Shine's Submission to the Department of Industry, Science and Resources

- ## Supplementing the Document by *Campaign for AI Safety*

Inquiry into Supporting Responsible AI

26 July 2023

Table of Content

Technology Strategy Branch
Department of Industry, Science and Resources
GPO Box 2013,
Canberra, ACT, 2601
Email: DigitalEconomy@industry.gov.au

By: Online

Supporting responsible AI:

Submission by Shine Lawyers

A. Introduction

1. Defects in software causing bad outcomes, whether caused by intrinsic or extrinsic reasons in relation to the software, have the potential to create large-scale real-world impacts, especially software that monitor and control physical industrial systems.

2. As part of its investigation into cyberattacks on India's electrical grid, Microsoft researchers reported a vulnerable open-source component in the Boa web server.[1] The researchers found in 2022 that, besides being used by India's power company, the server is still widely deployed in a range of routers and security cameras, as well as popular software development kits (**SDKs**). This despite the software's retirement in 2005.

3. As software technology increase in sophistication, and into the realms of artificial intelligence (**AI**), those inherent risks on the real-world associated with the use of software nevertheless remain.

4. In Canada, Toronto's use of AI predictive modelling which had replaced existing methodology as the only determiner of beach water quality had raised concerns about its accuracy, after allegedly conflicting results were found by a local water advocacy group using traditional means.[2]

5. With the ubiquitous nature of software in modern life, and the increasing pace of AI adoption, this submission discusses some of the liability issues relevant to the use of AI technology. Specifically, it looks at AI software makers' (**AI-makers'**) duty and liability towards both the end-user and non-contracting third party under Australia's present legal framework and explores options for legislative reform.

---

[1] Vulnerable SDK components lead to supply chain risks in IoT and OT environments | Microsoft Security Blog
[2] Incident 290: False Negatives for Water Quality-Associated Beach Closures (incidentdatabase.ai)

6.   This paper supplements the submission by *Campaign for AI Safety*. Specifically, this paper answers the consultation question under the heading "Potential Gaps in Approaches":

> *What potential risks from AI are not covered by Australia's existing regulatory approaches? Do you have suggestions for possible regulatory action to mitigate these risks?*

B. Background

7.   There is presently no universally adopted definition nor classification of what is AI.

8.   The diverse types of AI technology have variously been classified according to certain aspects of the technology. For example, it can be classified based on their evolutionary progress (to date and anticipated in the future):

a.   "**Reactive**" machines that have no memory and are task specific: meaning that an input always delivers the same output.

b.   "**Limited Memory**" machines that can investigate the past and monitor specific objects or situations over time: actions are based on both past and present moment data.

c.   It is anticipated that future categories will include "**Theory of Mind**" and "**Self-Aware**" machines: machines capable of thinking and acting for itself, just as higher beings do.

9.   Another way of classifying AI is based on their algorithms. Here, the three major categories are:

(1) supervised learning,

(2) unsupervised learning, and

(3) reinforcement learning.

10.   These algorithm categories differ in how the machines are trained and how they function.

(1) In **supervised learning algorithms**, the machine takes in clearly labelled data[3] while being trained to use that labelled data to predict outcomes for other data: it may include the use of various statistical formulae and decision trees.

---

[3] In general, clearly labelled data are factors, elements, or information that skilled humans have deemed relevant.

(2) In **unsupervised learning algorithms**, the machines are given data that are not labelled.[4] It then uses that unlabelled data to create models and evaluate the relationships between different unlabelled data points in order to give more insight to the data: for example, by clustering groups of data with closer relationships nearer to each other and uncovering main themes or patterns.

(3) Finally, the **reinforcement learning algorithms** takes in positive or negative feedback for every action the machine makes based on the stimuli/feedback it receives. Over time, the actions conform to what is deemed to be a preferred action for any particular set of stimuli. This is not dissimilar to the Pavlov's theory of conditioning in dogs and humans.

11.   There is a final way of classifying AI based on its capabilities:

(1) Weak/Narrow;

(2) General; or

(3) Strong/Super.

While this classification is useful for technology experts, it is <u>not</u> relevant for present purpose in relation to legal liability.

12.   To understand the present legal liability framework in relation to AI, understanding the first two classifications is important as they directly address the issue of legal causation.

13.   Knowing or not knowing what happens within an AI-"box" is fundamental in determining the chain of causative events: from the stimuli received by the AI to the action output or performance the AI makes. This understanding allows for the attribution of legal negligence. This is discussed in greater detail later in this paper.

14.   To begin, this submission will first outline the present legal framework within Contract, Consumer law, and Negligence.

15.   Because of the interaction of the three legal headings and to limit the length of this paper, only the broad legal issues of contractual abrogation of responsibility, unfair contract terms, and strict liability regime of consumer laws are examined in this paper as it relates to AI. Damages are not discussed in this paper in any detail.

---

[4] Unlabelled data means what relevant factors, elements or information are unclear. Each datum appears to be unique and individual. The clustering statistical process groups data that more alike closer together and help to uncover themes or patterns within the data set. For example, hypothetically, when a cohort of individuals' medical records are analysed, the AI may detect that individuals with certain electrocardiogram (ECG) waveforms have Chest X-rays that show similar radio-opacities patterns. The waveforms and patterns were not pre-labelled as it was not known to have any significance. The statistical association may prompt further research.

C. Present Legal Framework

Contract Law

16. Before discussing the various contractual terms and disclaimers found in software agreements, it should be noted at the outset that there are broadly two forms of software in the market: "Free" (including Open-source software, OSS), and "Non-Free" or Proprietary software.

