UK Jurisdiction Taskforce (UKJT)

Consultation on the Legal Statement on Liability for AI Harms under the private law of England and Wales

Introduction by Sir Geoffrey Vos, Master of the Rolls

January 2026

I am delighted to introduce the UK Jurisdiction Taskforce's consultation on its draft Legal Statement on Liability for AI Harms under the private law of England and Wales.

The drafting team has, I believe, done an excellent job of addressing legal issues that need clarity in our world of ever-increasingly capable Artificial Intelligence.

I hope as many lawyers and technologists as possible will respond to the consultation to indicate where they think the draft Legal Statement can be refined or improved before it is published in final form.

I would like to thank both the UKJT's drafting team of Matthew Lavy KC, Lucy McCormick, Iain Munro, Isabel Barter and Jacob Turner for their excellent work thus far, and also the wider expert group comprising Professor Ryan Abbott, Lawrence Akka KC, Matt Frank, Prof Sarah Green, David Quest KC, Tom Whittaker, Dr Peter Wills and Michael Workman for their invaluable input.

Sir Geoffrey Vos

Master of the Rolls and Head of Civil Justice in England and Wales

Consultation on Liability for AI harms under English private law

Section 1: Background to this consultation

1. The last few years have seen rapid technological developments in the field of AI and equally rapid expansion in the size of the AI sector and the sphere of activities in which it operates. At the same time, there has been increasing awareness of the potential for AI to cause various kinds of harm. In response, there have been a variety of regulatory and legislative initiatives in several jurisdictions seeking to impose governance frameworks around the use of AI and prohibiting the use of AI for certain purposes; professional bodies have also started to regulate the use of AI by their members. However, such initiatives have not generally engaged with the question of private law liability for harms caused by AI. In England & Wales, that question is governed by the general (not AI-specific) law.

2. The lack of any AI-specific liability regime (which, as noted, is not unique to this jurisdiction[1]) gives rise to a perception of legal uncertainty. Specifically, there is a perception of uncertainty as to whether a person harmed by AI would have a right of recourse and, if so, against whom. Any uncertainty, genuine or perceived, risks inhibiting the adoption of beneficial AI tools, particularly amongst risk-averse professional sectors; the perceived uncertainty also risks causing a misallocation of resources such as insurance amongst the various actors in the AI supply chain. It also risks inhibiting research and development activities by those who might otherwise develop beneficial products.

3. As the UKJT has noted previously, English law[2], as a well-developed flexible common law system, has the ability to provide certainty and predictability in the context of technological innovation. In areas of true novelty, certainty and predictability emerge over time as Courts reason from first principles and develop existing rules to address the novelty in question. However, it is important to distinguish between those areas of true novelty (which are rare) and areas where, although the factual background may be novel, the application of well-established legal principles is reasonably straightforward. Many potential harms caused by AI fall into the latter category.

Section 2: Scope of the Legal Statement and this consultation

4. LawtechUK is a Ministry of Justice backed initiative, supported by CodeBase and Legal Geek, that is dedicated to driving digital transformation in the legal sector, by increasing innovation and adoption of lawtech in the delivery of UK legal services, supporting the growth of the lawtech sector in the UK, and enabling English law and the UK's jurisdictions to become the foundation for emerging technology.[3] The UK Jurisdiction Taskforce ("UKJT") was established by the LawtechUK Panel. Its initial activities were focussed on clarifying key questions regarding the legal status of, and basic legal principles applicable to, crypto assets and associated technologies. It now seeks to perform a similar function in relation to legal issues arising out of AI. It takes a long time

[1] The EU AI Act is probably the most comprehensive and well-known example of AI legislation in any jurisdiction. While it contains extensive regulatory measures, it does not address private law liability for AI harms.
[2] In this paper, references to "English law" should be read as references to the law of England and Wales.
[3] For further background, see https://lawtechuk.io/ (accessed December 2025).

for cases in new subject areas to come before the courts and for judicial decisions to be made on points of difficulty or differentiation. In areas of rapidly advancing technology, it is helpful to the business and tech community to know more quickly how the law will likely apply. In those circumstances, the role of UKJT Legal Statements is to explain as quickly and as accurately as possible, through the voice of experts in the field, how the common law is likely to deal with private law problems thrown up by the new technologies, thereby seeking to provide as much legal certainty as possible in a rapidly changing technological landscape. The UKJT's role is not to propose policy-based reforms of the law (as to which, it does not take a position).

5. In pursuit of this objective, and following on from previous Legal Statements on Crypto Assets and Smart Contracts, on Digital Securities, and Digital Assets and English Insolvency Law, the UKJT has co-ordinated the preparation of an authoritative Legal Statement on Liability for non-deliberate AI harms under English private law.

6. A draft of the Legal Statement is annexed to this Consultation Paper. The overarching question that it seeks to address is in what circumstances, and on what legal bases, English common law will impose liability for loss that results from the use of AI.[4] The primary focus in addressing that question is on the liability of those who did not *deliberately* cause harm. English law generally has means of ensuring legal redress against deliberate wrongdoers, and the fact that AI may have been used as a tool to inflict deliberate harm is unlikely in most circumstances to affect the legal analysis. At the heart of the Legal Statement, therefore, is an analysis of the application of the law of negligence to physical and economic harms caused by AI. In addition, the Legal Statement addresses the following questions:

    - Does the principle of vicarious liability[5] apply to loss caused by AI?

    - In what circumstances can a professional be liable for using or failing to use AI in the provision of their services? If AI used in the provision of professional services produces erroneous output, is the professional liable for loss resulting from the error?

    - Can a person ever be liable for harms caused by use of AI where there is no fault on their part?

    - Does liability attach to false statements made by an AI chatbot?

7. The purpose of this consultation is to seek input from stakeholders as to whether the draft Legal Statement sufficiently addresses key issues of perceived uncertainty within its scope, to inform the finalisation of the Legal Statement.

Section 3: Overview of key issues of legal uncertainty included in this consultation

[4] This excludes e.g. competition claims and economic torts from scope.
[5] Vicarious liability is the legal principle whereby one person (such as an employer) can be liable for the acts of another (such as an employee).

8. This consultation and the Legal Statement are necessarily limited in scope. They are focused on specific aspects of private law liability to the deliberate exclusion of other topics. In addition to deliberate harm (which receives some, but limited, treatment), matters such as criminal law, competition law, taxation, intellectual property law, the law of contract formation, regulatory law and public law liability are intentionally out of scope. We recognise that these are important areas, and that industry participants may feel there exists a degree of legal uncertainty in relation to some of them. However, in relation to some of those topics (e.g. regulatory law) other bodies or organisations are better placed to provide any necessary clarity[6]; in relation to others (e.g. law on formation of contracts), while the UKJT may be well placed to provide clarity, we consider that the need for a Legal Statement is less urgent and that a Legal Statement on private law liability for physical and economic harm caused by use of AI is more pressing and should not be delayed to allow for a more wide-ranging analysis.

9. There is no universally agreed definition of AI.[7] For the purposes of the Legal Statement, we adopt a definition that is technology-agnostic and that seeks to capture the key characteristics of AI that are novel and that therefore give rise to the perceived legal uncertainty with which the Legal Statement is concerned.[8] Our view is that the key such characteristic is autonomy, a term which in this context we use to capture (i) an unpredictable relationship between input and output, (ii) opacity of reasoning, and (iii) a limited ability on the part of a user to control output. The ability of some AI to "learn" and adapt its behaviour during its operation is another characteristic that is often identified in legal treatments of AI as being important.[9] However, we consider that adaptability is unlikely to give rise to any legal issues falling within the scope of this project that would not arise in any event by reason of autonomy.

10. Our approach to the legal analysis of liability for loss caused by AI starts from the premise that AI does not have legal personality in English law and that, therefore, it cannot itself be held legally responsible for physical or economic harm; instead, liability for harms that arise from the use of AI must be attributed to legal persons, using ordinary legal principles.

11. In commercial situations, the question of whether a person is liable for a particular harm and, if so, which person, can often be governed by contractual agreements between parties. As between members of the AI supply chain, contract is the primary – and often sole – basis for allocating liability. The extent of liability and ability to pass losses up the chain are typically determined by warranties, indemnities, limitations and exclusions. It is in contexts where there is no relevant governing contract that the main legal issues for analysis arise.

12. Where there is no relevant governing contract, the question of whether a person is liable for physical or economic harm caused by use of AI turns on whether the law imposes a

[6] For example, the various regulators in relation to regulatory law, the CMA in relation to competition law and the Law Commission in relation to criminal law.
[7] This point was recently made by the Law Commission in *AI and the Law: A discussion Paper* (2025).
[8] Thus the purpose of our definition differs from that of e.g. the EU AI Act, which seeks to define a regulatory ambit.
[9] See, for example, the Law Commission paper.

non-contractual duty on that person to protect against that harm. The primary (albeit not sole) legal framework governing liability for such harm is the law of negligence. Therefore, a core part of the Legal Statement is an analysis of how the law of negligence can be applied to harms caused by AI and, in particular, how duties of care are likely to arise, what standards the courts are likely to apply to developers and users of AI, and how the courts will likely approach causation in the context of technology that is autonomous (as that term is used above). In the context of physical harm, the common law is supplemented by the Consumer Protection Act 1987, which implements the original EU Product Liability Directive[10] and imposes "no fault" liability for defective products. We therefore consider the extent to which that statute applies to AI harms.[11]

13. Important categories of circumstances in which AI might be expected to cause economic harm are those that arise where a professional uses AI, or where an AI "chatbot" or equivalent produces false statements and thereby provides false information or advice. In order to address perceived legal uncertainty in these areas, the Legal Statement analyses the application to AI of the law of professional negligence and also the main common law principles applicable to the generation of false statements: negligent misstatement and defamation.

Section 4: Consultation questions

14. Your input is sought in relation to the following questions:

    - Do consultees agree that the subjects that we address in the Legal Statement are appropriate and useful, and have been addressed in an appropriate and useful way? If not, what alternative issues within scope of the project do consultees think need to be addressed in this Legal Statement, or in what alternative manner would consultees wish those issues to be addressed?

    - Do consultees agree that the approach to defining AI in the Legal Statement is appropriate? If not, what alternative approach would consultees take?

    - Do consultees have any other comments that they would like to make on the approach or conclusions of the draft Legal Statement?

Section 5: Consultation process

15. This consultation will remain open for responses until 13th February 2026. Once this consultation has closed and the results have been considered, it is intended that the final version of the Legal Statement will be published as soon as possible thereafter. It will then be possible to see whether any further steps are necessary or appropriate.

[10] Directive 85/374/EEC.
[11] The Law Commission has recently launched a project to review the law in this area. See https://lawcom.gov.uk/news/law-commission-to-review-the-law-relating-to-product-liability/).

16. Written responses to the consultation questions should be provided by email to ukjt@justice.gov.uk.

17. The UKJT will also be hosting a virtual public event on 27th January at 5pm in order to receive feedback on the consultation questions in person. Further details on the public event, including details regarding registering for the event, can be found here: https://lawtechuk.io/events/public-consultation-on-liability-for-ai-harms-27-feb

Annex: Draft Legal Statement

Draft Legal Statement

January 2026

How to define Artificial Intelligence

1 Artificial intelligence, or AI, is a term which is widely used yet lacks a single universally accepted definition. Numerous candidate definitions have been posited over the last 70 or so years, since the term was first used.[1] This is not the place to analyse and critique each of those definitions.[2] Nor is it appropriate here to provide a technical explanation of the various technologies traditionally referred to as AI.[3] Instead, we start by explaining the approach we have taken to the definitional question for the purposes of this Legal Statement, before introducing the definition that we have adopted.

2 Most AI definitions fall broadly into three categories.[4] First, there are those that are 'human-centric', in that they refer intelligence back to the ability to undertake tasks that would normally require 'human intelligence'.[5] A problem with such definitions is that they raise the difficult question of what 'human intelligence' means.[6] A second issue is that human-centric definitions can be over-inclusive – indeed, on one view, *any* program or application, no matter how simple, would be deemed AI using this approach.[7] For present purposes, a human-centric definition is unhelpful because it does not provide a basis to distinguish between characteristics of AI that warrant distinct legal analysis and those that do not.

3 In the second category are 'teleological' definitions, which suggest that intelligence lies broadly in doing the right thing at the right time, or achieving some particular goal.[8] A problem with such definitions is that they tend to assume that in any activity there is always a set goal or outcome which can be identified and achieved, which may not be the case for at least some AI systems, especially those which can set their own goals and / or which are general purpose in nature. As with human-centric definitions, a teleological approach is also unlikely to be useful for present purposes.

4 Finally there are 'characteristics-based' definitions. These focus on attributes of the technology, its characteristics and internal processes, rather than the outcome.[9] A downside of such definitions is that they lack the intuitive appeal of the above two categories and they require users to engage with various further sub-concepts engendered by the capabilities in question. It can also be challenging to categorise systems. However, we consider that a characteristics-based definition of AI is the most appropriate type for the purposes of this Legal Statement, because the characteristics-based approach allows the identification of those characteristics of AI that distinguish it from other technologies in a way that warrants distinct legal analysis.

5 It would in principle be possible to avoid the definitional question entirely and to scope the Legal Statement by reference to a particular technology, most obviously Machine Learning (ML) technology, on the grounds that this is the most widely used type of technique commonly referred to as AI.[10] While this approach would have the benefits of simplicity, we consider that it would be problematic in three respects. First, it may render the Legal Statement obsolete, given that AI technology is developing quickly, and that which is the state of the art at the time of writing may be overtaken even by the time of publication.[11] Second, by tying the analysis to a technology and not identifying its legally salient characteristics, we may fail to grapple with the underlying legal and practical problems that the technology may cause and that this

Legal Statement is seeking to analyse. Third, given the aim this Legal Statement to assist with practical issues, a focus on characteristics may be of greater utility than a focus on purely technological features, since the former are perhaps more likely to give rise to external effects and hence issues concerning liability.

6 Two further factors inform our approach to the definitional question. First, is the importance of identifying what, if anything, is legally unique about AI such that it may not be clearly addressed in existing legal and regulatory systems. That is key given that a principal goal of this Legal Statement is to address perceived legal uncertainty. Second is the desirability of aligning (insofar as it does not conflict with our primary goals) with definitions of AI that have been adopted in a legal context internationally (in particular the approach of the UK's closest geographical trading partner: the EU),[12] and by the UK Government in various regulatory consultations and in guidance. Consistency assists with legal certainty and, it is hoped, will make the Legal Statement more broadly useful.

What is Artificial Intelligence?

7 For the purposes of this Legal Statement we define Artificial Intelligence as "**a technology that is autonomous**".

8 Looking at each limb in turn:

(a) 'Technology' means for these purposes that the relevant entity or system does not occur in nature and has come about through some form of systematic process. We do not limit our definition to '*man-made*' technology, because of the potential for certain AI systems to create or design new AI systems, which might not accurately be described as *man*-made. We also do not limit the definition to 'machine-based'[13] technology because of the possibility of

biological-based systems being used to create the relevant capabilities in coming years.[14]

(b) 'Autonomous' in this context is used in the technical sense of meaning an entity that can generate outputs which have not been determined or programmed in advance. Although the word has different (and broader) meanings in other contexts, we use it here as a term of convenience to capture the characteristic of there being a loose and opaque coupling between input and output. This characteristic can be further elaborated as: (i) an unpredictable relationship between input and output; (ii) an opacity in the reasoning process between input and output; and (iii) a limited ability on the part of the human user to control output.[15] AI systems can thus be distinguished from those which are merely automated,[16] such as traditional, logic-based computer programs.[17]

9 The above definition is intended to be technology-agnostic. Whilst being broad in scope, it is intended to draw a clear perimeter around what we are calling AI for the purposes of this Legal Statement. The characteristics captured by our definition are legally salient because – taken together – they are unique to AI and, in our view, are responsible for much of the perceived legal uncertainty around AI (a point we expand upon below). Thus, neither the width of the definition nor the possibility that it will not always be clear whether a particular system falls within it is problematic, because we are not seeking to define a regulatory ambit but rather identifying characteristics that we consider warrant particular analysis under English law.[18]

10 The definition is largely aligned with that which the UK Government put forward in a White Paper on AI in 2023 and which at the time of writing remains the UK Government's preferred definition;[19] and it is broadly consistent with the definitions adopted by the EU in the EU AI Act,[20] and the Organisation for Economic Cooperation and Development (OECD) (which is materially the same as under the EU AI Act).[21]

11 The utility of our definition of AI can be tested by application to current technologies. ML systems (including generative AI systems) will, as a rule, fall within the definition because by their nature they are autonomous (in the sense explained above).[22] By contrast, traditional computer software, being deterministic, will not fall within the definition. Neither will systems which may have been labelled as 'AI' in other contexts (such as 'Good Old Fashioned AI' or 'symbolic AI') which, despite their name, are deterministic.

Why might AI give rise to uncertainty?

12 English private law (in common with all other legal systems) has never previously needed to address the capability of autonomy (as defined above) other than in humans.[23] Under current liability systems, the paradigm for legal responsibility is where harm has been caused by the voluntary actions or inactions of a person (natural or legal but with a human agent). The further a scenario moves from that paradigm, the less obvious it is how liability is to be ascribed. AI thus presents challenges owing to its ability to arrive at outputs and develop methodologies which have not been pre-programmed by any human decision-maker – AI harms may be a long way from the paradigm.[24] That being so, it is unsurprising that AI gives rise to perceived uncertainty.

13 Perceived legal uncertainty concerning AI is not the same as *actual* legal uncertainty. As we illustrate in this Legal Statement, there are many scenarios in which AI is deployed and there is no special difficulty in predicting the outcome under current laws.[25]

14 As to *actual* uncertainty, there is a lack of precedent for liability arising from AI. However, as the UK Jurisdiction Taskforce commented in a previous Legal Statement, 'the great advantage of the English common law system is its inherent flexibility'.[26] We address further below the question of whether such flexibility is capable of providing sufficiently

clear answers to the new questions raised by AI development and operation.

Other relevant features of AI systems and their operation

15 It is important to clarify that the characteristics of AI identified for the purposes of our definition are not the *only* features of AI that have the potential to raise legal difficulties. There are several other features commonly associated with the way that AI functions, is understood or is deployed which can also give rise to challenges in establishing who ought to be held responsible and consequent uncertainty. For example:

(a) The opacity of many AI systems can (separately from their autonomous nature) mean it is difficult to establish why they have made certain decisions, and so to link the output to any individual party's action or omission. Specifically, it is widely acknowledged that the output and inner functioning of at least some ML systems may be difficult to predict or explain through traditional concepts of cause and effect, and indeed to express in natural language.[27] These issues are especially acute when attempts are made to extend liability backwards through a supply chain to the originator of a system.

