk activities such as nuclear power generation should use absolute liability standards, i.e., without any exceptions whatsoever, also denying the exception of the acts of God.

The usual benefits of using the strict fault standard would also translate to AI-related harms and include simplified litigation with lower costs and quicker resolution since the fault is not required to be proven, and thus could provide a more effective solution to the victims.[285] The strict liability would encourage AI developers to adopt the highest safety standards

[285] J X Rosenn, "Comment: litigation involving manufacturers' liability for defective medical products: judicial perspectives" (1976) 2:2 Am J L & Med 245–255.

proactively because, unlike less stringent standards discussed previously, the no-fault standard of liability does not provide any safe harbour from liability whatsoever. However, the strict standard of liability would make AI developers, in spite of all the due diligence and hence encourage AI developers not to develop AI for these activities, shutting out AI from the high-risk domains altogether, reducing innovation from essential economic activities including nuclear power generation and even its corresponding electrical grid applications.

Hence, while implementing AI into ultrahazardous activities, the developers who are not otherwise engaged in ultrahazardous activities may shy away from contributing to these critical infrastructures of the economy. This would deny the high-risk economic activities from even using AI to improve the safety itself and force them to develop AI in-house without external expertise, building rudimentary systems at exponential costs. While in the case of *Burlington Northern & Santa Fe Railway Co. v United States* [286], Shell was held to be not liable since even though Shell was aware that minor accidental spills were occurring, however, Shell took numerous steps to encourage its distributors to reduce the likelihood of spills. Nevertheless, this was a very fact-specific case in just one jurisdiction, and uncertainties would remain regarding liability in cases of participation of external AI developers in ultrahazardous activities being operated by another entity. In such cases, adopting less stringent liability for AI developers could encourage them to provide their input rather than the current vague position where they could be roped in at strict standards of liability as well.

2.7 Practical consideration based on corresponding stories in technology.

Policymakers have to be mindful of the selection of the standard of liability because there is a clear disparity in technical prowess and economic sizes of big tech companies, as

[286] *Burlington Northern & Santa Fe Railway Co. v United States* 520 F. 3d 918 (USSC).

compared to the nations themselves, which lowers the leverage the nations have to regulate these tech companies. For example, in 2023, Apple reportedly became the first company to reach the valuation of US$ 3 trillion (in market value), a value which was more than the GDP of nations like India, the UK and France.[287] In another example, recently, the chipmaker Nvidia, buoyed by its increasing role of chipsets in AI products, reached a market valuation of US$ 2 trillion, nearly equal to the GDP of Canada, which stands at US $2.12 trillion.[288]

This disparity is reflected in the attitude that big tech companies take towards regulatory changes. For example, Meta decided to strategically withdraw news content in response to laws made in Australia and Canada, showcasing the agility of tech companies in navigating regulatory tussles.[289] This was recently referred to as a "problem" by Canadian Prime Minister Trudeau where he feared it would have negative repercussions for not only journalism but also for infrastructure, which tech companies rely on and has been built on the back of publicly funded universities and social systems that provide freedom to encourage innovation.[290] However, this does not mean that all the nations have this issue since the Brussels effect and California effect have been recognised as enablers for legislators to exercise their market power to bring about desired public policy changes. In one instance, the EU seemed to have achieved its victory in enforcing USB-C standardisation against the wishes of giants such as Apple, which shows that regulatory pressures can bring the desired change.[291]

---

[287] "Apple is more valuable than entire countries — here's a closer look at the tech giant", online: <https://www.cnbctv18.com/business/companies/apple-is-more-valuable-entire-countries-heres-a-closer-look-at-the-tech-giant-16421141.htm>.
[288] Anis Heydari, "Chipmaker Nvidia is worth nearly as much as the entire Canadian economy. Here's why", *CBC News* (24 February 2024), online: <https://www.cbc.ca/news/business/ai-nvidia-valuation-1.7124435>.
[289] Jessica Mundie, "Canadians will no longer have access to news content on Facebook and Instagram, Meta says", *CBC News* (22 June 2023), online: <https://www.cbc.ca/news/politics/online-news-act-meta-facebook-1.6885634>.
[290] *Supra* note 250.
[291] Abby Hughes, "Apple adopting USB-C port for new iPhone 'a sigh of relief' for EU lawmaker", *CBC Radio* (13 September 2023), online: <https://www.cbc.ca/radio/asithappens/apple-adopting-usb-c-port-for-new-iphone-a-sigh-of-relief-for-eu-lawmaker-1.6966018>.

Whenever we discuss AI-related harms, one of the hottest topics is surrounding self-driving cars. In this regard, the recent story of General Motors-backed self-driving robo-taxi service Cruise is interesting. Cruise is one of the two major self-driving robo-taxi services, alongside Waymo. Cruise suffered a major financial blow when it failed to be upfront with regulators about one critical incident where a fully autonomous car hit a woman. In this incident, the AI wasn't at fault for the first collision; however, the car dragged the woman for 20 feet, mistakenly thinking she was on the side rather than beneath the car.[292] The company had this knowledge and failed to share the complete multimedia file with the regulators during the hearing, citing technical glitches.[293] Events like these also erode the trust of the policymakers and the public in AI. Hence, it is crucial that policymakers set the correct standards of liabilities that encourage these tech companies to be more transparent with the regulators about honest mistakes without any fear of repercussions.

Another case related to self-driving cars is a notable car accident in Florida wherein the Tesla car crashed while on autopilot because of many factors, killing the driver of the Tesla car. The human factors, as determined by the National Transportation Safety Board (NTSB), were that the truck failed to yield completely and the Tesla's driver wasn't paying attention.[294] NTSB also noted that other factors were "overreliance on automation" and "permitted disengagement by the driver".[295] Other notable factual issues, in this case, are that the Tesla's autopilot failed to detect a truck crossing the intersection perpendicular to the Tesla car and that the car was cruising just under 70 miles per hour in a 55 miles per hour maximum zone.[296]

---

[292] "In a single night, self-driving startup Cruise went from sizzling startup to cautionary tale. Here's what really happened—and how GM is scrambling to save its $10B bet", online: *Fortune* <https://fortune.com/2024/05/16/inside-gm-cruise-self-driving-car-accident-san-francisco-what-really-happened/>.
[293] *Ibid*.
[294] Trisha Thadani et al, "The final 11 seconds of a fatal Tesla Autopilot crash", online: *Washington Post* <https://www.washingtonpost.com/technology/interactive/2023/tesla-autopilot-crash-analysis/>.
[295] *Ibid*.
[296] *Ibid*.

