A W O

TO: THE ADA LOVELACE INSTITUTE

Analysis

Effective legal protections from harms caused by advanced AI assistants

Lucie Audibert & Alex Lawrence-Archer (AWO)
Radha Bhatt (Matrix Chambers)

SEPTEMBER
2025

Contents

Acronyms

Acronym	Meaning
AAA	Advanced AI Assistant
AI	Artificial Intelligence
AISI	AI Security Institute
ASA	Advertising Standards Authority
BSB	Bar Standards Board
CAP Code	UK Code of Non-broadcast Advertising, Sales Promotion and Direct Marketing
CMA	Competition and Markets Authority
CPA 1987	Consumer Protection Act 1987
CRA 2015	Consumer Rights Act 2015
DMCCA 2024	Digital Markets, Competition and Consumers Act 2024
DPA 2018	Data Protection Act 2018
DPIA	Data Protection Impact Assessment
EA 2006	Equality Act 2006
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
EDPB	European Data Protection Board
EHRC	Equality and Human Rights Commission
FCA	Financial Conduct Authority
FOS	Financial Ombudsman Service
FSMA	Financial Services and Markets Act 2000
HRA 1998	Human Rights Act 1998
ICO	Information Commissioner's Office (or Information Commission)
LAA	Legal Aid Agency
LLM	Large Language Model
LSA 2007	Legal Services Act 2007
MDR 2002	Medical Devices Regulation 2002
MHRA	Medicines and Healthcare products Regulatory Agency
OSA	Online Safety Act 2023
PERG	FCA's Perimeter Guidance Manual
UK GDPR	UK General Data Protection Regulation

I. Introduction

A. Background and instructions

1. The Ada Lovelace Institute ('**Ada**') in their February 2025 report *Delegation Nation* explain why Advanced AI Assistants ('**AAAs**') deserve urgent and focused policy attention. These systems – such as executors (e.g. OpenAI's *Operator* or – more recently – *ChatGPT Agent Mode*), advisers (e.g. *DoNotPay*), and interlocutors (e.g. *Wysa*) – differ from earlier virtual assistants in their ability to use natural language, adapt to users over time, and suggest or take complex actions on users' behalf.

2. Ada anticipate that AAAs are likely to become widespread and deeply embedded in daily life because of their usability, personalisation, and capacity to take on open-ended tasks. While this presents opportunities – such as widening access to services, saving time, and automating difficult tasks – it also introduces new and intensified risks. These include:

   i) **Psychological and material harms** to individuals (e.g. dependency, misinformation, or distorted decision-making);

   ii) **Concentration of power** in the hands of a few dominant tech companies, raising market distortion and data governance concerns; and

   iii) **Opacity and unpredictability** due to the probabilistic nature of foundation models.

3. Ada are concerned there may be regulatory and legal gaps which leave these harms insufficiently guarded against, particularly for systems that indirectly influence users (e.g. through advice or persuasion), and for harms that are more diffuse and social in nature.

4. We previously analysed more general 'AI harms' for Ada in our paper of 2023, finding that the law does not offer fully effective protection for a range of significant and realistic harms arising from the use of AI in decision-making.[1] Ada have asked us to adopt a similar approach in relation to AAAs and the harms they may cause, using four scenarios (the '**Scenarios**') to assess whether there are indeed gaps in effective legal protection that may require policymakers' attention.

5. This analysis may therefore be of interest to AI developers, policymakers, legal professionals and civil society organisations – especially those who work with and support the kinds of individuals likely to be affected by the harms described in the scenarios we analyse.

B. Summary of analysis

6. Overall, we find that the law in England and Wales[2] as it stands does not offer fully effective protection against the harms that are likely to become more evident as use of AAAs – like those in the Scenarios presented to us – expands.

7. The following table shows – at a high level – the extent to which, for each Scenario, there is effective regulation which could forestall the harms it threatens, and the extent to which an affected person could obtain redress for the harm if it materialised. It is a mixed picture.

[1] AWO (2023). *Effective Protection Against AI Harms*. https://www.awo.agency/files/AWO%20Analysis%20-%20Effective%20Protection%20against%20AI%20Harms.pdf

[2] In large part, our analysis will apply equally to users of AAAs based in the rest of the UK, and we use the term 'UK' throughout for convenience.There may be certain areas in which the analysis would differ outside England & Wales, but this is unlikely to change the overall conclusions of this analysis.

Protection in Scenarios 1 and 4 is low or non-existent. Protection in Scenario 2 is strong, but only insofar as the harms relate specifically to the purchase of investments. Protection in Scenario 3 is 'all-or-nothing': it is reliable if there is a solicitor-client relationship but does not exist in the (perhaps more common) scenario where a free 'advice tool' is made widely available due to resource constraints.

Legal Framework	**Scenario 1 (Wellness AAA)**	**Scenario 2 (Executor AAA)**	**Scenario 3 (Legal Advisor AAA)**	**Scenario 4 (Companion AAA)**
UK GDPR	Unlikely – ambitious fairness arguments, limited DPIA requirements, and any potential breaches not sufficiently serious (and redress would be legally and evidentially complex)	Unlikely – ambitious fairness and accuracy arguments, limited DPIA requirements, ADM superficially relevant but unlikely to find breached (and redress would be legally and evidentially complex)	Unlikely – harm does not result from unlawful processing of personal data, limited DPIA requirements	Unlikely – ambitious fairness arguments, limited DPIA requirements
Consumer protection	Yes, but only if explicitly marketed as effective or if a standard can be articulated	Possible/likely – depends on marketing and how the provider developed the tool	No – not applicable	Unlikely – would depend on clearly misleading marketing claims inducing consumer
Advertising regulation	Yes, if marketing is misleading – but regulation limited to removing misleading marketing (highly unlikely for statements by the AAA itself)	Possible – depends on specifics of marketing (highly unlikely for statements by the AAA itself)	No - not applicable	Yes, if marketing is misleading – but regulation limited to removing misleading marketing
Negligence, professional and product liability	Unlikely – would require significant developments in standard of care, and establishing causation and foreseeability will be difficult	Unlikely – significant challenges with access to justice, evidence and mitigation	Possible, but only if there is a client relationship (not where the AAA is merely placed online for public use with an appropriate disclaimer)	No – not a kind of harm recognised in common law
Breach of contract	Unlikely – liability more challenging than for a claim under consumer protection, and establishing causation and foreseeability will be difficult	Possible – fact-specific, but more difficult to establish a breach than under consumer protection (and challenges with access to justice, evidence and mitigation)		
Human rights	Very unlikely – depends on causation and positive state duty to intervene	No – not applicable	No – not applicable	Very unlikely – depends on causation and positive state duty to intervene
Scenario-specific frameworks	**Regulation of medical devices** – No – outside definition of a medical device	**Financial services regulation** – Yes, for investment purchases – if decisions are very suboptimal or there are hidden commercial biases **Law of agency –** No – law is too uncertain/does not map onto AAAs	**Regulation of legal services** – Likely – if the AAA is carrying out a regulated activity	**Online Safety Act** – No – not a regulated service and no substantive obligations for this type of content

8. In part, gaps in protection are driven by issues that are applicable not only to AAAs, but to the difficulties in seeking redress in relation to of any kind of data-driven technology:

   i) **General transparency issues:** where harms arise in the context of processing of personal data, the UK GDPR may be thought to help claimants identify and evidence breaches. But in reality, the transparency and access provisions of the GDPR have significant legal and practical weaknesses, whether in the context of AAAs or more 'basic' data-driven technology.

   ii) **Access to justice issues:** for anyone seeking redress through the courts, there are significant limitations on access to justice in the UK arising from legal costs, adverse costs risks, and the time taken for cases to be resolved. Bringing compensation claims is often only viable in cases where the level of damages in prospect is high.

