Artificial Intelligence Systems and Liability – An Inquiry into the Tort of Negligence Paradigm

Monidipa Sengupta[1]
Dr. Shreya Matilal[2]

Abstract

With the rapid development of artificial intelligence systems in the last decade, legal liability arising out of the interaction between humans and artificial intelligence systems may become commonplace in the near future. While there are many academic opinions demanding for a new legal framework to address the new technology of artificial intelligence systems, this paper endeavours to draw from traditional tort law and build a realistic and workable structure of ascribing liability arising from interactions between humans and artificial intelligence systems. However, it is limited to the liability arising from negligence and hence primarily focuses on the tort of negligence. To this end, first, it analyses the application of the reasonable person test on matters involving interactions with artificial intelligence systems. Second, it deals with the application of previously known methods of calculating risk in human interactions with artificial intelligence systems. Third, it deals with defences against the tort of negligence claims and their applications on matters involving interactions with AI systems. And finally, it briefly looks into causation and how to determine if a negligence claim has merit or is trivial. The paper concludes that the liability in a negligence claim for an injury occurring during a human interaction with an artificial intelligence system will be attached to one or more of three parties: the developer, the manufacturer, and/or the user. Any other parties simply would not have enough causal link with the injury for it to be a legitimate claim.

Keywords: *negligence, artificial intelligence, tortious liability, reasonable person test, calculus of risk.*

---

[1] *Research Scholar, Rajiv Gandhi School of Intellectual Property Law, IIT Kharagpur.*
[2] *Assistant Professor, Rajiv Gandhi School of Intellectual Property Law, IIT Kharagpur.*

I. Introduction

The world is witnessing an unparalleled advance of artificial intelligence systems permeating almost all spheres of contemporary social, cultural and economic existence. Not lagging much behind the technology forerunners, the lawyers, academics and policymakers seized the occasion to speculate on the legal problems - some real some imaginary- that AI had put forward or is likely to do so. Unsurprisingly, volumes of academic writings emerged across the jurisdictions suggesting solutions to AI-triggered troubles and mischiefs. Primarily, the existing writings on AI and Law focus on aspects ranging from conjectures as fantastic as ascribing legal personality to AI to concerns as crucial as envisaging new laws to create feasible liability frameworks for AI *vis-a-vis* their developers. This paper, however, endeavours to explore the possibility of construing from the traditional bounds of tort law a realistic structure of liability that arises or is likely to arise from human-AI interactions.

In the past, the contours of tort law were continually structured and restructured by adopting new interpretative paradigms whenever predicaments never witnessed before surfaced for legal determination. The liability framework of tort law, when historically seen, presents a strong time-tested model that fared well in accommodating newer technical and scientific predicaments to the legal system. Tort law does have to its standing the rare credit of surviving the onslaught of complex social and technical challenges through the ages including the most recent formidable encounter of the digital and communication revolution. In light of this belief, this study keeps aside the preconceived prejudice of assuming the fragility of tort law to the challenges posed by AI. Rather the paper looks deep into the traditional principles of tort law to find answers to the liability issues of AI systems.

Typically, the tort of negligence is set in motion when someone having a duty of care acts in a manner that is not expected from a reasonable person standing in his/her position. Construing a reasonable person, therefore, constitutes the first and most crucial step and then comparing and contrasting the defendant against that yardstick determines the outcome. This necessitates an in-depth scrutiny of how the understanding of reasonableness evolved, what it means and how far it can be made applicable on artificial intelligence systems. The unique challenge presented to the existing legal paradigms by artificial intelligence systems is that

they are apparently capable of causing damage to parties without human intervention. It is the authors' submission that the legal paradigms in tort law as they stand today are capable of dealing with all such tortious liability arising out of the application of an artificial intelligence system.

When it comes to fixing liability for the AI systems that have gone wrong, obviously, this traditional branch of tort law becomes the first choice. The coders who have programmed such systems, and companies who manufacture them come under this scanner of reasonableness inquiry. This section, in general, analyses the law of negligence in common law jurisdictions and then, turns to examining its applicability to the AI Systems generated legal issues.

For the purposes of this paper, the authors are using the term artificial intelligence to loosely mean any computer software that is trained using a data set to perform any action that requires human-like intelligence. This paper is divided into four parts. The first, deals with the application of the reasonable person test on matters involving interactions with artificial intelligence systems. Second, it deals with applying previously known methods of calculating risk in interactions with artificial intelligence systems. This part studies the application of the Hand Formula and Bayes Theorem as tools for risk determination. The third part of this paper deals with defences against the tort of negligence claims and their applications on matters involving interactions with AI systems. The fourth part briefly looks into causation and how to determine if a negligence claim has merit or is trivial.