17. OSS software is distributed with the source code disclosed, or open, at the point of distribution.

18. Proprietary software is distributed with no source code disclosed, meaning that anyone who wishes to discover that source code must engage in a difficult and time-consuming process of reverse engineering.

19. Some non-free software may incorporate OSS within it, but this hybrid will not be specifically discussed here in this paper.

I. Denying Duty

20. The types of indemnities, disclaimers and limitations of liabilities may differ between OSS and non-free software agreements.

Free Software

21. As an example, ChatGPT provided by OpenAI provides the following in its agreement:[5]

> **Indemnification; Disclaimer of Warranties; Limitations on Liability**
>
> (a) **Indemnity**. You will defend, indemnify, and hold harmless us, our affiliates, and our personnel, from and against any claims, losses, and expenses (including attorneys' fees) arising from or relating to your use of the Services, including your Content, products or services you develop or offer in connection with the Services, and your breach of these Terms or violation of applicable law.
>
> (b) **Disclaimer**. The services are provided "as is." Except to the extent prohibited by law, we and our affiliates and licensors make no warranties (express, implied, statutory or otherwise) with respect to the services, and disclaim all warranties including but not limited to warranties of merchantability, fitness for a particular purpose, satisfactory quality, non-infringement, and quiet enjoyment, and any warranties arising out of any course of dealing or trade usage. We do not warrant that the services will be uninterrupted, accurate or error free, or that any content will be secure or not lost or altered.
>
> (c) **Limitations of Liability**. Neither we nor any of our affiliates or licensors will be liable for any indirect, incidental, special, consequential or exemplary damages, including damages for loss of profits, goodwill, use, or data or other losses, even if we have been advised of the possibility of such

---

[5] Interestingly, under its mandatory arbitration and dispute resolution clauses, the ChatGPT agreement explicitly states that disputes must be brought on an individual basis only and may not be brought as a plaintiff or class member in any purported class, consolidated, or representative proceeding. Class arbitrations, class actions, private attorney general actions, and consolidation with other arbitrations are not allowed. If for any reason a dispute proceeds in court rather than through arbitration, each party knowingly and irrevocably waives any right to trial by jury in any action, proceeding, or counterclaim. This does not prevent either party from participating in a class-wide settlement of claims.

damages. Our aggregate liability under these terms shall not exceed the greater of the amount you paid for the service that gave rise to the claim during the 12 months before the liability arose or one hundred dollars ($100). The limitations in this section apply only to the maximum extent permitted by applicable law.

22. The broad abrogation of responsibility is understandable given its free/open-source nature. The exclusion may potentially be broader than would be found in the proprietary licences/non-free software for major software developers.

23. The bases upon which contractual end-users or non-contracting third parties can sue a software maker are not unique. Whether liability remains in relation to OSS-makers/AI-makers towards its end-users and/or non-contracting third party despite these disclaimers will be discussed later.

24. For now, it should be noted that while the statutory guarantees contained within consumer laws present a potential problem for commercial software developers trying to exclude their liability for their OSS, they do not necessarily assist users of the free software, nor offer protection to other developers that receive free software from a third party and then modify it as a part of a further release. This is because the statutory guarantees may not apply, or will be ineffective from a practical perspective, if: [6]

    (1) the initial developer licensing the software is not subject to the Australian legislation;

    (2) an exception to the guarantees applies — including where a licensee re-supplies the free software in a modified version to its own licensees; or

    (3) the original developer of the free software cannot be found or has insufficient assets or insurance to recover damages from.

25. This means that, despite any statutory guarantees, commercial software developers and business users alike will usually be the ones that bear the liability of the OSS in terms of its performance.

26. Further, there remains the usual difficulties in pursuing entities based purely outside the Australian jurisdiction or those financially unable to meet any compensation ordered by the courts.

Proprietary Software

27. In contrast to the OSS disclaimers, by way of examples, non-free software from international entities Adobe, Apple and Microsoft currently includes the following for its (non-AI) software:

> **Adobe end-user license agreement (EULA)**
> 13.2 Adobe specifically disclaims all liability for any actions resulting from your use of the Software. You may use and access the Software at your own discretion and risk, and you are solely responsible

---

[6] Mathew Baldwin. *Open source software — what to look out for* (2011) 14(6) Internet Law Bulletin 150

for any damage to your computer system or loss of data that results from the use of and access to any Software.

14. Limitation of Liability.

14.1 Adobe, its affiliates, suppliers, and certificate authorities <u>will not be liable to you or anyone else for any special, incidental, indirect, consequential, moral, exemplary, or punitive damages whatsoever, regardless of cause, including losses and damages</u>

(a) resulting from loss of use, data, reputation, revenue or profits,

(b) based on any theory of liability, including breach of contract or warranty, negligence or other tortious action; or

(c) arising out of or in connection with your use of or access to the Software.

Nothing in this Agreement limits or excludes Adobe's liability for gross negligence, intentional misconduct of Adobe or its employees, death or personal injury. Adobe's <u>total aggregate liability in any matter arising out of or related to this Agreement is limited to US $100</u>.

4.2 The foregoing limitations and exclusions apply to the extent permitted by applicable law in your jurisdiction. This limitation of liability may not be valid in some jurisdictions. You may have rights that cannot be waived under consumer protection and other laws.

20.3 Limitation of Liability

20.3.1 Except for German or Austrian consumers, Section 14 (Limitation of Liability) still applies. You are advised to take all reasonable measures to avoid and reduce damages, in particular by making a backup copy of the Software and backup copies of your data.