After 25 seconds of AI detecting that the user's hand was no longer on the wheel, it was supposed to display a warning; however, the crash occurred before that.[297] The crash could have been avoided by improvements in human factors as well as AI factors. In such cases, at what standard should liability be assessed for AI? Tesla always provides manuals and warns its users to pay attention to the road, and even has indemnification clauses in the contract, which make the driver solely liable in case of an accident.[298] However, if we were to apply an utmost good faith-based due diligence standard of liability to the AI developer in this case, the developer should undertake additional measures to minimise risks. These measures could include ensuring that the autopilot system always strictly adheres to the speed limit, thereby reducing the risk of accidents; autopilot can slow down the vehicle when it detects that the user has removed their hands from the steering wheel; and the AI developer can implement a mandatory training mechanism to promote more responsible driving behaviour amongst its users. At the same time, the AI developer can counter-argue each of these additional measures.[299] Nevertheless, a good faith-based due diligence standard would require balancing these considerations with the overarching duty to minimise harm and ensure the safety of all road users.

Recent controversies related to Boeing, wherein the aircraft manufacturer has been accused of cutting costs at the cost of safety with the objective of boosting its profits.[300] This shows how technology, if monopolised, can result in big corporations with little accountability. This situation is further exacerbated by the fact that aviation regulators themselves lack

[297] *Ibid*.
[298] *Ibid*.
[299] For example, a. It was the user who selected the cruising speed, and what if, in case of emergency situations, accelerating is the only way to avoid a collision? Further, is it constitutionally valid for AI developers to take freedom away from their users to limit the speed limits in such cruise modes? b. Abrupt changes immediately after the driver removes their hands could compromise safety. c. They may also contend that overly stringent measures might not be well-received by consumers, potentially affecting sales and could be circumvented by users.
[300] LastWeekTonight, "Boeing: Last Week Tonight with John Oliver (HBO)" (07 March 2024), online (video): <www.youtube.com> [perma.cc/Y9VL-KM9K].

technical expertise and often rely on Boeing employees for safety assessments of the aircraft.[301] Hence, the regulation alone wouldn't be sufficient; to enact effective public policy, the governments need to have independent institutional capacities and technical understanding of emerging technologies, especially AI.

2.8 Selecting the appropriate standard of liability.

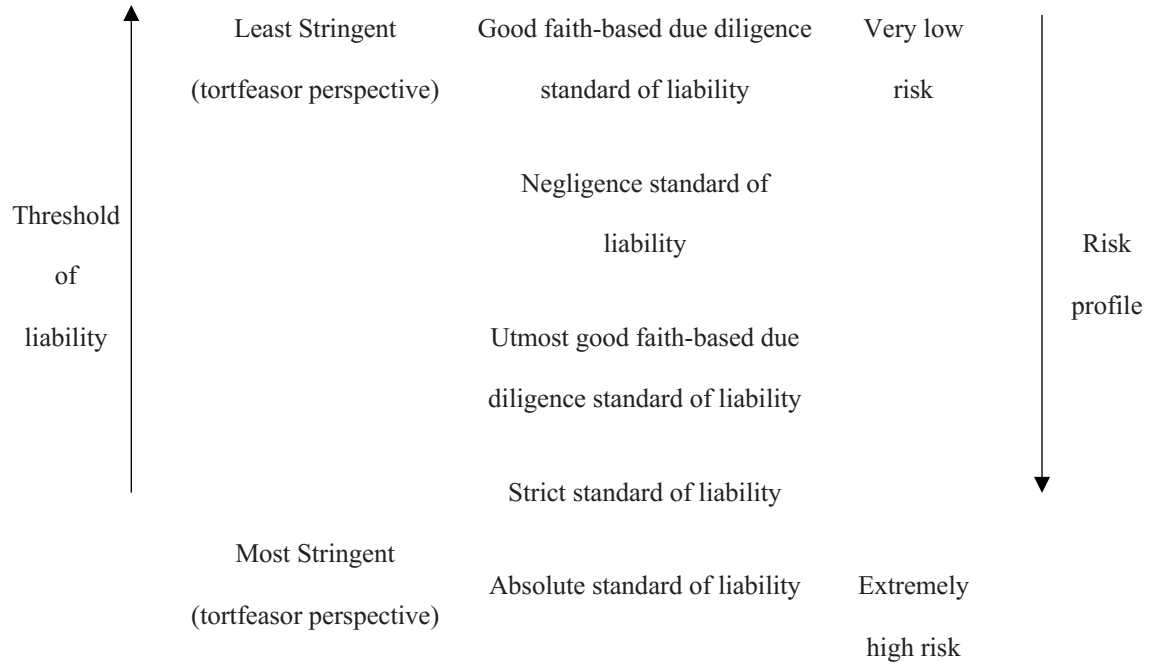
Selecting an appropriate standard of liability can be challenging, and there may not be a single correct answer. I would argue that the best approach is to apply a broad framework initially, then continuously monitor the effects of the adopted standard of liability in a particular domain, and thereafter keep fine-tuning it, as has been the approach of the EU AI regulation. The present research attempts to decipher existing standards of liability and delineates considerations that should be kept in mind while selecting the appropriate standard. The most important factors in selecting the appropriate standard of liability are the risk levels of a particular economic activity, the importance of economic activity for a particular jurisdiction, the degree of control exercised by different stakeholders, geopolitical circumstances and the financial, social, and political capacity of the entities involved.

However, additionally, while selecting the appropriate standard of liability, it is critical to consider not only risk but also the various traits and the pros and cons of each standard of liability. Further, the policymakers must be honest about their own capacities before making such a choice. For example, the utmost due diligence standard may be very lucrative to apply for the policymakers since, in such a case, policymakers do not have to iron out the details of compliances or due diligence measures that are required to be followed by the AI developers and deployers. However, an absence of policy efforts and public discourse will make the utmost due diligence standard vague and can turn AI into either a self-regulated sector with no

[301] *Ibid.*

accountability or a strict standard of liability, depending on how courts react. It would also reduce the certainty in the regulatory environment, which is the very problem the present thesis attempts to resolve.

Based on the discussion of this chapter, here is a figure which provides a simple representation that may be used for a quick reference while assigning a standard of liability for economic activity based on risk profile, as shown in **Figure 1** below.



From the perspective of the developers and operators of AI, a strategy can be to empower end users through training and transparency to distribute their responsibility through the democratisation of AI.

2.9 Distribution of liability amongst stakeholders for AI-related harms.

The selection of an appropriate standard of liability for an economic activity only provides us with an optimal criterion for determining liability. However, determining the

distribution of liability amongst different stakeholders can be rather challenging since the process of AI development and deployment is collaborative, and the supply chains therein are very complex.[302] [303] This is likely to be further complicated by the complex and incoherent web of contracts between the multiple stakeholders, which can make litigation and claim settlement a significant challenge for the victim with less financial resources. This will also put a strain on the judicial resources since the judiciary will then be tasked to discern the complex interconnections of these contractual obligations, as well as the technical complexities and real-life practicalities to unpack other non-contractual relationships.[304]

In this chapter, we will examine some of the principles and mechanisms already in place for some of the other domains of law that can help determine the distribution of liability for AI-related harms.