II. Meaning and Contours of the Reasonable Person Test and Applying it to Interactions with Artificial Intelligence Systems

Tort law does place immense importance on the concept of care expected from a reasonable person, however, what ideally should be the traits of such a person varies with the social, economic, and cultural backdrop from which the facts of the case emanate. This open-ended issue poses a situation-sensitive ad-hoc mechanism to determine the liability that leads to fluctuating fortunes of justice emanating from different jurisdictions.

In its broadest sense negligence is defined as the failure to act with the prudence of a reasonable person. The understanding of reasonable standards gives a sense of objectivity as it does not depend on the specific characteristics of the defendant

before the court. But at the same time, it cannot be summarised or standardised as a simple cost-benefit analysis. The standard is based on a typical person's ability to make long-term plans in a manner that minimises the risk they impose on others in carrying out said plans.

As early as 1850, Justice Shaw had defined 'reasonable care' as the care and caution a reasonable, prudent man would take against probable danger.[3] Even this vague definition was sufficient to establish that a reasonable man was one who took into account the harm his action could bring to others and accordingly modified his actions to minimise the probable harm. There is no measuring scale to determine just how much caution a reasonable person is to take in order to minimise harm to others. It is largely understood that a reasonable person would place equal weight on the harm or injury of another person as he would place on his own harm and injury. The reasonable person, then, should be an "impartial spectator"[4] who treats the burdens and probable harms of the potential injurer and potential victim as worthy of the same level of consideration and does not arbitrarily give preference to one over the other. In *Bolton v. Stone*[5], Justice Radcliffe opined that a reasonable person was someone who treats the probable harm to others resulting from their conduct as if it were their own.

So, a reasonable person would analyse both the burden to himself and the foreseeable harm to everyone, including himself, and determine if he has to exercise care and how much. An important question here is how much a reasonable person takes into account a person's individual capacity or lack thereof. In *Vaughan v. Menlove*[6], the court laid down two important points.

In the instant case, the defendant insisted that the standard of care should not be a reasonable person standard but a good faith standard. Under the good faith standard, the court would decide whether the defendant did their best as per their capacity and level of intelligence, therefore, making the standard for care subjective. In this case, the defendant contended that he should not be penalised

[3] Brown v. Kendall, 1850 WL 4572 (Mass.).
[4] ADAM SMITH, THEORY OF MORAL SENTIMENTS 129-137 (D.D.Raphael & A.L.Macfie eds., Clarendon Press, 1976)(1759).
[5] 1951 AC 850.
[6] 132 Eng. Rep. 490 (C.P. 1837).

for his inability to assess the risk. One must keep in mind that the reasonable standard test is objective and expects the amount of care that an average person in society would have exercised.

The defendant went so far as to claim that if the court continued to use the objective reasonable person standard then people would have trouble figuring out what an average person would do, and failing to meet this seemingly arbitrary requirement, they would be penalised as under a strict liability standard.

The court rejected the good faith standard on two grounds. First, it would be impossible for the court to assess subjectively the capacity of each defendant that stood before it. Also, by the very nature of the standard, no defendant would ever be found liable as the court would be hard put to find a defendant who actually admits that they failed to care for a risk that they foresaw. Second, that others should be able to feel secure in the knowledge that the defendant would be tested against an expected standard of care and not a lower one, in which case the injured party would have to compensate for the defendant's lack of care. A third reason was also touched upon by the court, that the objective reasonable person standard is more predictable than the subjective good faith standard. The subjective standard makes it extremely difficult for potential victims to know the individual defendant's capacity for care. In summary, the objective standard forces some guidance to potential tortfeasors.

The above case law is, in general, credited with introducing the objective standard of care in negligence cases. Yet most suggestively, the defendant in Vaughan insisted the appropriate rules for liability should be drawn from the law of bailment- cases where one party delivers a chattel with a promise for its return at some future date. The unavoidable element of divided control in bailment cases makes the simple boundary-crossing rules used in boundary disputes and highway accidents a poor guide for the ultimate decision.[7]

In the case of *Coggs v. Bernard*, [8]Justice Holt adopted the six types of bailments in Roman Law. Every one of these had a specific standard of care. These are gratuitous bailment for safekeeping, bailment for the bailee's use, a simple pawn,

[7] Epstein, The Many Faces of Fault in Contract Law: Or How to Do Economics Right, Without Really Trying, 107 Mich. L. Rev. 1461,1464 (2009).
[8] 92 Eng. Rep. 107 (Q. B. 1704)

bailment for hire, bailment whereby a bailee agrees for a fee to operate or manage the thing bailed, and bailment of a thing to be managed (not merely stored) by the bailee without compensation. This ties the bailee's standard of care to the benefit he derives from the bailment. So he is liable for the slightest negligence if the loan is for his benefit. If the loan is for safekeeping for the bailor, then the bailee is only liable for gross negligence.