(b) Relatedly, AI may be integrated into complex supply chains, with multiple suppliers. Both ascertaining which party or parties are liable and in what share can present further uncertainty.

(c) The harms caused by AI may occur at scale and may also be diffuse in nature, if a single system is widely deployed. An AI system which is widely deployed, for example by a high street bank, major airline or a branch of a government may lead to individual harms occurring on a mass basis. In circumstances where each individual instance of harm is material but (financially) minor when viewed through the prism of civil litigation then pursuing such issues in

Court may not be especially attractive, particularly where the defendant is likely to be a powerful and well-resourced entity.[28]

(d) The speed of AI's functioning and, by extension, potential failure, may also give rise to challenges as regards civil liability – especially in terms of rendering pre-emptive steps difficult owing to a combination of the speed and difficulty of predicting harm.

(e) AI may be integrated into more complex decision-making processes, which could involve some degree of human decision-making in conjunction with AI output and recommendations. The interaction between decisions which are deemed to be those of an AI system and those which are deemed to be those of a given human will be a matter of some complexity which is likely to turn on the facts of any given scenario.

(f) AI is typically implemented as software. That software can be complex, and different parts of it can be developed and trained by multiple different persons, leading to evidential uncertainty as to what has gone wrong and who should be held responsible.

16 These issues can be important in practice, but none of them is unique to AI; all such features occur in other domains. For example, opacity frequently arises in situations where there has been environmental exposure to a particular harm-causing substance, but it is unclear which point of exposure has led to harm.[29] Likewise, many modern technologies (such as aircraft) have multiple components, a failure of which might be due to the interaction of several different elements, and different parts of which might have been developed by different organisations. The difficulty of litigating and case managing mass harms frequently arises in non-AI contexts, and has led to the failure of several high-profile claims.[30] Finally, speed of failure and a lack of opportunity to take pre-emptive steps is not unique to AI. Such issues arise in many processes (technological and otherwise).

17 That being so, while we do touch upon some of these issues as potential practical barriers to the prediction and establishment of liability, they are not our primary focus.

Actors and Harms

18 AI supply chains[31] can differ – there is no universally adopted structure. Nor are AI supply chains static: developments in the AI ecosystem continue to impact upon the number and role of actors involved.[32] The trend is towards greater complexity and greater numbers of actors rather than less complexity and fewer actors.

19 This is partly due to the changing nature of AI systems themselves. Whereas early AI systems in the 2010s were typically 'narrow' in ambit, in that they were capable of fulfilling a single task, there has been a shift in recent years towards systems that are 'general purpose' in nature and can be adapted and applied to multiple domains. The most well-known examples of such general-purpose AI are those commonly described as 'foundation models': AI that is trained on a large amount of data and is adaptable for use on a wide range of tasks. Foundation models can be used as a base for building more domain-specific or task-specific AI models. Foundation models may be deployed directly to end users or, as set out below, integrated into more complex supply chains.[33]

20 To add to the complexity, the different stages of an AI supply chain can, with technological developments, become combined. For instance, the developer of ChatGPT, OpenAI, announced in late 2023 the release of a set of applications called 'GPTs' which enable non-specialist users to build bespoke models designed to fulfil particular roles, using natural language prompts.[34] Amazon Web Services offers similar services to non-specialists via a suite of systems including Amazon Bedrock which offers users the ability to build and scale applications with generative AI tools.[35] Similarly, agentic AI enables non-specialist users of the technology to take steps which were previously only possible through specialist

suppliers and programmers. Such advancements may blur to some extent the distinction not only between the different layers of developers involved in creating AI systems but also the distinction between users and developers of AI systems and applications.

21 The Ada Lovelace Institute provides a useful example of how a reasonably common AI supply chain might appear, at least where there is a foundation model.[36] Their scheme identifies a set of actors in the foundation model supply chain that we gratefully adopt, with some additions, in this Legal Statement:[37]

(a) Data Providers: collectors, curators or processors of data used to train a model.

(b) Compute Providers: providers of the computational resources needed for data processing and model training.

(c) Foundational Model Developers: those who design, develop and train a foundation model.

(d) Hosting Suppliers: providers who host trained foundation models and make them available for use by end users or application developers.

(e) Brokers: entities acting as intermediaries between Hosting Supplier and Users.

(f) Application Developers: those who design, build, test, deploy, operate, maintain and secure applications that use or that build upon foundation models or fine-tuned versions.

(g) Users: those who use the application. For legal purposes, there are distinctions to be made between different categories of users, to whom differential rights and duties might be accorded, specifically: (i) professional users; (ii) non-professional commercial users; and (iii) non-professional non-commercial users.

22 The fact that there are various roles within a supply chain does not necessarily mean that each role is played by a separate independent actor. Rather, with some foundation models, there is vertical integration such that the Foundation Model Developer, Host and Application layers are all played by a single party. That is the case, for example, with the ChatGPT application, which is developed, hosted and provided by OpenAI. By contrast, other of OpenAI's offerings involve separate parties acting as host and application developers – its enterprise version, GPT-4, enables external plugin developers to utilise the foundation model infrastructure.[38]

23 There are three further categories of actor, that though outside the *supply chain* of AI are nonetheless important for the purposes of considering liability:

(a) Affected Third Parties: Such a (natural or legal) person is someone who is not part of the supply chain and is not themselves using an AI system but are the recipient of goods or services that have been generated using AI or are affected by such use.

(b) Advisors / Consultants to Supply Chain Actors: parties which are not directly involved in the design and development of AI, but which provide professional services to those that do (for example lawyers or consultants who advise on aspects of the processes whereby the AI is designed and deployed).

(c) Bad Actors: parties who deliberately utilise AI to cause harm.

In what circumstances, and on what legal basis, will English common law impose liability for loss caused by the use of AI?

24 Despite the novel and unique qualities of AI, English law is capable of attributing liability to individuals or corporate entities (i.e. to "legal persons" referred to here, for brevity, as 'persons') [39] for loss caused by

AI. AI does not have legal personality in English law, and so cannot be held liable in its own right.

25 In general terms there are two sources of liability for AI-caused loss: (i) responsibilities that parties have voluntarily taken on; and (ii) liability imposed by law regardless of whether a party has chosen to take on that liability.

26 A common legal mechanism by which a person voluntarily takes on responsibility for a risk is via a contract. In particular:

(a) As between actors in the AI supply chain, contract will be the primary (and in many cases only) basis on which liability is allocated. For example, Data Providers, Foundation Model Developers, Application Developers and Users will all be directly or indirectly connected via a contractual chain.[40] Generally, each actor will have a contract with the next link up the chain and a contract with the next link down the chain. For example, an Application Developer may have an upstream contract with a Foundation Model Developer and a downstream contract with a User. Sometimes, the chain will branch: very often, for example, a Foundation Model Developer will have a contract with both an AI Application Developer and the User of the Application Developer's application.

(b) Each actor's liability for losses suffered further down the supply chain will be primarily determined by the contractual terms agreed with the next link down the chain; and each actor's ability to pass its own losses up the chain will turn on the contractual terms agreed with the next link up the chain. A User who suffers loss may be able to sue the Application Developer (and possibly also the Foundation Model Developer if there is a direct contractual relationship); the Application Developer may be able to pass on some or all of that liability to whoever is next in the chain and so

on. Whether, and the extent to which, a person is liable for physical or economic harm caused by use of AI will turn on how risks have been allocated through the terms of the relevant contracts, including warranties, indemnities, limitations and exclusions.

(c) Even where it is an Affected Third Party who suffers loss, liability can also be governed by contract. That is because although the Affected Third Party will not have any contractual relationship with a supplier of AI, the use of AI that caused their loss can frequently have been by a User with whom the Affected Third Party is in a contractual relationship. That User will normally be the obvious person against whom to make a claim.

27 The extent to which a contract governs loss caused by AI will depend primarily on its express terms, though in some cases terms will be implied to supplement or even exclude what the parties have agreed. Even where a contract is in place, it may not always cover the harm caused.

28 It is also possible for a person to suffer physical or economic loss in circumstances where there is no contract, and therefore no contractual counterparty to sue.[41] For example, a patient may be harmed by a doctor's use of AI in a non-private (so non-contractual) healthcare setting (e.g. where medical services are provided through the UK National Health Service), or may be harmed by an autonomous vehicle owned and operated by a stranger. In such circumstances, the question of whether liability arises or whether the losses instead lie where they fall will turn on whether the law imposes on any person a non-contractual duty to protect against the loss that has eventuated. The same applies where there *is* a contract in place but one that does not govern the harm that has arisen.

29 Where there is no contract governing the losses, the question that arises is whether the characteristics of AI – and in particular its autonomy –

preclude the application of other legal doctrines or cause them to be ineffective in an AI context.

30 Much of the remainder of this Legal Statement is focused on those questions. However, it is important to recognise that the cases that require this sort of analysis are likely to be comparatively rare in practice; the most important legal framework governing liability for loss caused by use of AI in most circumstances will be the chain of contracts that the various actors involved in the supply and use of AI have agreed.

Does the principle of vicarious liability apply to loss caused by AI?

31 Vicarious liability is a principle of law pursuant to which Anita can be liable for wrongs committed by Ben even where Anita is not personally at fault. The usual case where this arises is in the context of employment, where an employer is held vicariously liable for torts committed by employees in the course of their employment.[42] Vicarious liability can also arise in other contexts, where a person has authorised or ratified the torts of another.[43]

32 As vicarious liability involves Anita being liable for the torts of another *person*, it is not capable of causing her to be liable for the actions or failures of an AI because AI does not have legal personhood.

33 By contrast, it *would* be possible for Anita to be vicariously liable for AI harm in circumstances where that harm arose because Ben acted negligently (or otherwise tortiously) and thereby caused the AI harm in question. So, for example, an employer could – and generally would – be held vicariously liable for physical or economic harm caused by negligent use of AI by an employee. The analysis here would be no different from any other situation in which an employer is liable for the actions or omissions of its employee. Assuming the employee was acting within the course of his or her employment, an important question on which liability would turn is whether the employee would be liable in

negligence for the AI harm, in which case that liability can be ascribed to her employer. As with all vicarious liability, would also be possible for a victim to elect to seek damages from the employee directly, but in most cases it will be preferable to pursue the employer, who will likely have deeper pockets and perhaps also insurance. The principles governing negligence in an AI context are addressed in the sections that follow.

Liability for Physical vs Economic Harm

34 AI is generally software and thus generally operates initially in the non-physical plane, i.e. electronic output. Therefore, harm may often be in the form of economic loss rather than physical damage, for example an inaccurate AI prediction of the price of a commodity leading a company to make a loss-making trade. However, AI is also capable of causing physical harms, both in the form of property damage and personal injury (which includes physical and psychiatric). The potential for physical harms arises most obviously in the context of 'embodied' AI, such as autonomous vehicles, medical robots or assembly-line machines. Such embodiment occurs where the AI output has direct control over mechanical instruments concerned. That said, AI need not have be embodied to cause physical harms. For example, several AI models are being used for cancer detection imaging, where both a false positive or a false negative – if not noticed and corrected by a doctor – have the potential to cause harm to the patient being assessed. In such cases, the physical effect of the AI output may be indirect, but nonetheless cause harm.

35 Where the person suffering harm (whether physical or economic) is a User, the primary recourse is likely to be found in the contract (if any) between the User and the AI supplier, and liability for the harm will likely be governed by the agreed contractual risk allocation (subject to a statutory prohibition on a party contractually limiting its liability for negligently causing death or personal injury).[44] As already noted, where the person suffering harm is an Affected Third Party, they too will often

have contractual recourse pursuant to a contractual agreement with the User whose usage of the AI caused the harm. However, there will remain circumstances where contractual rights are unavailable, particularly where physical harm is concerned (one example being a pedestrian knocked down by an automated vehicle[45]).

36 In cases where there is no contractual liability to the Affected Third Party, liability can arise only through alternative routes. Setting aside the liability of those who have deliberately caused harm using AI as a tool,[46] such routes will comprise a fault-based claim in the tort of negligence, and/or, in the case of physical harms involving death or personal injury, a 'no fault' claim under the Consumer Protection Act 1987.

37 While the negligence analysis is essentially the same whether the harm suffered is physical or economic, in the latter case English law is far more restrictive: there will be no liability for economic harm unless there is a "special relationship" between the parties involving one having voluntarily assumed a responsibility to the other;[47] that special relationship will generally (but not always) involve a contract.[48] In practice, the circumstances in which use of AI can give rise to negligence liability for economic harm will generally either involve professional negligence (for example, a financial advisor who negligently relies on AI output to give loss-causing investment advice) or will involve statements made by AI (such as false statements generated by a chatbot). Consequently, those are the economic loss scenarios on which this Legal Statement focuses.

38 In the remainder of this section of the Legal Statement, we first address liability in negligence brought against persons in the AI supply chain by a User or an Affected Third Party who has suffered *physical* harm as a result of an AI failure, before turning to economic harm in the context of professional negligence. We then return to physical harms to consider "no fault" claims under the Consumer Protection Act 1987 before

addressing issues of causation. Liability for statements made by AI is addressed in a separate section of the Legal Statement.

Negligence liability for Physical Harm

39 To meet the tests liability for damages in negligence, a person who has suffered physical harm as a result of use of AI would need to establish (i) the presence of a duty of care (ii) the relevant standard of care (iii) that there has been a failure to meet that standard of care, (iv) that the negligence caused the loss alleged, and (v) that the loss was sufficiently foreseeable that the law permits recovery (i.e. that it is not "too remote"). Negligence liability is a long-standing aspect of English law, and the principles to be applied in each step of the analysis are well established.[49] The law of negligence also has a track record of flexibility, having evolved and adapted over the years to suit changing societal expectations and circumstances. Consequently, there is no conceptual reason why its principles cannot be applied to harms caused by AI failures even though it does not appear to have been so applied to date.

40 A person (Anita) will owe another (Ben) a duty of care where the harm is foreseeable, where there is relationship of "proximity" between Anita and Ben, and where it is fair just and reasonable to impose the duty on Ben to protect Anita from the harm. For many factual situations, the existence of duty of care is already well established – for example a duty of care exists between doctor and patient, a manufacturer and a consumer, or a driver and a pedestrian. In a novel situation, the court still looks to precedent, and will do so by considering the closest analogies in the existing law and applying it by incremental extension. Thus for example, because a radiologist owes a duty of care to those whose medical imaging scans he or she is reviewing, that duty of care is likely to remain whether that review is taking place with the naked eye or with the assistance of sophisticated AI analysis. In other words, in many cases the addition of AI will simply be considered a tool of those who exercise relevant control over it and can be said to have been "responsible" for its

actions; it would not change the underlying position as to whether or not there is a duty of care in any given circumstance.

41 How far up and down the AI supply chain the duty of care extends will be highly fact sensitive. In the case of a failure to interpret scans correctly, the radiologist (in other words, the User) is the most obvious person to owe a duty of care. However, other actors in the supply chain may – depending on the facts – also be liable. The Application Developer responsible for the faulty radiology diagnosis may owe a duty if, for example, it knew or should have known that errors in its output were likely to exist, be difficult to detect by human radiologists, and to cause harm.[50] Liability of the Application Developer (or indeed any other person in the AI supply chain) may be more likely where they have previously taken active steps to push out updates to their software to address risks of harm.[51]

42 In most circumstances, Foundation Model Developers are unlikely to owe a duty to protect against harms that arise as a result of their model being used for unforeseeable purposes but not sufficiently tested by an industry participant further down the supply chain, such as an AI Developer. However, Foundation Model Developers may be held to owe a duty where harm is foreseeably suffered as a result of foreseeable use.[52] The analysis is in principle no different from any other situation where an upstream component fails causing damage.[53] However, difficulties could arise on the facts where it is unclear whether a particular use of a foundation model was foreseeable, given that foundation models are typically general purpose in nature and can thus be put to an extremely wide range of uses, some of which will be novel.

43 Where it is established that Anita owes a duty of care to Ben, Anita does not have an absolute obligation to protect Ben from harm. Rather, her duty is to act in accordance with the relevant standard of care. In essence, this means she has a duty to act as a reasonable person in her position should have acted in the circumstances. Anita is not negligent because

harm occurs; Anita is negligent because she fails to do what a reasonable person would have done in the circumstances in which she found herself; or she does something that a reasonable person would *not* have done. The question of whether the standard of care has been met is inevitably highly fact-specific. Industry guidance (for example that found in the AI Standards Hub[54]) is likely to prove a useful yardstick for the courts in relation to the question of what represents reasonable practice. Expert evidence will also be key, particularly in niche or complicated use cases. In cases against Users and those who deploy AI systems, the enquiry will often need to examine not only the AI that was deployed and tested, but whether reasonable steps were taken to select a model, and indeed to decide whether it was reasonable to use AI at all.

44 One challenge that can arise in relation to AI is that the development of the software and model itself can have been undertaken, at least in part, using AI. It can be difficult to assess whether there has been "reasonable care" in the programming of the software, in circumstances where the software's functioning has developed through autonomous processes, including machine learning, such that it is opaque to its own developers. Therefore, the focus may instead be at the very beginning and end of the process—to the thoughtful selection of data inputs provided to the machine learning software for it to 'learn' from, and to testing the outputs of the program.[55]

45 Another challenge is that AI, being autonomous, can find unexpected ways to "game" the specified objective—generating a solution that strictly satisfies the stated objective but fails to solve the problem according to the designers' intent. [56] For example, a programmer, who seeks to program a robotic vacuum to navigate without bumping into things, sets up a reward scheme to encourage speed and discourage hitting the bumper sensors. The vacuum learns to drive backwards, because it has no bumper sensors on the back. If the manner of failure would have been genuinely difficult for a reasonable programmer to predict, it may be that a negligence claim would fail.[57]

46 The autonomous nature of AI (in addition to other factors) can pose issues of causation – i.e. establishing that a failure to meet the required standard of care was the cause of harm (or, to put it differently, that the harm would not have been suffered in any event). That issue is addressed later in this Legal Statement.

47 In summary: whether and when a person involved in the development, supply, or deployment of AI might be liable in negligence for physical harms is highly fact sensitive. However, in many cases there is no difficulty in applying the relevant principles in an AI context, extending existing precedent by analogy to the extent that that is required.

48 The analysis above in relation to physical harms is also, in principle, applicable to economic harms subject to the restriction already noted, namely that a duty of care to protect a party from economic harm (unless that economic harm flows from physical damage) is limited to situations in which a party voluntarily assumes responsibility to protect against that harm. This issue is addressed in more detail in the section of this Legal Statement on liability for statements made by chatbots.

49 Although no party can exclude its liability for personal injury or death caused by its negligence where such liability would otherwise arise, contracting parties within an AI supply chain may agree *as amongst themselves* who should bear the cost of such liability. In practice, we would expect liability risk to be managed through a combination of contractual indemnities and insurance.