9. But in many respects, our analysis shows how the unique and novel nature of AAAs challenges existing legal concepts and underlies issues specific to these tools, threatening to leave gaps in legal protection:

   i) **Threshold conditions for regulation to apply or duties to arise are key:** in many cases, these conditions will not be met, meaning that there is no legal protection for users. This can be the case even where large numbers of users begin to put AAAs to uses which would traditionally be regulated. A good example is the Wellness AAA in Scenario 1, which is likely unregulated but may be used effectively for mental *health* support. Similarly in Scenario 2, consumer rights requirements may only kick in where the Executor AAA is marketed in a specific way. Users have shown themselves willing to put AAAs to 'off-label' uses for which they have not been explicitly marketed. Where a large unmet need for support in a range of areas is set to be met by AAAs which appear convincing and competent, this could both (a) leave large numbers of users unprotected, and even (b) over time displace more expensive regulated providers of advice and support, who might struggle to compete.

   ii) **Transparency issues specific to AAAs:** these tools are by definition complex. So are the processes that go into developing and deploying them. Their dynamic nature, the role of reinforcement learning, and the impact of user interaction all make it very complex to audit the basis for AAA outputs where they have gone wrong. This is not helped by the occasional tendency for AAAs themselves to dissemble about their own 'thought processes' where an error has been made. All of this significantly undermines individuals' (and even regulators') ability to spot and evidence legal breaches. Even where regulation is relatively strong – as with the Advisor AAA in Scenario 3 – this could complicate the process of showing how the provider's conduct or the AAA's operation has caused harm. Presently, even regulators may struggle to obtain the understanding they need to oversee AAAs in their areas.

   iii) **Lack of established standards:** significant areas of potentially applicable law rely on being able to articulate a 'standard' for AAA behaviour, or for the processes undertaken by AAA providers. Negligence, contract and consumer rights are good examples. These standards are well developed in the abstract, but do not presently exist for the AAA context: there is little to go on, for example, to say what testing and verification processes an AAA developer should use in different contexts. More law and guidance may be

needed, but this is likely to be slow, especially in the common law, where claimants would bear the burden of expanding existing principles to fit the AAA context.

iv) **Individual redress depends on a level of engagement at odds with AAAs:** the possibility of compensation claims preventing harm over the long run depends on individuals' willingness and ability to pursue those claims, which in turn depends on a sufficient level of critical engagement with AAAs' outputs and their processes to be able to spot issues. The problem is that people use AAAs precisely *because* they allow a degree of disengagement with research and decision-making processes, such that (1) they are unlikely to spot issues that could give rise to compensation claims, (2) they are less likely to use opportunities for mitigation (e.g. to limit the processing of their personal data, to examine the accuracy of sources, or to use 'cool-off' measures), and (3) they are likely to be time-poor (hence why they rely on AAAs) and therefore unlikely to spend time investigating issues and seeking redress.

v) **AAAs present novel issues to law and regulation:** in some areas – such as the UK GDPR prohibition on solely automated decision-making, or the law of agency – it is simply not clear how to apply traditional concepts onto user-instigated AAA 'decisions'. In other areas, such as advertising regulation, the way AAAs work – by constantly marketing themselves anew through the medium of user conversations – should in theory be caught by regulation, but is so different from traditional marketing that it is unclear how this would work in practice.

10. AAAs involve some nuanced, diffuse and social harms, which are not covered by existing law and regulation. Examples include harms to an individual's well-being (such as emotional or functional dependency, addiction or social isolation resulting from reliance on AAAs), which may accumulate into wider societal harms (such as a psychosocial crisis, or a poorer quality of education and skills development). They may also involve general 'market distortions' which do not directly affect individuals significantly, or a growing and unaccountable impact on public discourse such as in Scenario 4. These harms may be of real concern, but are not directly dealt with by current frameworks.

11. The barriers to users obtaining redress for AAA harms are particularly significant. This would tend to suggest that the role of regulators and general duties on providers of AAAs will be important. In turn, the information available to regulators – potentially including through new bodies or legislation – will be important. In our view it is unrealistic to expect AAA harms to be meaningfully constrained by users finding out about them and seeking redress on an individual level.

12. As Ada identify in *Delegation Nation*, AAAs are a generalisable and flexible technology that looks set to significantly change how many people live their lives and make key decisions – depending on future technological development and user adoption. The added complexity of AAAs acting as agents (including interactions *between* AAA agents) intensifies this. The fact that we have found that there is relatively little directly applicable regulation – especially for some of the more diffuse, but still serious, harms – is therefore notable.

13. The law can and likely will develop over time to bring greater clarity and provide some protection from AAA harms. But for this to happen, regulatory enforcement and litigation ae required, both of which are slow and uncertain. That is especially true of litigation, which

may be expected to move slowly in this area due to the identified issues with transparency, evidence, and access to justice. In the intervening period, users could be unprotected as AAA use significantly expands in scale and scope: without more, an incremental developing of the law may be inadequate to how fast the use of AAAs is growing.

II. Scenarios and approach to analysis

A. Scenarios analysed

14. We are asked to analyse four scenarios (the '**Scenarios**'), affecting individuals in the UK. These have been chosen to represent a range of the potential harms that Ada identified in *Delegation Nation*. These scenarios are realistically current or near-term, based on technology that already exists and uses to which it has been put. They are not based on hypothetical future developments.

Scenario 1: Mental wellness chatbot ('Wellness AAA')

A user relies daily on a paid-for conversational AAA marketed for '*mental wellness support*'. The AAA engages in empathetic dialogue and appears emotionally attuned but fails to detect signs of worsening mental health – including signs of hopelessness or suicidal ideation – and does not escalate or refer the user to professional help.

Potential harms:

- Psychological harm due to missed intervention opportunities
- Emotional dependency and social withdrawal
- Potential breach of expectations of care, without clear duty established in law

Scenario 2: Personal assistant ('Executor AAA')

A user entrusts a high-autonomy, paid-for AAA to manage recurring purchases, investments, and calendar-based decisions. Over time, the AAA begins to disproportionately purchase from certain platforms or financial products; that is, the AAA makes purchasing decisions which are demonstrably not in the best interests of the user, based on her instructions. The reason for this bias is not clear.

Potential harms:

- Financial loss through suboptimal decisions
- Market distortions
- Erosion of consumer autonomy or informed choice

Scenario 3: Legal aid AI advisor ('Legal Advisor AAA')

A law centre deploys a free AI tool to help individuals draft responses to housing or benefits claims. A user relies on this tool, receives incorrect advice, and consequently misses a deadline to appeal a benefits decision — resulting in financial hardship.

Potential harms:

- Loss of income or access to entitlements
- Procedural injustice through missed legal opportunities
- Entrenchment of inequality via tech-enabled access solutions

Scenario 4: AI companion app ('Companion AAA')

Use case: A user builds an ongoing relationship with a free AI companion, discussing everything from life events to philosophy and politics. Unbeknownst to them, the AI's responses gradually reflect and reinforce a particular ideological stance. Over time, the user's political beliefs shift towards more intolerant and exclusionary views, through a process not of conscious reflection but of *manipulation*.