20.3.2 If you obtained the Software in Germany or Austria, and you usually reside in that country, then Section 14 does not apply. Instead, subject to the provisions in Section 20.3.2, Adobe's statutory liability for damages shall be limited as follows:

(a) Adobe shall be liable only up to the amount of damages as <u>typically foreseeable</u> at the time of entering into the license agreement, with respect to damages caused by a <u>slightly negligent</u> breach of a material contractual obligation; and

(b) Adobe shall not be liable for damages caused by a slightly negligent breach of a non-material contractual obligation.

20.3.3 The aforesaid limitation of liability shall not apply to any mandatory statutory liability, in particular to liability under the German Product Liability Act, liability for assuming a specific guarantee, or liability for culpably caused personal injuries.

Apple (Applications)

The services and content accessed through the app is at a customer's sole risk (specific disclaimers of liability are set out for third party content, financial information, location data and stock information);

- the customer must not infringe any intellectual property rights in the app, or use the app in a manner that infringes any other person's rights;
- <u>the app is provided "as is", and to the extent permitted by law, no warranties are given with respect to the app;</u>
- the app developer may modify, remove or disable access to any service at any time and without liability; and
- <u>all liability for personal injury and numerous other broad heads of damages (including loss of data, loss of profits, business interruption or any other commercial damages or losses) are excluded to the maximum extent permitted by law, and that liability of the app developer for all claims is limited to $50.</u>

Microsoft end-user license agreement (EULA)

11. Exclusion of incidental, consequential and certain other damages.

To the maximum extent permitted by applicable law, in no event shall Microsoft or its suppliers be liable for any <u>special, incidental, indirect, or consequential damages whatsoever (including, but not limited to, damages for loss of profits or confidential or other information, for business interruption, for personal injury, for loss of privacy, for failure to meet any duty including of good faith or of reasonable care, for negligence, and for any other pecuniary or other loss whatsoever) arising out of or in any way</u> related to the use of or inability to use the software product, the provision of or

failure to provide support services, or otherwise under or in connection with any provision of this EULA, even in the event of the fault, tort (including negligence), strict liability, breach of contract or breach of warranty of Microsoft or any supplier, and even if Microsoft or any supplier has been advised of the possibility of such damages.

12. Limitation of liability and remedies.

Notwithstanding any damages that you might incur for any reason whatsoever (including, without limitation, all damages referenced above and all direct or general damages), the entire liability of Microsoft and any of its suppliers under any provision of this EULA and your exclusive remedy for all of the foregoing (except for any remedy of repair or replacement elected by Microsoft with respect to any breach of the limited warranty) <u>shall be limited to the greater of the amount actually paid by you for the software or US. $5.00</u>. The foregoing limitations, exclusions and disclaimers (including sections 7, 8, and 9 above) shall apply to the maximum extent permitted by applicable law, even if any remedy fails its essential purpose.

28. The disclaimers of these non-free software have largely focused on the contracting end-user of the software, although the language have attempted to apply the terms to non-contracting third parties also.

29. The issue of a software maker's liability to a non-contracting third party will be explored later.

II.   Exclusion Terms

30. Australian courts will give an exclusion clause in a contract its natural and ordinary meaning, but if the drafting of any of the software agreement terms are ambiguous, an exclusion clause may be construed strictly *contra proferentum*, that is, against the party who put the clause forward as a defence to an action for breach and for whose benefit the clause operates.[7]

31. Any loose wording of exclusion clauses found in software licences may not be effective in limiting liability for negligence and consequential damages.

32. The courts will look at the provision as a whole and, if the exclusion attempts to limit liability for the very purpose of the contract, it will need to be clearly and unambiguously drafted to survive challenge.[8]

33. However, where the parties are businesspersons, a court will seek to construe the documents broadly, resolving any ambiguity or uncertainty of expression to give a sensible and business-like meaning to the contract having regard to the commercial objective of the transaction.[9]

---

[7] For examples, *Davis v Pearce Parking Station Pty Ltd* (1954) 91 CLR 642 at 649 (liability for negligence - exempting clause must be construed strictly, and that clear words are necessary to exclude liability for negligence); *McRae v Commonwealth Disposals Commission* (1951) 84 CLR 377 at 398 per Dixon and Fullagar JJ (McTiernan J concurring) (exclusion of liability for breach of warranty did not apply to breach of condition).

[8] Peter James "Open Source Software: An Australian Perspective" in B Fitzgerald and G Bassett (eds) *Legal Issues Relating to Free and Open Source Software* (2003) at 80.

[9] *Australian Broadcasting Commission v Australasian Performing Right Assn Ltd* (1973) 129 CLR 99 at 109 per Gibbs J (the court should construe commercial contracts fairly and broadly, without being too astute or subtle in

34. The business-like approach may sometimes also be extended to particular type of contracts, not just between businesspersons, because it is necessary to give that contract 'business efficacy'.[10]

35. This may be relevant when comparing agreements for an OSS, in contrast to non-free software: A court may potentially hold a software company less accountable for coding errors resulting in performance issues in an OSS, and thus more likely to accept (in the context of negligence) exclusion of liability in that regard, as compared to a non-free software.[11]

III. Unfair Terms

36. Australia has laws relating to Unfair Contract Terms (**UCT**)[12] that apply to all standard form contracts, such as the software agreement examples discussed above.