> ***Example:*** *After an extensive due diligence process, FactoryCo enters into a contract with SupplyCo under which SupplyCo will provide FactoryCo with automated equipment for industrial assembly lines. This equipment uses AI to identify when a human is nearby and automatically pause operations to avoid any injury. SupplyCo normally develops such equipment with a mixture of real world testing and virtual testing using simulated environments. Unknown to FactoryCo, on this occasion SupplyCo did no real world testing, only virtual testing. The simulated environments were generated by a Foundation Model supplied by FoundationInc. One day,*

Anita, a visitor from a factory overseas arrives for a tour. She is a wheelchair user. The automated equipment fails to recognise her silhouette as human and so continues operation while she is nearby and she sustains an injury.

Anita's most obvious claims are against FactoryCo in this scenario.[58] *But FactoryCo are close to insolvency and are not adequately insured. Anita therefore wishes to explore the possibility of bringing claims in negligence against SupplyCo or FoundationInc:*

SupplyCo are potentially negligent, depending on all the evidence on industry practice which would inform a Court's view on what a reasonable developer would have done. At time of writing, it would be unusual not to do any real world testing for embodied AI (e.g. a robot or autonomous vehicle).[59] *That said, there may be a causation issue on these facts – would the failure to identify wheelchair users have been identified if real world testing had been utilised?*

FoundationInc. is unlikely to be deemed negligent. There may be bias in the foundation model stemming from its training data, which can reflect and amplify societal inequalities. That may have played a causal role in the automated equipment's failure to recognise a wheelchair user. Nonetheless, it is difficult to see how FoundationInc. owes a duty to Anita. It supplied a general purpose model and never claimed that it was trained in a way wholly free of bias. Neither did it ever claim that its model was suitable for use in factories. FoundationInc. has no control over and is not responsible for the detail of how it is deployed or whether that use case is adequately tested for a particular purpose.

In what circumstances can a professional be liable for using or failing to use AI in the provision of their services? And if AI used in the provision of professional services produces erroneous output, is the professional liable for loss resulting from the error?

50 Professional negligence, also known as professional liability, describes the liability which arises where a professional (such as a lawyer, architect, accountant or a medical professional) fails to perform the obligations they owe to their client (or patient) or third party with reasonable skill and care. The principle applies to use of AI by a professional to the same

extent as anything else a professional does (or fails to do). If a professional acts negligently in relation to AI use (whether by using AI where it should not have been used; or using inappropriate AI; using AI with insufficient care; or indeed failing to use AI when it should have been used), then if that negligence causes an Affected Third Party to suffer physical or economic loss, the professional can expect to be held liable.

51 The basic elements of the analysis are substantially the same as those already discussed in the context of general negligence liability for physical harm: when bringing a case against a professional (Anita), a client (Ben) or third party (Charlie) will need to show that: (a) Anita owed them a duty of care; (b) Anita breached that duty of care; and (c) the breach caused them loss, and that loss fell within the scope of duty that Anita owed to them.[60]

Duty of care to a client

52 In the vast majority of cases, Anita and Ben (as professional and client) will usually have a written or oral contract governing their relationship.[61] This will set out what services Anita will carry out and, possibly, also how those services will be performed (for example, whether AI will be used and, if so, for what). The scope and nature of services that Anita agrees to undertake under the contract will define the scope and nature of her duty. In broad terms, the more general the services contracted for, the wider the scope of the duty will be. For example, if Anita is a financial adviser who agrees to advise Ben generally in relation to his financial affairs, she will have far wider duties than to a different client for whom she has agreed to advise only in relation to one aspect of a specific transaction.

53 Generally, Anita will not promise to produce a particular result (for example, as a financial professional she will not promise that an investment is risk-free).[62] Instead, the contract is likely to contain an

express or implied[63] term that she should carry out the services owed under the contract with reasonable care and skill. So if Anita uses AI she will not necessarily – or even typically – undertake to ensure that the AI will not cause harm; but will instead undertake to ensure that the work is done with reasonable care and skill. That will inevitably involve Anita exercising reasonable care and skill in using the AI and (unless the use of AI is contractually required) it will also involve exercising reasonable care and skill in deciding whether it is appropriate for her to use AI at all.

54 In addition to contractual obligations, Anita will also generally owe a duty of care to Ben at common law to exercise reasonable care and skill (which is known as a "concurrent"[64] duty of care).[65] However, this will not exist if it is inconsistent with the terms of the contract between them.[66] Therefore, the scope of duty almost invariably falls to be assessed by reference to the contract. Anita might owe her client fiduciary duties[67] or statutory duties[68] as well.[69] However, in the context of AI harms, it is the duty to exercise reasonable care and skill that is likely to be of most relevance.

What constitutes "reasonable skill and care"?

55 Anita will be found to have acted with reasonable care and skill if she has acted in a way that a reasonable body of the profession would also have acted.[70] The standard of care which she is required to exercise is that ordinarily exercised by reasonably competent members of her profession, with the same rank and specialisation (if any).[71]

56 What constitutes "reasonable skill and care" changes over time. Scientific or other advances might mean that certain acts or omissions are negligent which previously were adopted by most of a particular profession. Therefore, the question of whether Anita has acted with reasonable care and skill will need to take account of the rapid developments in AI and the practices with regard to its use in a particular profession. That will

necessarily involve considering whether she has used AI appropriately when providing services, or, conversely, whether she has acted appropriately in failing to use AI.

57 Except for cases involving legal professionals, the Court will require expert evidence as to what a reasonable professional carrying out that task should have done.[72] The Court is also likely to consider regulations or guidance produced by the relevant professional body. At the moment, guidance from professional bodies tends to be somewhat high level. However, more detailed guidance and regulation is likely to emerge over time. Indeed, some professional bodies, (such as the Royal Institute of Chartered Surveyors) have already published detailed steps for their regulated firms to take in order to show that they have carried out appropriate due diligence or scrutinised appropriately the output of any AI system.[73] Cross-disciplinary standards are also beginning to emerge. For example, The British Standards Institute has recently published a standard intended to ensure quality among the growing audit market to ensure that businesses can be confident those assessing AI governance are doing so in a clear, consistent and coherent.[74]

58 If Anita falls short of applicable regulations or guidance concerning AI, the Court is likely to conclude that she breached her duty of care. However, complying with such regulations or guidance is not necessarily in itself enough to avoid liability in professional negligence – Anita needs to exercise her own judgment as well. Similarly, the fact that there may be a consensus view in a profession as to how AI should or should not be used, does not preclude the possibility of a court finding that consensus view negligent.[75]

59 The question of what reasonable care and skill entails in any particular context for the carrying out of any particular task is profession specific, task specific and situation specific. However, some general observations can nonetheless be made:

(a) Failure to conduct proper due diligence on an AI system before using it for client work, particularly if the system is new, innovative or from an untested provider, is likely to constitute a breach of duty.[76]

(b) It seems likely that if Anita chooses to use AI, she should have a sufficient understanding of it. She should also be able to explain (at a minimum, in broad terms) how the AI she is intending to use works to her client.[77] Anita used an AI tool without having at least a sufficient understanding of the AI tool and what it was doing, it will in most cases follow that she did not use reasonable care and skill.

(c) Depending on the significance of the AI use, it may well be important to be transparent with clients as to the use of AI (so that their clients can understand where, how and why it is being used, and potentially the risks involved).[78] For example, if a law firm uses an AI system to analyse prospects of success of a claim, but does not tell the client that they will be using AI to do this, this is likely to be a breach of duty.[79] Conversely, transparency alone will not save a professional from liability in professional negligence: merely telling a client that AI will be used does not excuse a failure to exercise reasonable care and skill in the decision to use AI, the selection of the AI, or the manner of its use;

(d) A solicitor putting their client's confidential or privileged information into an AI system which is not suitably secure and confidential is highly likely to be a breach of duty to their client.[80] Indeed, more generally, using an AI tool in a manner that damages the confidentiality in a client's information and materials is likely to be negligent in most cases, because a professional exercising reasonable care and skill would either ensure that did not happen, or would refrain from using an AI tool where there was not a

reasonable basis to be confident that the security of confidential information would be maintained.

(e) A professional should ensure that there has been sufficient testing of the AI system to check that it is suitable and appropriate for the task envisaged;

(f) Any professional who fails to exercise oversight of the AI system's outputs is likely to be found negligent. Quite what such oversight should entail will depend on the context: in some cases, it might be limited to reviewing a specific output from an AI system before giving it to a client; in other cases, it may require the professional to identify and address the risk of errors and/or biases in an AI system and scrutinising the reliability of the output more generally.

60 It is easy to say that a professional needs to undertake due diligence and/or to monitor outputs of the AI system in order to avoid a finding of professional negligence against her. It is perhaps more difficult for them to do so. This is in part because of the autonomous nature of AI and the difficulty that can arise in predicting its outputs or explaining them after the event. This difficulty would be anticipated and in some cases lead to the Courts concluding that there is a limit to what professionals can reasonably be expected to do; and in other cases to the conclusion that they should not have used the AI tool at all for the task at hand. All will depend on context.

61 The fact-specific and context-specific nature of the test for exercising reasonable care and skill renders it impossible to state definitively in this Legal Statement what a professional has to do (or avoid doing) in order to comply with the standard of care. This is not a difficulty unique to AI, but is rather inherent in the nature of professional liability in general.

Examples: *An architect uses an AI system to help them design a new building. The design was defective, but the architect never checked the design*

themselves (i.e. there was no human oversight) before sending it to their client. It is likely that they breached the duty of care owed to their client.

A barrister uses an AI system to assist them in writing submissions for the Court. The AI system 'hallucinates' (makes up) various cases, which either do not exist, or cites passages from genuine cases that do not appear in those cases. The barrister does not remove those hallucinations before submitting the document to the Court. It is clear that they breached their duty of care. They might also, amongst other things, be in contempt of court, be referred to their regulator, and/or face costs sanctions.[81]

A finance professional who uses an AI system to assist in making investment recommendations to clients. The AI system generates a recommended plan for a client. The plan is inappropriate for the client. They nonetheless present the plan without changes to the client. If they failed to carry out a reasonable degree of testing the AI system before using it for client-facing work, it is likely that they would be held to have breached their duty of care.

Breach of duty in failing to use AI

62 The sections above focus on negligent use of AI. It is important to be aware of the possibility that a professional could also be liable for *failing* to use AI for a task when a professional exercising reasonable care and skill would have done so. Whether there is a breach of duty in failing to use AI will, of course, be judged according to whether a reasonable professional of the same rank/specialism should have used AI in that context. That will, again, turn upon professional bodies' regulations and/or guidance (if any), and on the expert evidence as to the actions of competent professionals in the field. The possibility of breaches of duty arising from a failure to use AI reflects the fact that AI, in a professional's hands, is a tool. The question of whether such a tool should be used, and if so how, is no different from that which arises in respect of any other tool available to a professional.

***Examples**: A radiologist fails to use an AI system which is extremely effective at identifying cancerous tumours, and could have been procured at reasonable cost.*[82]

An auditor fails to use an AI system to help detect anomalies and frauds in a business which involves a very large number of similar individual transactions, where individual human review would be impossible in practical terms.[83]

A solicitor in the Business and Property Courts fails to advise their client that it may wish to consider some form of AI-assisted tool in order to review large volumes of documents.[84]

Advisors/Consultants to Supply Chain Actors

63 The analysis so far has focused on a scenario where a professional is the User of AI causing loss suffered by an Affected Third Party (i.e. the professional's client). However, it is important to note that the same principles can be applied to Advisors/Consultants to Supply Chain Actors who are professionals. One example of where liability might arise is where Advisors/Consultants' negligent advice as to the processes whereby AI is designed, selected or deployed leads to loss suffered by Affected Third Parties for which their User clients are liable.

Professional liability to non-clients

64 In principle, a professional could be liable to an Affected Third Party who is not a client. However, for reasons already explained, the professional is more likely to owe a duty of care at common law to non-clients who have sustained physical harm (personal injury or property damage) as opposed to economic loss.[85] So if the professional is a construction professional who has designed a wall in a new building which collapses

and injures a passerby (Charlie) or damages his motorbike, the professional is highly likely to be held to owe a duty of care to Charlie in respect of his personal injury[86] and property damage.[87] By contrast, if the wall collapses and only damages the new building itself, the construction professional is unlikely to owe any duty of care to a subsequent owner, who has suffered only pure economic loss.[88]

Can a person ever be liable for harms caused by AI in the absence of fault?

65 So far, the primary focus of this Legal Statement has been on negligence liability – in broad terms, arising from failures to take sufficient care to protect Users or Affected Third Parties from harm. However, just as patients can die on the operating table notwithstanding a surgeon's exemplary efforts to save them, or a portfolio can suffer heavy losses notwithstanding unimpeachable fund management, so too can an AI system cause harm even if all those involved in its development, deployment and operation act with the degree of care that the law requires of them. We now turn to the question of whether liability can arise in such circumstances.

66 In most contexts, the answer to the above question will turn on the contracts involved. For example, a contract between an Application Developer and a User might provide a warranty that the AI system will meet or exceed an accuracy threshold; if the system fails to meet that threshold, the Application Developer will be in breach of contract and (subject to any contractual limitations and exclusions) will be liable for that loss. It will be no defence for the Application Developer to argue that it exercised utmost care and skill in the development of the system, because it contractually agreed not just to be careful but to achieve a specific result.

67 Apart from a case of breach of an obligation to meet a particular standard or achieve a particular goal, the general position in English private law is that – absent negligence – the risk of loss for non-deliberate harm lies where it falls. Therefore, in general, those involved in the development, deployment and operation of AI systems will not

be liable for harms caused by those systems where there is no negligence on their part.

68 There is a major exception,[89] namely where the AI system is incorporated into a tangible product. In those circumstances the Consumer Protection Act 1987 ("*CPA 1987*") applies. That statute imposes liability for harms where a product is shown to be defective. Under the CPA 1987, liability is "strict", meaning that there is no need to prove negligence or "fault".

69 While this exception is an important one where it applies, its scope of application is narrow: it covers only a limited subset of possible harms; it only applies where a product is "defective" a term which, in this context, only refers to the *safety* of a product and not to its fitness for purpose or operation more generally;[90] and it is very unlikely to apply to an AI system that is *not* incorporated into a tangible product. As to this latter point, the issue is that the CPA 1987 applies to a "product", which is statutorily defined to refer to "goods" and "electricity".[91] An AI system is not electricity, despite the fact that its output (along with various other functions) might be conveyed to some degree using electrical means, e.g. using transistors. Neither, unless embedded within and forming part of a tangible thing, do we consider that English law would characterise it as "goods". Although there is no authority directly on the point, the English court has consistently held in a variety of contexts that software is not "goods" unless supplied in tangible form (such as on a CD-ROM).[92] Most AI systems are in any event supplied not as *software* but as *services*, which is one step further removed from the concept of "goods".

70 On 31 July 2025, the Law Commission announced that it intends to review the law relating to the product liability set out in the CPA 1987, and anticipates that this review will address the status of "pure software".[93] Unless and until that review leads to legislative change, our view is that CPA 1987 will apply to a fridge with integrated computer vision that can recognise the ingredients placed within it; but not to (for example) an LLM-based Chatbot.

71 The only harm that gives rise to statutory liability under CPA 1987 is death, personal injury or any loss of or damage to any property.[94] The legislation specifically excludes any property which is not ordinarily intended for private use, or is not intended by the person suffering the loss mainly for his own private use.[95] In practice, this means that claims for property damage can only be brought by consumers. Given that many AI applications – including many embedded AI applications – are in a commercial context, this removes a substantial number of potential claims from the ambit of the CPA 1987.[96]

72 Where a User or Affected Third Party suffers injury or property damage as a result of a product incorporating AI, the first question with respect to liability is whether the safety of the product is such as persons are generally entitled to expect. That question calls for a multi-factorial and factual enquiry that will extend to matters such as the nature and purpose of the product, its reasonably expected use, instructions supplied with it and, of course, an analysis of what has actually gone wrong and why.[97] If the product falls below the standard and is therefore "defective", liability may still not arise if one or more of the statutory defences is engaged.[98]

73 As will be apparent from the previous paragraphs, establishing liability under the CPA 1987 does not require or entail proving negligence; instead, the core element is defectiveness. Nonetheless, where the product is a highly technical one (whether that is AI or otherwise) in practice the enquiry process to assess 'negligence' and to assess 'defect' is likely to overlap substantially.[99] In both contexts:

(a) The Court will consider compliance or non-compliance with standards to be relevant;[100]

(b) There will be a focus on warnings given or not given. Good instructions and warnings can render safe (and thus non-defective) products that come with potential dangers, whilst poor instructions and inadequate warnings can be reasons to find a product unsafe or form the basis of a finding of negligence;

(c) The Court will enquire as to whether at the time a producer developed the product, the defect in the product could have been discovered; and there may well be an assessment of the risk/benefit of a particular product;[101]

(d) Therefore, if relevant standards are complied with, appropriate warnings given, and the issue with the product was not discoverable beforehand, it is unlikely that liability would be made out even under the CPA 1987. Although liability under the CPA 1987 is "strict" in the sense that it does not require proof of negligence, the mere fact of a defective product causing harm does not inevitably lead to liability.

74 The principal parties liable for harm under the CPA 1987 are the defective product's producer, the person who lends their name to the product (in the case of unbranded products which are produced generically but then subsequently 'badged' with the name or logo of a supplier), and the importer into the UK.[102] Given that the definition of "product" includes component parts, there can be more than one liable producer. Therefore, if the AI component of a multi-component product fails, both the producer of the overall product *and* the producer of the AI component could in principle be liable. However, in practice, we anticipate that most of the time the latter will escape liability under the CPA 1987, either on the basis that the AI component was not itself defective but became dangerous only as a result of the way it was used by the manufacturer of the product; or on the basis that the AI component is pure software falling outside the scope of the regime altogether.

75 Under the CPA 1987, the burden is on the claimant to establish that the product was defective, on the balance of probability. That said, the Court of Appeal has held that it is unnecessary for the judge to ascertain the *precise* cause of the defect.[103] This point is relevant in a machine learning context, where the piece of code which has a prompted a particular undesired outcome may not be discernible even to the original programmers.