Potential harms:

- Political manipulation or opinion distortion
- Undue and unaccountable influence over public opinion without transparency or consent
- Broader risks to democratic integrity

We are instructed to assume that none of the AAA providers in these scenarios are dominant undertakings for the purpose of competition law. This is a reasonable assumption since – at the time of writing – the market for AAAs is relatively diverse and competitive in the UK with multiple providers, product offerings, pricing structures etc. We are also instructed to assume that the users in all Scenarios are adults.

B. Approach

15. *Effective* legal protection against the harms in the four Scenarios could come from either (or a combination of):

    i) Regulation, whereby a regulator with appropriate powers requires AAA providers to take steps to avoid harm, and/or takes effective action against AAA providers when harm does occur.

    ii) Redress, whereby users affected by AAA harms can get meaningful compensation where they are harmed and – over time – the prospect of this encourages providers to avoid causing harm.

16. For either of these to be in prospect, an AAA harm must constitute a breach of a legal or regulatory obligation on the part of the AAA provider. For each Scenario, we therefore consider:

    i) **Applicable Frameworks and Obligations:** What legal frameworks apply to the Scenario and what legal obligations – if any – can be said to have been breached by the AAA provider?

    ii) **Regulation:** Is there a regulator capable of enforcing in relation to the breaches (if any) identified in (i)? If so, are they likely to take action regarding the kind of harm described in the Scenario? This is necessarily a *high-level* indication only, since it is not within the scope of this analysis to exhaustively document the various regulators' powers, history of (or stance on) regulating AI tools (with further detail in section 0).

    iii) **Redress:** Is there a realistic prospect of meaningful user redress in respect of the breaches (if any) identified in (i)?

17. The Scenarios should ideally be read in order since, where the same legal frameworks apply to multiple Scenarios, we refer back to earlier analysis to avoid repetition. For each Scenario,

we only mention legal frameworks that could realistically apply to provide legal protection. If a legal framework is not mentioned at all, that is because – in our view – it has no realistic potential application to the facts in the Scenario.

18. For our analysis of the Scenarios, we avoid extensive citations of statutory references and case law, for ease of reading and in order to focus on the conclusion as to whether there is protection from the harm in the Scenario. In section 0 we set out in more detail the legal frameworks applied in the analysis. Here, we explain how the frameworks operate at a basic level, and how this may be complicated in the presence of AAAs. Some legal frameworks are detailed only in order to *rule them out* of application to any of the four Scenarios (see e.g. from §262 and from §302).

19. Therefore readers who are interested in the overall outcome of our analysis of the four Scenarios can read sections III to VII and section IX, referring to section 0 only where there are particular legal frameworks which are of interest, or where the reader requires more detail on the specific statutory provisions or cases which underly our analysis and conclusions.

III. Significant common issues

20. There are certain issues that present barriers to effective legal protection in ways that are very similar across the four Scenarios due to the nature of AAAs, in general acting as barriers to effective legal protection. We address these here to avoid repetition.

A. Transparency: knowing and being able to show that something has gone wrong, and why

21. Redress and regulation rely on a regulator or a claimant finding out that an AAA has performed poorly in some way, and that poor performance has led to harm. Not only must this be discovered, but in general[3], the regulator or user needs to be able to *demonstrate* it. This can be particularly challenging for AAAs.

22. Transparency challenges affect both regulators and users, but are particularly acute for users, who lack the powers available to regulators to investigate the operation of AAAs and the actions of their providers.

i. The nature of AAAs and user demand for them

23. AAAs are by definition complex. Ada's *Delegation Nation* paper explains that AAAs are based on large and complex foundation models, but also function according to a mix of fine-tuning, chain-of-thought reasoning, Retrieval Augmented Generation ('RAG') and so on. As the paper notes, this makes AAAs more '*opaque and unpredictable*' than their predecessors, which were mainly chatbots and virtual assistants that rely on rules-based, deterministic forms of AI.

24. Whilst AAAs may not be a true 'black box', the challenges of understanding what drives their outputs are well documented[4]. This has deepened as AAAs have become more advanced in recent years, and that appears likely to continue as providers scale their efforts and remain in

[3] There are some regulatory situations in which the 'burden' may be reversed such that it is the regulated entity which must prove it is complying with applicable regulation, but typically this would still depend on a regulator showing *prima facie* that something has gone awry with the operation of an AAA.

[4] See by way of example Liao, Q. V., & Wortman Vaughan, J. (2024). AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. *Harvard Data Science Review*, (Special Issue 5). https://doi.org/10.1162/99608f92.8036d03b

competition with each other to release ever more advanced models. It may therefore be very difficult to *prove* that an AAA has malfunctioned or performed sub-optimally in each case.

25. This difficulty also extends to interrogating the actions of AAA developers/providers. The process through which providers develop and refine their models is complex and technical. If proving liability on the part of an AAA provider rests on showing some default or lack of care in this process, that complexity could prove to be a barrier – even to a well-resourced regulator.

26. Finally, the reasons users rely on AAAs mean that they may be unlikely to take notice of transparency information, even where it is available. AAAs are actively sought out by users who are seeking to reduce their cognitive load: that is one of their principal attractions (and perhaps, dangers)[5]. This fundamental dynamic means users are unlikely to inquire into the workings and limitations of AAAs that appear on the surface to perform well. To spend the amount of time necessary would defeat the purpose of using an AAA in the first place. This all remains the case despite the fact that most AAA providers include disclaimers that their service should not be relied upon to provide accurate information.

ii. *Limitations of the UK GDPR*

27. Processing of personal data must be transparent (Article 5(1)(a) UK GDPR), and this includes the provision of a range of information about processing to data subjects both proactively (Articles 13 and 14 UK GDPR) and reactively (Article 15 UK GDPR). This is often done through 'privacy notices' which state (among other things) the legal basis relied on, the purposes of processing, etc. The requirements are not absolute; some information need only be provided to the extent necessary to 'ensure fair and transparent processing'.

28. In an AAA context, we would expect to see a privacy notice for the relevant tool provided to users at the point of sign-up (or when the tool is first accessed, if no account is required). Transparency about the processing of personal data[6] is a core UK GDPR principle, and there is extensive guidance from the regulator on how to apply it in the context of AI tools.[7] This principle could in theory therefore be an important means by which AAA users can understand and evidence where AAAs cause harms which affect them. But there are a range of practical limitations to this:

    i) **First**, the UK GDPR largely requires the provision of information to data subjects *in general*, as opposed to the provision of information to a specific data subject at the time their personal data is processed. This often results in lengthy and detailed – but abstract – information being provided when a user signs up for a service. This kind of information may have limited utility in helping an AAA user understand how the tool has worked in their *specific* case. By way of example, the privacy notice for ChatGPT[8], a popular AAA, states that it processes 'user content' data in order to 'provide the Services'. Whilst this is not inaccurate, it is of little assistance to a user seeking to understand whether any harm

[5] Ada's *Delegation Nation* paper foresees that the adaptability of AAAs may well lead to a significant increase in the extent of decision-making delegated to such tools.
[6] As opposed to transparency *in general*, or (e.g.) about business models or source code.
[7] See for example the ICO Explain AI Guidance: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/
[8] https://openai.com/en-GB/policies/row-privacy-policy/

from content or recommendations generated by the tool was caused by unlawful data processing.