The polluter pays principle

The polluter pays principle is an economic principle that is used in the context of environmental harms, wherein the external costs (negative external effect of pollution) are internalised (polluter is made liable) in a bid to create a balance between private cost and social cost.[305] The polluter pays principle recognises that the end user generally cannot be individually liable and can only be collectively liable.[306] Another rationale behind the polluter pays principle is to create incentives to take proactive action rather than to allow the creation of victims and then compensate them.[307] The best use of the polluter pays principle, thus, is to identify the

[302] "Expert explainer: Allocating accountability in AI supply chains", online: <https://www.adalovelaceinstitute.org/resource/ai-supply-chains/>.
[303] Gloria J Miller, "Stakeholder roles in artificial intelligence projects" (2022) 3 Project Leadership and Society 100068 at 8.
[304] *Supra* note 142.
[305] OECD, *The Polluter Pays Principle: Definition, Analysis, Implementation* (Paris: Organisation for Economic Co-operation and Development, 2008) at 23.
[306] *Ibid* at 26.
[307] *Ibid*.

person or entity who has the effective economic and technical power to combat pollution and make them liable rather than targeting end consumers.[308] When we apply the core principles of the polluter pays principle to AI-related harms, we return to the principle of identifying the degree of control, as stressed in the previous chapter, to apply the correct standard of liability.

The deep pocket theory

Normally, the person or entity is held liable for their own misdeeds.[309] However, sometimes, there are circumstances when the party at fault cannot be identified or doesn't have the capacity to compensate for the harm.[310] In such cases, the courts can tend to look at the deep pocket theory, wherein a stakeholder with good financial resources is made liable for the harm.[311] This approach ensures that victims can receive adequate compensation.[312] This theory can be helpful in circumstances where there is an involvement of very large corporations when AI is being developed and/or deployed. However, there are criticisms of this theory since it runs a risk of violating the cardinal principles of the rule of law, and more specifically because it fails the foreseeability criteria under tort law, and the damage caused is too remote for the tortfeasors to be made liable.[313] Nevertheless, the deep pocket theory can be useful in circumstances wherein the following four conditions are met, i.e., the victim is truly innocent, the victim's injuries are severe, the true wrongdoer is unavailable or lacks the financial strength to compensate the victim, and that there was some degree of foreseeability for the large corporation.[314] One can wonder whether the calls to make Tesla liable in the previously discussed Banner's crash would have been could have stood the test of deep pocket theory and

[308] *Ibid*.
[309] Victor E Schwartz, Phil Goldberg & Christopher E Appel, "Deep Pocket Jurisprudence: Where Tort Law Should Draw the Line" (2018) 70 Okla L Rev at 359.
[310] OECD, *supra* note 305 at 26.
[311] *Ibid*.
[312] *Ibid*.
[313] *Ibid*.
[314] Schwartz, Goldberg & Appel, *supra* note 309 at 404.

its respective criticisms. It is being observed that AI industries are becoming more and more concentrated, and these corporations are further benefitting from greater scale, higher revenue and cost reduction.[315] Hence, the principle of deep pocket theory can be used by the courts in cases related to AI-related harms, in the interests of justice where the harm caused is very severe. However, it would be wise to exercise caution and not award damages against AI developers or deployers who are too remote, against principles of the rule of law, and have a deterrent effect on the innovation and adoption of AI.

In this regard, it is pertinent to note that vicarious liability already makes the principals liable for the fault of their agents, as seen usually in the employer-employee relationships where the employers are required to compensate for mistakes of their employees, which are done as a normal part of their duties.[316] One of the most important justifications for vicarious liability is the maxim '*qui facit per alium facit per se*', or 'He who acts through another, acts for himself'.[317] Interestingly, though, it is recognised that the main policy reason why vicarious liability exists is also based on the principle of deep pocket.[318]

Lessons from Intellectual Property (IP) Law

As seen in the previous chapter, an entity's degree of control is a major factor in determining a correct standard of liability. Similarly, the degree of control is crucial in determining the distribution of liability when there are multiple stakeholders. However, determining the degree of control may also require an assessment of the complex web of contractual obligations and even confidential and proprietary information. While financial

[315] "Survey: AI adoption proves its worth, but few scale impact | McKinsey", online: <https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact>.
[316] Paula Giliker, *Vicarious Liability in Tort: A Comparative Perspective*, Cambridge Studies in International and Comparative Law (Cambridge: Cambridge University Press, 2010) at 1–20.
[317] *Ibid* at 14.
[318] *Ibid* at 230.

information has been used to recalibrate the scales of justice through the principles of the account of profits (or "accounting for profits" or "accounting"), which originates in the cases of fiduciary liabilities[319], has also been adopted in the world of IP disputes. Account of profit leads to an accounting of profits, and the wrongfully gained profits or a portion thereof can be used to compensate the parties which suffered the loss.[320]

It is estimated that, on average, almost 40% of a company's market value is not shown in the balance sheets.[321] IP Audit is a systematic review of the IP owned, used or acquired by an entity, and is primarily used to assess and manage risks in relation to IP ownership.[322] However, an IP audit can reveal meaningful insights regarding the distribution of not just a technical degree of control (through assessment of IP), which would be crucial to determine who has technical control over the AI system but also information regarding financial control. This approach thus ensures that those who have the most degree of control over the functioning of the AI system are held accountable for any harm, thereby creating an incentive to create a system in a manner that is less prone to harm.

As such, however, AI developers may not be that inclined to share confidential and proprietary information to protect their competitiveness in the market. However, confidentiality-related issues are also faced in commercial and specifically IP disputes, wherein the courts have formulated the mechanism of confidentiality club to tackle this issue. A confidentiality club or confidentiality ring is a mechanism which was postulated by the English courts wherein an exhaustive list (of a ring of members who are deemed trustworthy) is created

[319] Garner, *supra* note 205 at 22.
[320] "Account of profits | Practical Law", online: <https://content.next.westlaw.com/practical-law/document/I09b673c0fd6b11e8a5b3e3d9e23d7429/Account-of-profits?transitionType=Default&contextData=%28sc.Default%29>.
[321] "IP Panorama - Module10 IP Audit", online: <https://www.wipo.int/export/sites/www/sme/en/documents/pdf/ip_panorama_10_learning_points.pdf> at 6.
[322] *Ibid* at 4.

from a few members representing each of the parties[323] and the judicial members to inspect the specific evidence or documents by this ring of members to facilitate with the trial without compromising the confidentiality.[324] Hot tubbing is another mechanism that is used in IP disputes by the courts wherein the experts are sworn in together and provide opposing testimonies, which enables the courts to quickly identify the correct technical position.[325] As such, hot tubbing has been formalised in some jurisdictions like India through the rules of the court and has been used in Standard Essential Patent disputes where the technology is complex.[326]

Conclusion

The distribution of liability for AI-related harms requires a holistic approach, wherein all the relevant factors, such as financial capacity, degree of control, the severity of harm (or its potential owing to industry-specific risks) and also the extent of due diligence measures applied by the AI developers or deployers on *bona fide* or *uberrinma fide* basis. The deep pocket theory, vicarious liability, and polluter pay principle all provide a mechanism to ensure that victims receive adequate compensation. However, there is a need for caution while applying some of these harsher principles wherein the entities complying with due diligence measures should not be put on the same pedestal as the entities that are careless or are not taking enough due diligence measures to ensure that the incentive developing and deploying responsible AI is not taken away.