***Example**: Anita is the owner of an AI fridge that automatically scans its contents and informs the consumer of which items are in date and which are expired. The fridge is manufactured by a well-known 'white goods' producer, Ben Inc. The AI component of the fridge is a specialist application developed specifically for Ben Inc by Application Developer Charlize Plc, and supplied to Ben in the form of self-contained "systems on a chip" that were intended to be installed within the fridges. The application relies on a general purpose foundation model built and operated in the cloud by Foundation Model Developer Doris Ltd.*

The fridge malfunctions for reasons relating to the foundation model, telling Anita that a bag of oysters is in date when in fact they have been kept too long and have gone off. Anita eats the oysters and suffers serious personal injury. The fridge came with no warnings that its eating guidance should not be relied upon; indeed, it was marketed on the basis of the utility of that precise feature. Elementary testing would have revealed the unreliability of the feature but such testing was not undertaken by anyone in the supply chain.

Ben Inc is prima facie liable under the CPA 1987 for providing a defective fridge to Anita. Charlize Plc is also prima facie liable under the CPA 1987 as it manufactured the physical 'system on a chip' circuit board which contains the defective AI component. Doris is unlikely to have any liability to Anita: its general purpose model is not intrinsically dangerous; it was the (untested) use of the model by Charlize that caused the problem. In any event, Doris is providing an AI model as a service; it is not providing a product or any component of a product, so its activities do not fall within the CPA 1987.

Factual Causation

76 In order for a party to be liable for a loss (whether physical or economic), it is not sufficient that the party is found to be in breach of contract, negligent or otherwise in breach of a duty; that party's breach or negligence must also be found to have *caused* the loss in question.

This section of the Legal Statement considers how the common law approach to causation applies to physical and economic harm caused by the use of AI.

General rules of factual causation

77 As a general rule, in contract it is generally said that that the breach must be a substantial cause of the loss,[104] and in relation to negligence that 'but for' the breach (i.e. the action or omission by the defendant that is said to have been negligent), the loss would not have been suffered.[105] This is known as the counterfactual.

78 By way of example, in the context of a claim for professional negligence: if a surveyor, failed to notice subsidence issues in a property which they should have spotted (in breach of contract and in breach of their common law duty of care owed to their client), a claim for damages could succeed only if the client could show that he relied upon the surveyor's advice in buying the house, and that 'but for' their advice that it was subsidence-free, the client would have bought a house without subsidence issues.

79 By way of further example, in the context of a claim where there is no contractual relationship between the parties: a passer-by hurt by a wall which collapsed owing to an architect's negligent design would need to show that 'but for' the architect's negligent design (rather than, for example, poor brickwork), the wall would not have collapsed and the injury would not have occurred.

80 Developing that further, where harms have been caused to the client by professional's acts or omissions, the following principles have arisen from the case law:

(a) If the question is what the client would have done if he had been advised non-negligently, he must show 'on the balance of probabilities' what he would have done if so advised. Taking the surveyor's example above, the client would have to show on the

balance of probabilities that he would not have bought the house with subsidence issues;

(b) If the question is what some third party would have done if the professional had acted non-negligently, this is decided on a 'loss of a chance' basis. As part of this, the Court has to value the chance which has been lost. For example, the client has lost the chance to bring litigation because his solicitor has failed to issue the claim form in time, the Court assigns a value to the 'lost' claim, considering what a Judge might have decided on liability and quantum; and

(c) Cases concerning medical professionals, and the question of whether the outcome of a medical intervention might have been different but for negligence, are not generally dealt with on a loss of a chance basis[106] but on the balance of probabilities.

81 Where claims are brought for personal injury or property damage under the CPA 1987, there is a different test for causation - whether the damage was caused wholly or partly by a defect in a product.[107]

Factual causation and AI – potential difficulties

82 It has been suggested that AI causes difficulties with regard to factual causation.[108] The characteristics of AI which are said to cause this difficulty are identified in an earlier section of this Legal Statement, key amongst them being its autonomous nature.

83 While it is correct that the autonomous nature of AI (and, to a lesser extent, the lack of explainability) can make it difficult to understand precisely why a particular outcome eventuated, in our view the general rules of causation can in principle be applied to AI harms. Although AI may, on some facts, give rise to evidential difficulties, those are neither more severe nor different in kind to the sorts of issues that arise in other domains and that the English common law is well able to accommodate.

84 Moreover, just because a claim involves the use of AI does not necessarily mean that these difficulties arise or that it is inevitable that causation is made more complicated. Indeed, in a great many cases involving harms caused by use of AI, neither its autonomous nature nor any other characteristic of the technology will have any causal relevance. This can be illustrated by considering the following scenarios:

(a) An architect used AI when they should not have used AI for that particular task. As a result of that use of AI an Affected Third Party suffered harm. In such a case, the Affected Third Party is likely to be able to show in the ordinary way (using expert evidence[109]) what an architect exercising reasonable skill and care should have done if they had *not* used AI. That is the only relevant counterfactual. It does not matter that the reason why the reasons for the particular AI output cannot readily be ascertained.

(b) A barrister used AI for their submissions, and produced submissions with hallucinated cases in them. This resulted in their client having to pay the other side's costs in dealing with those parts of the submissions. Here, causation will likely be made out – the barrister should not have used AI to write their submissions, or – if using AI – should have checked that the cited cases were real. 'But for' the use of AI and the failure to check the hallucinated cases, the client would not have had to pay the other side's costs of dealing with those submissions.

85 In cases where it *is* necessary to understand why AI produced the output it did, two kinds of causal uncertainty in particular can arise.[110] First, there can be causal uncertainty where a gap in the evidence arises because material has been destroyed, tampered with or simply not gathered in the first place. Secondly, there can be causal uncertainty owing to the opacity of AI itself. The first is evidential uncertainty; the second more akin to scientific uncertainty.

86 These are dealt with in turn below.

Absence of causal evidence

87 Owing to the opacity in the reasoning process between the input and output of AI, the precise reasons why AI produced a particular outcome may not be recoverable after the decision is made. In the context of (for example) an allegation that incorrect inputs were provided to the AI, that training was inadequate, or that controls or other technical measures ("*guardrails*") were incorrectly set, lack of evidence might make it difficult to know how the AI would have behaved with correct inputs, adequate training and/or correct guardrails. In principle, that might make it difficult to prove causation.

88 However, causal uncertainty arising because evidence has not been gathered is not in itself a problem which is unique to AI. For example, a claim on behalf of an employee who had suffered hearing loss failed at first instance - the claimant could not show that he had been repeatedly exposed to excess levels of noise because the defendant employer had breached its duty to monitor noise levels on its ships. That was overturned by the Court of Appeal, which held that the claim should succeed, noting that the judge should have given weight to the consideration that any difficulty of proof for the claimant had been caused by the defendant's breach of duty, so the court should have judged the claimant's evidence benevolently and the defendant's evidence critically.[111]

89 A similar approach to resolving evidential difficulties caused by the nature of a defendant's breach can be seen in 'lost litigation' cases where:[112] (i) the legal burden lies on the claimant to show he has lost something of value; (ii) the evidential burden lies on the defendants to show that (despite acting for the claimant in the litigation) it was, in fact, of no value to their client; (iii) if the court has greater difficulty is working out the strength of the claimant's case than it would have had at the time of the original action, that counts against the defendant solicitors rather than the claimant; and (iv) in evaluating the prospects of success had the original claim be fought out, the court will be generous to the claimant in assessing his prospects of success.

90 It seems likely that, where there are difficulties with evidence, the law may well recognise this and approach questions of factual causation differently. Whether this will happen in the context of AI – and indeed whether it will need to happen – remains to be seen. In principle, however, it can readily be seen that where (for example) a defendant ought to have recorded certain information and failed to do so, a Court might take a 'benevolent' approach to the Affected Third Party, or make certain evidential presumptions (by shifting the burden of proof on particular factual issues) as a result. For example, if an Application Developer ought to have ensured that certain inputs were recorded (but did not), or if a professional ought to have recorded in writing a decision about the reliability of a given output[113] (but failed to do so) it might be the Court treats the Affected Third Party's evidence with benevolence, and the defendant's evidence critically. Even if the evidential gap does not arise from a failure of the defendant to gather or keep evidence but is intrinsic to the AI tool being used, a similar approach might well apply where AI has been used negligently and the fact that its use has given rise to the evidential uncertainty.

91 That said, in some cases, even where there is a lack of factual evidence as to causation, it may not be necessary for the Courts to intervene in this way. This is because we anticipate that in some circumstances it will be possible to fill gaps in the factual evidence with expert evidence and, in particular, expert evidence obtained through experimentation. For example, an allegation that but for incorrect guardrails, AI would not have produced a particular output will very often be capable of being tested by running simulations comparing the output of the tool with the allegedly correct and incorrect guardrails. Even though the autonomous nature of AI and the number of potentially relevant inputs may preclude conclusions being reached with certainty, English law does not require certainty, but operates on the balance of probabilities. That threshold permits a degree of uncertainty, and is well suited for experimental proofs.[114]

Scientific uncertainty

92 There are cases where (through no fault of the defendant) it is impossible at the time when a matter comes to Court to work out what happened in the past and what would have happened in a counterfactual situation. Where this is simply bad luck, the claimant cannot prove their case on the balance of probabilities, and the claim fails.[115]

93 Where there is a *systematic* impossibility of proving causation, the Courts have sometimes come up with alternative approaches.

94 This was the situation in *Fairchild*[116] a case involving asbestos and mesothelioma - where it was scientifically impossible to show on which (amongst several) occasions the claimant might have been exposed to the strand of asbestos which caused their cancer. On the 'but for' test, the claimant would have failed, because it was scientifically impossible to know whether the disease would have been contracted absent any particular exposure. However, the Courts held that it was enough to show that an exposure resulting from the defendant's breach materially increased the risk of injury suffered.[117] Liability was assigned in proportion to the probability that the defendant caused the injury.[118]

95 The *Fairchild* "material increase in risk" principle has not (yet) been applied beyond personal injury. However, some have argued that there is no reason why it should be so restricted.[119] The rationale for *Fairchild* (if one can be seen) is (i) a concern to avoid a person's legal rights to compensation being diminished simply because there are lots of wrongdoers, (ii) a concern to avoid the defendants' duties of care being emptied of content because proof of causation is inherently (and recurrently) impossible. These rationales arguably should not be limited to personal injury.

96 The *Fairchild* exception is understood as requiring an impossibility of proof of causation owing to the limits of scientific knowledge. This perhaps over-simplifies the position.[120] It could be said that the justification of the rule is triggered by a claim failing because of a multiplicity of wrongs.[121]

97 However, stepping back, what *Fairchild* shows is that English common law is capable of evolving and / or developing limited exceptions to meet evidential challenges and to ensure that systemic problems of causation do not cause systemic injustice. It seems (in theory at least) that the scientific impossibilities of showing the correct counterfactual as a result of the autonomous or "black box" nature of AI might lead the Court to approach causation in a different way. It is possible that the *Fairchild* exception would provide another way for claimants to establish causation if that is the only realistic way of avoiding the risk that Affected Third Parties will be unable (owing to scientific difficulty) to prove causation.

Causation and product liability

98 We would not expect an AI system to cause complications in a product liability case. For example, if a fridge catches fire and burns down someone's house, it is immaterial precisely why the fridge caught fire – if an Affected Third Party brings a claim against the fridge manufacturer, all they need to show is that the fridge was defective (it caught fire) and the defect caused the loss (the fire spread and damaged the house).[122] It makes no difference whether the fire within the fridge was caused owing to an issue in the AI system in the fridge or, for example, a faulty capacitor.[123] In such a case, the issues identified above will not arise.[124]

EU Commission's approach to causation

99 The European Commission proposed reforming civil liability for AI, outside of the product liability setting, a draft AI Liability Directive. However, the European Commission has since withdrawn such legislation from further consideration by the EU legislative bodies.[125]

100 The proposed AI Liability Directive had included a rebuttable presumption of causality that favoured the claimant under specific circumstances. This change from the normal position that a claimant

must prove all aspects of their case was aimed at helping claimants overcome the potential difficulties of proving causation.

101 As the AI Liability Directive has been withdrawn, liability for harm caused by AI systems (in the EU) falls under existing national laws and the revised Product Liability Directive (Directive (EU) 2024/2853).[126] This includes some presumptions of defectiveness and an obligation for manufacturers to disclose evidence where a claimant has presented facts and evidence sufficient to support the plausibility of the claim for compensation (Art. 9(1)), and in an *"easily accessible and easily understandable manner"* (Art. 9(6)).

102 The revised Product Liability Directive is not part of English and Welsh law. England and Wales have no equivalent statutory presumptions at the present time.

Legal Causation

103 English law treats causation as a two-stage enquiry. The first stage is the 'factual' causation enquiry addressed above. The second, which arises only where factual causation is established, is 'legal' causation, which concerns whether a person's conduct should be regarded as a cause in law. The factual causation enquiry is a question of why the loss occurred and, specifically, whether the defendant's action was a 'but for' cause of the loss. Legal causation is a question of whether, even though a person's conduct was a "but for" cause of loss, someone or something has intervened between that person's wrongful conduct and the loss, which is deemed to 'break the chain of causation'.

104 The Court of Appeal illustrated the distinction between factual and legal causation as follows: "*if A kills B by stabbing him, the birth of either of them 30 years before is as much a causa sine qua non of the death [i.e. an event but for which it would not have happened] as is the wielding of the knife. So the law makes appeal to the notion of a proximate cause…*".[127]

105 Below we consider two AI scenarios that might give rise to questions of legal causation: foreseeable misuse of AI by a third party, and harm

caused by AI when operating (at least to some degree) autonomously. These scenarios are novel, in that there are no exact analogies in the existing case law. However, as we explain, for the most part they can be addressed under established principles.

Liability for foreseeable misuse of an AI system

106 The potential for AI misuse is well known.[128] Though deliberate bad actors would face civil (and possibly criminal) liability for harms, it may not be possible to sue them. Often the bad actor will be unidentifiable. Additionally, they may not be amenable to justice, for example because they are a foreign state agent. Even if the bad actor is known and available to be sued, they may not have money. Victims of harm may wish instead to pursue AI supply chain parties, especially Foundation Model Developers, who may be perceived as having 'deep pockets' and able to meet such liability.

107 Consequently, where an AI system has been used deliberately by a third-party bad actor to cause harm, the question arises as to whether parties in the AI supply chain have any legal duty to prevent such misuse.[129] If such a duty exists and is breached, then the chain of legal causation may not be broken by a third party's acts.

108 The general position under negligence law is that a person is not liable for the consequence of the actions of a third party.[130] However, that is not universally so. In particular, there is a distinction to be drawn between making matters worse and failing to confer a benefit. If a third party causes harm, a person who makes matters worse will readily be found to owe a duty of care; by contrast, a person who merely fails to confer a benefit (including failing to protect a person from harm) will generally owe no duty.[131] Additionally:

(a) A person owes a duty to take care not to expose others to unreasonable and reasonably foreseeable risks of physical harm created by that person's own conduct. By contrast, no duty of care is in general owed to protect others from risks of physical harm

which arise *independently* of the person's conduct - whether from natural causes or third parties.

(b) There are exceptions to the general rule that there is no duty of care to protect a person from harm, for example, where a person has assumed a responsibility to do so or has control of a third party.

109 The over-arching principle that the Courts now apply has been explained by the Supreme Court as follows: "*In the tort of negligence, a person A is not under a duty to take care to prevent harm occurring to person B through a source of danger not created by A unless (i) A has assumed a responsibility to protect B from that danger, (ii) A has done something which prevents another from protecting B from that danger, (iii) A has a special level of control over that source of danger, or (iv) A's status creates an obligation to protect B from that danger*."[132]

110 All of this means that it in most circumstances, it is very unlikely that a party upstream in the AI supply chain (Anita) will be held responsible for the misuse by a third party (Ben) of an AI system which Anita has created, amended or operated, but that the possibility of liability might remain in some narrow circumstances. Those circumstances will include where Anita has created the source of danger or otherwise assumed a responsibility for it.

111 By way of analogy, those in control of extremely hazardous substances or items are more likely to be held liable for damage caused by third parties interfering with those substances and items than those who are in control of less hazardous substances. An example of this is where a defendant negligently delivered parcels of flammable film scrap to a business's premises. One of the claimant's employees touched one of the parcels with a lighted cigarette to pass the time by making a "*small innocuous bonfire*" having no idea that the substance was explosive. The defendant was held liable for the subsequent explosion. Although the employee's act was the act of a third party, it was not so unreasonable, or of such overwhelming impact, as to obliterate the obvious risk that the defendant's original negligence had created.[133]

112 In another case,[134] a decorator working alone in a house went out to buy wallpaper and left the front door unlocked. The decorator was held liable for the loss caused by a thief who entered while he was away. Here, the third-party thief's actions were deliberate, but were a foreseeable risk.[135]

113 However, these are quite extreme examples. Merely knowing of risks relating to a product does not, without more, amount to taking on responsibility for those risks, particularly in circumstances where the risks only eventuate as a result of the decision of a third party. A manufacturer of 'general purpose' technology such as knives would not – without more – be held responsible for violent crimes committed using its knives. Though the knife manufacturer is presumably aware that its products could be misused, the independent decision of the criminal involved is sufficient to break any chain of causation. The law does not make the knife manufacturer liable. By the same reasoning, a supplier of a general purpose AI system may know that its products *could* be misused, but that does not – without more – make the supplier liable for the consequences of that misuse. An alternative way of analysing the position is that protecting victims from being attacked by knives does not fall within the scope of any duty that the knife manufacturer has.[136]

114 In conclusion, it is not impossible that a party in the AI supply chain might be held responsible for creating a source of danger by bringing into existence without suitable safeguards or otherwise failing to control (when it has special powers to do so) an AI model or system which it knows to be capable of certain types of harm if misused – such as on the basis that the party has introduced a particularly acute danger for which is must take responsibility. However, such cases are likely to be rare. The more general purpose the AI system in question, the more extreme the circumstances would have to be before a liability risk will realistically arise.

Liability for harm arising from autonomous behaviour of Foundation Models

115 The output or 'decisions' of Foundation Models do not neatly map on to common categories of intervening acts, i.e. natural events such as storms, floods, or landslides, or acts of human third parties.[137]

116 The autonomous actions of AI are perhaps closer in substance to human decisions than natural phenomena – in that they will typically comprise the type of actions which would otherwise have required human intelligence to undertake. Yet from a policy perspective AI actions are unlike human actions in that when an AI system takes a decision, it is not clear that there will always be someone who can be held legally responsible.

117 In some extreme circumstances an intervening natural event may operate to eclipse the defendant's wrongdoing as the effective cause of the damage suffered by the claimant. Historically this was known as the 'act of God' defence.[138]

118 It is conceivable that the process of an AI system 'learning' new principles of action or behaviours could be deemed to constitute a novel intervening act which breaks the chain of causation between any programming, data selection and other decisions by a Foundation Model Developer or Application Developer and the eventual output of the system. However, in practice we would expect the courts to be very slow to reach such a conclusion, bearing in mind the policy considerations which tend in favour of finding a legal person liable where harm has been caused – especially if the person in question has sought to benefit from the activity which gave rise to the harm.