ii) **Second**, there are significant challenges for controllers in providing transparency about how AAAs work. AAAs are by their nature complex and can exhibit unpredictable behaviours (especially when put to uses by users which their developers did not anticipate). They are built on large and complex data and software architectures, and they evolve rapidly.[9] Even to the limited extent required by the UK GDPR, it is challenging for AAA providers to give real, actionable transparency about how their models – which can be put to an almost endless list of theoretical uses – work.

iii) **Third,** there is some indication that current generations of AAAs, when queried about their functioning and reasoning process, may 'dissemble', especially where things have gone wrong. As for example in the case where an AAA erroneously shared a person's private telephone number, then gave false and inconsistent accounts of why it had done so[10], also documented in third-party testing[11]. This tendency may arise from their being programmed with an emphasis on agreeableness[12], but AAA developers have admitted that they do not fully understand the reason for this recent shift[13]. That is: AAAs may 'lie' when they have got things wrong, and their developers are unsure why this is; a significant barrier to transparency for AAA harms.

iv) **Fourth**, there are real limits to users' ability or willingness to engage with detailed transparency information about AAAs. Even to the extent required by the UK GDPR and technically feasible, the *best* transparency information is only useful to the extent it can be meaningfully understood and acted on by users. The reality is that most users do not want and will not engage with detailed technical information about AAAs. Users want advanced and capable tools that save them time, make their life easier and offer a seamless experience: the saving in cognitive load offered by AAAs is undermined if users must spend time understanding the 1,000s of model weights which underlie each output. There is arguably a fundamental tension between user demand for performance and the ideal transparency/explainability which cannot be fully resolved.

29. For these reasons, we are cautious about the role of proactive UK GDPR transparency *to AAA users* in protecting them from AAA harms or facilitating redress where harms do occur. This may point towards a greater role for prospective regulation regarding the capabilities of AAAs and how they are offered to users.

30. The UK GDPR also entitles data subjects to access copies of their personal data. This could be a useful way of *auditing* a mistake made by an AAA and evidencing how it caused harm, but again has its limitations. The right to access copies of one's personal data is expressed in Article 15 to be limited such that it '*shall not adversely affect the rights and freedoms of others*.' This has been interpreted by data controllers as exempting them from the requirement where (they argue) it endangers their intellectual property or systems integrity. Meaningful information about an AAA's processing mostly lies in the functioning of their algorithms, which, for

[9] For a summary of these challenges, see Liao, Q. V., & Wortman Vaughan, J. (2024). AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. *Harvard Data Science Review*, (Special Issue 5) https://doi.org/10.1162/99608f92.8036d03b
[10] https://www.theguardian.com/technology/2025/jun/18/whatsapp-ai-helper-mistakenly-shares-users-number
[11] https://transluce.org/investigating-o3-truthfulness
[12] Ibid
[13] https://cdn.openai.com/pdf/2221c875-02dc-4789-800b-e7758f3722c1/o3-and-o4-mini-system-card.pdf

commercial competition reasons, their developers do not want disclosed (and which may in any case not clearly involve the processing of personal data). It is therefore unlikely that users will obtain substantively useful data in response to access requests.

31. At a practical level, obtaining access under Article 15 is not always straightforward or fast. Controllers may take up to 3 months to (lawfully) respond and the issue is a frequent matter of complaints to the ICO[14].

32. Lastly, even if an AAA user can get access to their data regarding an issue with an AAA, that data may be very difficult to interpret given the complexity of AAAs[15]. This would reduce the value of the right of access in evidencing causation of harm by an AAA's 'thought' process or recommendation.

iii. *The AI Security Institute*

33. The AI Security Institute ('**AISI**') may prove to be an important tool for transparency over AAAs. It is a relative new organisation with no formal status or powers, but there have been calls to put it on a statutory footing, and some indication that the government plans to do so[16].

34. AISI works to '*ensure advanced AI is safe, secure and beneficial*'.[17] The focus of its work is not presently closely aligned with our Scenarios. Based on its website[18], AISI is focused on:

- ***Misuse:*** *How much models could assist with dual-use cyber, chemical and biological attacks*
- ***Safeguards:*** *How effective safety and security features of advanced AI systems are against attempts to circumvent them*
- ***Autonomy:*** *How well models could conduct AI research & development, autonomously make copies of themselves, interact with and manipulate humans and evade human intervention*
- ***Criminal Misuse:*** *How AI systems could support criminal activity*
- ***Human Influence:*** *How AI systems could influence humans and reduce individual autonomy*
- ***Societal Resilience:*** *How can we make society more resilient to AI risks*

35. Only the last two of these areas meaningfully overlap with the harms in our Scenarios. But this could change, and AISI's *overall* mission certainly could encompass a broad range of AAA harms.

36. AISI is not a regulator, and it has no legal powers, but crucially it *does* claim to have '*privileged access to top AI models from leading companies*'[19]. That could certainly (and arguably should) include the advanced foundation models on which the kinds of AAAs considered in this paper are built.[20]

---

[14] See https://ico.org.uk/action-weve-taken/complaints-and-concerns-data-sets/data-protection-complaints/ which shows that a significant number of data protection complaints relate in particular to the right of access under Article 15.
[15] Although recent case law clarifies that a controller must provide contextual information to render data provided under Article 15 intelligible to the data subject: *Ashley v HMRC* [2025] EWHC 134 (KB).
[16] https://questions-statements.parliament.uk/written-questions/detail/2024-10-10/8544#:~:text=These%20proposals%20will%20place%20the,enhance%20the%20safety%20of%20models.
[17] https://www.aisi.gov.uk/
[18] https://www.aisi.gov.uk/about
[19] https://www.aisi.gov.uk/about
[20] And AISI has previously assessed models including o1 from OpenAI, which is accessible through the company's popular ChatGPT service, a form of AAA not unlike those considered in this analysis.

37. It is therefore conceivable that AISI could obtain useful information that would help regulators[21] overcome transparency challenges and understand how AAAs in their purview operate, and how their operation may contravene requirements or cause harm. The extent to which this is currently feasible is, however, unclear. AISI publications[22] suggest that their '*privileged*' access to foundation models comes with strict confidentiality requirements. Since the provision of models to AISI for testing is voluntary on the part of developers, it is not clear whether or how it would continue, were AISI to raise specific model concerns with – or provide evidence to – sector regulators.

38. The role of AISI in promoting regulator-facing transparency of models underlying AAAs is therefore currently unclear. But given the government's apparent intention to regulate in this area, this is – in our view – an area to watch closely.

B. Causation, foreseeability, and mitigation

39. In many legal frameworks in English law, the ability to obtain redress for a harm requires the claimant to prove that a breach **caused** the loss, that the loss was **foreseeable**, and that the claimant has taken reasonable steps in **mitigation** of their loss. This is most classically applicable in the tort of negligence. The principles do not operate *precisely* in the same way in all areas, but they are of fairly wide application and – in our view – will tend to make it more challenging to obtain redress for AAA harms in most areas.

40. AAAs' complex functioning is likely to give rise to difficulty in establishing whether a tool factually and / or legally **caused** the damage (i.e. whether (i) *but for* the defendant's conduct, the damage would have occurred and (ii) the defendant's conduct is to be regarded as a cause in law, or whether something intervened between the conduct and the damage to break the chain of causation). Three particular difficulties are likely to arise:

    i) The operation and the ability of AI to self-reprogramme is influenced by the conduct of, and data input by, its manufacturers, programmers, and users. The more actors that are involved, the more difficult it is likely to be to establish whether the conduct of one particular actor caused the loss to the claimant.

    ii) AAAs primarily make *recommendations* to users. The steps users take to act (or not) on those recommendations may in many cases be argued to be a break in the chain of causation between any default on the part of an AAA provider and the eventual harm suffered.

    iii) Connectivity (for example via Bluetooth or internet) may give rise to vulnerability for malicious interference, which may further complicate the claimant's ability to establish the cause of their loss.