[323] Either external members or counsels which are not directly related to the parties but can act in their interest.
[324] "Methods of management of confidential data in the context of national judicial proceedings", online: *CURIA* <https://curia.europa.eu/jcms/upload/docs/application/pdf/2020-11/ndr_2018-007_neutralisee-en.pdf> at 15.
[325] "Debate Over Hot-Tubbing In Patent Litigation", (4 October 2011), online: *IPOsgoode* <https://www.yorku.ca/osgoode/iposgoode/2011/10/04/debateoverhottubbinginpatentlitigation/>.
[326] Rachna Bakhru, "New Patent Suit Rules pave the way for expediting patent lawsuits", (31 March 2022), online: *Lexology* <https://www.lexology.com/library/detail.aspx?g=3bf9603f-ea00-41d9-8b22-fd3edf712f8d>.

Findings and Conclusion

AI regulation is in its nascent stage, and there are many uncertainties, especially regarding liability law in the case of AI-related harms. AI's autonomous nature, coupled with the opacity of its algorithms and lack of human-like intent, presents significant challenges for traditional liability frameworks. Hence, as observed in Chapter 1, the US and the EU have begun their regulatory efforts by adopting broad definitions of AI which have been adopted to accommodate future regulatory needs and technological advancements.

Traditional tort laws such as intentional torts, negligence, and strict liability face substantial hurdles when applied to AI-related harms. Intentional torts require proof of intent and are generally inapplicable to AI due to the absence of human-like intent, and it is also difficult to make the entities liable who are largely engaged in a *bona fide* manner. Negligence is one of the most flexible options and can be applied very broadly. The application of negligence to AI requires proving that the AI developer or deployer had a duty to exercise reasonable care, breached this duty, and caused harm as a result. However, it will also face difficulties in establishing duty, breach, and causation due to the complexity and unpredictability, which are integral parts of AI systems. AI can be autonomous and opaque, which can complicate this process, making it challenging to trace liability back to specific actions or inactions.

If applied, strict liability, characterised by its no-fault basis, will reduce the burden of proof on the injured party. A no-fault liability framework is particularly relevant for AI-related harms, as it does not require establishing intent or negligence. However, strict liability is criticised for potentially being too harsh on developers, as it holds them liable regardless of the due diligence measures adopted by the AI developers and deployers. This could stifle innovation by imposing excessive burdens on AI developers and deployers.

The EU adopted a risk-based approach to regulating AI, categorising AI applications based on their risk levels: minimal, limited, high, and unacceptable risk, with an additional category for GPAI. This approach allows for targeted regulatory measures to address specific risks associated with different AI applications. For example, unacceptable category AI applications have been altogether banned, high-risk AI applications face stringent compliance requirements, and minimal-risk applications encounter minimal regulatory scrutiny. This tiered regulation aims to balance the benefits of AI innovation with the need to protect public interests.

In Chapter 2, the present research involved a comparative study using the approach of analogical reasoning and extracted different standards of liability, including good faith-based due diligence standard of liability and utmost good faith-based due diligence, which offer different degrees of suitability for AI applications. It was identified that the concept of the standard of liability is not merely a standard of care but encompasses the whole framework to define a threshold to determine if a party is liable or not under such framework, even though the expected standard of care does remain an important factor in the overall framework. The standard of criteria also considers contractual and quasi-contractual norms, including statutory safe harbours and other defences.

The good faith-based due diligence standard

The due diligence standard of liability was explored in Chapter 2, which was largely based on a Narrow application in the banking sector by setting a reference to India's Negotiable Instruments Act as a case study, even though almost identical provisions also exist in other jurisdictions like the UK, Canada, and New Zealand. It was observed that this standard shields bankers from liability if they act in good faith (simply said, without bad intention) and without

negligence, adhering to ordinary care rather than a moderate care standard needed in negligence law. The discussion was then extended to the context of AI through analogical reasoning.

The due diligence standard is a bare minimum standard marked with the absence of express ill intentions and doing bare minimum due diligence at an ordinary care level. It was observed that there are challenges in defining due diligence measures required for AI owing to the complexity and evolving nature of the field, which makes it difficult to establish stable, long-lasting standards. Expertise and resource barriers also play a role, as standards development organisations (SDOs) like ISO and IEEE are often industry-led and may not fully address the socio-technical intricacies involved in AI development.

AI encompasses a wide range of applications, each with unique requirements and risks, making it challenging to create comprehensive and universally applicable standards since specificity is two-pronged, i.e., with respect to the type of AI and type of application. The due diligence standard relates to only the non-maleficence requirement of ethical guidelines. However, the lack of uniform regulations across different countries and regions adds another layer of complexity, leading to inconsistencies in what constitutes adequate due diligence. This situation creates challenges for AI developers operating in multiple jurisdictions as they must navigate a complex regulatory landscape, especially since compliance with the mandatory laws is necessarily a part of the due diligence standard. At the same time, the due diligence standard largely relies on the existence of a statutory safe harbour, and hence, it requires active steps from the policymakers if this standard has to be brought into existence.

Despite these challenges, a good faith-based due diligence standard could promote innovation in AI. By providing a safe harbour for developers who act in good faith and comply with basic minimum legal requirements, this standard can create a more supportive environment for experimentation and innovation. This approach encourages developers to stay

updated with common minimum risk management frameworks (RMFs) and adopt evolving safety measures, thus balancing the need for innovation with the imperative of minimising AI-related harms. This standard is particularly suitable for low-risk AI, where the severity of harm is likely to be very low and not immediate.

The utmost faith-based due diligence standard

Chapter 2 of the thesis also explores the concept of "utmost good faith" (*uberrima fide*) and its application in respect of AI-related harms. *Uberrima fide* principle (not the overall standard) originates from insurance law, wherein utmost good faith requires full disclosure of all material facts by one party (usually the insured) to another (the insurer). This principle contrasts with the caveat emptor ("let the buyer beware") principle in general contract law. Misrepresentation and non-disclosure are usually differentiated, wherein misrepresentation is accompanied by negative intent and is not condoned under the law. However, under *uberrima fide* requirements, non-disclosure and failure to volunteer material information, even in the absence of specific queries, cannot be condoned.

The application of the utmost due diligence standard, which is more stringent than negligence, is justified by the asymmetry in "degree of control", specifically in terms of AI-related harms wherein the technical control and understanding largely rests with the AI developers and deployers and not the end user. By ensuring compliance with the most stringent regulations of various regulations in different jurisdictions, adhering to technical and ethical standards in completeness, and having dedicated R&D teams or at least some designated individuals involved in ensuring ethics and safety, AI developers and deployers can meet the utmost good faith requirement. The chapter compares this standard with the bona fide (good faith) standard and outlines a novel five-step test for assessing compliance with utmost due diligence, inspired by the inventive step assessment in patent law.