III. A Brief Understanding of the Application of the Reasonable Person Test on the Elderly and on Children

The reasonable person test has some more implications for the very young and very old. In *Roberts v. Ring*[9], the court sought to determine if the child exercised the degree of care commonly exercised by the ordinary child of the defendant's age and maturity. In the case of an old man, the reasonable person test is said to provide no special breaks based on age.[10] This difference in metric is easily understandable if we remember that elders have more foresight. Young children are unable to assess risk that an average adult is capable of. Also, it must be kept in mind that it is unreasonable to expect a child to meet the standard of care that an average adult is subjected to because of both physical and mental shortcomings. Therefore, it falls on the adults around them to compensate for the children's lack of ability to care to the proper standard of an average adult.

For the old who have lost their physical agility and dexterity to age, the burden of care works differently. They are expected to take more care because they are able to foresee potential harms at par with, if not better than, an average adult. It is expected they would take additional precautions for their physical shortcomings except in cases of senility.

This standard of care decreases when there is an obvious physical impairment like an amputation or a visible disability or injury. In that case, the standard of care is reduced to a suitable level. It is obvious that a person who has met with an accident and has sustained major injuries in many parts of their body will have difficulty moving around and people around them cannot reasonably expect them to behave and move the same way an average adult would. For example, in

[9] 173 NW 437 (Minn.1919).
[10] Keith N. Hylton, *Torts and Choice of Law: Searching for Principles,* 56 J. LEGAL. EDUC. 551, 554 (2006).

Fletcher v. City of Aberdeen,[11] the court held that the reasonable person test adopts a lower standard of care for the blind and requires compensating care on the part of those who can see. But also, should the said blind person try to engage in an unsuitable activity such as riding a cycle then they will be subjected to the original objective standard.

So far, we have discussed the soundness of the body and now we will discuss the soundness of the mind. It needs to be analysed whether insanity is a valid defence to negligence. The answer is that only a very specific kind of insanity would serve as a defence to a negligence claim. This would be the type of insanity that prevents a person from foreseeing potential harm and taking reasonable precautions to avoid it. We can therefore say that the tort of negligence treats insanity as incapacity similar in nature to a seizure or a cardiac arrest, sudden and debilitating. The test to determine negligence in matters involving sudden incapacity is as follows:

1. Was the actor aware of the risk of sudden incapacity? If not,
2. Did the actors put themselves in a risky setting on the assumption that they did not foresee the risk? If yes,
3. Was the incapacity such that it prevented the actor from meeting the reasonable care standard?

Imagine a cardiac arrest in a young person. They would not have foreseen the cardiac arrest if they had no history of heart disease. Their decision to go on a cycling trip might be considered reasonable. If they have a cardiac arrest while riding the cycle, in this situation, then they have not been negligent. But if the arrest is not too severe they may still be held liable for not meeting the reasonable standard of care. In case the person already was at high risk of a heart attack upon physical exertion and then had an attack while riding the cycle, then they would be held liable as they could foresee the risk.

Abraham[12] disputes the conventional view that in simple negligence cases, such as slip-and-fall or motor vehicle accidents, the finder of fact, given his familiarity with the activity, can simply invoke his conscience and thereby accurately

[11] 338 P.2d 743 (Wash. 1959).
[12] Abraham, *The Trouble with Negligence*, 54 VAND. L. REV. 1187, 1196 (2001).

enforce a community norm of appropriate behaviour. According to Professor Abraham, Vaughan defies this notion:

"The principal issue was whether an aperture should or should not have been built in the stack. One would think that if there was such a thing as a community norm regarding haystack construction, it would encompass whether and when to build apertures. But right at this seminal moment in the development of negligence law, the report of the decision in *Vaughan*- apparently recounting the evidence- makes a deeply revealing statement. The defendant "…made an aperture or chimney through the rick; but in spite or perhaps in consequence of this precaution. The rick at length burst into flames…". Just as in my hypothetical slip-and-fall case, the community norm in *Vaughan*, if there was one, did not come all the way down to the ground. The conscience of the community in *Vaughan* apparently was divided about apertures.

Here we must come to a very important discussion made by Oliver Wendell Holmes Jr. The premise is what would a court do if a defendant were to plead that after exercising all their intelligence to determine what standard of care a prudent person would exercise, they came to the wrong conclusion and failed. Would the court now be obligated to judge them on the subjective metric of their own intelligence? Holmes stated that the question was not whether the defendant thought his conduct was that of a prudent man but whether the people on the jury thought so. The law does not account for the multitude of differences in people's temperaments and nor is it capable of doing so. This is what makes the metric of the reasonable person so extremely important. Although as discussed before a person physically or mentally incapable of taking reasonable precautions would not be penalised for not taking them.

IV. Applying the Reasonable Person Test on Interactions with Artificial Intelligence Systems

With the growing number of artificial intelligence systems making their way into the day-to-day lives of average humans, the probability that claims of negligence would arise keeps increasing. It might, therefore, be prudent to ask the question as to how exactly the Reasonable Person Test to ascertain tortious liability be applied.