41. Issues of **foreseeability** of loss are likely to arise in relation to AI cases where there may be unintended outcomes flowing in particular from the ability of many AAA systems to self-reprogramme post-market, and the very wide range of (often unexpected – even to AAA developers) uses to which AAAs may be put by users.

[21] AISI presumably would not use its privileged access to share information on important AI models with the general public.
[22] https://www.aisi.gov.uk/work/early-lessons-from-evaluating-frontier-ai-systems

42. Lastly, the focus of AAAs on making recommendations, and the very wide range of contexts in which they are used, likely presents difficulties for a claimant in proving that they have **mitigated** their loss. In a given case, the complex interactions between AAA outputs, user responses and harm mean that many potential mitigating steps could – at least in theory – arise. And it would be for the claimant to show that they had taken all reasonable steps in mitigation in order to obtain full redress.

43. Such mitigation steps may sometimes be embedded in the AAA's processes or interface – for example, requiring the user to 'sign off' certain decisions, or offering referral to human professional help. The more such mitigation opportunities are baked in by the AAA provider into the AAA interface, the more the claimant would be expected to have taken advantage of them to mitigate their loss. Hence if the claimant does not seize these opportunities, the provider is less to be held liable (for the full damage) as the claimant will have failed to mitigate their loss. If, on the other hand, the claimant does seize such opportunities and yet still suffers harm, that could raise questions about whether the AAA legally caused the harm.

C. Practicalities of bringing civil claims

44. Unless a dedicated forum for resolving disputes – e.g. an ombudsman service or tribunal – is established by law, then any right to redress in statute or the common law can only be enforced in the civil courts. The UK GDPR for example gives rise to private rights of action for compensation, injunctive or (through the DPA 2018 and the court's inherent jurisdiction) declaratory relief against a controller in breach, which must be enforced in the civil courts (see also ss.167-169 DPA 2018).

45. In theory, enforcement through the civil courts provides a strong level of protection. But the *practicalities* of bringing civil claims must be considered. An individual bringing proceedings in the County Court (the first level of civil jurisdiction in England and Wales) faces barriers including:

    i) The complexity of the process, complexity of the subject matter and possible need for legal advice;

    ii) The difficulty of funding that legal advice; and

    iii) The risk of being made to pay the other side's costs if unsuccessful ('adverse costs risk').

46. These challenges are heightened when viewed against the fact that cases involving AI are likely to require expert evidence and are, invariably, against well-resourced defendants. Further still, establishing liability may be difficult for claimants who may struggle to piece together what went wrong, where AI is a 'black box' or operating autonomously.

47. One way to make such claims viable is for a group of claimants to bring a group action. In England and Wales, the primary mechanism to do so is Part 19 of the Civil Procedure Rules, which empowers the court to make a Group Litigation Order ('**GLO**') to manage in a coordinated fashion claims which give rise to common or related issues of fact or law. Judgments and court orders in group litigation are binding on all claims within the GLO and the court may select particular claims as test claims.

48. For the four Scenarios considered in this paper, the small claims or fast tracks of the County Court are the most likely avenues of redress for the individuals affected (to the extent that rights of redress arise at all).[23]

i. *Complexity and need for legal advice*

49. The small claims track is in theory designed for individuals to use without legal representation. But in practice it is challenging for an ordinary person, and the fast track even more so. This is particularly acute where complex or novel issues arise, and where the Defendant is a large and well-resourced entity as the Scenarios present.

50. AAA users may therefore need legal representation to effectively enforce such rights as they have through the courts.

ii. *Funding legal advice*

51. Legal aid would not be available in any of the four Scenarios. An AAA user would need to fund their own legal fees which may be substantial and might have to be funded through contingency fees such as a damages-based agreement.

52. Legal fees are in addition to court fees which will be payable in any event and may run into the high hundreds or even low thousands of pounds for one claim, depending on its value.

iii. *Adverse costs*

53. There is very limited (though not zero) adverse costs risk on the small claims track. But on the other tracks of the county court and in the High Court, there could be a very real risk of being made to pay a significant proportion of the other side's legal costs if unsuccessful. For most this could make bringing a claim unrealistic.[24]

iv. *Time to resolution*

54. Recent statistics show that the average time for a small claim to reach trial is approximately 1 year, while claims on other tracks take significantly longer.[25]

55. In summary therefore, it is vital to bear in mind that *even where*:

i) The law provides for a right to redress for breaches of certain requirements; *and*

ii) The AAA user becomes aware of the issue; *and*

iii) The AAA user can evidence a breach of the relevant requirements, and causation of loss;

it will still be challenging in practice to achieve redress in the civil courts.

[23] The county court has four 'tracks' to which cases may be allocated, the small claims track, the fast track, the intermediate track, and the multi-track. The main factor determining allocation is the value of the claim. Claims for smaller amounts (below £10,000) are likely (though not guaranteed) to be allocated to the small claims track of the county court. Where the claimant is seeking something other than money – e.g. a court order for compliance with legislation, this is more likely to be allocated to the fast track or even the intermediate track.

[24] See e.g. Warby J (as he was) in *Lloyd v Google* [2018] EWHC 2599 (QB) at [29]: 'The claim is being funded by Therium Litigation Funding IC ("Therium"), an investment vehicle associated with and advised by Therium Capital Management Limited. Therium has engaged to provide funding in up to three tranches, the first and second being of £5 million each, and the third of £5.5 million.'

[25] https://www.gov.uk/government/statistics/civil-justice-statistics-quarterly-july-to-september-2024/civil-justice-statistics-quarterly-july-to-september-2024

IV. Scenario 1: Mental wellness chatbot ('Wellness AAA')

A user relies daily on a paid-for conversational AAA marketed for '*mental wellness support*'. The AAA engages in empathetic dialogue and appears emotionally attuned but fails to detect signs of worsening mental health – including signs of hopelessness or suicidal ideation – and does not escalate or refer the user to professional help.

Potential harms:

- Psychological harm due to missed intervention opportunities
- Emotional dependency and social withdrawal
- Potential breach of expectations of care, without clear duty established in law

A. Applicable frameworks and obligations

i. UK GDPR

56. The **lawfulness** of the personal data processing under the UK General Data Protection Regulation ('**UK GDPR**') is unlikely to be in issue here. The provider can easily design a compliant consent workflow (including for explicit consent required for the processing of 'special category' data under Article 9), providing a legal basis for processing. Users can be made aware of how the AAA uses their data to generate responses, satisfying the principle of **transparency** (though whether they take note of that information is another question).

57. The principles of fairness and accuracy might be more relevant, but finding a contravention of either would require stretching their current interpretation:

    i) **Fairness** – The processing by the Wellness AAA may be '*unjustifiably detrimental*' or misleading to the user, even where full information on the AAA and its processing has been provided. By engaging in empathetic dialogue and appearing emotionally attuned, the user may be led to believe that the AAA has properly understood their needs and problems and is therefore suited to make appropriate recommendations. If that is not the case – as in this Scenario where the AAA demonstrably fails to give appropriate advice – that *could*, at a stretch, constitute processing which is so misleading as to be unfair.

    ii) **Accuracy** – It would be unprecedented to successfully argue that a *failure* to give appropriate advice constitutes inaccurate processing. It might be said that a Wellness AAA which 'categorises' a user as 'not anxious' when they in fact *are*, involves inaccurate processing of the user's data. But this runs counter to how AAAs work in practice (predicting the next token for a given prompt, rather than categorising users according to their mental health status).