Utmost due diligence standard has many challenges; however, it is highly subjective with potential of hindsight analysis, can be vague and over-encompassing without any escape for potential tortfeasor from the liability and thus being very close to strict liability standard, is very expensive since it requires extensive expert testimonies when being administered, very burdensome for the courts to discern, and very expensive and difficult for small developers. This standard can be effectively applied in various domains, including insurance, medical malpractice, and transportation, and is suitable for high-risk AI applications, wherein there is a constant need for innovation to improve safety and ethics.

While the utmost due diligence standard imposes higher obligations than negligence, it is still less strict than no-fault liability. In the context of AI-related harms, the strict liability standard is suitable for extremely high-risk economic activities such as healthcare, critical infrastructure, and nuclear power generation. Strict liability simplifies litigation by not requiring proof of fault, thus encouraging AI developers to adopt the highest safety standards. However, it may also deter developers from engaging in high-risk areas, potentially stifling innovation.

Practical considerations

The disparity between the regulatory power of nations is a factor versus the economic might of big tech companies, exemplified by cases like Apple and Nvidia, which shows that the big tech and monopolistic economic activities will be harder to regulate for nations with smaller market sizes. In such cases, collaboration and diplomacy can be the way forward for smaller nations. Policymakers need to be vigilant in setting appropriate standards of liability that balance innovation with safety and be wary of their own limitations, including red tape and a lack of technical understanding of AI. Building domestic industrial infrastructure and public institutional capacities could be the key to addressing the challenges posed by the likes

of Facebook's removal of Canadian and Australian news from the platform and Boeing's abnormal laxity in safety requirements, respectively.

There is a need for a flexible, evolving approach to the standard of liability tailored to the risk profile of specific economic activities. Additionally, while selecting the appropriate standard of liability, it is critical to consider not only risk but also the various traits and the pros and cons of each standard of liability since some pros or cons may vary very heavily in the case of a particular jurisdiction. Other relevant considerations include existing legal frameworks in the jurisdiction, precedents, religious nuances, cultural and socio-economic nuances, the extent of regulatory and statutory compliance by the tortfeasor, public institutional capacity, and political climate.

Distribution of liability

Chapter 3 of the thesis addresses the distribution of liability among stakeholders for AI-related harms, emphasizing that selecting an appropriate standard of liability is only the first step. The complexity of AI development and deployment involves numerous stakeholders and intricate supply chains, compounded by a web of contracts that complicates litigation and claim settlement. This complexity can overwhelm judicial resources as courts navigate these contractual obligations and technical details.

The polluter pays principle, traditionally used in environmental law, internalizes external costs to create a balance between private and social costs, making those with effective economic and technical power responsible rather than end users. Applying this principle to AI-related harms involves identifying the degree of control over the AI system to apply the correct standard of liability.

The deep pocket theory is another way to ensure victims receive adequate compensation when the party at fault cannot be identified or lacks the financial capacity. This theory suggests holding more financially capable stakeholders liable, especially large corporations involved in AI development and deployment, especially in cases when there are no other resources available for the victims. However, the deep pocket theory has criticisms, such as potentially violating the rule of law and failing the foreseeability criteria under tort law. Vicarious liability, where principals are liable for their agents' faults, is also relevant. This principle is justified by the maxim '*qui facit per alium facit per se*', i.e., 'He who acts through another, acts for himself' and is also based on the deep pocket theory (with an additional factor of relationship between agent and principle), ensuring that employers compensate for their employees' mistakes.

Lessons from Intellectual Property (IP) law again highlight the importance of the degree of control in determining liability. IP audits can provide insights into both technical and financial control, which is essential for identifying who should be accountable for AI-related harms. However, confidentiality concerns may arise, which can be addressed using mechanisms like confidentiality clubs and hot tubbing. Distributing liability for AI-related harms requires consideration of financial capacity, degree of control, severity of harm, and due diligence measures adopted by the relevant stakeholders. While principles like the deep pocket theory, vicarious liability, and the polluter pays principle can ensure victims receive compensation, caution is needed to differentiate between entities that comply with due diligence and those that do not, maintaining incentives for responsible AI development and deployment.

Findings of this thesis

This thesis highlighted that AI's technical nature, including unpredictability and inscrutability, complicates the application of traditional liability systems and, specifically, tort

laws for AI-related harms. The US has initiated regulatory efforts, beginning with formulating definitions of AI and implementing executive monitoring for military-grade AI. The definitions of AI in both the US and the EU have been kept broad to accommodate future advancements and regulatory needs. The EU has finalised an extensive overarching framework to regulate AI that will create due diligence requirements; however, the same has not been integrated with liability regimes at this stage.

Through a comparative study, this thesis explored different standards of liability, including good faith-based due diligence and utmost good faith-based due diligence, and discussed their suitability for AI-related harms. This thesis found that a good faith-based due diligence standard of liability can promote innovation by providing a safe harbour for developers and operators of AI acting in the absence of any bad intentions. The utmost good faith-based due diligence standard of liability imposes more burden on developers and operators of AI and requires them to take proactive measures beyond state-prescribed due diligence measures and do their best to ensure safety.

The more stringent standards of liability are desirable where there is larger imbalance of degree/locus of control, high-risk applications, and/or high-stake applications. This thesis further addressed the distribution of liability among stakeholders, emphasising the importance of the degree/locus of control and financial capacity in ensuring fair compensation for AI-related harms. Overall, this thesis delineates various standards of liability and legal principles available in other domains of law that policymakers and lawmakers can choose to meet their policy objectives and balance fostering innovation with compensating the victims.

Bibliography

Legislation

National Artificial Intelligence Initiative Act of 2020, H.R.6216 (12 March 2020).

AI Bill of Rights, (October 2022).

Civil liability regime for artificial intelligence, 2020.

Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), (2022).

Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on liability for defective products, (2022).

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, (2021).

Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) (Text with EEA relevance), OJ L 2022.

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance), OJ L 2022.

Regulation (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023 on general product safety, amending Regulation (EU) No 1025/2012 of the European Parliament and of the Council and Directive (EU) 2020/1828 of the European Parliament and the Council, and repealing Directive 2001/95/EC of the European Parliament and of the Council and Council Directive 87/357/EEC (Text with EEA relevance), OJ L 2023.

REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL AND THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, (2020).

The Artificial Intelligence and Data Act, C-27 (44–1) (16 June 2022).

Cheques Act of 1957 (UK), s 4.

Bills of Exchange Act, RSC 1985, c B-4, s 175.

Cheques Act 1960 (NZ), 1960/17, s 5.

Jurisprudence

Andrews v United Airlines, 24 F (3d) 39 (9th Cir 1994).

Burlington Northern & Santa Fe Railway Co. v United States 520 F. 3d 918 (USSC).