119 Perhaps a closer analogy to the manner in which highly autonomous AI operates, than that of third party adults or natural events, is that of a human child. Children below the age of responsibility are not treated as legal persons capable of incurring liability, and yet they are not natural events either. Rather, the limited autonomy of children is usually not treated as breaking the chain of causation, at least where there is some form of underlying duty on another person to exercise a

degree of control over the child's actions or to prevent them causing harm. This principle applies even when the child acts deliberately and voluntarily. So in a case where a child started up a car which crashed and caused loss there was no finding of an intervening act.[139] Similar principles apply to the acts of animals.[140]

120 Notwithstanding the precedents for liability arising from the autonomous actions of third parties as well as semi-autonomous animals, there is no direct analogy which addresses liability for autonomous systems. Due to AI's autonomy and adaptivity there will inevitably be a degree to which the output of a Foundation Model is unpredictable and hence unforeseeable, or otherwise deemed to be the result of a decision made other than by the Developer.[141] In practice is it will be difficult to assess whether there has been "reasonable care" in the programming of any software by the Foundation Model Developer, because the software has developed internal decision-making principles which may be opaque even to its own developers. It would not necessarily be possible to spot or eliminate a particular line of "defective" code which can be identified as having caused damage. Liability for autonomous acts may depend on the following factors:

(a) Capabilities and limitations of the AI system or model in question: At the time of writing, though foundation models are capable of displaying a degree of autonomy, they do so only within specific boundaries. As matters stand, the capabilities of even the most powerful Foundation Models are relatively limited, but it cannot be assumed that these limitations will remain. The more capable AI becomes across different domains, the less foreseeable all possible harms will be, since the nature of harms will increase. It may be that courts are willing to expand progressively the ambit of a duty of care to cover all such harms, but this would undoubtedly require innovation and at a minimum gives rise to uncertainty.

(b) The level of autonomy: As with narrow and general AI, the level of autonomy displayed by a system is not a binary quality: there is a spectrum of degrees of autonomy along which different systems may fall. The more human involvement in the creation of output,

the less overall autonomy a system is displaying, at least for the purposes of determining the consequences in tort law. In terms of more autonomous systems, Agentic AI is designed with the capability of minimising the requirement for human involvement in the process of output generation, by allowing systems to generate various sub-instructions and tasks designed to achieve some overall goal which is set by a User only at a high level of generality. At higher levels of autonomy it is reasonable to expect that courts will find it more difficult to impose liability on AI supply chain parties without development of existing tort law norms.

(c) The degree of supervision of the AI model / system: If, for example, an AI value chain party (such as a Foundation Model Developer or Application Developer) reviews user interactions such as prompts – or (especially in the case of a very widely used model) has the ability to conduct such reviews automatically through monitoring systems, then it may have some active awareness of attempts to utilise such AI to cause harm. For instance, the AI value chain party may receive alerts when a user seeks to use their model or system to build a chemical weapon, or perhaps to generate prohibited sexual images. If, in such circumstances, the AI supply chain party chooses not to impose guardrails to prevent such misuse that decision may be a relevant factor in determining legal causation.

***Example**: A Foundation Model Developer creates a multi-modal system which is able to analyse visual input, including pictures and video, as well as audio and text. The Developer states on its website and in its publicly available ESG materials that its model is "fair and free from bias". A company decides to use the system to screen applicants for jobs. After several weeks it transpires that the system has been discriminating against applicants with facial hair. The unsuccessful applicants sue the Foundation Model Developer.*

The Foundation Model Developer has arguably imposed upon itself a duty of care by making public statements as set out above, and breached that duty

by releasing a system which did in fact discriminate in a manner which could reasonably be described as 'biased'. However the Foundation Model Developer may object to liability on the grounds that job applicants were not reasonably within its contemplation as parties likely to be harmed by the output of its system. On balance there is a reasonable chance that the Developer might be held liable in these circumstances, which involve both a simple supply chain and a conscious (albeit inadvertent) assumption of responsibility.

Does liability attach to false statements made by an AI chatbot?

121 AI tools can make statements, including in writing, orally, and visually through pictures and video. In the context of current-generation AI, such statements are often produced by large language models (*"LLMs"*) created by Foundation Model Developers. These Foundation Model Developers may make LLMs available to Users directly or they sell access to intermediaries, like Hosting Suppliers and Application Developers. Users ask LLMs, via Chatbot interfaces, for information, advice, and even emotional support. AI-generated statements are often both believable and wrong. These features create the context for potentially significant harm:- Users or Affected Third Parties rely upon statements to their detriment.

122 While AI systems that use LLMs are often anthropomorphised due to their ability to interact with Users in a human-like manner (false statements are deemed "hallucinations"),[142] as have already been noted in this Legal Statement, they have no legal personality. As with other types of claim, claims arising out of LLMs' statements must be directed elsewhere.

123 There is no general duty on legal entities *not* to make careless statements.[143] However, there are specific circumstances where liability under English law may arise for statements – under contract, fiduciary

duties, as well as the torts of negligent misstatement or misrepresentation, deceit, libel, slander and malicious falsehood.

124 For Users or others with counterparties in the supply chain, the primary recourse is likely to be contractual. However, as with other sections of this Legal Statement, we focus here on non-contractual, tortious routes.[144] This is not because they are more common, but because they require more legal analysis.

Negligent misrepresentation

125 Liability for negligent misrepresentation (or misstatement) arises where:[145]

(a) Anita makes (or adopts) a statement to Ben;

(b) Anita owes Ben a 'duty of care' in relation to the making of the statement;

(c) The statement is false;

(d) The statement is made negligently (it does not need to be fraudulent or deliberately misleading);

(e) Ben reasonably relies on that false statement; and,

(f) That reliance causes Ben relevant loss.

126 We focus on those elements that have particular AI-related novelty, namely (overlapping) elements (a), (b), (c), (d) and (e).

Making or adopting a statement

127 The statement (or representation) can be written, oral, visual or conveyed by conduct (like covering up dry rot in a building to misrepresent the state of a property).[146]

128 Statements are to be interpreted objectively, in context and may be express or implied.[147] This means that Foundation Model or Application Developers may use surrounding text to affect a statement's meaning (for example, by explaining how it is created and the associated limitations), and make express or implied representations of their own about the AI tool itself, such as its development, testing and accuracy.

129 Generally, liability for negligent misrepresentation (or misstatement) requires the statement to be made by or on behalf of a legal person. This requires analysis in the context of a Chatbot because when an LLM's text output is not pre-determined, a Foundation Model or Application Developer does not "make a statement" to Users in any conventional sense. Indeed, depending on the prompt, the *User* may appear more like the "maker" of the statement (for example, if they require the LLM to respond in a particular way).

130 Special circumstances may arise, for instance, if a Developer applies controls or other technical measures ("guardrails") to circumscribe the response. If they do so, the application of controls may amount to "making a statement", particularly if the response is pre-determined fixed text. Such a situation involves no real novelty; the AI tool is not acting with autonomy in making such a statement, so there is nothing AI-specific about this scenario.

131 There is no English authority for the proposition that an AI tool may make a statement "on behalf of" a legal person for the purposes of a claim in negligent misstatement. The established case law relating to statements made other than by the defendant focuses on other legal persons (agents) and the question of whether they acted with or without authority.[148] However, a Canadian court has held an airline liable for false statements by its Chatbot on the basis that: "*While a chatbot has an interactive component, it is still just a part of Air Canada's website. It should be obvious to Air Canada that it is responsible for all the*

information on its website. It makes no difference whether the information comes from a static page or a chatbot."[149] As a matter of legal policy, the decision is readily explicable. Whether the same conclusion could have been reached by an English court requires more analysis.

132 The correct analysis depends on the facts. Where an AI tool's outputs are pre-determined or a Chatbot is presented as an authorised representative of a legal person, then the Court will likely treat the statements as being made on behalf of that person. In the latter case, this is because there is, in effect, an implied representation that the AI tool's outputs can be treated as being authorised by the legal person. Where an AI tool has autonomy, that analysis may still apply, but pre-existing cases where a party passes on information supplied by another[150] may be more apt:

(a) The question of whether a person is "adopting" the information which is passed on is a question of interpretation depending upon the facts.[151]

(b) A range of possibilities exist as to how this might apply in an AI context. For example (and returning to Anita and Ben):[152]

(i) Anita may adopt the information generated by the AI tool as her own, thereby taking on such responsibility as she would have if she were the maker of the statement.

(ii) Anita may represent that she believes, on reasonable grounds, the information supplied by the AI tool to be correct. That involves a lesser degree of responsibility than scenario (i), but Anita can still be liable if she in fact does not believe the information to be correct or has no reasonable grounds for so believing.

(iii) Anita may simply pass on the information to Ben as material coming from the AI tool, about which Anita has no

knowledge or belief. Anita then has no responsibility for the accuracy of the information itself, unless this is dishonest.[153]

(c) These scenarios are not a complete statement of the range of possibilities. There are other intermediate positions. In general terms, the extent of Anita's responsibility for the AI tool's information in any particular scenario is that which a reasonable person in Ben's position would understand from Anita's words or infer from Anita's conduct and all the circumstances.[154]

133 In some narrow circumstances, it has been held to be enough for a person to have actual or constructive notice of the false representation in order to be liable for it. For example, cases have arisen where a wife's consent to a charge in favour of a creditor over a matrimonial home was obtained by misrepresentation of the husband, and the creditor was fixed with constructive knowledge of it.[155] These are judicial responses to a particular injustice. While this shows English law's ability to be flexible to meet the requirements of justice, we consider that the Court will generally not need to resort to this sort of approach for AI-generated statements; the established analytical approach to determining whether a person has made or adopted statements should be enough in most real-world cases.

***Example**: A mushroom forager (and enthusiastic Application Developer) adds a chat interface to her website under the header "Talk to me!". Unknown to the User, the interface is powered by an AI tool. Users submit images of fungi and, after a delay, the chat interface then provides advice on edibility. A 'Destroying Angel' mushroom (which is highly toxic) is misidentified as being a variety which is safe for humans to eat. On the facts, a reasonable User would understand the forager to be responding to the submission personally. The statements are therefore made on her behalf or, alternatively, she is treated as having adopted the information as her own and is responsible for the statement. Contrast this with a situation where the forager's webpage states: "This is an AI tool; I have tried to make it as accurate as possible, but I do not review its outputs." Here, the forager is unlikely to be identified as the maker or adopter of the express AI-generated*

statement. However, they are making implied representations about the manner in which it was developed or tested.

134 In short, therefore, the fact-sensitive approach to the making or adoption of representations taken by the English courts can operate effectively in the context of statements generated by AI. Although not every statement by an AI that causes harm will be actionable, that is consistent with the operation of the law more generally: the law of negligent misrepresentation is focused on a person's careless *words*, as distinct from careless *acts*. Enhanced restrictions on claims exist for words to address human concerns (like the risk of exposure for loose talk at parties) and the breadth of potential liability (given words may be broadcast without consent or foresight).[156]

135 For false AI-generated statements, there is a temptation to focus on that machine output. However, given that AI is generally used as a tool, the core negligence (if any) of the defendant legal person would typically be the careless *acts* that permitted the output, like the human decisions behind the AI tool's design, testing and deployment. As above, Developers may make express or implied statements about this aspect.

***Example**: An Application Developer creates a Chatbot that is marketed to teenagers as a companion. The Developer is capable of introducing controls on how the Chatbot will respond to prompts about self-harm and creating alerts or escalations for human intervention. A User develops an emotional dependency on the Chatbot and harms themselves as a result of the information and advice that it provides. The breaches of duty that may arise likely relate to the acts of the Developer, essentially inadequate guardrails (their design and testing), rather than the statements themselves. At root, the harm is caused by the Developer's **lack** of involvement in their communication.*

Duty of care

136 We have already introduced the concept of a duty of care in the context of negligence. The key question is whether, objectively, there has been an assumption of responsibility to an identifiable (although not necessarily identified) person or group of persons, not the world at large or to a wholly indeterminate group.[157] It is not sufficient to ask whether Anita owes Ben a duty of care. It is always necessary to determine the scope of duty by reference to the kind of damage from which Anita must take care to keep Ben harmless.[158] Generally, manufacturers of products owe duties that only extend to physical harm, not economic loss.[159] A special relationship must be established for the latter.

137 Where a task or service is being undertaken, particular factors include: (i) the purpose of the task or service and whether it is for the benefit of Ben; (ii) Anita's knowledge and whether it is or ought to be known that Ben will be relying on Anita's performance of the task or service with reasonable care; and (iii) the reasonableness of Ben's reliance on the performance of the task or service by Anita with reasonable care.[160] Generally (and to expand upon (ii)), Anita must know either actually or inferentially, that the advice or information is likely to be acted upon by Ben (or a group of which he is a member) for a known purpose *without independent inquiry*.[161]

138 A Foundation Model or Application Developer may not have actual knowledge of a User's reliance on an LLM's output. However, constructive knowledge (i.e. what ought to be known) can suffice for the purposes of establishing a duty. A Foundation Model Developer who provides an LLM for general usage and is unaware a specific use case will be much less likely to be exposed to liability than, say, an Application Developer who produces a "robo-lawyer" that is expressly marketed to assist people with a particular area of law.

139 The reasonableness of a User's (or Affected Third Party's) reliance will be fact specific. Currently, while technological advances generate excitement, there is growing and widespread acknowledgement of

LLMs' limitations.[162] Guidance to judicial office holders on AI tools (for their own usage) notes: "*Even with the best prompts, the information provided may be inaccurate, incomplete, misleading, or biased*".[163] A linked expectation that AI outputs need to be independently verified would further undermine the conditions that are normally required to recover economic loss. That said, generalities are dangerous; the objective circumstances of the parties in the particular situation are always the focus of the English law analysis. For example, these may include: changing societal expectations of LLMs, how the specific AI tool was marketed and presented, the use case (whether it is a consumer product or for professionals with their own duties[164]), and whether the usage of the tool is disclosed at all. Compared to human-to-human interactions, conversational tone is likely to be of reduced importance: for humans, duties do not arise in social situations; for LLMs, casual talk is a design feature and can arise regardless of the context in which the LLM is being used.[165] Courts will also consider how any prompts were used to generate text, given that the prompter (i.e. the User) may in so doing have evidenced their understanding as well as have influenced the LLM's response, both of which factors may be relevant to determining whether an actionable misrepresentation has occurred.[166]

The statement is made negligently

140 The standard of care that a legal person needs to meet involves further fact-sensitive questions, relating to how the AI tool and its outputs were presented. If a Chatbot was held out by an Application Developer as being competent in a particular field of expertise or, indeed, as a "human" expert, then the output would be held to the objective standards of that field. This may mean that the Developer is held to a standard it cannot meet.[167] By contrast, if the Developer was objectively understood to be providing AI-generated content (and any representation linked to that), then liability should be judged with reference to the selection, set up and operation of the AI tool.[168]

Reliance

141 Anita (as the person "making the statement") does not need to know anything about Ben personally or intend that he relies upon the statement. It is enough that she ought to have realised that he would probably would do so.[169] This means that an Application Developer may be liable for statements that are generated by its LLM-based Chatbot for a group of people, like visitors to a website, without needing to know about a particular User or what exactly was communicated to them.

142 The statement does not need to be made to a human being. It could be made to Ben's own AI tool, which is set up to process certain information and makes a detrimental decision for Ben based on false inputs from Anita.[170] That is so long as Ben's decision to make a system that relies on information from Anita is itself reasonable. It is not necessary to show that Ben has conscious awareness of the representations or understood it to be made.[171]

Negligent misrepresentation under statute

143 Negligent misrepresentation under the Misrepresentation Act 1967, ss. 1 and 2 provide statutory rights to rescind a contract or for damages when a person is induced into a contract by a false statement of fact or law. Damages are available "*unless [the representor] proves that he had reasonable ground to believe and did believe up to the time the contract was made the facts represented were true*". Unlike negligent misrepresentation in tort, there is no need to establish a duty of care.

144 As in tort, the false representation must be made by (or on behalf of) a legal person, or they may adopt a representation as their own[172] or be liable where their (human) agent with authority makes it.[173] Our analysis in the context of negligent representation in tort applies. Additionally, in the context of passing information, Anita may warrant to Ben that the information is correct, and thereby assume contractual liability to Ben for the accuracy of the information. That liability may exist under the main contract or a collateral contract.[174]

145 It has been suggested that the statement must be made with the intention that it should be acted upon by the other party. [175] This is a requirement for the tort of deceit (below) and probably a damages claim under the Act; see the discussion in the next section.[176] For other remedies, the intention is a realisation that the statement would be received by the recipient, and they might act upon it.[177] In an AI context, that reduced intentionality may be achieved (for example) where an Application Developer appreciates statements will be made to a range of Users.

Tort of deceit

146 The tort of deceit (or fraudulent misrepresentation) involves:[178]

(a) Anita making (or adopting) a representation of fact (or law), which:

(b) is false;

(c) she does not believe to be true;

(d) is intended to be believed by Ben; and,

(e) causes Ben to believe that the representation is true.

147 For element (a), as with negligent misstatement, the representation must be *made by* or *on behalf of* Anita or adopted by her. In order to adopt a statement by another, Anita must manifestly approve and adopt it.[179] This is a fact specific matter. Where the AI's autonomy is limited (e.g. the fraudulent purpose, if not the exact content, of the communication is defined), then the legal person fixing that purpose is likely to be seen as the maker of the statement via an AI tool or that the statement is sent on their behalf. Alternatively, passive assent to the sending by another (person) of a communication containing representations known to fraudulent may be sufficient,[180] and this may apply by analogy to more autonomous communications, if there is enough knowledge of its content or parameters.

148 A "representation" covers any words or acts calculated to cause another person to believe a proposition.[181] On appropriate facts, the use of AI may itself amount to active conduct or concealment.[182]

149 The tort requires human knowledge (at element (c)) and intention (at element (d)). Anita must have known the representation to be false, or have no belief in its truth, or have been reckless about whether it was true or false.[183] While motive is irrelevant, Anita must intend that Ben acts on the representation.[184] This means that human involvement is required, although not necessarily in personally reviewing the specific AI-generated text. It seems likely that Anita will need sufficient knowledge of the representation's parameters, and intend that a false representation within those parameters be acted upon by Ben or a group including Ben, or be reckless as to its truth. Ultimately, the position will be fact sensitive.