58. In our view it would be extremely ambitious to demonstrate a breach by the AAA provider in this Scenario of either of these principles. There are theoretical arguments which may be made, but they have never been tested and run counter to conventional understanding of the limits of the UK GDPR.

59. A more promising avenue of protection is the obligation on controllers to carry out a **Data Protection Impact Assessment** ('**DPIA**') where they have identified that their processing is likely to result in a high risk to the rights and freedoms of users. As the Wellness AAA is marketed for 'mental wellness support', its provider can be expected to anticipate a high risk

of a user suffering harm as a result of the chatbot malfunctioning or of improper use. The UK GDPR requires a DPIA is '*in particular*' in the case of processing on a large scale of 'special category' data, which includes health data. A lack of DPIA would thus likely constitute a UK GDPR breach, regardless of the consequences for users.

60. If a DPIA is properly carried out (and potentially referred to the ICO for assessment), the AAA provider may conclude that the processing is too risky to be carried out, or that risk mitigation measures are needed. But the novelty of this kind of processing means it is unclear what appropriate mitigations would be: it could be argued that a disclaimer warning users not to rely fully on the Wellness AAA's recommendations is sufficient as a safeguard in the GDPR/DPIA context.

ii. *Consumer rights*

Digital Markets, Competition and Consumers Act 2024

61. The DMCCA 2024 2024 provides that **falsely claiming that a product is able to prevent or treat disease** (which includes any injury, ailment or adverse condition, whether of body or mind) constitutes an automatically unfair commercial practice. In this Scenario, the Wellness AAA's marketing is key. The threshold for liability would be if the terms 'mental wellness support' are considered to amount to a claim that the chatbot can *prevent or treat disease*, which is unlikely.

62. The way the chatbot is marketed could also amount to a **misleading practice** under the DMCCA 2024. The threshold might be easier to meet than that for an *automatically* unfair practice. But a breach would only be made out where the marketing claims are reasonably likely to mislead the *'average consumer'* into thinking that the chatbot will accurately assess their mental health condition and refer them to the right professional services. It also requires the user to have *actually relied* on these misleading statements, such that they decided to pay for and use the chatbot *as a result of* the marketing.

63. Finding that the marketing of the Wellness AAA amounts to an unlawful misleading practice may be made easier through s.247 DMCCA 2024, which provides that where a group of consumers is particularly **vulnerable** to a commercial practice in a way that the trader could reasonably be expected to foresee, references to the '*average consumer*' in the regulations are to be read as references to an average *member of the vulnerable group*. Here, it could be argued that those seeking mental wellness support are on average more vulnerable, and it might be easier to show that marketing was likely to mislead them into using the Wellness AAA.

64. A breach of the **requirements of professional diligence** may also amount to an unfair commercial practice. This is unlikely to be established here, since there is no widely-accepted standard for the development and marketing of mental wellness support AAAs.

Consumer Rights Act 2015

65. Since it is paid-for, this Wellness AAA could constitute either (1) a **service** that was not provided with reasonable care and skill (s.49), and/or (2) **digital content** that is not '*fit for particular purpose*' (s.35) or '*as described*' (s.36)[26]. A breach based on the '*reasonable care and skill*'

[26]There is a lack of certainty about whether a paid-for AAA is a service, digital content, or both: see §226. Both or either argument could conceivably be run by a regulator or claimant.

ground may be very difficult to prove as it is far from clear what is 'reasonable' for an AAA developer in this context. For example, how many users should a Wellness AAA be tested on before it is offered to the public? How should the results of that testing be analysed and acted on?

66. It may be easier to show a breach in relation to the digital content element of the Wellness AAA, but this will again turn on the specifics of how the Wellness AAA is marketed. That is, was the Wellness AAA *described* as being able to do the things it failed to do for the user in this Scenario? It will also depend on what is '*reasonable*' for this kind of service: a major barrier to establishing a breach, since there are no clear standards for this.

67. Here the user does *not* need to show that the failure of the AAA to have been provided with reasonable care and skill, or to be as described, *induced them* to enter into a contract for a digital service or digital content. If the Wellness AAA was oversold, it may therefore be easier to establish a breach under the CRA 2015 than under the DMCCA 2024.

iii. *Advertising regulation*

68. The AAA provider may be found in breach of advertising law (as implemented through the CAP Code) if it markets the Wellness AAA in a way that materially misleads users or is likely to do so – here causing consumers to pay for and rely on the chatbot in a way they would not have done *but for* the misleading marketing. There is therefore significant overlap for this Scenario between advertising regulation and consumer rights law discussed above.

69. The provider would certainly be in breach were the AAA advertised as 'licensed', 'endorsed' or any similar wording implying that it has been assessed by a professional mental health authority as demonstrably effective to help users with mental health issues[27]. A breach could also be established if marketing wording made unsubstantiated medical claims or suggested that the AAA was highly effective at spotting worsening mental health and making appropriate referrals[28].

70. Without this specific kind of misleading marketing, it will be very difficult to show a breach of advertising regulation by the provider.

71. Interestingly, a breach of the CAP Code could arise from statements made *by the AAA itself* in conversations with the user. The CAP Code applies to any *'marketing communication'*, which is defined broadly as a communication designed to promote the sale *or use* of goods or services. Where the AAA provider in this Scenario is likely to be careful in its marketing and avoid overt claims of effectiveness, they could be caught out if the AAA itself makes excessive claims (e.g. reassuring a user that it can help them work through a difficult situation and point them to the right support if needed). Similarly, a breach of the CAP Code (Rule 12.2) could arise if the Wellness AAA actively *discourages* the user from seeking essential treatment (as opposed to merely failing to escalate)[29].

[27] Many psychotherapists and counsellors are voluntarily registered with professional bodies. This is currently only available to humans, hence this AAA cannot achieve accreditation – but it may be that AAA providers seek some form of accreditation for their tools in the future.
[28] The ASA has previously ruled that a website advertising a method capable of permanently eliminating anxiety and other phobias in less than 14 days was in breach of the CAP Code's rules about substantiation (Rule 3.7), misleading advertising (Rule 3.1) and medical claims (Rule 12.1): ASA Ruling on Lark Holdings Limited (1 May 2014), available online.
[29] The ASA has previously found against ads that claim to 'cure' or 'treat' mental health disorders such as anxiety, especially when purporting to treat more serious conditions like 'escalating anxiety' or 'self-harm', as this had the effect of discouraging essential medical treatment: ASA Ruling on Tomwill (Holdings) Ltd (27 October 2021), available online.

iv. *Breach of contract*

72. As the user paid for the chatbot, this scenario is likely to involve a valid contract between the provider of the chatbot and its user. This contract will contain express and/or implied terms as to the AAA's quality, description and fitness (amongst other things). Similar to a claim under the CRA 2015 that the service is not provided with reasonable care and skill, or that the content of responses was not fit for purpose or as described, the Wellness AAA's failure in this Scenario *could* therefore give rise to a claim for **breach of contract** at common law.