Carter v Boehm (1766), 3 Burr 1905 (King's Bench).

F. Hoffmann-La Roche Ltd. and Ors. v Cipla Ltd., 2016(65) PTC 1 (Del), (2016) (HC India)

Helling v Carey, 83 Wn. 2d 514, 519 P.2d 981 (Wash 1974)

M.C. Mehta and Another v Union of India and Others, AIR 1987 SC 1086 (1986) (SC India).

Pradeep Kumar and Another v Post Master General and Others, 2022 SCC Online SC 154, (2022) (SC India).

Rylands v Fletcher (1868) LR 3 HL 330 (HL (Eng)).

Secondary Materials

Colyer, J S, *A Modern View of the Law of Torts* (Elsevier Ltd, 1966).

Cross & Frank B, *Business law : negligence and torts*, Great courses (Teaching Company, 2015).

Davis, J L R & Rosalie P Balkin, *Law of torts*, 5th ed (LexisNexis Butterworths, 2013).

Dobbyn, John F & Christopher C French, *Insurance Law in a Nutshell*, 5th ed (West Academic Publishing, 2015).

Garner, Bryan A, ed, *Black's Law Dictionary*, 9th ed (St. Paul, MN: West Group, 2009).

Giliker, Paula, *Vicarious Liability in Tort: A Comparative Perspective*, Cambridge Studies in International and Comparative Law (Cambridge: Cambridge University Press, 2010).

Goldberg, John C P et al, *Tort Law: Responsibilities and Redress* (Aspen Publishing, 2021).

Hard Fork, "Justin Trudeau Wants A.I. For Good | EP 86" (07 June 2024), online (video): <www.youtube.com> [perma.cc/FQK8-485N]

Hodgin, Ray, *Insurance Law: Text and Materials*, 2nd ed (London: Routledge-Cavendish, 2002).

Koszowski, Maciej, *Analogical Reasoning in Law*, 1st ed (Newcastle upon Tyne: Cambridge Scholars Publishing, 2019).

LastWeekTonight, "Boeing: Last Week Tonight with John Oliver (HBO)" (07 March 2024), online (video): <www.youtube.com> [perma.cc/Y9VL-KM9K].

Lim, Ernest & Phillip Morgan, eds, *The Cambridge Handbook of Private Law and Artificial Intelligence* (Cambridge University Press, 2024).

OECD, *The Polluter Pays Principle: Definition, Analysis, Implementation* (Paris: Organisation for Economic Co-operation and Development, 2008).

Rose, Francis, *Marine Insurance: Law and Practice*, 2nd ed (London: Informa Law from Routledge, 2012).

Russell, Stuart & Peter Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed (USA: Prentice Hall Press, 2009).

Schauer, Frederick, *Thinking Like a Lawyer: A New Introduction to Legal Reasoning* (Harvard University Press, 2009).

Alqodsi, Enas Mohammed & Dmitry Gura, "High tech and legal challenges: Artificial intelligence-caused damage regulation" (2023) 9:2 Cogent Social Sciences 2270751.

Bathaee, Yavar, "The Artificial Intelligence Black Box and the Failure of Intent and Causation" 31:2 : Spring 2018 Harv JL & Tech.

Buiten, Miriam, Alexandre de Streel & Martin Peitz, "The law and economics of AI liability" (2023) 48 Computer L & Security Rev 105794.

Cerka, Paulius, Jurgita Grigienė & Gintarė Sirbikytė, "Liability for damages caused by artificial intelligence" (2015) 31 Computer Law & Security Review.

Coleman, Jules, "The Morality of Strict Tort Liability" (1976) 18:2 Wm & Mary L Rev 259.

Epstein, Richard A, "Vicarious Liability of Health Plans for Medical Injuries" (2000) 34 Val U L Rev 581.

Fenwick, Mark, Wulf A Kaal & Erik P M Vermeulen, "Regulation Tomorrow: What Happens When Technology Is Faster than the Law" (2016) 6:3 Am U Bus L Rev 561–594.

Fischer, Lukas et al, "AI System Engineering—Key Challenges and Lessons Learned" (2020) 3 Machine Learning and Knowledge Extraction 56–83.

Forcier, Mélanie, Lara Khoury & Nathalie Vézina, "Liability issues for the use of artificial intelligence in health care in Canada: AI and medical decision-making" (2020) 46 Dalhousie Medical J.

Gerstner, Maruerite E, "Liability Issues with Artificial Intelligence Software Comment" (1993) 33:1 Santa Clara L Rev 239–270.

Giuffrida, Iria, "Liability for AI Decision-Making: Some Legal and Ethical Considerations" (2019) Fordham Law Review, online: <https://www.semanticscholar.org/paper/Liability-for-AI-Decision-Making%3A-Some-Legal-and-Giuffrida/7b0e4436bffa0a627e3efc33c1f3f2a3480d74a3>.

Hallevy, Gabriel, "The Criminal Liability of Artificial Intelligence Entities - From Science Fiction to Legal Social Control" (2010) 4:2 Akron Intell Prop J 171–202.

Hasson, R, "The Doctrine of Uberrima Fides in Insurance Law—A Critical Evaluation" (2011) 32 The Modern L Rev 615–637.

Heydari, Anis, “Chipmaker Nvidia is worth nearly as much as the entire Canadian economy. Here’s why”, *CBC News* (24 February 2024), online: <https://www.cbc.ca/news/business/ai-nvidia-valuation-1.7124435>.

Hubbard, F Patrick, “‘Sophisticated Robots’: Balancing Liability, Regulation, and Innovation” (2015) 66 Fla L Rev.

Hughes, Abby, “Apple adopting USB-C port for new iPhone ‘a sigh of relief’ for EU lawmaker”, *CBC Radio* (13 September 2023), online: <https://www.cbc.ca/radio/asithappens/apple-adopting-usb-c-port-for-new-iphone-a-sigh-of-relief-for-eu-lawmaker-1.6966018>.

Jobin, Anna, Marcello Ienca & Effy Vayena, “The global landscape of AI ethics guidelines” (2019) 1:9 Nature Machine Intelligence 389–399.

Kang, Cecilia & Adam Satariano, “As A.I. Booms, Lawmakers Struggle to Understand the Technology”, *The New York Times* (3 March 2023), online: <https://www.nytimes.com/2023/03/03/technology/artificial-intelligence-regulation-congress.html>.

Karnow, Curtis E A, “Liability for Distributed Artificial Intelligences” (1996) 11 Berkely Tech LJ 147.

Kipnis, Jonathan, “Immune system: The ‘seventh sense’” (2018) 215:2 J Exp Med 397–398.

Korteling, J E (Hans) et al, “Human- versus Artificial Intelligence” (2021) 4 Front Artif Intell 622364.

Koszowski, Maciej, “Analogical Reasoning in Statutory Law” (2017) 8:2 J of Forensic Research 372.