In the process that brings an artificial intelligence system to the market and eventually into the hands of the end user, there are three parties against whom a claim of negligence can be brought. These are the developers of the artificial intelligence system, the manufacturers of the artificial intelligence system, and the users of the artificial intelligence system.

At the outset, it must be determined which of the three potential parties has committed the negligent act. The following sections will shed more light on the various types of negligence claims that can aid one in determining the correct party to bring the negligence claim against.

A. Negligence with Respect to Operation

Negligence is an operation very simply claims that the defendant was negligent in the day-to-day conduct of his operation. For example, a person who is running a poultry farm and has not devised a proper method for waste disposal and is contaminating the public spaces around his shop can be said to be operating his business with negligence.

Using this understanding, it can be deduced that a user of an artificial intelligence system can operate the system in a negligent manner that may cause harm to another party. As a hypothetical example, consider the example of a facial recognition AI system that was trained on data points from a racially divided society. For the sake of this hypothesis let us assume that a majority of this society's incarcerated population is made up of persons of colour and that most of the data points provided for the AI system's training were faces with Eurocentric features.

Consider a user who uses this system as part of their household alarm system and sets it such that when the system assesses a person to be the risk of being a criminal, it will alert the nearest police station of an emergency situation. This would, in all likelihood, lead to the harassment of some people and the wrongful arrests of some. In such a case, the user of the AI system has negligently used the system to the detriment of persons of colour, assuming they knew of the shortcomings in the training of the AI system.

B. Negligence with Respect to Design

Similarly, a negligent design claim is one where the plaintiff claims that the defendant was using a product or process with a design that the defendant was negligent in making. In the poultry farm example, the plaintiff might make a negligent design claim by asserting that the waste disposal system followed by the owner of the poultry was not properly designed. A negligence in warning claim is one where the plaintiff asserts that the defendant failed to give adequate warning to the plaintiff so that they could avoid the potential harm that can be caused. For example, again referring to the poultry situation, the plaintiff could claim that the defendant did not provide enough warning regarding the contamination of the public spaces around his shop. These can be further studied with the aid of case laws.

In Blyth v. Birmingham Water Works,[13] the plaintiff suffered harm due to a large amount of water escaping out of the water pipeline and getting into the plaintiff's house. The said leak was caused by a frost. The court ended up finding in favour of the defendant as the accident was considered to be highly unlikely so much so that no reasonable prudent person would think to take precautions against it.

In *Adams v. Bullock*, [14]a child was injured due to an electric shock. The child was playing with a wire that he was swinging and which came into contact with electricity lines utilised by a trolley company. The plaintiff claimed that the trolley company's design was negligent. The court, however, found otherwise. It was found that the accident, like in Blyth, was unforeseeable enough for the trolley company to not have taken precautions against it. It was further held that the accident had such little probability of happening that it was not fair to force the trolley company to change its design of the electricity supply wires.

Further, in *Cooley v. Public Service Co.*,[15] the plaintiff upon being injured by an electric charge that travelled through the wires maintained by the defendant company alleged that the design of the wiring and their placement was negligent. The incident in question had happened during a particularly heavy storm. The plaintiff went on to suggest that a wire mesh basket structure should be made so

[13] 156 Eng. Rep. 1047 (Ex. 1856).
[14] 125 N.E. 93 (N.Y. 1919).
[15] 10 A.2d 673 (N.H. 1940).

that power lines do not come in contact with the defendant company's wires as had happened during the accident that injured her.

The court acknowledged that the design suggested by the plaintiff would indeed reduce accidents of the nature that injured her. But at the same time, it would increase the risk of electrocuting pedestrians. The court finally came to the conclusion that the accident that injured the plaintiff was of such a rare nature that no ordinary or even prudent person would have foreseen it to take precautions against it. The defendant's company was found to be not negligent.

If we refer to the hypothetical AI system in the previous section, we can deduce that the person who selected the data points that were used to train the system was negligent in their selection. In failing to provide a representative set of data points, the developer of the AI system would be found negligent in the process of developing (or designing) the artificial intelligence system and an action for the same may be brought against them.

C. Negligence with Respect to Warning

In *Rinaldo v. McGovern,*[16] the glass of the plaintiff's car was shattered by a golf ball hit by the defendant. The defendant was on a golf course and the plaintiff was driving past it on a public road. The plaintiff alleged that the duty that the defendant breached was the duty to warn passers-by of errant golf balls. The court rejected the argument stating that a warning needs to be effective to be valid. In the instant case, even if the defendant had somehow managed to warn the plaintiff of the errant ball after hitting it, the plaintiff would not have had enough time to change course to avoid getting hit. This renders the warning expected by the plaintiff ineffective. Moreover, balls are hit erratically on a golf course all the time and if each of them were to be accompanied by a warning of some kind, people would learn to ignore them simply due to how frequent they are. At least in this scenario warnings were ineffective because they greatly overstated the potential harm to possible victims.