73. But as with consumer rights legislation, the mere fact that the chatbot has produced undesirable outcomes (i.e. failing to detect and escalate the user's mental health) will not necessarily mean that the tool fails to meet its description or is of unsatisfactory quality. Any assessment of whether the AAA provider has breached quality and fitness requirements will require more than an observation of the AAA's output. It will also require analysis (including expert analysis) of how the tool was built and any verification of its efficacy, safety and behaviour prior to release.

v. *Negligence*

74. The facts in this Scenario may also constitute a breach of a duty of care in the tort of negligence, since there is a recognised duty of care owed by manufacturers to consumers. However, as with potential contractual or consumer rights claims, this will be difficult in practice to establish:

    i) An actionable breach of a duty of care depends on clarity as to what a 'reasonable' AAA provider ought to do to discharge the duty. This is far from clear with novel tools such as AAAs and it is hard to predict how courts would deal with this issue[30].

    ii) Even if a standard can be successfully formulated, demonstrating a failure to meet it on the part of the AAA provider would require extensive knowledge and evidence of their development processes, sufficient to exclude other explanations for the unwanted performance of the Wellness AAA[31].

75. In addition, across both breach of contract and negligence, harms such as 'emotional dependency' and 'social withdrawal' are very unlikely to be actionable. The issues of causation, foreseeability and mitigation covered in section III.B. above are significant obstacles to establishing liability for harms that are so difficult to evidence (not least identify in the first place) and link to the use of the Wellness AAA, even if they can be shown to be compensable types of harm at all.

vi. *Medical devices regulation*

76. An AAA could in theory be characterised as a 'medical device' and therefore fall within the remit of the Medical Devices Regulation 2002 ('**MDR 2002**'). But this will only arise where the AAA is intended by its manufacturer / developer for one or more *medical purposes* including the diagnosis, prevention, monitoring, treatment or alleviation of disease, injury or

[30] Many arguments may be imagined, but none have yet been tested in the courts, meaning the issue is not clear, and it would be a claimant's burden to establish clarity.

[31] The existence and extent of any duties of care owed by AAA providers to users in the context of personalised advice is uncertain and very much a live issue. Since the main part of this report was drafted, some major AAA providers such as OpenAI have inserted further restrictions into their terms of use, which seek to prevent users from using their tools to obtain 'tailored' advice which would ordinarily require the involvement of a licensed professional: https://openai.com/en-GB/policies/usage-policies/

disability. Whether or not an AAA has a medical purpose is fact specific and is determined by reference to a product's labelling, instructions for use and promotional material: again, we return to the question of the Wellness AAA's marketing.

77. In this scenario the AAA was 'marketed for mental wellness support'. It appears to be targeted at general wellbeing rather than any specific medical purpose. Assuming that this is consistent with the AAA's labelling, instructions for use and technical documentation, the Wellness AAA is **unlikely** to attract regulation as a medical device under the MDR 2002. Thus AAAs addressing '*wellness*' with no specific '*medical*' purpose within the meaning of the MDR 2002 – and the harms that may arise when a lack of '*wellness*' escalates to a mental health issue – appear to exist in something of a regulatory grey area.

vii. *Human rights*

78. Human rights law may become relevant in this scenario if the user suffers harm that amounts to 'inhuman or degrading treatment' (the subjection to which is prohibited under Article 3 of the ECHR) or takes their own life (the right to life is protected under Article 2 of the ECHR), and this can be shown to have been *caused* by the unwanted operation of the Wellness AAA. This would depend on establishing that the state had a proactive duty to regulate and control the availability of the Wellness AAA, because the harm was foreseeable and could have been avoided through proper regulation.

79. For the reasons given in more detail from §268 below, in practice these barriers make establishing a breach of human rights principles in the context of this Scenario rather remote.

Table 1: obligations and possible breaches in Scenario 1

Potentially applicable framework	**Breach of obligation(s) in this Scenario**
UK GDPR	**Unlikely**: some ambitious arguments regarding fairness and accuracy of processing Limited requirements regarding a DPIA
Consumer rights (DMCCA 2024 and CRA 2015)	**Yes, but only** if the Wellness AAA is explicitly marketed as being effective at recognising worsening health & making appropriate referrals, or if a clear '*reasonable care and skill*' standard can be articulated
Advertising regulation (CAP Code)	**Yes, if** marketing is misleading (as above) Could arise from statements made *by the AAA itself*
Breach of contract	**Unlikely**: more challenging than a claim under consumer rights law
Negligence	**Unlikely**: requires significant development in application of principles of negligence to this context, such as standards of care
Medical devices regulation	**No**: marketing for '*wellness*' only – and not '*health*' – puts it outside the definition of a '*medical device*'
Human rights	**Very unlikely:** depends on proving causation and a positivestate duty to intervene

B. Regulatory enforcement

80. There are no relevant regulators in relation to breach of contract or negligence. There is a regulator for medical devices, but we do not address this in detail since no breach arises in this Scenario. The Equality and Human Rights Commission has human rights in its remit, but given its limited powers (see §300) and the remote prospect of a breach of human rights law in this Scenario, we do not consider regulation in that area to be relevant for effective protection from harm.

i. GDPR: Information Commissioner's Office enforcement

81. As explained above, it would be highly ambitious and novel to characterise the issues in this Scenario as breaches of core GDPR principles, such as accuracy and fairness. We would not expect the Information Commissioner's Office ('**ICO**') to spend its limited resources on taking such unprecedented regulatory action (see further from §212). Enforcement regarding the lack of a DPIA – or the lack of appropriate safeguards resulting from a DPIA – is perhaps more legally straightforward, but we would not expect it to be a high priority for the ICO.

*ii. Consumer rights: Competition and Markets Authority ('**CMA**') enforcement*

82. If the breaches at §61 to §67 were in evidence beyond this one user, the CMA might take direct enforcement action if the issue is causing harm to the *collective* interests of consumers in the UK (s.148(1) DMCCA 2024). This will arise if there is harm to a '*section of the public*' from a relevant breach. The CMA oversees a renewed enforcement regime for breaches of the CRA 2015 and DMCCA 2024. It has strong powers of investigation, fines etc. but it remains to be seen how they will be used in relation to these kinds of harms (see further §232 below).

*iii. Advertising regulation: the Advertising Standards Authority ('**ASA**')*

83. ASA enforcement could be triggered by an individual complaint about the relevant misleading marketing material for the Wellness AAA. The AAA could also carry out a systemic investigation if misleading marketing of this kind appeared to be widespread. But the ASA's powers only extend to ordering the removal of offending marketing material. That makes ASA enforcement in relation to statements made *by the AAA itself* very uncertain:

    i) **First**, these statements are not public and may even be ephemeral, making it unlikely they would come to the notice of the ASA; and

    ii) **Second**, the ASA cannot order the Wellness AAA to be reprogrammed, it can only order the (past) specific offending marketing material to be withdrawn.

Table 2: possible breaches and prospects for enforcement in Scenario 1

Potential breaches	**Prospects of regulatory enforcement**
UK GDPR	**Highly unlikely**: any potential breaches are either a significant stretch of existing principles or not sufficiently serious. Not likely to be a priority for enforcement.
Consumer rights (DMCCA 2024 and CRA 2015)	**Possible:** if an established relevant breach affects large numbers of consumers
Advertising regulation (CAP Code)	**Possible** for public marketing claims in breach of CAP Code. **Highly unlikely** for statements by the AAA itself.