37. Under the legislations,[13] an UCT is one that meets all three of the following UCT threshold criteria:

(1) would cause a substantial[14] disadvantage to the customer (or advantage to the software maker) because it:

   a. is weighted significantly in favour of the software maker;[15] or

   b. grants to the software maker a beneficial option or discretion or power, or imposes on the customer a disadvantageous burden or risk or duty;[16]

(2) is not reasonably necessary to protect the legitimate interests of the software maker, with which the onus of proof lies.[17]

---

finding defects); *Cohen & Co v Ockerby & Co Ltd* (1917) 24 CLR 288 at 300 per Isaacs J (a mercantile contract, which is usually expressed shortly, and leaves much to be understood, ought to be construed fairly and liberally for the purpose of carrying out the object of the parties. That does not mean you are to stretch its terms in favour of one party against the other; but, it means that the expressions, and particularly any elliptical expressions, in a mercantile contract are to be read in no narrow spirit of construction, but as the Court would suppose two honest business men would understand the words they have actually used with reference to their subject matter and the surrounding circumstances).

[10] *Wellington v Huaxin Energy (Aust) Pty Ltd (formerly Cuesta Coal Ltd)* [2020] QCA 114 at [71]; *Bensons Property Group Pty Ltd v Key Infrastructure Australia Pty Ltd [2021] VSCA 69* at [107]–[108]. *Hughes Aircraft Systems International v Airservices Australia (1997) 146 ALR 1* at 38; *Vodafone Pacific Ltd v Mobile Innovations Ltd* [2004] NSWCA 15 at [190]; *Concrete Pty Ltd v Parramatta Design & Developments Pty Ltd* (2006) 229 CLR 577 at [59].

[11] *Australian Broadcasting Commission v Australasian Performing Right Assn Ltd* (1973) 129 CLR 99 at 110 per Gibbs J (the business-like approach to interpreting contracts is not limited to agreements drawn by businessmen for themselves and without legal assistance).

[12] Under both the *Competition and Consumer Act 2010* (Cth) and the *Australian Securities and Investment Commission Act 2001* (Cth).

[13] Ibid.

[14] *Jetstar Airways Pty Ltd v Free* [2008] VSC 539, [105] (Cavanough J).

[15] *Australian Competition and Consumer Commission v CLA Trading Pty Ltd* [2016] FCA 377, [54].

[16] *ACCC v ACN 117 372 915 Pty Ltd (in liq) (formerly Advanced Medical Institute Pty Ltd)* [2015] FCA 368, [950].

[17] ACL s 24(4); ASIC Act s 12BG(4).

What is reasonably necessary for a software maker will, amongst other potential factors, depend on:

    i.   its own facts and commercial circumstance;

    ii.   the nature of the particular business of the software maker and the context of the contract as a whole;[18]

    iii.   the proportionality of the term or condition against the potential loss sufferable;[19]

    iv.   whether other options that might be available to the software maker in relation to protecting its business interests are available;[20] and

(3) would cause detriment (whether financial or otherwise) to the software maker's customer if it is allowed to remain in effect.

38. It remains to be tested in court whether the various terms within software agreements excluding liability against its users or non-contracting third parties would be deemed UCTs.

39. It also remains to be seen if contractual terms that deny claimants the option to participate in a representative action is legally enforceable.[21]

Consumer Law

40. In addition to the UCT laws found within the *Competition and Consumer Act 2010* (Cth) discussed above, the Australian Consumer Law (**ACL**)[22] imposes provisions that require that any entity engaged in a commercial or a related supply of software to consumers meet certain minimum levels of quality upon any software it provides, as well as regulating the way that software is represented.

---

[18] *Poole v Australian Pacific Touring Pty Ltd* [2017] FCA 424; *Australian Competition and Consumer Commission v Ashley & Martin Pty Ltd* [2019] FCA 1436, [48] - [49], [51], [53].

[19] *Australian Competition and Consumer Commission v Ashley & Martin Pty Ltd* [2019] FCA 1436, [55].

[20] *Australian Securities and Investments Commission v Kobelt* [2019] HCA 18, [98], [264], cited in *Australian Competition and Consumer Commission v Ashley & Martin Pty Ltd* [2019] FCA 1436, [54].

[21] See class actions exemption for ChatGPT in footnote 5. In relation to the judicial determination of such exemptions, the appeal of *Carnival plc v Karpik (The Ruby Princess)* [2022] FCAFC 149 will be heard by the High Court in September 2023 where ground 2 of the appeal relates to class action waiver clauses under s 23 of the Australian Consumer Law and/or contrary to Pt IVA of the *Federal Court of Australia Act 1976* (Cth). The question for special leave seeks the High Court's determination on whether a foreign corporation that chooses to carry on business in Australia and enter into contracts of adhesion with consumers (wherever resident) for the supply of services in and from Australia, can rely on a class action waiver clause and an exclusive jurisdiction clause to deny such consumers access to the Federal Court of Australia under Pt IVA of the FCAA, or at all.

[22] Schedule 2 of the *Competition and Consumer Act 2010* (Cth).

41. Under the ACL, computer software is explicitly considered "goods"[23] and the software maker guarantees that the:

(1) software will be of acceptable quality;[24]

(2) software is fit for a particular purpose;[25]

(3) software will match their description;[26]

(4) software will match the sample or demonstration model;[27] and

(5) the supplier will honour any express warranties.[28]

42. These implied conditions and warranties cannot be excluded by contract.[29] This makes any exclusion terms in a contract for a consumer, as discussed earlier, void under the ACL.

43. All non-free software would be deemed to have been supplied in trade or commerce and fall under the ACL if such software are normally sold to consumers for personal, domestic or household use.

44. OSS may or may not be deemed to be supplied in trade or commerce in order to be covered by the ACL. For example,

(1) A volunteer individual OSS community software maker that provided the free software to the public at large, outside of any business relationship, will likely attract no liability towards an end-user or non-contracting third party;

(2) An established software company providing a free OSS as part of a suite of other non-free software to their paid customers will in contrast likely have a duty towards an end-user or non-contracting third party under the ACL.