Marchant, Gary & Rachel Lindor, “The Coming Collision Between Autonomous Vehicles and the Liability System” (2012) 52:4 Santa Clara L Rev 1321.

McCarthy, John, “WHAT IS ARTIFICIAL INTELLIGENCE?”

Miller, Gloria J, “Stakeholder roles in artificial intelligence projects” (2022) 3 Project Leadership and Society 100068.

Moylan, Christina M, “In Pursuit of the Appropriate Standard of Liability for Defective Product Designs” (1990) 42 Me L Rev 453.

Mundie, Jessica, “Canadians will no longer have access to news content on Facebook and Instagram, Meta says”, *CBC News* (22 June 2023), online: <https://www.cbc.ca/news/politics/online-news-act-meta-facebook-1.6885634>.

Oh, Jee-Sun, Moon-Koo Kim & Duk Hee Lee, “A study on the selection of future AI+X promising fields and the direction to strengthen competitiveness” (2021) 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) 371–374.

Priel, Dan, “The Indirect Influence of Politics on Tort Liability of Public Authorities in English Law” (2013) 47:1 Law & Soc’y Rev 169–198.

Rosenn, J X, “Comment: litigation involving manufacturers’ liability for defective medical products: judicial perspectives” (1976) 2:2 Am J L & Med 245–255.

Sachs, Reynold M, “Negligence or Strict Product Liability: Is There Really A Difference in Law or Economics?” 8 Ga J Intl & Comp L 259.

Samanta, Ash, Jo Samanta & Michael Gunn, “Legal considerations of clinical guidelines: will NICE make a difference?” (2003) 96:3 J of the Royal Society of Medicine 133–138.

Satariano, Adam & Cecilia Kang, “How Nations Are Losing a Global Race to Tackle A.I.’s Harms”, *The New York Times* (6 December 2023), online: <https://www.nytimes.com/2023/12/06/technology/ai-regulation-policies.html>.

Schwartz, Victor E, Phil Goldberg & Christopher E Appel, “Deep Pocket Jurisprudence: Where Tort Law Should Draw the Line” (2018) 70 Okla L Rev.

Selbst, Andrew D, “Negligence and AI’s Human Users” (2019) 100 BUL Rev 1315.

Turing, A M, “I.—COMPUTING MACHINERY AND INTELLIGENCE” (1950) LIX:236 Mind 433–460.

Vladeck, David C, “Machines without Principals: Liability Rules and Artificial Intelligence Essay” (2014) 89:1 Wash L Rev 117–150.

Zipursky, Benjamin C, “Reasonableness In and Out of Negligence Law” 163 U Pa L Rev.

“Basel III: international regulatory framework for banks” (2017), online: <https://www.bis.org/bcbs/basel3.htm>.

“Torts. Negligence. Liability of Owner of Automobile for Negligent Act of Driver. Section 282-e of New York Highway Law Interpreted” (1927) 27:7 Colum L Rev 886–887.

“Watch: Geoffrey Hinton tells BBC of AI dangers”, *BBC News*, online: <https://www.bbc.com/news/av/world-us-canada-65453192>.

“What Happens When You Click ‘Agree’?”, *The New York Times* (23 January 2021), online: <https://www.nytimes.com/2021/01/23/opinion/sunday/online-terms-of-service.html>.

Bakhru, Rachna, “New Patent Suit Rules pave the way for expediting patent lawsuits”, (31 March 2022), online: *Lexology* <https://www.lexology.com/library/detail.aspx?g=3bf9603f-ea00-41d9-8b22-fd3edf712f8d>.

Capraro, Valerio et al, *The impact of generative artificial intelligence on socioeconomic inequalities and policy making* (arXiv, 2023).

Field, Hayden, “AI lobbying spikes 185% as calls for regulation surge”, (2 February 2024), online: *CNBC* <https://www.cnbc.com/2024/02/02/ai-lobbying-spikes-nearly-200percent-as-calls-for-regulation-surge.html>.

Harris, Laurie & Chris Jaikaran, “Highlights of the 2023 Executive Order on Artificial Intelligence for Congress (R47843)”, (17 November 2023), online: *Congressional Research Service (CRS)* <https://crsreports.congress.gov/product/pdf/R/R47843/3>.

Henshall, Will, “Why Biden’s AI Executive Order Only Goes So Far”, (1 November 2023), online: *TIME* <https://time.com/6330652/biden-ai-order/>.

House, The White, *Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (2023).

Khanmohammadi, Reza et al, *An Introduction to Natural Language Processing Techniques and Framework for Clinical Implementation in Radiation Oncology* (arXiv, 2023).

Kristen Thomasen, “AI and Tort Law” in Florian Martin-Bariteau & Teresa Scassa, eds, *Artificial Intelligence and the Law in Canada* (New York: Rochester, 2021).

McCarthy, J et al, “A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE”, (31 August 1955), online: <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>.

Nasr-Azadani, Mohamad M & Jean-Luc Chatelain, *The Journey to Trustworthy AI- Part 1: Pursuit of Pragmatic Frameworks* (arXiv, 2024).

Parker, Dr Lynne, *National Artificial Intelligence Initiative* (2022).

Stock-Homburg, Ruth et al, *Evaluation of the Handshake Turing Test for anthropomorphic Robots* (2020).

Strauss, Jason David, *Uberrimae Fidei and Adverse Selection: the equitable legal judgment of Insurance Contracts* (2008).

Sunstein, Cass R, *Analogical Reasoning* (New York: Rochester, 2021).

Thadani, Trisha et al, “The final 11 seconds of a fatal Tesla Autopilot crash”, online: *Washington Post* <https://www.washingtonpost.com/technology/interactive/2023/tesla-autopilot-crash-analysis/>.

Venkat, Aruna, *“Polluter Pays” Principle: A Policy Principle* (New York: Rochester, 2012).

WestJEM, “The Standard of Care: Legal History and Definitions: the Bad and Good News”, (24 February 2011), online: *The Western Journal of Emergency Medicine* <https://westjem.com/articles/the-standard-of-care-legal-history-and-definitions-the-bad-and-good-news.html>.

“17 fatalities, 736 crashes: The shocking toll of Tesla’s Autopilot”, (10 June 2023), online: *Washington Post* <https://www.washingtonpost.com/technology/2023/06/10/tesla-autopilot-crashes-elon-musk/>.

“A definition of Artificial Intelligence: main capabilities and scientific disciplines | Shaping Europe’s digital future”, (18 December 2018), online: <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>.

“A pro-innovation approach to AI regulation”, online: *GOVUK* <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>.

"Account of profits | Practical Law", online: <https://content.next.westlaw.com/practical-law/document/I09b673c0fd6b11e8a5b3e3d9e23d7429/Account-of-profits?transitionType=Default&contextData=%28sc.Default%29>.

AI liability in the EU and the US: stifling or securing innovation? (2023).

"All change to the EU's strict product liability regime", (25 November 2022), online: <https://www.taylorwessing.com/en/insights-and-events/insights/2022/10/all-change-to-the-eus-strict-product-liability-regime>.