Now, the authors will apply this understanding to the same hypothetical artificial intelligence system in the previous two sections. If the manufacturer of this AI system neglects to put a warning regarding the risks of using this facial

[16] 587 N.E.2d 264 (N.Y. 1991).

recognition system in a security system, they would be considered negligent in providing adequate warning and a negligence action may be brought against them.

V. Calculus of Risk

A. Hand Formula for Determining Risk

While the reasonable person standard provides some groundwork, Judge Learned Hand in *Unites States v. Carroll Towning Co.,* [17]laid down an algorithm (while discussing the matter where it was to be decided whether the owner of a barge was negligent in not preventing it from breaking away from its moorings):

"Since there are occasions when every vessel will break from her moorings, and since, if she does, she becomes a menace to those about her; the owner's duty, as in other similar situations, to provide against resulting injuries is a function of three variables:

1. The probability that she will break away;
2. 2. The gravity of the resulting injury, if she does;
3. The burden of adequate precautions.

Possibly it serves to bring this notion into relief to state it in algebraic terms: if the probability be called P; the injury, L; and the burden, B; liability depends upon whether B is less than L multiplied by P: i.e., whether B<PL"[18]

This is popularly known as the Hand Formula and it basically states that a person can only be held negligent if B<PL and not if B>PL. this means if the burden of taking adequate precautions is greater than the product of the probability of the event happening and the gravity of the injury it would cause then a person should not be found negligent.

This, however, has limited application. It cannot be reasonably or even mathematically applied in most cases. It works as a base to determine which factors to give the most importance to while determining negligence, but it is nearly impossible to put a numerical value to it while taking all elements into consideration. Also, the Hand Formula, while landmark, is not the only way to

[17] 159 F. 2d 169 (2d Cir. 1947).
[18] *Id.at 173.*

determine if a person acted up to the reasonable person standard in a matter of negligence.

In fact, in some cases, the courts have made decisions that are not in keeping with Judge Hand's recommendation. For instance, in *Bird v. Holbrook*[19], the plaintiff was injured by a trip gun installed in the defendant's garden to catch thieves without any warning that the garden had a trip gun installed. The court did not attempt to make a comparison between the value of property, in this case, the garden, and the value of a life because that would have been unreasonable.

Similarly, if a person who owns a very expensive motorcycle swerves to avoid damage to their motorcycle and hits a pedestrian instead, they would be held negligent as no court would find it prudent to compare the value of the motorcycle to the value of the pedestrian's life using the Hand Formula.

The reasonableness of a given risk may depend on the following factors:

i. The magnitude of the risk. A risk is more likely to be unreasonable the greater it is.
ii. The value or importance of that which is exposed to the risk, which is the object that the law desires to protect, and may be called the principal object. The unreasonableness of a risk means its reasonableness with respect to the principal object.
iii. A person who takes a risk of injuring the principal object usually does so because he has some reason of his own for such conduct, - is pursuing some object of his own. This may be called the collateral object. In some cases, at least, the value or importance of the collateral object is properly to be considered in deciding upon the reasonableness of the risk.
iv. The probability that the collateral object will be attained by the conduct that involves risk to the principal; the utility of the risk.
v. The probability that the collateral object would not have been attained without taking the risk; the necessity of the risk.[20]

[19] 130 Eng. Rep. 911 (C.P. 1825).
[20] Henry Terry, Negligence, 29 Harv. L. Rev. 40, 42-44 (1915).

At this stage, we should analyse some cases to understand how exactly the courts decide on finding a defendant negligent. In *Osborne v. Montgomery*, [21]the plaintiff, riding a bicycle, collided with the defendant's open car door when the defendant stopped his car to run an errand. The handle of the plaintiff's bicycle hit the open door of the defendant's car and tipped it over, causing injury to the plaintiff. The court's ruling was as follows:

"1. By ordinary care is meant that degree of care which the great mass of mankind, or the type of that mass, the ordinary prudent man, exercises under like or similar circumstances.

2. Negligence is the want of ordinary care.

3. Every person is negligent when…he does such an act, or omis to take such a precaution, tat, under the circumstances present, he ought reasonably to foresee that some injury or damage might probably result from his conduct.

He is in duty bound to foresee all such natural consequences of his conduct as an ordinarily prudent an intelligent person would ordinarily foresee under the then present circumstances."

This ruling is faulty in as much that a great mass of humanity does not regularly indulge in activities that can potentially result in harm to others. While a vast number of harms and injuries are caused on a regular basis, most of them fall in the category of harm without legal injury or damnum sine injuria.

The liability for a wrongful act is based on the balancing of social interests. For example, a public transport vehicle such as a bus in the city of Kolkata in the height of monsoon may spray a pedestrian passerby with dirty or muddy water. But the benefit of allowing the public transportation system to function in the height of monsoon so far outweighs the probable injury that the conduct is never questioned. Also in the above case, the court offered no objective standard of care, or even a way to determine an objective standard of care that the defendant may have taken to prevent being found negligent.