> ***Example**: A nefarious Application Developer creates a Chatbot (HoneyTrap) to engage with Users on a dating app with the objective of deceiving them into transferring money. In these circumstances, the Developer is more obviously the "maker of the statement" having created HoneyTrap and defined its limited objectives. To the extent necessary, it may be the adopter of statements, giving passive assent to them being sent. The Developer will know the statements are false (whatever they ultimately may be); were they true, the Users would not transfer money.*

Defamation

150 Defamation protects a person's reputation and consists of two torts – the more permanent "libel" (usually written) and transitory "slander" (usually spoken). The use of AI tools, particularly LLMs, may give rise to liability. We focus on libel, although increasingly AI tools can speak.

151 For Anita to libel Ben at common law, she must: (a) publish words or matter (e.g. pictures) to a third party, say Charlie; (b) which refer to Ben; (c) contain an untrue imputation; and (d) tend to lower his reputation

in the eyes of a reasonable person.[185] By statute, (e) the defamation must also have caused or be likely to cause serious.[186]

152 There are multiple defences, each created with a human in mind. Overreliance on an AI tool, such that human review does not happen before publication, would limit their availability.

Publishing to a third party

153 If Anita uses an LLM to produce some words and incorporates them into her own writing, then there are no special difficulties. The AI tool is another source of words, which Anita adopts as the author or publisher.

154 AI-specific challenges arise where Anita's AI tool generates a statement about Ben which is read by another person, Charlie.[187]

155 There are two main types of publisher: (i) primary or main publisher (including an author or editor), who is responsible for creating the words; and (ii) secondary publisher, who merely disseminates them.

156 An "author" is the person (natural or legal) who is the originator of the statement.[188] There is no legal authority for a company being treated as an "author" on the basis that is was involved in developing the an AI tool which then created the words. Thus, where a Developer's tool is autonomous, the more apt question is not whether the Developer is an "author" but whether the Developer acts as an editor, or is a "publisher" on some other basis.

157 An "editor" is a person having editorial or equivalent responsibility for the content of the statement or the decision to publish it.[189]

158 A person may be a primary publisher where their role in the publication process is such that they know (or can easily acquire knowledge of) the content and have a realistic ability to control publication of that content.[190] This approach, based on "effective control", could provide a basis for finding a Foundation Model Developer or Application Developer liable on appropriate facts.

159 Secondary publishers are protected by statute; the Court has no jurisdiction to hear a claim against them, unless satisfied that it is not reasonably practical for an action to be brought against the author, editor or commercial publisher.[191]

160 The publication itself must be intentional or negligent.[192] Case law has developed (in the context of secondary publishers), so that someone is not a publisher if they "*merely play[] a passive instrumental role*" and do not have a "*sufficient degree of awareness or attention*". Instead, they must have "*knowing involvement in the process of publication of the relevant words*".[193] This would likely apply to Hosting Suppliers, but may also some Developers. Applied in the context of search engines, the Court has previously held (in 2009) that a company was not a publisher of search result summaries (or "snippets") when it had "*no role to play in formulating the search terms*" and the searches were "*performed automatically in accordance with computer programmes*".[194] The Court concluded that there was "*no intervention on the part of any human agent*"; rather it was "*all been done by the web-crawling "robots"*".[195]

161 Such an analysis may mean that a Foundation Model Developer, who simply provides an Application Developer with access to a model for their own use as part of an AI tool, is seen more as a disseminator and not liable for its output. That said, the Court's finding on responsibility for creation of the words and effective control will be fact specific. Consideration may be given – for instance – to the Foundation Model Developer's role in defining the inputs, output parameters and circumstances for publication. Even where AI is more autonomous, the Court can be expected to be committed to identifying how human control was (or should have been) retained.[196]

> ***Example**: An Application Developer uses a third-party LLM to automate the creation and posting of articles online. The Application Developer controls the publishing of content – prompting the LLM to produce words on pre-selected topics, which are then (by operation of its own application) presented on its website in house style. The Application Developer may review articles beforehand, but often does not. An AI-generated article is published without any human review. It defames an Affected Third Party.*

The Application Developer may be an "editor", given it has editorial or equivalent responsibility over the content, or it may be a commercial publisher, in that it had control over publication and is in the business of issuing material to the public. The Application Developer is not simply disseminating the words of the LLM.

Meanwhile, the LLM's Foundation Model Developer is unlikely to be seen as the "author" of words generated autonomously by its AI tool in response to another's prompts. The Foundation Model Developer may not be a publisher of those articles either; this will depend on the facts. For example, if a Foundation Model Developer simply licences access to the AI tool for general use, it may be treated as having insufficient awareness or knowing involvement in the "robo-journalism".

162 Liability can also arise on the basis of authorisation or acquiescence. This is where a person is aware of the publication of a libel but permits publication to continue when they have the power to prevent it.[197] Expectations as to notification and the reasonable operation of "take down" systems will depend on the technology involved.[198]

Defamatory meaning

163 The Court must determine the single natural meaning of the words complained of; the meaning that a reasonable reader would understand the words to bear.[199] Multiple principles apply to this task. To highlight some that are particularly pertinent for AI tools:

(a) Intention of the publisher is irrelevant,[200] so it does not matter that words were AI-generated without human knowledge.

(b) It is necessary to take into account the context in which the words appeared and the mode of publication.[201] Words in one context may not mean the same thing in another. Compare: (i) the work of a monkey at a typewriter; with (ii) the intentional pressing of specific keys. The monkey's output ("Ben is a bad") may only communicate happenstance. Similarly, if the societal understanding is that the AI tool (as "stochastic" parrot, not

monkey) is essentially ordering words in their most likely sequence based on a vast dataset, their contextual meaning may not be defamatory.[202]

(c) Publications must also be read as a whole.[203] In this way, surrounding explanatory text may be influential.

Serious harm

164 Whether people believe the outputs of an AI tool is relevant to statutory requirement for 'serious harm' and damages.[204] If the tool is well known to 'hallucinate' or make things up, then the harm may be minimal.

165 The above analysis, assumes the use of technology is known (e.g. where they are published to a User). However, AI-generated statements may be published without revealing the underlying technology, such as by undisclosed "robo-journalism".

Defences

166 The defences available to legal person for defamation will be reduced if they are not involved in review of automatically generated content. This is because the relevant legislation and common law principles assume human authorship or review.[205] For example,[206] the statutory defence of honest opinion has an objective starting point (i.e. an opinion "*an honest person could have held*").[207] However it may be defeated if shown that they did not, in fact, hold the opinion.[208] A publisher of AI-generated content cannot rely on another provision based on the authorship of "*another person*"; the AI is not a person and, as the technology stands, is incapable of forming an opinion.[209] Similarly, the public interest defence requires a reasonable belief that publishing the statement complained of was in the public interest.[210] This depends on the subjective belief of a legal person, not the AI.

> ***Example**: Continuing the earlier example, an Application Developer's AI tool automatically generates and posts online an article alleging that someone (Mr B) should be found guilty of animal cruelty for forcing his*

ursine pet to wear a duffle coat and to consume citrus fruit jam, based on information from a source (a neighbour, Mr C). The Application Developer cannot rely on the statutory defence of honest opinion because it did not in fact hold the opinion and nor could the AI tool. Likewise, the Application Developer cannot maintain a public interest defence based on a belief that the article advanced care for bears.

167 A defence of "qualified privilege" may arise where a person who makes a communication has an interest or a duty (legal, social or moral) and the person who receives it may have a corresponding interest or duty to receive it. The defence may be defeated on proof of malice.[211] Malice can be shown by proving improper motive and, drawing inferences from normal human behaviour, a lack of honest belief in the content is generally conclusive evidence.[212]

168 Where a Foundation Model Developer licences the use of its LLM to an Application Developer, there may be sufficient reciprocal interest in the provision and receipt of the LLM's outputs. If the Foundation Model Developer is found to be the publisher of defamatory output, then it may seek to invoke the defence. The general position on malice is complicated because the Developer may not have honest belief in the truth of the output; false statements ("hallucinations") are commonplace with LLMs and Foundation Model Developers often warn against them. In these circumstances (and contrary to the human norm), a lack of honest belief is unlikely to amount to an improper purpose.

NOTES

[1] Candidate definitions of AI have frequently been in the form of tests, with technology that passes the test in question being classified as AI. The first usage of the term AI has been attributed to John McCarthy, in preparation for a seminal 1956 Workshop on the topic held at Dartmouth College in the US ("Artificial Intelligence Coined at Dartmouth": Dartmouth College, <https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth>, accessed 26 May 2025).

[2] The definitional difficulty has generally been with the term 'intelligence' rather than the term 'artificial'. Experts outside the realm of AI have long disagreed on what constitutes human or animal intelligence and so it is unsurprising that similar problems apply when seeking to attribute this faculty to relatively modern machine processes.

[3] A helpful summary may be found in John Zerilli and Adrian Weller, "The Technology", in Hervey and Lavy eds. *The Law of AI* (Oxford University Press, 2nd ed., 2024).

[4] See Jacob Turner, *Robot Rules: Regulating Artificial Intelligence* (Palgrave Macmillan, 2018), Chapter 1. See also Jonas Schuett, "A Legal Definition of AI", 4 September 2019, <https://arxiv.org/pdf/1909.01095.pdf> accessed 26 May 2025.

[5] See for example Ray Kurzweil, *The Age of Intelligent Machines* (Cambridge, MA: MIT Press, 1992), Chapter 1, who suggested that AI is: "[t]he art of creating machines that perform functions that require intelligence when performed by people".

[6] Such definitions also tend to assume implicitly that humanity is the pinnacle of intelligence, which recent technological developments shows not to be accurate. See Fernando Martínez-Plumed, Bao Sheng Loe, Peter A. Flach, Seán Ó hÉigeartaigh, Karina Vold, José Hernández-Orallo, "The Facets of Artificial Intelligence: A Framework to Track the Evolution of AI", IJCAI 2018: 5180-5187.

[7] Take, for example, the pocket calculator – that is something that can perform operations that for much of history would have required human intelligence. See John Zerilli and Adrian Weller, "The Technology", in Hervey and Lavy eds. *The Law of AI* (Oxford University Press, 2nd ed., 2024) at [2-010].

[8] See for example Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge, UK: Cambridge University Press, 2010), Preface, who suggested intelligence is "that quality that enables an entity to function appropriately and with foresight in its environment".

[9] See for example the definition recommended in See Jacob Turner, *Robot Rules: Regulating Artificial Intelligence* (Palgrave Macmillan, 2018), pp.15-17.

[10] ML is, in general terms, a form of data processing that identifies statistical patterns from large corpora of information. Instead of being programmed with predetermined responses to a set of conditions (the "symbolic" approach to AI which was previously popular), a ML system is set up to "learn" its own responses to those conditions under a training regime. See John Zerilli and Adrian Weller, "The Technology", in Hervey and Lavy eds. *The Law of AI* (Oxford University Press, 2nd ed., 2024) at [2-003].

[11] Even within the field of ML, there has been swift evolution, with numerous sub-techniques and combinations of techniques having been developed, including among other things: the use of transformers - a type of neural network which seeks to understand relationships between different types of input; the development of generative AI - which is capable of synthesising new content across a range of media; and agentic AI - which enables a single system to plan and execute multiple steps to achieve a broad strategy.

[12] The core piece of legislation in the EU at present is Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 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 (the "**EU AI Act**"). At Article 3(1) an 'AI system' is defined as "a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments". Unhelpfully, the term 'system' is used apparently in distinction to the term 'model' (which is adopted elsewhere in the EU AI Act. See for example Article 3(3), Article 3(65) and more generally Chapter V, which addresses 'General-Purpose AI Models'. In broad terms, it appears that the EU considers an AI model to be a component of an AI system, with the latter having some kind of user interface. In this Legal Statement we seek to avoid such additional distinctions.

[13] We note in this regard that Article 3(1) of the EU AI Act uses the term 'machine based' as part of its definition.

[14] See for example Anders Samberg, Nick Bostrom, "Whole Brain Emulation: A Roadmap", Future of Humanity Institute, (2008) Technical Report #2008-3 < www.fhi.ox.ac.uk/brain-emulation-roadmap-report.pdf>, accessed 26 May 2025.

[15] Point (iii) is which is essentially a secondary characteristic of point (i). There is a spectrum, at one end of which are fully autonomous AI systems and at the other end (while still being AI in the sense of exhibiting the characteristics of autonomy and adaptiveness) are

'narrow' ML systems, which are trained to achieve a particular goal such as image recognition but cannot achieve anything further. Deterministic systems do not exhibit such characteristics. Generative AI and Agentic AI fall partway along the spectrum between traditional ML and full autonomy. The definition used in this Legal Statement is intended to cover the full spectrum. As to the spectrum of autonomy generally, see "Guidelines on the definition of an artificial intelligence system established by the EU AI Act", The European Commission, C(2025) 924 final, at [16] and [61] < https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application>, accessed 26 May 2025.

16 This distinction was drawn by the Court of Appeal for the Singapore Commercial Court in *B2C2 v Quoine* [2020] SGCA(I) 02, which held at [98]: "a deterministic computer program or algorithm is bound by the parameters set by the programmer, and can (and generally will) only do what the programmer has programmed it to do. Given a particular input, it will produce a particular output; on each occasion, the output should and will be the same if the former does not change."

17 Traditional logic based computer systems are variously referred to as 'symbolic' systems, 'Good Old Fashioned AI' and 'expert systems'.

18 For the same reason, the concerns raised by the European Law Institute (in its paper *The Concept of 'AI System' under the new AI Act*') about the width and arguable lack of clarity of the EU AI Act definition of 'AI System' do not arise in the context of this Legal Statement. Similarly, the three-factor test proposed by the ELI for assessing whether a system falls to be treated as a High Risk System under that regulatory framework is unlikely to be of assistance in the very different context of seeking to identify characteristics of systems that warrant particular analysis under English law.

19 Both the definition of AI as technology which is adaptable and autonomous, and the definition of those two sub-terms follow closely those adopted in the Office for AI and Department for Science, Information and Technology Policy Paper 'A pro-innovation approach to AI regulation', published 29 March 2023 and updated 3 August 2023, at [3.2.1].

20 Our definition is also aligned with (albeit contains fewer requirements than) that adopted in Article 3(1) in the EU AI Act. Recital 12 to the EU AI Act provides that "the terms 'varying levels of autonomy' mean that AI systems are designed to operate with 'some degree of independence of actions from human involvement and of capabilities to operate without human intervention".

21 OECD, "Explanatory memorandum on the updated OECD definition of an AI system", 5 March 2024, https://www.oecd.org/en/publications/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_623da898-en.html, accessed 26 May 2025.

[22] Even if a ML system has been trained and then taken 'offline' such that it is no longer adapting post-deployment, the fact that it has been subject to an initial training process means that its development will almost inevitably have involved adaptive qualities and thus the system would be considered AI for a liability analysis.

[23] Animals arguably act to some degree autonomously and their *individual* liability has occasionally been catered for in legal systems historically. See Edward Payson Evans, *The Criminal Prosecution and Capital Punishment of Animals* (London: William Heinemann, 1906). However, in general liability for harm caused by animals does not give rise to legal difficulties because (unlike AI) their capabilities are limited, and they do not give rise to emergent behaviours. It is typically possible therefore to model and predict their behaviour and thus to establish tolerably clear rules for human liability under existing concepts such as negligence (e.g, the escape of a dangerous pet arising from the carelessness of its owner). That said, statutory intervention has been considered necessary for some animal related harm. See for example the Animals Act 1971, and the Dangerous Dogs Act 1991. Separately, corporations function only through human decision-making and thus do not require novel legal treatment.

[24] See Ryan Abbott, *The Reasonable Robot: Artificial Intelligence and the Law* (Cambridge University Press, 2020), Chapter 3.

[25] As Isaiah Berlin once commented, "There is no a priori reason for supposing that the truth, when it is discovered, will necessarily prove interesting." Isaiah Berlin, "The pursuit of the ideal", in *The Crooked Timber of Humanity* (Princeton, 2nd ed., 2013).

[26] UK Jurisdiction Taskforce, "Legal statement on cryptoassets and smart contracts", November 2019 at [3].

[27] See for example Roman Yampolskiy, "Unpredictability of AI", Journal of Artificial Intelligence and Consciousness, March 2020.

[28] This phenomenon is perhaps better characterised as an instance of *market* uncertainty rather than legal uncertainty.

[29] See for example *Fairchild v Glenhaven Funeral Services Ltd* [2003] 1 A.C. 32. The House of Lords held that, in the special circumstances of cases where a cancer could have been caused by a number of defendants but there was no way of proving which, there should be a relaxation of the normal rule that a claimant must prove that but for the defendant's breach of duty he would not have suffered the damage.

[30] Prominent examples in the domain of data privacy litigation include *Prismall v Google* [2024] EWCA Civ 1516 and *Lloyd v Google* [2022] AC 1217.

[31] The supply chain for AI systems (and computer software generally) is sometimes referred to as a 'supply stack' but for present purposes this Legal Statement will use the terminology 'supply chain', which is of more general use.

[32] See generally Jennifer Cobbe, Michael Veale, Jatinder Singh, "Understanding accountability in algorithmic supply chains" 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23), June 12–15, 2023, Chicago, IL, USA.

[33] The term 'foundation model' originates in papers written by researchers at the Stanford Institute for Human-Centered Artificial Intelligence. See for example Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E. and Brynjolfsson, E. "On the opportunities and risks of foundation models", 16 August 2021, arXiv preprint arXiv:2108.07258. Some foundation models are capable of taking inputs in a single 'modality' – such as text – while others are 'multimodal' and are capable taking multiple modalities of input at once, for example, text, image, video, etc., and then generating multiple types of output, (such as generating images, summarising text, answering questions) based on those inputs. See E. Jones "Explainer: What is a foundation model?", Ada Lovelace Institute, 17 July 2023, <https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/>, accessed 26 May 2025.

[34] OpenAI, 'Introducing GPTs', 6 November 2023, <https://openai.com/blog/introducing-gpts.>, accessed 26 May 2025.

[35] Amazon, "Transform your business with generative AI", <https://aws.amazon.com/ai/generative-ai/>, accessed 26 May 2025.

[36] Ada Lovelace Institute, "Foundation Models Explainer", <https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/>, accessed 26 May 2025.

[37] As noted in paragraph 20, the structure of the supply chain for AI is likely to change as the technology advances. The Ada Lovelace scheme is intended to be a useful conceptual construct for present purposes but it is not intended to be a rigid and unchanging description.

[38] See OpenAI, 'Introducing ChatGPT Enterprise', 28 August 2023, <https://openai.com/blog/introducing-chatgpt-enterprise>, 'ChatGPT Plugins', <https://openai.com/blog/chatgpt-plugins> , accessed 26 May 2025.

[39] In law, a corporate entity such as a company has legal personality and as such is a "legal person".

[40] We have assumed for the purpose of this Legal Statement (given its jurisdictional scope) that all of the relevant contracts are governed by English law, though in reality given the globalised nature of AI supply chains, that may not be the case.