"Apple is more valuable than entire countries — here's a closer look at the tech giant", online: <https://www.cnbctv18.com/business/companies/apple-is-more-valuable-entire-countries-heres-a-closer-look-at-the-tech-giant-16421141.htm>.

"Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world", online: <https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>.

"Artificial intelligence liability directive", online: <https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI%282023%29739342_EN.pdf>.

"CED: An Overview of the Law - Torts: Principles of Liability | Westlaw Canada Portal", online: <https://www.westlawcanada.com/blog/insider/ced-an-overview-of-the-law-torts-principles-of-liability-808/>.

"Commission welcomes political agreement on AI Act", online: *European Commission - European Commission* <https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473>.

"Consumer product safety", online: <https://commission.europa.eu/business-economy-euro/product-safety-and-requirements/product-safety/consumer-product-safety_en>.

"Criticisms & Defenses of Reasoning by Analogy in Law", online: <https://cyber.harvard.edu/bridge/Analogy/analogy4.htm>.

"Crosswalk AI RMF (1.0) and ISO/IEC FDIS 23894 Information technology - Artificial intelligence Guidance on risk management", online: <https://www.nist.gov/system/files/documents/2023/01/26/crosswalk_AI_RMF_1_0_ISO_IEC_23894.pdf>.

"Debate Over Hot-Tubbing In Patent Litigation", (4 October 2011), online: *IPOsgoode* <https://www.yorku.ca/osgoode/iposgoode/2011/10/04/debateoverhottubbinginpatentlitigation/>.

"Defective products: revamped rules to better protect consumers from damages | News | European Parliament", (12 March 2024), online: <https://www.europarl.europa.eu/news/en/press-room/20240308IPR18990/defective-products-revamped-rules-to-better-protect-consumers-from-damages>.

"Expert explainer: Allocating accountability in AI supply chains", online: <https://www.adalovelaceinstitute.org/resource/ai-supply-chains/>.

"FTC Authorizes Compulsory Process for AI-related Products and Services", (21 November 2023), online: *Federal Trade Commission* <https://www.ftc.gov/news-events/news/press-releases/2023/11/ftc-authorizes-compulsory-process-ai-related-products-services>.

"High-level summary of the AI Act | EU Artificial Intelligence Act", online: <https://artificialintelligenceact.eu/high-level-summary/>.

"In a single night, self-driving startup Cruise went from sizzling startup to cautionary tale. Here's what really happened—and how GM is scrambling to save its $10B bet", online: *Fortune* <https://fortune.com/2024/05/16/inside-gm-cruise-self-driving-car-accident-san-francisco-what-really-happened/>.

"Intel Corp. Investment Policy Committee v. Sulyma", (26 November 2019), online: *LII / Legal Information Institute* <https://www.law.cornell.edu/supct/cert/18-1116>.

"intentional tort", online: *LII / Legal Information Institute* <https://www.law.cornell.edu/wex/intentional_tort>.

"IP Panorama - Module10 IP Audit", online: <https://www.wipo.int/export/sites/www/sme/en/documents/pdf/ip_panorama_10_learning_points.pdf>.

"Lawsuits test Tesla claim that drivers are solely responsible for crashes", (28 April 2024), online: Washington Post <https://www.washingtonpost.com/technology/2024/04/28/tesla-trial-autopilot-lawsuit/>.

"Methods of management of confidential data in the context of national judicial proceedings", online: *CURIA* <https://curia.europa.eu/jcms/upload/docs/application/pdf/2020-11/ndr_2018-007_neutralisee-en.pdf>.

"New Product Liability Directive", online: <https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739341/EPRS_BRI%282023%29739341_EN.pdf>.

"New Product Liability Directive | Think Tank | European Parliament", online: <https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2023)739341>.

"nonmaleficence", online: *Oxford Reference* <https://www-oxfordreference-com.proxy3.library.mcgill.ca/display/10.1093/oi/authority.20110803100237642>.

"Product liability and safety in the EU: overview", online: *Practical Law* <http://uk.practicallaw.thomsonreuters.com/w-013-0379?transitionType=Default&contextData=%28sc.Default%29>.

"Products liability law as a way to address AI harms", online: *Brookings* <https://www.brookings.edu/articles/products-liability-law-as-a-way-to-address-ai-harms/>.

"Responsible and trusted AI - Cloud Adoption Framework", (28 July 2023), online: <https://learn.microsoft.com/en-us/azure/cloud-adoption-framework/innovate/best-practices/trusted-ai>.

"sovereign immunity", online: *LII / Legal Information Institute* <https://www.law.cornell.edu/wex/sovereign_immunity>.

"Survey: AI adoption proves its worth, but few scale impact | McKinsey", online: <https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact>.

"TechDispatch #2/2023 - Explainable Artificial Intelligence | European Data Protection Supervisor", (2 April 2024), online: <https://www.edps.europa.eu/data-protection/our-work/publications/techdispatch/2023-11-16-techdispatch-22023-explainable-artificial-intelligence>.

"Texts adopted - Artificial Intelligence Act - Wednesday, 13 March 2024", online: <https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html>.

"The Bridge: Economic Analysis of Alternative Standards of Liability in Accident Law", online: <https://cyber.harvard.edu/bridge/LawEconomics/neg-liab.htm>.

"The Bridge: How Reasoning By Analogy Works in Law", online: <https://cyber.harvard.edu/bridge/Analogy/analogy3.htm>.

"The United States' Approach to AI Regulation: Key Considerations for Companies", online: <https://www.morganlewis.com/pubs/2023/05/the-united-states-approach-to-ai-regulation-key-considerations-for-companies>.

"tort", online: *LII / Legal Information Institute* <https://www.law.cornell.edu/wex/tort>.

"ubi jus ibi remedium", online: *Oxford Reference* <https://www-oxfordreference-com.proxy3.library.mcgill.ca/display/10.1093/oi/authority.20110803110448446>.

"Understanding the Interplay Between Strict Liability and Product Liability", online: <https://www.lexisnexis.com/community/insights/legal/b/thought-leadership/posts/understanding-the-interplay-between-strict-liability-and-products-liability, https://www.lexisnexis.com/community/insights/legal/b/thought-leadership/posts/understanding-the-interplay-between-strict-liability-and-products-liability>.

"What is AI? / Basic Questions", online: <http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html>.

"What Is Artificial Intelligence (AI)?", online: *Google Cloud* <https://cloud.google.com/learn/what-is-artificial-intelligence>.

"What will the role of standards be in AI governance?", online: <https://www.adalovelaceinstitute.org/blog/role-of-standards-in-ai-governance/>.

"White Paper on Artificial Intelligence: a European approach to excellence and trust", online: <https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en>.

"Why we launched DeepMind Ethics & Society", (3 October 2017), online: *Google DeepMind* <https://deepmind.google/discover/blog/why-we-launched-deepmind-ethics-society/>.