---

[23] Section 2(1) ACL. At present there is no distinction or clarification made in the legislation in relation to software that is installed onto a user's machine (as was traditionally the case) and software that is wholly web-based and utilised solely online via any non-affiliated and commonly available third-party browsers (e.g. Google Chrome, MS Edge, Mac Safari). It therefore remains at least arguable that in the latter example of a web-based non-executable digital content is neither a "service" nor a "goods": see *Gammasonics Institute for Medical Research Pty Ltd v Comrad Medical Systems Pty Ltd* (2010) 77 NSWLR 479 where software was not deemed a good under the old *Sale of Goods Act 1923* (NSW) but under the ACL, in contrast, the decision in *Australian Competition and Consumer Commission (ACCC) v Valve Corp (No 3)* (2016) 337 ALR 647 at [126]-[157] (decision not challenged on appeal in *Valve Corp v Australian Competition and Consumer Commission (ACCC)* (2017) 258 FCR 190 at [83]) held that although the supply of online games might include the supply of services, as the 'core' of the transaction was the supply of computer games by means of software, the transaction as a whole was characterised as the supply of goods. See also discussion on the gaps in the ACL in relation to non-executable digital content does not clearly fit within the categories of goods or services: David Lindsay, Genevieve Wilkinson and Evana Wright *Responding to the challenges of Consumer Internet of Things devices: The case for reforming the Australian consumer guarantees* (2022) 29 CCLJ 226
[24] Section 54 ACL.
[25] Section 61 ACL.
[26] Section 56 ACL.
[27] Section 57 ACL
[28] Section 59 ACL
[29] Section 64 ACL.

45. A business can also be a consumer and be covered by the guarantees under the ACL if the software bought for business use meets at least one of the following conditions:

    (1) it costs less than $100,000 including GST; or

    (2) it the software is commonly bought for personal, domestic or household use.

46. A business will <u>not</u> be covered under the ACL when the software is purchased:[30]

    (1) for resupply;

    (2) for use or transformation in production or manufacturing; or

    (3) for use or transformation to repair or treat other goods.

47. A specialised software or AI technology, not of a kind ordinarily acquired for personal, domestic or household use; or acquired to be used in trade or commerce, in the manufacture or production or repair of something else will <u>not</u> be covered by the ACL. For example, specialised software programmes, including AI, used for physical industrial systems such as electricity grids and territorial water supply quality are excluded from the guarantees of the ACL.

48. It should also be noted that while statutory remedies available under the ACL are more generous than under the law of contract, and applies to both contracting and non-contracting parties, a claimant is <u>not able</u> to claim damages in respect of *misleading or deceptive conduct* if the misled person suffered personal injuries or death.[31]

49. However, claims for personal injury and death *against a manufacturer* by both contracting and non-contracting parties <u>are still available</u> under sections 138-139 of the ACL for goods with "safety defects".

50. A product has a "safety defect" if its safety is not such as persons are generally entitled to expect.[32] This is an objective standard which is based upon what the public at large, rather than what any particular individual, is entitled to expect.[33]

---

[30] Meaning of consumer: s 3(2) ACL.

[31] Under *Competition and Consumer Act 2010* (Cth) ss 137C, 137D, 137E. Note that each state and territory have enacted legislation to apply the ACL as a local law through the: *Fair Trading (Australian Consumer Law) Act 1992* (ACT); *Fair Trading Act 1987* (NSW); *Consumer Affairs and Fair Trading Act* (NT); *Fair Trading Act 1989* (Qld); *Fair Trading Act 1987* (SA); *Australian Consumer Law (Tasmania) Act 2010*; *Australian Consumer Law and Fair Trading Act 2012* (Vic); and *Fair Trading Act 2010* (WA). As a result, there is a uniform set of rules across all jurisdictions in relation to consumer protection and a uniform articulation of acceptable conduct in the commercial sphere. There are, however, exceptions to this uniformity. For example, there is no such restriction in relation to death or personal injury for misleading or deceptive conduct in *Australian Consumer Law (Tasmania) Act 2010* (Tas). However, it is unlikely that liability in relation to software or AI will be based on state or territory jurisdiction.

[32] Section 9 ACL

[33] Based on cases decided under the former *Trade Practices Act 1974* (Cth): *Cheong v Wong* (2001) 34 MVR 359; *Morris v Alcon Laboratories (Aust) Pty Ltd* (2003) ASAL ¶55-104 at [22]; *Roots v Trussmaster Pty Ltd* [2003] QSC 348; *Thomas v Southcorp Australia Pty Ltd* [2004] VSC 34; *Laws v GWS Machinery Pty Ltd* (2007) 209 FLR .

51. It is unclear how "safety defect" will be construed in relation to any personal injury and death caused by a software or AI performance, however, it appears at least arguable that the fact that a plaintiff cannot identify the safety defect with any particular level of precision would still require the manufacturer to prove its non-existence.[34] In other words, once the allegation is made that there is a safety defect in the software/AI and shown to have resulted in the harm, the evidentiary burden then shifts to the software manufacturer to prove that it was _not_ a safety defect that caused the personal injury and death. How precise a plaintiff must be with regards to identifying the safety defects will likely be determined on a case-by-case basis. The fact that it may be more difficult for a defendant to establish a defence if the safety defect is not or cannot be identified with precision may not matter under the ACL.[35]

52. A court will also not consider whether it was likely that a software failed due to a safety defect, but rather whether it was more likely that the failure was due to the safety defect than some other cause, such as improper use.[36] This means that any statistical or historical evidence of high safety (or low defect) rate provided by the manufacturer will be irrelevant if it cannot prove that a safety defect did not cause the specific incident of personal injury or death. The manufacturer will need to prove the cause of the personal injury or death by proving, for example, improper use by the user.