[21] 234 N.W. 372 (Wis. 1931).

In the case of *Cooley v. Public Service Co.*,[22] the injury to the plaintiff was caused by the cables that the defendant's telephone company maintained above ground in the city of Manchester. The plaintiff raised the issue that their maintenance of cables was negligent but not that the act of maintaining cables above ground itself was negligent.

Consider by way of illustration the problem of pedestrian-automobile accidents (and, as we are now discussing the unilateral case, let us imagine the behaviour of the pedestrians to be fixed). Suppose that drivers of automobiles find it in their interest to adhere to the standard of due care but that the possibility of accidents is not thereby eliminated. Then, in deciding how much to drive, they will contemplate only the enjoyment they get from doing so. Because (as they exercise due care) they will not be liable for harm suffered by pedestrians, drivers will not take into account that going more miles will mean a higher expected number of accidents. Hence, there will be too much driving; an individual will, for example, decide to go for a drive on a mere whim despite the imposition of a positive expected cost to pedestrians."[23]

This is where negligence differs from strict liability as under the principles of strict liability the injurer must compensate the losses even when they arise out of an accident. The person would therefore have to consider both the nature of his activity and the manner in which he performs them so as to minimise the potential harm to others.

While the reasonable person test is nearly universally accepted, the Hand Formula is one of the methods to apply the reasonable person test.

B. Using Bayes Theorem to Determine Risk

Because of the nature of the tort, most often the evidence in a matter of the tort of negligence is circumstantial. The res ipsa loquitor doctrine permits and regulates the submission of such evidence in negligence cases.

The res ipsa loquitor doctrine can be summarised in three rules:

[22] 10 A. 2d 673 (N.H. 1940).
[23] Shavell, *Strict Liability Versus Negligence*, 9 J. LEGAL STUD. 1, 2-3 (1980).

1. The event must be of a kind that ordinarily does not occur in the absence of someone's negligence;
2. It must be caused by an agency or instrumentality within the exclusive control of the defendant; and
3. It must not have been due to any voluntary action or contribution on the part of the plaintiff.[24]

For rule one, to determine if an event occurs ordinarily without someone's negligence, one must fix the baseline probabilities, meaning, how frequently the said event happens in the presence and absence of someone's negligence. This is a prerequisite to reaching a conclusion regarding the likelihood of the event's occurrence in the presence of someone's negligence.

As per Bayes' Theorem the conditional probability of event A given event B is equal to the product of the conditional probability of event B given event A and the ratio of probability of event A to probability of event B.

$$P(A|B) = P(B|A) \times P(A)/P(B)$$

Consider a situation where heavy machinery is being lowered from the top of a building and it hits a passer-by pedestrian. The probability of the principal event of the machinery hitting the passer given the negligence of the person lowering it is equal to the product of the probability of the person lowering the barrel negligently (principal event) resulting in an accident and the ratio of the probability of accidents resulting from lowering of machinery to the probability of negligent lowering of machinery.

$$P(accident\ due\ to\ lowering\ of\ heavy\ machinery \,|\, negligent\ lowering\ of\ machinery) = P(negligent\ lowering\ of\ machinery|\ accident\ due\ to\ lowering\ of\ machinery) \times [P\ (accident\ due\ to\ lowering\ of\ machinery) /P(negligent\ lowering\ of\ machinery)$$

It is important that the principal event in question is not one that can easily happen without negligence on someone's part. If res ipsa loquitor doctrine is applied in matters where accidents frequently occur in the absence of negligence

[24] WILLIAM L. PROSSER, HANDBOOK OF THE LAW OF TORTS 214 (West Publishing Co. 4th ed. 1971).

it would not remain distinguishable from no-fault liability doctrines like strict liability.

For rule two, it would be simpler to compare two case laws where the event causing injury was an object being thrown out of a hotel. In the first, *Larson v. St. Francis Hotel*[25], a customer of the hotel threw a chair out of the window in sudden celebration and injured the plaintiff. The plaintiff's claim that the hotel had exclusive control over this event was rejected by the court on the grounds that the hotel had no prior notice or reason to believe that one of their customers could throw a chair through the window. But in the second, *Connolly v. Nicollet Hotel*[26], the defendant hotel was hosting a convention that was expected to get rowdy. So when the plaintiff claimed that the hotel had exclusive control over the object that was thrown from the hotel that injured the plaintiff, the court accepted it. The rationale was that the hotel in Connolly would have and should have been aware of the greater threat to its customers and nearby pedestrians and taken appropriate measures to safeguard them.

The third rule points towards contributory negligence. There is no set rule to determine the contribution of a plaintiff in a matter of negligence. Although some scholars have suggested algorithms and formulas for determining contributory negligence, they have not been put to practice.