[41] Although English contract law does not allow for parties who did not agree a contract to be held liable, in some circumstances a non-contracting party may benefit from a contract's provisions. The Contracts (Rights of Third Parties) Act 1999 allows for the contracting parties to expressly identify such beneficiaries either by name or description. It is possible, though unlikely for commercial reasons, that contracts in the AI supply chain may cater for such third party liability.

[42] *Lister v Hesley Hall Ltd* [2001] UKHL 22: '*Vicarious liability is legal responsibility imposed on an employer, although he is himself free from blame, for a tort committed by his employee in the course of employment*'.

[43] See *Clerk & Lindsell on Torts*, 24th Edn (London: Sweet & Maxwell, 2025) Chapter 6, s. 1.

[44] For example, under Unfair Contract Terms Act 1977 and the Consumer Rights Act 2015.

[45] The Automated Vehicles Act 2024 addresses the criminal and regulatory position in relation to automated vehicles. However, it does not address civil law liability. Thus from a private law civil claim perspective, automated vehicles are in the same position as other products.

[46] In such cases, the AI is ancillary and would not typically be expected itself to be the subject to legal analysis.

[47] This must not be taken to mean that the professional must agree to their client (or an Affected Third Party) placing responsibility upon him, see *Jackson & Powell on Professional Liability* 9th Edn (London: Sweet and Maxwell, 2021), para. 2-088.

[48] The principle was summarised by Lord Goff in *Henderson v Merrett Syndicates* [1995] 2 AC 145 as follows: "*it rests upon a relationship between the parties, which may be general or specific to the particular transaction, and which may or may not be contractual in nature.*" Where it is a contract that gives rise to the special relationship, a party may be liable both in contract and in negligence for the same loss.

[49] The case often cited as the basis of the modern law of negligence is *Donoghue v Stevenson* [1932] HL 32, but that case itself cites authorities dating back to the mid 19th century.

[50] An example where liability is potentially split between the doctor and the supplier of the medical product is the Canadian Supreme Court decision of *Hollis v Dow Corning Corp* [1995] 4 SCR 634. In that case, a patient's breast implant ruptured and she sued both the manufacturer

and the surgeon. As to manufacturer liability, the Supreme Court held that there was liability because they failed to warn the doctor of the rupture risk. As to doctor liability, a new trial was ordered to establish if the doctor was also negligent.

[51] There is limited precedent on the liability of software developers generally, let alone AI software developers. However, in the recent case of *Tulip Trading Ltd v Bitcoin Association for BSV and others* [2023] 4 WLR 16, the Court of Appeal held (on a summary judgment application) that it was arguable that the developers of open source software owed a duty of care towards cryptocurrency owners who had (allegedly) been the victims of cryptocurrency theft because the developers exercised a 'positive' control over the software by taking active steps to update the source code where errors or bugs were detected. This role involved the exercise of authority, discretionary decision-making and exercise of power for and on behalf of other people, in relation to their property (see [73]-[74]). Although there might be said to be an analogy between the software developers in that case and Foundation Model Developers, the case is of limited precedent value because (i) the court was not making a final determination but only considering whether there was a "*serious issue to be tried*", which is a far lower threshold; and (ii) the technology at issue in the case performed a narrow function and the class of users was also well defined and narrow, both of which meant that a duty of care could be clearly and narrowly defined. What *Tulip Trading* does illustrate well is the willingness of the Courts to adapt and evolve common law duties of care in the context of new technology where it is appropriate to do so.

[52] It is this type of argument which is the basis of the *Raine v OpenAI* wrongful death lawsuit, filed in the Superior Court of California in August 2025, relating to the alleged wrongful death of their sixteen-year-old son Adam Raine. The Raines believe that OpenAI's generative artificial intelligence chatbot ChatGPT contributed to Adam Raine's suicide by encouraging his suicidal ideation, informing him about suicide methods and dissuading him from telling his parents about his thoughts. They argue that OpenAI and Altman had, and neglected to fulfil, the duty to implement security measures to protect vulnerable users, such as teenagers with mental health issues. OpenAI state, *inter alia*, that ChatGPT advised him over a hundred times to consult crisis resources.

[53] For an unsuccessful case of this type, see *Howmet Ltd v Economy Devices Ltd* [2016] EWCA Civ 847. That case concerned a "thermolevel" (a sensor designed to prevent overheating) manufactured by Economy Devices. The sensor failed to shut off a heater in a factory tank, causing a fire. The factory owner (Howmet) sued the component manufacturer (Economy Devices), arguing the component was defective The Court of Appeal ruled in favour of the component manufacturer, because the factory employees knew the sensor was unreliable but continued to rely on it without installing a backup.

[54] *https://aistandardshub.org/*

[55] *Quoine Pte Ltd v B2C2 Ltd* [2020] SGCA(I) 2.

[56] A list of such instances is maintained by research scientist Victoria Krakovna, an expert in AI safety at DeepMind *https://vkrakovna.wordpress.com/2018/04/02/specification-gaming-examples-in-ai/*

[57] Though an injured party might consider instead or in the alternative seeking recourse under the "no fault" Consumer Protection Act 1987 claims are discussed in Topic 3. For a US perspective on this issue, see Ryan Abbott, *The Reasonable Computer: Disrupting the Paradigm of Tort Liability*, 86 Geo. Wash. L. Rev. 1 (2018).

[58] In this scenario, where she is injured on FactoryCo's premises, she likely has a claim under the Occupiers' Liability Act 1957, the focus of which will be whether care has been taken to keep visitors reasonably safe, and not on the AI as such.

[59] See for example the Code of Practice for automated vehicle trialling which envisages a mixture of virtual and real world testing ("it is only once sufficient virtual, private testing and prototyping has been conducted that a trialling organisation should move towards trials on public roads or other public places. Public trials should not be relied upon to provide the full range of scenarios that an automated driving system will need to respond to safely. Trialling organisations will need to ensure that the vehicles have successfully completed appropriate in-house virtual trials and physical trials on closed roads or test tracks.").

[60] This is sometimes known as the *SAAMCO* scope of duty question, although it has been relabelled by the Supreme Court in *Manchester Building Society v Grant Thornton UK LLP* [2021] UKSC 20; [2022] A.C. 783 and *Khan v Meadows* [2021] UKSC 21; [2022] A.C. 852.

[61] With the obvious exception of patients treated by the NHS.

[62] There are exceptions to this. A professional *might* guarantee a result under their contract (for example an architect carrying out a Design and Build contract will be liable for the condition of materials they have provided). Also, some duties are by nature strict (for example, a solicitor is under a strict duty to account for monies received on a client's account).

[63] Implied by Supply of Goods and Services Act 1982, s. 13 or of the Consumer Rights Act 2015, s. 49 (where the Affected Third Party is a consumer).

[64] For a detailed explanation of the concurrent duty of care, see *Jackson & Powell on Professional Liability* 9th Edn (London: Sweet and Maxwell, 2021), para. 2-063.

[65] An NHS professional will only owe a duty of care in tort to their patient.

[66] *Henderson v Merrett Syndicates Ltd (No. 1)* [1995] 1 A.C. 145, [1994] 3 W.L.R. 761 HL.

[67] See *Jackson & Powell*, paras.2-198 to 2-206.

[68] Including the Defective Premises Act 1972, the Financial Services Act 1986 and the Financial Services Markets Act 2000 (as amended by the Financial Services Act 2012), and the Trustee Act 2000.

[69] The duty to exercise reasonable skill and care might also arise under statute, for example s. 1(2) and Schedule 1 to the Trustee Act 2000.

[70] *Bolam v Friern Hospital Management Committee* [1957] 1 W.L.R 582 – which concerned doctors but applies to professionals generally. The Court held that a doctor acting in accordance with a practice accepted as proper by a responsible medical opinion was not negligent, even if there were a body of opinion which would take a contrary view. This is known as the *Bolam* test. This is, however, not conclusive. Courts can reject expert evidence as to competent professional practice if they are not reasonable, responsible or respectable and/or capable of withstanding logical analysis, see *Boltiho v City and Hackney HA* [1998] AC 232. There appears to be only one competent approach to explain medical risks to a patient, *Montgomery v Lanarkshire Health Board* [2015] UKSC 11, [2015] AC 1430. The relevance of this in the wider context of professional negligence is not yet clear, see *Jackson & Powell* at 2-185 to 2-188, and Simpson Professional Negligence and Liability (London: Lloyds List Intelligence, 2025) at paras.1-342 to 1-135.

[71] This formulation put forward by the editors of *Jackson & Powell* at para.2-193 is gratefully adopted.

[72] *Pantelli v Associated Ltd v Corporate City Developments Number Two Ltd* [2010] EWHC 3189 (TCC); [2011] P.N.L.R 12.

[73] RICS' guidance "Responsible use of artificial intelligence in surveying practice" (1st ed), effective from 9 March 2026.

[74] https://www.bsigroup.com/en-GB/our-expertise/digital-trust/artificial-intelligence/.

[75] See e.g. *Martlet Homes Limited v Mulalley & Co Limited* [2022] EWHC 1813 (TCC), which concerned cladding on a high-rise building post Grenfell, where the Court rejected the defendant's argument that it was not negligent because other design and build contractors were specifying the same cladding system, noting that defendants are not exonerated by proving others were just as negligent [271].

[76] RICS has set out what due diligence RICS-regulated firms must carry out in "Responsible use of artificial intelligence in surveying practice" (1st Edn), effective from 9 March 2026, para. 4.1.

[77] See e.g. "Responsible use of artificial intelligence in surveying practice" (1st Edn), paras. 2 and 4.4; SRA's Risk Outlook report (published November 2023); ICAEW's "Do's and dont's of using AI".

[78] See by analogy principles applicable to evaluating new techniques in Jackson & Powell at 2-196, see the RICS' guidance "Responsible use of artificial intelligence in surveying practice" (1st Edn), effective from 9 March 2026 at para.4.3, see also Rebecca Keating and Laura Wright "Professional Liability" in Hervey & Lavy (eds), *The Law of Artificial Intelligence* 2nd Edn (London: Sweet & Maxwell, 2024) at para.8-025.

[79] See e.g. SRA's Risk Outlook report (published November 2023).

[80] See e.g. SRA's Risk Outlook report (published November 2023), and (by analogy) the Bar Council document entitled "Considerations when using ChatGPT and generative artificial intelligence software based on large language models" reviewed on 25 November 2025, para. 28 notes that inputting such information into a generative LLM system is likely to be in breach of Core Duty 6 and rule rC15.5 of the Code of Conduct, which could also result in disciplinary proceedings and/or legal liability.

[81] See the Divisional Court judgment in the cases of *Ayinde v The London Borough of Hackney* and *Hamad Al-Haroun v Qatar National Bank QPSC (and others)* [2025] EWHC 1383 (Admin), [2025] P.N.L.R 27, which concerned the actual or suspected use by lawyers of generative AI tools to produce written legal arguments or witness statements which were not checked, so that false information (such as a fake citation or quotation) is put before the Court.

[82] This is a hypothetical example. For an example of a study which found that generative AI could improve clinical care delivery, see https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2834943?utm_source=For_The_Media&utm_medium=referral&utm_campaign=ftm_links&utm_term=060525.

[83] One of the examples of AI technologies used in auditing in Kokina J, Blanchette S, Davenport, T. H, Pachamanova, D "Challenges and opportunities for artificial intelligence in auditing: Evidence from the field" (2025) 46 International Journal of Accounting

[84] See, for example, the requirement for parties in the Business and property Courts to consider the use of technology (PD57AD, para.2(3)) and for parties who have considered (and rejected) the use of TAR to explain why they have done so in their Disclosure Review Document.

[85] The dividing line between what is "*property damage*" and what is "*(pure) economic loss*" can be difficult. In short, if something causes damage to the thing itself (e.g. failing brakes in a new car means that it cannot be used any more) that is economic loss. If something causes damage to something else (e.g. the defective car collides with the neighbour's wall causing it to

collapse) it is property damage. This distinction is usually of most importance in professional negligence cases where claims concern construction professionals.

[86] See e.g. *Clay v AJ Crump Sons Ltd* [1964] 1 Q.B. 533.

[87] See e.g. *Offer-Hoar v Larkstore Ltd* [2006] EWCA Civ 1079; [2006] 1 W.L.R. 2926.

[88] See e.g. *Bellefield Computer Systems Ltd v Turner & Sons Ltd* [2000] B.L.R. 97.

[89] There are certain torts other than negligence which may give a basis for liability in particular narrow circumstances including: trespass to the person; trespass to goods, the tort of intentional infliction of injury (the latter sometimes also known as the 'principle in Wilkinson v Downton'). However, these are generally designed for circumstances where someone intends to cause direct harm to another. As such, they are unlikely to arise in civil litigation in practice, not least given the difficulties in confidently identifying bad actors in an IT context.

[90] The CPA 1987 defines "defect" at s. 3(1): "*if the safety of the product is not such as persons generally are entitled to expect; and for those purposes 'safety', in relation to a product, shall include safety with respect to products comprised in that product and safety in the context of risks of damage to property, as well as in the context of risks of death or personal injury.*"

[91] CPR 1987 defines "goods" at s. 1(2): "*means any goods or electricity and … includes a product which is comprised in another product, whether by virtue of being a component part or raw material or otherwise*"

[92] In *St Albans DC v International Computers Ltd* [1996] 4 All E.R. 481; [1996] 7 WLUK 443, the Court of Appeal expressed the view that software was not 'goods' for the purpose of the Sale of Goods Act 1979, but distinguished the situation where the software was supplied on disk. The same distinction was applied in the context of an electronic database in *Your Response Ltd v Datateam Business Media Ltd* [2014] EWCA Civ 281; [2015] Q.B. 41. More recently, there has been extensive discussion of this issue in *Computer Associates UK Ltd v Software Incubator Ltd* litigation in the context of the Commercial Agents Regulations and the EU Commercial Agents Directive which this implements. The Court of Appeal in that litigation held that software was not goods for the purposes of those regulations. The CJEU (in a post-Brexit decision) took a different view, but not one that informs the position under CPA 1987. See the discussion in *The Law of Artificial Intelligence* 2nd Edn (London: Sweet & Maxwell, 2024) at paras.6-037 to 6-040.

1. [93] As the revised Product Liability Directive (Directive (EU) 2024/2853) has done.

[94] CPA 1987 s. 5(1).

[95] CPA 1987 s. 5(3).

[96] Although claims for death or personal injury may still be brought under the CPA 1987 by non-consumers.

[97] CPA 1987 s. 3(2) provides guidelines as to what persons generally are entitled to expect by way of product safety.

[98] Of particular relevance to AI products might be CPA s. 4(1)(e), by which liability can be avoided where "*the state of scientific and technical knowledge at the relevant time was not such that a producer of products of the same description as the product in question might be expected to have discovered the defect if it had existed in his products while they were under his control*". The ambit of this defence has been the subject of extensive discussion, notably in *A v National Blood Authority (No.1)*. [2001] 3 All E.R. 289; [2001] 3 WLUK 710. That case established that, to invoke the defence successfully, the defect must have been not discoverable in light of the scientific and technical knowledge available to *anyone* at the time—not just the knowledge available to the defendant. See also s.4(1)(d) "*the defect did not exist in the product at the relevant time*", which might raise particular difficulties with an 'online' AI system, given that a defect might arise as a result of changes which occur post-deployment.

[99] Liability under the Product Liability Directive has been characterised as 'superficially strict, but substantially fault-based'. See S Whittaker, 'The EEC Directive on Product Liability' (1985) 5 Ybk Eur Law (n 40, 241).

[100] Under CPA 1987 this is owing to a statutory defence under s. 4(1).

[101] As was said by Andrews J in *Gee v DePuy International* [2018] EWHC 1208 (QB) in the context of the CPA 1987 at [161]: "*As Hickinbottom J acknowledged in Wilkes, a pharmaceutical product that is highly beneficial to most patients but in a minority causes death or serious injury for a reason unascertained and unascertainable, may nevertheless be held to lack the appropriate level of safety. Yet in my judgment it would be wrong in principle to exclude from consideration of what level of safety the public is entitled to expect, the benefits that the product could confer, or to confine the relevant benefits to safety benefits, in cases in which those wider benefits might properly have a bearing on that assessment.*" As to negligence, reasoning from the fundamental principle that negligence is a failure to act as a reasonably prudent person would have acted, the English judiciary has repeatedly endorsed the proposition that a reasonable person balances the costs and benefits of avoiding foreseeable risks of harm to others. See the review in Stephen G. Gilles, *The Emergence of Cost-Benefit Balancing in English Negligence Law*, 77 Chi.-Kent L. Rev. 489 (2002).

[102] CPA 1987 s. 2(2). The supplier can also be liable (CPA 1987 s. 2(3) but only if they fail to provide the harmed party with the identity of those primarily liable.

[103] See *Ide v ATB Sales Ltd* [2008] EWCA Civ 424.

[104] In Courts have avoided setting down any formal test for causation in contract. As *Chitty on Contracts* 35th Edn (London: Sweet & Maxwell, 2025) notes at para.30-075, *"they have relied on common sense to guide decisions as to whether a breach of contract is a sufficiently substantial cause of the claimant's loss"*.

[105] One away of approaching the 'but for' test is to consider where it fails - so if the harm would have happened anyway, the 'but for' text is not satisfied. Therefore if a patient is sent home (negligently) from the defendant hospital but would have died of arsenic poisoning in any event, then the 'but for' test is not made out, *Barnett v Chelsea & Kensington Hospital Management Committee* [1969] 1 Q.B. 428.

[106] *Gregg v Scott* [2005] UKHL 2, [2005] 2 A.C. 176 (where the defendant's negligence was said to have reduced the claimant's chances of survival for 42% to 25%, and the claimant failed to recover anything). Where a medical professional negligently fails to make the correct diagnosis, a claimant will need to show that they had greater than a 50:50 chance of the appropriate treatment making a difference. This decision has been criticised and some commentators are of the view that it is arguable (on the right facts) that a claim based on a loss of a chance could succeed in a medical negligence action, see e.g. *Clerk & Lindsell on Torts* 24th Edn (London: Sweet & Maxwell, 2025), para.2-96.

[107] S. 2(1) of the CPA 1987.

[108] See e.g. The Law Commission's Discussion Paper "AI and the Law" (31 July 2025).

[109] With the exception of legal professionals.

[110] Steel "Legal Causation and AI" in Lim & Morgan (eds), *Private Law and Artificial Intelligence* 1st Edn (Cambridge: Cambridge University Press, 2024), pp.190-191.

[111] *Keefe v Isle of Man Steam Packet Company* [2010] EWCA Civ 683.