53. It is therefore possible that should a software or AI cause personal injury or death as a result of its performance, that a court may deem it to be a safety defect under the ACL unless the software or AI-maker proves otherwise. We do note that this view is far from settled and may be subject to contested litigation should this issue arise.

Negligence

54. Should a claimant not have recourse to the ACL guarantees and remedies, it is still possible for them to rely on other causes of action. This may occur, for example, in a non-consumer context such as where a specialised software or AI technology is not of a kind ordinarily acquired for personal, domestic or household use; or where a software or AI was acquired to be used in trade or commerce, in the manufacture or production or repair of something else.

55. Since the common law case of _Donoghue v Stevenson_,[37] non-contracting third parties affected by a product have a right against the product manufacturer. This is called the

---

[34] Based on cases decided under the former _Trade Practices Act 1974_ (Cth): _Cheong v Wong_ (2001) 34 MVR 359; _Batchelder v Holden Ltd_ [2009] VSC 29; _Effem Foods Ltd v Nicholls_ (2004) ATPR ¶42-034; _Saint-Gobain Abrasives Pty Ltd v McPherson_ [2009] NSWCA 214.

[35] Based on the case decided under the former _Trade Practices Act 1974_ (Cth): _Batchelder v Holden Ltd_ [2009] VSC 29 at [14].

[36] Based on case decided under the former _Trade Practices Act 1974_ (Cth): _Saint-Gobain Abrasives Pty Ltd v McPherson_ [2009] NSWCA 214.

[37] [1932] AC 562.

neighbour principle where the law attaches a duty of care to the situation where one party has a responsibility to ensure that other persons that may be affected by their actions are safe from harm or loss.

56. This right of non-contracting parties has since been codified under the various Civil Liability Acts enacted in the Australian states and territories.[38]

57. Alternatively, in appropriate circumstances, the *Hedley Byrne & Co. Ltd v. Heller & Partners Ltd*[39] "special relationship" principle, supplemented, again in appropriate circumstances, by relevant consumer protection legislation may also apply. This case established that it may be possible to make a claim in negligence for pure economic loss where there is a special relationship assuming responsibility between two parties, despite them not being in a contract. This is not discussed in any further detail in this paper.

58. It should be noted that under both the common law and the standard of care pursuant to the Civil Liability Acts, strict liability does not attach to the software manufacturer. This means that liability for personal injury, loss of property or financial or economic loss will be avoided, for example, if established procedures were followed during programming or coding.[40]

59. Furthermore, assuming that there is a duty on (non-consumer) software and AI makers to ensure that their products are safe against the harm caused to the third-party non-user plaintiffs, and that a breach of that duty had occurred, there is still the onus on the harmed plaintiff to prove, on the balance of probabilities, that the actions or inactions of the defendant software/AI-maker resulted in harm or loss to the plaintiff. Here, knowing what happens within the software or AI-"box" is fundamental in determining the chain of causative events.

60. For AI-boxes that are "Reactive" or "Limited Memory" machines, or which rely on "supervised learning" using clearly labelled data (collectively, "**Transparent-boxes**"), causation is likely to be straightforward. The outcome of these machines' actions can be worked backwards to determine what happened. If it is determined that the coding by the software was clearly in error, for example, simplistically, one plus one was coded to equal three, then the chain of causation is sheeted home to the software/AI-maker.

61. However, if the coding by the software was accurate and it was the input stimuli data provided by the end-user of the software/AI machine that caused the performance that caused harm (i.e. improper use), then the faulty input stimuli data had broken the chain of

---

[38] QLD: *Civil Liability Act 2003*; NSW: *Civil Liability Act 2002*; Victoria: *Wrongs Act 1958*; WA: *Civil Liability Act 2002*; SA: *Civil Liability Act 1936*; Tasmania: *Civil Liability Act 2002*; NT: *Personal Injuries (Liabilities and Damages) Act 2003*; ACT: *Civil Law (Wrongs) Act 2002.*

[39] [1964] AC 465.

[40] Note that it is also an absolute defence to liability under the ACL if the product had a "safety defect"' due to a manufacturer's compliance with a mandatory standard: s 142(b). This may become relevant if a mandatory standard for software/AI coding becomes relevant for software/AI-makers.

causation, rendering the software/AI-maker non-liable. This is since the software/AI machine performed as it was designed or coded to do.

62. In contrast, for what is anticipated to occur for the future, "Theory of Mind" and "Self-Aware" machines (where these machines are capable of thinking and acting for itself); or for machines with algorithms of unsupervised or reinforcement learning (such that the input stimuli are not clearly labelled, or where which past stimuli were incorporated or rejected in the future actions may not always be known), such machines (collectively, **"Black boxes"**) present greater difficulty under the present negligence legal framework.

63. This is because the black box thwarts any attempts to examine if a golden thread ties the input stimuli data to the ultimate performance emanating from the black box that caused the harm. It is for this type of black-box AI systems that should be the focus of legislative reforms.

D. Proposed Legislative Reform Overseas

64. The present focus of legislative reform by governments overseas, in relation to liability, have largely been on transparent boxes.

65. The spotlight has also been largely in relation to cybersecurity rather than harm or consequential harm from the software operating normally, albeit opaquely.