C. Applying the Calculus of Risk Tools to the Interactions with Artificial Intelligence Systems:

In matters of interactions with artificial intelligence systems, it is easier to use Bayes Theorem to determine the risk. The current state of the art in artificial intelligence systems follows a fairly predictable learning pattern based on the data points provided for it by the developer. This makes the application of Bayes Theorem on these interactions much easier than the application of the more abstract Hand Formula for the determination of risk.

This is so because most artificial intelligence systems at this stage follow predictable mathematical models during their training and therefore the factors required to apply Bayes Theorem would be more accurately determinable.

[25] 188 P.2d 513 (Cal. App. 1948).
[26] 95 N.W.2d 657 (Minn. 1959).

VI. Defences

A. Contributory Negligence

The defendant in a negligence claim can assert that there was contributory negligence on the part of the plaintiff. It is essentially a claim that the plaintiff has failed to take reasonable care to avoid the injury in question. It is important to understand that not every failure to take reasonable precautions amounts to contributory negligence. Only the failure of reasonable care that compounds the negligence of the defendant is considered contributory negligence.

For example, if a negligent driver of a car hits a tree and a branch falls off due to the impact and injures the plaintiff pedestrian, then the defendant cannot claim that the plaintiff was also negligent in placing themselves in such close proximity to the tree.

But if a negligent car driver collides with a rash moped driver who falls and injures their head, the driver can claim contributory negligence as the moped driver could have easily avoided the injury by simply driving safely.

Therefore, it can be understood that a cautionary or precautionary measure that actually affects the probability of the incident that caused the injury itself is the only one that can be considered for a contributory negligence claim. A measure that would potentially reduce the severity of the injury without impacting the probability of the accident at all, is not one on which a claim for contributory negligence can be based.

In most cases, a person has the responsibility to take reasonable care to avoid injury. There are two exceptions to this general rule. First, a person rescuing another from a dangerous situation[27] is not expected to take reasonable care to protect themselves. Second, the owner of a property[28] is not expected to take precautions against trespass and subsequent injury.

The reasonable person test is applied to determine whether the plaintiff has in fact breached their duty to self-care. In a claim of contributory negligence,

[27] Eckert v. Long Island R.R., 43 N.Y. 502 (1871).
[28] LeRoy Fiber Co. v. Chicago, Milwaukee & St. Paul Ry., 232 U.S. 340 (1914).

proximate causation is determined from the answer to the question "whether the injury was a foreseeable conclusion of the plaintiff's negligence?" and factual causation is determined by answering the question "whether the injury would have happened even if the plaintiff had taken reasonable care?"

A clear defence to a contributory negligence claim is the "last clear chance" rule. This rule applies in cases where the injurer has a clear opportunity to prevent the injury but does not take it. For example, in *Fuller. v. Illinois Central R.R.*[29], the victims negligently got stuck in the tracks of the defendant's railroad as a train was approaching. The train was at a negligently high speed but had enough time to stop to avoid running over the plaintiff's victim. In this case, the defendant had to pay damages despite clear negligence on the part of the plaintiff.

B. Act of God

An Act of God (vis major) is generally understood to be a direct, sudden, violent act of nature that could not be reasonably foreseen and if it could be foreseen it could not have been avoided by any amount of human care. In Nichols v. Marshland[30], the artificial lakes maintained by the defendant faced extraordinary rainfall due to which their banks burst and the water that escaped from them washed away four bridges. It was held that no amount of precautions taken by the defendant could have prevented the damage and therefore it was an Act of God and the defendant was not liable to pay damages.

Indian courts have taken a higher standard to the Act of God defence than other jurisdictions in matters of absolute liability. In India, an Act of God is not a valid defence for absolute liability claims in India.

VII. A Brief Insight into Causation

It is a simple deduction that liability can only be attached to the defendant when a causal link has been established between the defendant's negligence and the plaintiff's injury. In other words, only the "actual cause" of injury is considered while attaching liability to a defendant.

[29] 65 So. 783 (Miss. 1911).
[30] 1876 2 Ex. D. 1.

For this determination, the court asks the question "if not for the actions of the defendant, would the injury of the plaintiff have occurred"? If the answer to that question is 'no', then the defendant is held liable.

In *Barnett v. Chelsea & Kensington Hospital Management Committee,*[31] the plaintiff's husband had gone to the defendant's hospital when he felt ill after drinking some tea. Despite the husband looking very sickly, the orderly (due to medical reasons of their own) turned him away and asked him to consult other doctors. The husband died in a few hours due to arsenic poisoning. In the action against the defendant hospital, the judge found that there was evidence of negligence but that this negligence could not cause injury or death to the plaintiff's husband. In the humble opinion of the author, a far better action may have been brought against the person who sold the plaintiff tea which had arsenic in it.