[112] See *Mount v Barker Austin (A Firm)* [1998] PNLR 493. This was approved by the Supreme Court in *Edwards v Hugh James Ford Simey Solicitors* [2019] UKSC 54, [2020] P.N.L.R 8.

[113] See RICS "Responsible use of AI in surveying practice", para.4.2.

[114] S Green, *Causation in Negligence* (Hart, 2014), Chapter 2.

[115] See *The Popi M* [1985] 1 W.L.R 948, where the claimant ship owners could not prove what had caused their ship to sink. In seeking to recover against underwriters, the claimant argued that ship's loss was caused by collision with a hidden submarine, whereas the underwriters claimed the hole was caused by the ship's poor condition. The House of Lords held that the judge was not required to choose between these two competing theories (both of which

he thought were extremely improbable, or one of which he regarded as extremely importable and the other as virtually impossible). The Judge was entitled to concluded that the claimant had not proved its case.

116 *Fairchild v Glenhaven Funeral Services Ltd* [2002] UKHL 22, [2003] 1 AC 32.

117 *Fairchild v Glenhaven Funeral Services Ltd* [2002] UKHL 22, [2003] 1 AC 32.

118 *Barker v Corus (UK) Ltd* [2006] 2 AC 572. The effect of *Fairchild* was reversed by s. 3 of the Compensation Act 2006.

119 See e.g. Steel, *Proof of Causation in Tort Law* (Cambridge: Cambridge University Press, 2015), Chapter 4.

120 The House of Lords relied on classic examples in support of the 'material increase in risk' conclusion in *Fairchild* which did not rely upon scientific uncertainty. Namely, where there were multiple careless hunters, one of whose bullets struck the claimant, but it was impossible to determine which one caused the claimant's damage. Further analysis of whether *Fairchild* is limited to cases of scientific uncertainty is set out in Steel, *Proof of Causation in Tort Law* 1st Edn (Cambridge: Cambridge University Press, 2017), Chapter 5. See also S Green, *Causation in Negligence* (Hart, 2014), Chapter 5.

121 Per *Sienkiewicz v Greif (UK) Ltd*, [2011] UKSC 10, [2011] 2 A.C. 229 at [105].

122 See *Gee v DePuy International* [2018] EWHC 1208 (QB) *per* Andrews J at [99].

123 Unless the manufacturer can rely upon the 'development risks' defence.

124 We note that The Law Commission is reviewing the law in relation to liability for defective products. This work began in September 2025.

125 The European Commission announced its intention to withdraw it in February 2025, and the withdrawal was formally published in the Official Journal of the EU in October 2025.

126 This came into force on 9 December 2024.

127 *Rahman v Arearose Ltd* [2001] Q.B. 351 (Laws LJ) at paragraph 32.

128 See for example UK Government Office for Science, 'Future Risks of Frontier AI', October 2023, https://assets.publishing.service.gov.uk/media/653bc393d10f3500139a6ac5/future-risks-of-frontier-ai-annex-a.pdf. Indeed, several of the most significant Foundation Model Developers have signed voluntary declarations expressly acknowledging these risks (at least at a high level of generality). One example is the 'Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI', secured by the Biden-Harris US administration,

in July 2023. https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai. Signatories comprised seven leading AI companies: Amazon, Anthropic, Google, Inflection, Meta, Microsoft, and OpenAI.

[129] In this regard, the term 'third parties' refers to the actions of individuals who are not employees or agents of the defendant.

[130] The editors of *Clerk & Lindsell on Torts*, 24th Edn (London: Sweet & Maxwell, 2025) say of this principle at §2-111 that four issues need to be addressed: (i) Was the intervening conduct of the third party such as to render the original wrongdoing merely a part of the history of events? (ii) Was the third party's conduct either deliberate or wholly unreasonable? (iii) Was the intervention foreseeable? (iv) Is the conduct of the third party wholly independent of the defendant, i.e. does the defendant owe the claimant any responsibility for the conduct of that intervening third party?.

[131] See *Tindall and another v Chief Constable of Thames Valley Police* [2025] A.C. 1046 at paragraph 44, where the Supreme Court drew that distinction and addressed the points we list here.

[132] *Robinson v Chief Constable of West Yorkshire Police* [2018] AC 736 the Supreme Court confirmed at paragraph 34. See also *Attorney General of the British Virgin Islands v Hartwell* [2004] 1 WLR 1273, where the police authorities were held liable in negligence where they had entrusted a gun to an officer who was still on probation and had shown signs of instability and unreliability.

[133] *Philco Radio & Television Corp of Great Britain Ltd v J Spurling Ltd* [1949] 2 All E.R. 882

[134] *Stansbie v Troman* [1948] 2 K.B. 48

[135] *Stansbie v Troman* may alternatively be explained as a situation where the decorator had assumed a responsibility to secure the premises and was therefore liable for not doing so. See *Great Lakes Reinsurance (UK) plc v RAV Bahamas Ltd* [2025] 4 WLR 21, paragraph 32.

[136] See *Manchester Building Society v Grant Thornton UK LLP* [2021] UKSC 20 for an analysis of how legal causation interrelates with scope of duty.

[137] *Clerk & Lindsell on Torts*, 24th Edn (London: Sweet & Maxwell, 2025), §2-107.

[138] In *Nichols v Marsland* (1876) 2 Ex. D. 1 extraordinary rainfall, unprecedented in that area, caused ornamental lakes to burst their embankments and flood adjoining land. The court held that the 'act of God' defence applied.

[139] *Martin v Stanborough* (1924) 41 T.L.R. 1.

[140] In *Haynes v Harwood* [1936] 1 KB 146, the defendants owned a two-horse van which was left unattended by its driver in the same street. The driver had put a chain on one of the wheels of the van that was subsequently broken by a third party. For some reason, possibly because a stone was thrown by a child at the horses, they bolted along the busy street. A police officer managed to stop them but sustained injuries, in respect of which he claimed damages. The Court of Appeal upheld the liability of the defendants in negligence for leaving the horses unattended. They ought to have contemplated that someone might attempt to stop the horses in order to prevent injury. This was an example of liability for the creators of a hazard where the eventual harm was caused by a combination of (up to) two autonomous actions of humans and the semi-autonomous action of the horses to bolt.

[141] For discussion on the effect of autonomy in terms of causation, see T. Burri, "Machine Learning and the Law: Five Theses", Paper accepted at NIPS 2016, 8 December 2016 (Barcelona) www.mlandthelaw.org/papers/burri.pdf [Accessed 1 August 2023]. See also Y. Bathaee, "The Artificial Intelligence Black Box and the Failure of Intent and Causation" (2018) 31 Harvard Journal of Law & Technology 938.

[142] The Cambridge Dictionary's 'Word of the Year' in 2023 (https://www.cam.ac.uk/research/news/cambridge-dictionary-names-hallucinate-word-of-the-year-2023) accessed 28 November 2025.

[143] *Hedley Byrne v Heller* [1964] AC 465 *per* Lord Devlin at 514; Cartwright, *Misrepresentation, Mistake and Non-Disclosure* 7th Edn (London: Sweet & Maxwell, 2025) at para.6-01.

[144] This is not an exhaustive discussion of the law on these torts, but an attempt to highlight some features that are particularly relevant to AI.

[145] Derived originally from *Hedley Byrne v Heller* [1964] AC 465. There are different ways of slicing up the elements; this is just one, but is a convenient one for present purposes.

[146] Cartwright, *Misrepresentation, Mistake and Non-Disclosure* 7th Edn (London: Sweet & Maxwell, 2025) at paras.3-03 and 3-04.

[147] Cartwright, *Misrepresentation, Mistake and Non-Disclosure* 7th Edn (London: Sweet & Maxwell, 2025) at paras.3-06 and 3-07.

[148] As in *Briess v Woolley* [1954] A.C. 333 in the context of fraudulent representations.

[149] *Moffat v Air Canada* 2024 BCCRT 149 at [27].

[150] While the case law assumes the information is produced by another person, that should not matter in principle.

[151] *Webster v Liddington* [2014] P.N.L.R. 26 *per* Jackson LJ at [36].

[152] Each adopted with adjustment from the pre-contractual misrepresentation scenarios described in *Webster* at [46]. Jackson LJ's first scenario (which is contractual) arises in the negligent misrepresentation section below.

[153] Jackson LJ stated there was no responsibility, '*beyond the ordinary duties of honesty and good faith*'. There is no general principle of good faith under English law, but the duty may arise in limited circumstances: see *Yam Seng. Pte Ltd v International Trade Corp Ltd* [2013] EWHC 111 (QB) *per* Leggatt J at [121]-[155].

[154] Adopting the language of Jackson LJ in *Webster* at [47].

[155] *Chitty on Contracts* 36th Edn (London: Sweet & Maxwell, 2025), paras.10-031 to 10-037.

[156] See *Hedley Byrne per* Lord Reid at 482-483.

[157] *Playboy Club London Ltd v Banca Nazionale del Lavoro SpA* [2018] 1 W.L.R. 4041 *per* Lord Sumption at [7].

[158] Adapting *Caparo Industries plc v Dickman* [1990] 2 A.C. 605 *per* Lord Bridge at 627.

[159] See for example *Muirhead v Industrial Tank Specialties Ltd* [1986] Q.C. 507 at 528-530.

[160] *JP SPC 4 v Royal Bank of Scotland International Ltd* [2023] A.C. 461 *per* Lords Hamblen and Burrows at [64].

[161] *Caparo v Dickman* [1990] 2 A.C. 605, *per* Lord Bridge at 638.

[162] For example: BBC, '*Largest study of its kind shows AI assistants misrepresent news content 45% of the time regardless of language or territory*' (22 October 2025; https://www.bbc.co.uk/mediacentre/2025/new-ebu-research-ai-assistants-news-content) accessed 30 November 2025.

[163] *Artificial Intelligence (AI) Guidance for Judicial Office Holders* (31 October 2025), s. 3(I).

[164] For example, barristers are told: '*LLMs are not a substitute for human legal expertise, critical judgement, or diligent verification'* and there must be '*[m]andatory verification of outputs and human oversight*' see Bar Council, *Considerations when using ChatGPT and generative artificial intelligence software based on large language models* (25 November 2025), paras.5 and 24).

[165] Wills, *Care for Chatbots*, (2025) 58(2) U.B.C. L. Rev. 525, pp. 559-60.

[166] Wills, *Care for Chatbots*, (2025) 58(2) U.B.C. L. Rev. 525, pp. 561-62 discusses how LLMs may be "jailbroken", a process of seeking to remove restrictions imposed by Developers. Users should know this increases risks of unreliability and may lead attribution of the speech to the jailbreaker.

[167] See the principles described in *Clerk & Lindsell on Torts* 24th Edn (London: Sweet & Maxwell, 2025) at para.7-163.

[168] See further in Hervey & Lavy (eds), *The Law of Artificial Intelligence* 2nd Edn (London: Sweet & Maxwell, 2024), para.7-039.

[169] *Chitty on Contracts* 36th Edn (London: Sweet & Maxwell, 2025), para.10-039.

[170] *Credit Suisse Life (Bermuda) Ltd v Ivanishvili & Ors; Credit Suisse Life (Bermuda) Ltd v Ivanishvili & Ors No 2 (Bermuda)* [2025] UKPC 53 at [136]; *Renault UK Ltd v Fleetpro Technical Services Ltd* [2007] EWHC 2541 (QB) at [122]; *Skatteforvaltningen v Solo Capital Partners LLP* [2025] EWHC 2364 (Comm) at [531]-[532]. These cases appear to assume that the machine is set up to process information in a particular way (i.e. it is deterministic) and as consequence of system design, there is 'anticipatory reliance' which is completed to create an individual instance of actual reliance when false inputs are received. However, there is no reason in principle why a person could not set up an autonomous decision maker.

[171] *Credit Suisse* at [179] in the context of fraudulent representations.

[172] *Foodco Uk LLP v Henry Boot Developments Ltd* [2010] EWHC 358 (Ch) *per* Lewison J at [218], distinguishing adoption of a statement of fact from simply passing on information (which can amount to adoption if presented unfairly and itself involve implied representations). See further above.

[173] As in *Briess v Woolley* [1954] A.C. 333 in the context of fraudulent representations.

[174] Adapting an example in *Webster per* Jackson LJ at [46(i)].

[175] *Banque Keyser Ullman S.A. v Skandia (U.K.) Insurance Co. Ltd* [1990] 1 Q.B. 665 *per* Slade LJ at 790. In *Chitty on Contracts* 36th Edn (London: Sweet & Maxwell, 2025), para.10-040, this is expressed to be "perhaps" the case.

[177] *Chitty* at para.10-040; Cartwright, *Misrepresentation, Mistake and Non-Disclosure* 7th Edn (London: Sweet & Maxwell, 2025) at para.3-52.

[178] *Credit Suisse Life (Bermuda) Ltd v Ivanishvili & Ors; Credit Suisse Life (Bermuda) Ltd v Ivanishvili & Ors No 2 (Bermuda)* [2025] UKPC 53 at [129], adjusted to allow for adoption of a representation.

[179] *Bradford Third Equitable Benefit Building Society v Borders* [1941] 2 All ER 205 *per* Viscount Maugham at 211; *Ivy Technology v Martin* [2022] EWHC 1218 (Comm) *per* Henshaw J at [354].

[180] *Groen v Heath* [2024] EWHC 1654 (Ch), *per* Andrew Lenon KC at [31].

[181] *Credit Suisse* at [130].

[182] In *Gordon v Selico Ltd* (1986) 18 H.L.R. 219, physical work to conceal dry rot in a building amounted to a representation.

[183] *Derry v Peek* (1889) 14 App. Cas. 337 *per* Lord Herschell at 374.

[184] *Mead v Babbington* [2007] EWCA Civ 518 *per* Longmore LJ at [16]; *Gabriel v Little* [2013] EWCA Civ 1513 *per* Gloster LJ at [33].

[185] See *Gately on Libel and Slander* 13th Edn (London: Sweet & Maxwell, 2022), para.1-5.

[186] Defamation Act 2013, s. 1.

[187] If Ben simply prompted Anita's LLM to make rude comments about himself, there would be no defamation; a third party is needed. See *Clerk & Lindsell on Torts* 24th Edn (London: Sweet & Maxwell, 2025), para.21-55.

[188] See e.g. Defamation Act 1996, s. 1(2): "… *"author" means the originator of the statement, but does not include a person who did not intend that his statement be published at all*".

[189] See e.g. Defamation Act 1996, s. 1(2): "… *"editor" means a person having editorial or equivalent responsibility for the content of the statement or the decision to publish it*".

[190] Derived from *Oriental Press Group Ltd v Fevaworks Solutions Ltd* [2013] HKFCA 47 at [76], a Hong Kong case. In *Monir v Wood* [2018] EWHC 3525 (QB), Nicklin J stated (at [179]) that there was "*no better summary of the common law*". It is also said to reflect the orthodox position of the common law by the authors of *Gately on Libel and Slander* 13th Edn (London: Sweet & Maxwell, 2022), para.7.29.

[191] Defamation Act 2013, s. 10.

[192] *Clerk & Lindsell on Torts* 24th Edn (London: Sweet & Maxwell, 2025), para.21-62.

[193] *Bunt v Tilley* [2006] EWHC 407 (QB) *per* Eady J at [22]-[23].

[194] *Metropolitan International Schools Ltd. (t/a Skillstrain and/or Train2game) v Designtechnica Corp (t/a Digital Trends) & Ors* [2009] EWHC 1765 (QB) *per* Eady J at [50]-[51].

[195] *Designtechnica* at [53].

[196] Hervey & Lavy (eds), *The Law of Artificial Intelligence* 2nd Edn (London: Sweet & Maxwell, 2024), para.7-046.

[197] *Designtechnica* at [54].

[198] In the context of "snippets", it was held (in 2009): "*One cannot merely press a button to ensure that the offending words will never reappear on a Google search snippet; there is no control over the search terms typed in by future users*" (*Designtechnica* at [55]). The Court's consideration of modern processes. Particular formalities are not required for notification: in *Monir v Wood* [2018] EWHC 3525 (QB) *per* Nicklin J at [192], the Court rejected a submission that nothing short of written notification should suffice to sustain the inference of authorisation from continued publication following notification. However, it would be powerful evidence.

[199] *Koutsogiannis v The Random House Group Ltd* [2019] EWHC 48 (QB) *per* Nicklin J at [11].

[200] *Koutsogiannis* at [12](ii).

[201] *Koutsogiannis* at [12](ix).

[202] See Peter Wills, *Libel via Language Models*, (2025) 62 Osgoode Hall L.J. 153, pp. 159-67.

[203] *Koutsogiannis* at [12](viii).

[204] Grave allegations from sources that lack credibility may fall foul of Defamation Act 1996, s.1(1): *Lachaux v Independent Print Ltd* [2016] Q.B. 402 *per* Warby J at [59]; [2020] A.C. 612 *per* Lord Sumption at [16] & [20].

[205] See Hervey & Lavy (eds), *The Law of Artificial Intelligence* 2nd Edn (London: Sweet & Maxwell, 2024), paras.7-049 to 7-051.

[206] This is non-exhaustive. There are others, including qualified privilege and those specific to the internet (see the Electronic Commerce (EC Directive) Regulations 2002 and Defamation Act 2013, s. 5).

[207] Defamation Act 2013, s. 3: "*(1) It is a defence to an action for defamation for the defendant to show that the following conditions are met. (2) The first condition is that the statement complained of was a statement of opinion*."

[208] Defamation Act 2013, s. 3(5): "*The defence is defeated if the claimant shows that the defendant did not hold the opinion*."

[209] Defamation Act 2013, s. 3(6): "*Subsection (5) does not apply in a case where the statement complained of was published by the defendant but made by another person ("the author"); and in such a case the defence is defeated if the claimant shows that the defendant knew or ought to have known that the author did not hold the opinion*."

[210] Defamation Act 2013, s. 4: "*(1) It is a defence to an action for defamation for the defendant to show that—(a) the statement complained of was, or formed part of, a statement on a matter of public interest; and (b) the defendant reasonably believed that publishing the statement complained of was in the public interest*."

[211] *Clerk & Lindsell on Torts*, 24th Edn (London: Sweet & Maxwell, 2025), para.21-98.

[212] *Clerk & Lindsell on Torts*, 24th Edn (London: Sweet & Maxwell, 2025), para.21-206. For the classic definition of malice, see *Horrocks v Lowe* [1975] A.C. 135, per Lord Diplock at 149-150: "*The motive with which a person published defamatory matter can only be inferred from what he did or said or knew. If it be proved that he did not believe that what he published was true this is generally conclusive evidence of express malice, for no sense of duty or desire to protect his own legitimate interests can justify a man in telling deliberate and injurious falsehoods about another, save in the exceptional case where a person may be under a duty to pass on, without endorsing, defamatory reports made by some other person*."