66. For example, in the US, the Biden administration has recently proposed a strategy that opens up software makers to potential lawsuits for software that are poorly developed to withstand cyberattacks and shifts the blame away from the end-users when a system is breached.[41]

67. The European Union (**EU**) is similarly currently debating the Cyber Resilience Act (**CRA**) which potentially hold software makers, even for free software, liable.[42]

68. This paper does not concern itself with cybersecurity or actions of malicious actors on a software. Rather, it focuses on software/AI makers' liability for their software or AI not acting the way it was advertised or supposed to do, in other words, a focus on non-cybersecurity related AI-harms.

69. For non-cybersecurity related AI-harms, the EU is currently considering a three-part regulatory system[43] for preventing harm and regulating AI in the EU through the

---

[41] Biden Administration, National Cybersecurity Strategy, March 2023: National-Cybersecurity-Strategy-2023.pdf (whitehouse.gov); see pinpoint Strategic Objectives 3.3, p 20-21.
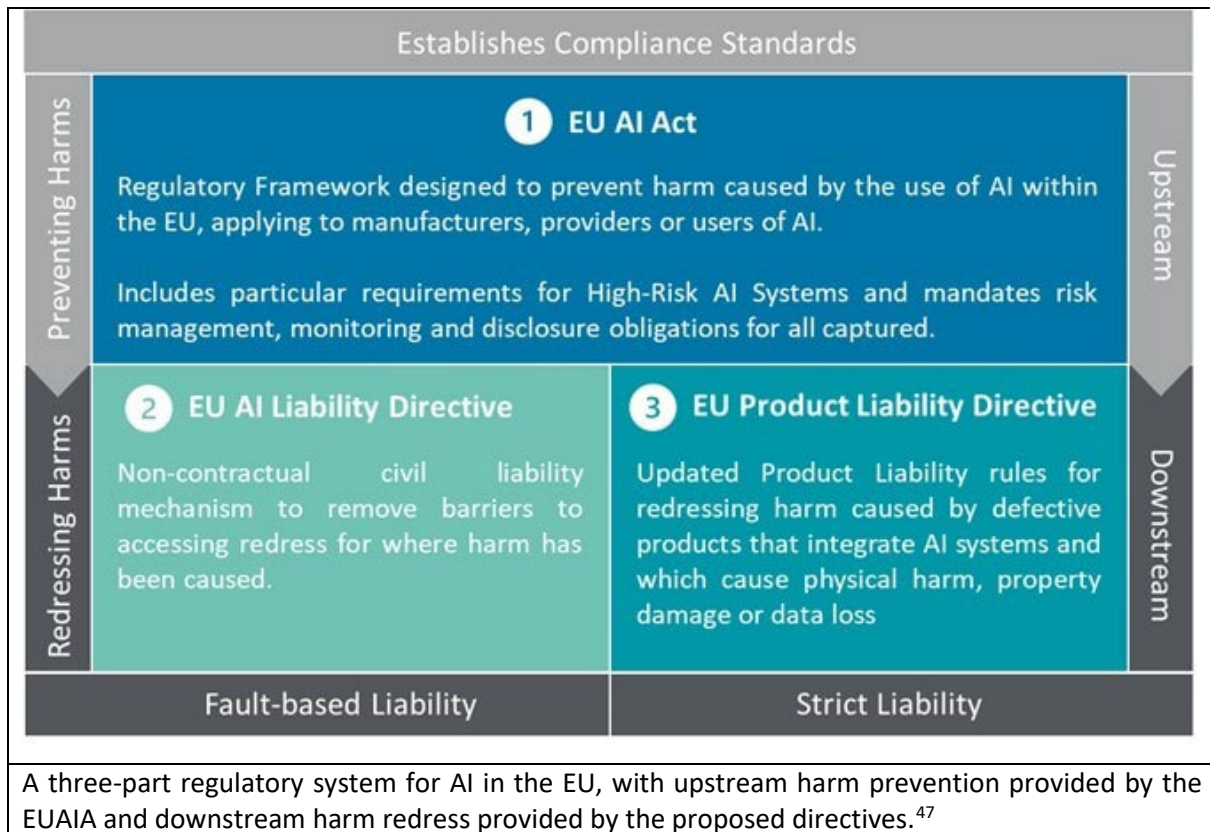
[42] EU cyber-resilience act (europa.eu)

[43] Every action taken by the EU is founded on the treaties. These binding agreements between EU member countries set out EU objectives, rules for EU institutions, how decisions are made and the relationship between the EU and its members. "Regulations" are legal acts that apply automatically and uniformly to all EU countries as soon as they enter into force, without needing to be transposed into national law. They are binding in their entirety on all EU countries. A "directive" is a legislative act that sets out a goal that all EU countries must achieve. However, it is up to the individual countries to devise their own laws on how to reach these goals.

(1) EU AI Act (EUAIA),[44]

(2) AI Liability Directive (AILD),[45] and

(3) Updated Product Liability Directive (UPLD).[46]

70.   The interaction of the three regulations is represented in the figure below.

A three-part regulatory system for AI in the EU, with upstream harm prevention provided by the EUAIA and downstream harm redress provided by the proposed directives.[47]

EUAIA

71. The proposed EUAIA assigns applications and systems of AI to three risk categories:

(1) **Unacceptable risk:**

- E.g. government-run social scoring of the type used in China.

- Banned.

(2) **High-risk:**

- E.g. resumé-scanning tool that ranks job applicants.

- Subject to specific legal requirements.

---

[44] The Act | The Artificial Intelligence Act and Annexures: Annexes | The Artificial Intelligence Act. The EUAIA is a Regulation.

[45] 1_1_197605_prop_dir_ai_en.pdf (europa.eu)

[46] New Product Liability Directive (europa.eu)

[47] Cited from AI: updated EU liability legislation | Deloitte UK

(3) **Not explicitly banned or listed as high-risk**:

- largely left unregulated.