But despite its near-universal acceptance, this test suffers some limitations. There is often a tenuous relation between actions. For example, if a person in the city of Kolkata had to take an alternate route to their workplace due to the construction of the flyover by the Kolkata Municipal Corporation, and they get robbed on this new route. They can probably claim that if the Kolkata Municipal Corporation had not started construction on the flyover they would not have been robbed. And while this claim is factually accurate, the causal link between the action and injury is too slim to be of any impact. This brings us to the second issue with the but-for test. Because of the many tenuous relations of actions and outcomes, this test can be misused easily to add parties to a claim in a frivolous manner. In the previous example, this may even include the elected representative from that constituency in the claim if this test were strictly followed.

In cases where damage is caused by interaction with AI systems, it is absolutely necessary to remember that despite many parties being capable of having some frivolous link to the injury, one must stay focused on the three parties mentioned above.

[31] 1969 1 QB 428

An example of such frivolous attempt at establishing causal link may be by including people who provided data for the training of the AI system instead of focusing on the people who actually selected what data points to use when training the system.

VIII. Conclusion

The three main questions that need to be answered in this ascertaining liability in matters where damage or injury is caused by an artificial intelligence system due to negligence are:

1. Who do we apply the reasonable person test to?
2. What duty of care does this person owe the plaintiff or the injured party?
3. When can this person claim any defence to the assertion of negligence?

To answer the first question, there are three prospective candidates on whom the liability can possibly be ascribed: the person who curates the training data for the artificial intelligence system, the author of the system, or the user of the system. The liability of the user of the system will be discussed while answering the third question. The person(s) who curates the training data for the artificial intelligence system and the person who writes the system can both be liable for negligence under different circumstances depending on which part of the process was conducted negligently. That is to say, if the artificial intelligence system was programmed negligently, then the author of the system would be held negligent; and if the data that was used to train the artificial intelligence system was curated negligently, then the person who curated the data would be held negligent.

For example, an artificial intelligence system that is designed for a self-driving car to be launched in a mountainous country like Nepal is trained using only data collected from planar terrain would inevitably cause accidents because the system was trained using a set of data that was negligently curated. In such a situation, the person who curated the data would be held negligent.

If the same system was trained using the data from proper terrain but was never programmed to make the winding blind turns that are signature to mountainous terrain, then the author of the system would be held liable for negligence. Such liability can be a joint liability if the persons in question collaborate in the designing and training of the artificial intelligence system. It is entirely possible

that the system is written by and trained using the data curated by the same person, in which case this distinction need not be made.

Secondly, the duty of care that the persons writing and training the artificial intelligence system owe to the user or any other injured party is of three varieties. They have a duty to warn any other party about any shortcomings or foreseeable damage that can be caused by the artificial intelligence system. This disclosure should be made to the best of the knowledge of the persons involved in writing and curating the training data for the artificial intelligence system.

They also owe the duty to ensure that the designing and operation of the system is such that it causes no damage to any party as far as practicable and foreseeable. This means they should avoid using any method for writing or training the artificial intelligence system that might cause damage to the health, safety, and reputation of any party or violate the rights of any party in any manner.

Contributory negligence can be claimed if it can be proven that the user or any other party used the system negligently. It may be by not following the guidelines to use the system or misusing it for purposes it was not designed or trained for. Also, it can be assumed that the person agreed to the risks that led to the damage if sufficient warning and instruction were provided as to the use of the said system.

Who do we apply the reasonable person test to?	**What duty of care does this person owe the plaintiff or the injured party?**
Developer	That they have taken reasonable care in the design of the system.
Manufacturer	They have taken reasonable care in warning the user of the potential injuries that can be caused by the system
User	That they have operated the system with reasonable care

It can therefore be seen that the law of the tort of negligence as it has evolved over time is enough and capable of addressing the questions of liability that the negligent use design or training of an artificial intelligence system may pose, as the technology currently exists.

Although the trend of making new laws to deal with liability caused due to interaction with AI systems has already taken root, it is not a necessary action for dealing with negligent damage caused during interaction with an artificial intelligence system.

The tort of negligence has evolved over centuries and withstood the tests of major technological advancements like the Industrial Revolution and the Internet. It is clear from the above analysis of the law of tort of negligence that it is beautifully malleable and capable of dealing with injuries caused during interaction with AI systems.

In summary the liability in a negligence claim for an injury occurring during an interaction with an AI system will be attached to one or more of three parties: the developer, the manufacturer, and/or the user. Any other parties simply would not have enough causal link with the injury for it to be a legitimate claim.

In the current state of the art for AI systems, they can only be treated as tools or products. They are not sentient and any attempt to treat them as such is faulty. This is why the liability for negligent injury during interaction with an AI system has to squarely lie on the humans or natural persons behind the training, manufacturing, and usage of the system. And the law of tort of negligence is more than capable of being applied to natural persons while taking into account the increased risks to injury that an AI environment poses.