72. Article 24 of the proposed EUAIA explicitly places a duty (and therefore liability) on software/AI-makers of high-risk applications.

AILD

73. The proposed AILD adapts existing non-contractual EU civil liability rules to AI.

74. It recognised that when an injured person seeks compensation for damage suffered, general fault-based liability rules usually require that person to prove a negligent or intentionally damaging act or omission ("fault") by the person potentially liable for that damage, as well as a causal link between that fault and the relevant damage.

75. However, when AI is interposed between the act or omission of a person and the damage, the specific characteristics of certain AI systems, such as opacity, autonomous behaviour and complexity (I.e. Black-boxes), may make it excessively difficult, if not impossible, for the injured person to meet this burden of proof.

76. In particular, it may be excessively difficult to prove that a specific input for which the potentially liable person is responsible had caused a specific AI system output/performance that led to the damage at stake.

77. The AILD thus has two elements to address these concerns.

(1) The first element (article 3), the right to access, empowers the courts to order disclosure of information by persons who provide (i.e. the software maker) or use[48] high-risk AI systems.[49] A failure to provide the information will be presumed to mean the defendant did not comply with a relevant duty of care that the evidence was requested in relation to. This presumption would be rebuttable by the defendant demonstrating the duty of care was in fact, satisfied.

(2) The second element (article 4) is a rebuttable presumption of a causal link between a failed duty of care and harm caused by the AI system once a breach of a duty of care is established.

78. The second element of the AILD thus sheets home to the software/AI maker liability, which obviates the requirements to examine what happens inside Black boxes. It is up to the

---

[48] This will apply in relation to a non-contracting third party non-user suing a software/AI maker for the outcome created by the use of their software/AI by a contracting second party.

[49] In common law jurisdictions like Australia, the process of discovery is embedded in our legal rights and privileges. However, under the civil law system in Europe, discovery is not a guaranteed right, and is instead, at the courts discretion.

software/AI-maker to incorporate data logs within their systems to provide the rebutting evidence.

79. Unlike the proposals set out in the AILD, which relate to fault-based claims, the UPLD bring AI systems within the scope of the product liability consumer protection regime by including AI systems and AI-enabled goods and services into the regime, and by making software/AI providers and providers of digital services liable for compensation if the defect causes damage, without the injured person having to prove the manufacturer's fault under this strict liability regime.

80. The proposed amendments also extend the nature of damage from medically recognised harm to psychological health and loss or corruption of data.

81. OSS/Free software developed or supplied outside the course of commercial activity (as well as the source code of those software) are excluded from the definition of products covered under this EU proposal.

E. Recommendations for Australia

82. Although the current laws appear to provide some protections and legal recourse against harmed caused by AI, the laws appear inconsistent across consumer and non-consumer use and give rise to real difficulties regarding causation in respect of the way in which AI works through its black boxes. As such, there are enough uncertainties in the current framework that warrant legislative reform to ensure sufficient protections.

83. It is therefore recommended that all AI developers have a presumed duty to its end-users and non-contracting third parties. This duty can be rebuttable if the harm caused is too remote to have been foreseeable. This then provides to the general public (i.e. non-contracting third parties) better legal protection in the event a software or AI product causes them harm.

84. However, to prevent the situation of requiring software to be error-free (especially for OSS), which is unrealistic and impractical for software or AI-makers (and thus stifle development), the harm under negligence laws, as amended, should be restricted only to those that caused significant harm. What is significant harm would require further deliberation that is outside the scope of this paper.[50]

---

[50] For example, section 87S of the *Competition and Consumer Act 2010* (Cth) prescribes that the court must not award personal injury damages for non-economic loss where the loss the plaintiff suffers is less than 15% of a most extreme case. in Victoria, *Wrongs Act 1958* (Vic) s 28LB, injury threshold requirements stipulate that in personal

85. Such deliberation should consider explicitly including pure mental harm, such as embarrassment and mental distress, short of a medically recognised condition as harm for which a person is able to seek compensation.

86. The deliberation should also consider expressly including pure economic loss. This given the use of the technology in investment, shares and derivative trading will likely gain currency.

87. A further recommendation is that, similar to what was recently proposed in the EU, there should be a rebuttable presumption against an AI developer, of a causal link between a failed duty of care and harm caused by the AI system once a breach of a duty of care is established. This overcomes the difficulties posed by the uncertainties regarding the workings of black boxes and the lack of transparency. It also brings AI products more in line with consumer protections.

88. By placing the burden on software/AI-developers to rebut the presumption, the manufacturers will be encouraged to incorporate some kind of data log or safety systems that can track or retrace the steps of how a software or AI had come to enact a performance that caused the harm alleged.

89. Finally, it is a recommendation that any proposed AI-related laws should specifically state that any terms that exclude liability within user-agreements for AI-related software, or which excludes an individual's right to participate in a class actions, are to be deemed unfair and voided.

90. The explicit incorporation of such exclusion terms in user agreements within laws relating to UCT for AI will complement those currently found within the *Competition and Consumer Act 2010* and *Australian Securities and Investments Commission Act 2001*. This addresses the issue of an abrogation of duty through contract or making it economically unviable for individuals to seek justice given the costs associated with bringing about individual legal action may be prohibitive, and/or unnecessarily burden the courts with multiple individual actions where there are common issues of law or fact that can more efficiently be addressed through a class actions.

---

injury cases (other than spinal injury or psychiatric), the impairment threshold for claiming general damages requires a degree of permanent impairment exceeding 5%.