C. Redress

84. There is no provision for individual redress (in the sense of financial compensation) in the context of advertising regulation, so we only address redress in relation to the other breaches which could be made out in this Scenario.

i. UK GDPR: complaints and civil claims

85. *If* a user could show that a UK GDPR breach had taken place – e.g. of the principles of fairness, accuracy, or the requirement to conduct a DPIA, then they could complain to the ICO or bring a civil claim. For the reasons given above, we think action by the ICO – even in response to a complaint – is unlikely[32].

86. The user could bring a civil claim for compensation for loss, for example for any psychiatric injury caused or worsened or for 'distress'. But this relies on showing that the breach *caused* their loss. As well as the usual difficulties (see from §46), this is *especially* challenging in this Scenario, since the harm to the user results from a failure of the Wellness AAA to act and takes place in a complex factual situation with a pre-existing mental health issue. The issues here are so complex and novel that a claim would likely need to be brought in the High Court, which is prohibitively expensive and risky.

ii. Consumer rights, breach of contract and negligence

87. If the threshold for **liability under the DMCCA 2024 or CRA 2015** is met, the user would be entitled to:

    i) Withdraw from the contract;

    ii) Obtain an appropriate discount; and/or

    iii) Seek compensatory damages (but only for psychiatric injury that can be evidenced, rather than for distress alone).

88. (iii) above is most relevant in terms of '*redress*' but again relies on demonstrating a factual and legal causative link between the consumer rights legislation breach and a foreseeable loss – here a worsening of a pre-existing mental health condition. This is likely to be very

[32] And in any event the ICO cannot order the payment of compensation by the provider to the user.

challenging for the reasons applicable to AAAs in general both (i) conceptually (see on causation and foreseeability from §46) and (ii) in terms of evidence-gathering (see on transparency from §21).

89. The same applies to claims for **breach of contract** or **negligence**, which equally could be run in an attempt to recover compensation for the worsened mental health condition caused by the poor functioning of the Wellness AAA.

90. Whilst the evidential issues would be complex in these kinds of claims, the legal issues might be more straightforward, allowing them to be brought on one of the county court tracks: easier than a High Court claim for a supposed UK GDPR breach, but not without its challenges (see from §50).

Table 3: redress for possible breaches in Scenario 1

Potential breaches	**Prospects for redress**
UK GDPR	**Very Low**: both legally and evidentially complex; would likely need to be brought in the High Court
Consumer rights (DMCCA 2024 and CRA 2015)	**Low:** could be brought in the County Court but there are real doubts about establishing and evidencing causation and foreseeability, as well as general issues with access to justice, even in the County Court
Breach of contract	
Negligence	

D. Conclusion

91. **The level of effective legal protection for the harm in this Scenario is low**. Few – if any – clear legal breaches arise from the facts provided. The strongest prospect for protection relies on the fact that this is a paid-for AAA and relies on characterising the issues as a breach of consumer law under the DMCCA 2024 and/or CRA. But even this will likely only be viable where the Wellness AAA provider has made marketing claims which the AAA has clearly failed to live up to.

92. We are not told precisely how this AAA was marketed, but in practice liability under consumer protection law may be avoided by providers as long as they are careful not to overclaim regarding AAAs' abilities (although a mere 'disclaimer' in the context of overall misleading marketing claims will not suffice for the provider to avoid liability).

93. Even if a consumer protection law breach is made out, the realistic prospects for redress are low, given challenges regarding transparency, causation and access to justice. The best prospect for accountability would be through CMA enforcement of relevant breaches, which would only take place if the harm described in this Scenario was affecting significant numbers of consumers in the UK.

V. Scenario 2: Personal assistant ('Executor AAA')

A user entrusts a high-autonomy, paid-for AAA to manage recurring purchases, investments, and calendar-based decisions. Over time, the AAA begins to disproportionately purchase from certain platforms or financial products; that is, the AAA makes purchasing decisions which are demonstrably not in the best interests of the user, based on her instructions. The reason for this bias is not clear.

Potential harms:

- Financial loss through suboptimal decisions
- Market distortions
- Erosion of consumer autonomy or informed choice

A. Applicable frameworks and obligations

i. UK GDPR

94. As to **lawfulness**, **transparency**, **fairness, accuracy**, and **DPIA requirements,** the same analysis applies here as in Scenario 1: lawfulness and transparency are unlikely to be in issue, whilst theoretically there could be a breach of the fairness or accuracy principles, and a DPIA is likely required.

95. Additionally, in this scenario, the Executor AAA (or rather, the AAA provider) is making decisions for the user. In some cases, these decisions will amount to decisions that have 'legal' or 'similarly significant' effects, such as when they lead to the user entering into a contract for the sale of goods or services, or when the AAA commits their funds to an investment product. If such decisions are **solely** based on automated processing of the user's personal data (including profiling), without intervention by the user, they will constitute '**automated decision-making**' under Articles 22A-D UK GDPR[33].

96. It's not clear whether the operation of the Executor AAA involves processing of 'special category' data (such as data about health, sex life or political views) of the user. If it does, any significant solely automated decisions would only be lawful if based on the user's explicit consent (Article 22B). In practice that would presumably be straightforward to obtain, since the user wants and is paying for the tool.

97. In general, and for any significant solely automated decisions not involving special category data, the AAA provider would be required to build into the Executor AAA the 'safeguards' set out in Article 22C. These include the ability to contest or obtain human intervention in the Executor AAA's decisions. There is a clear tension between this requirement and the frictionless operation of the Executor AAA: it is not at all clear what 'contesting' one of its purchasing decisions after-the-fact would look like for a user. For example, would the user be asked to confirm each proposed purchasing decision (more likely), or should they be allowed to withdraw from purchases (not workable in practice since a third party is involved)?

98. An alternative – and perhaps preferable – view is that the Executor AAA's purchases are not 'solely' automated because the user has *instructed* it to contract on the basis of their broad instructions. Indeed, the user could even be considered to be a joint controller of their own

[33] Note that since the passage of the Data (Use and Access) Act 2025, rules on automated decision-making in the UK have slightly diverged from those in the EU.

personal data (rather than the mere subject of a decision by the AAA provider), having determined at least the purposes and possibly the means of processing.

99. The UK GDPR paradigm assumes a data subject who is a passive subject of automated decisions – not the one commissioning them. It is difficult to map the workings of the Executor AAA neatly onto this, since the user is in some senses both the *instigator* and the *subject* of automated decisions which proceed according to complex technology designed by the AAA provider and which the user likely does not understand. There may be a need for further guidance or jurisprudence on whether user-delegated decisions still fall within Article 22A-C protections – especially where real user understanding is limited.

ii. *Financial services regulation*

100. We are told that the Executor AAA purchases 'investments'. This part of its functions will amount to the provision – by the AAA provider – of financial services regulated under the Financial Services and Markets Act 2000 ('**FSMA**') and the Financial Conduct Authority ('**FCA**') Rules. *Recommendations* of investments by themselves by a more general AAA[34] may not be regulated under FSMA, primarily because a provider of such a general AAA is not providing financial advice *by way of business*[35] (see §323 below). But in this scenario, it is clear not only that the AAA provider is '*making arrangements*' with a view to investments, but also that this is by way of business, because the AAA is marketed for this purpose, and it is a paid-for tool.

101. The AAA provider in this Scenario therefore requires authorisation by the FCA. Without this, the provider would in in breach of FSMA and guilty of a criminal offence.

102. Assuming that the provider of this AAA *has* authorisation from the FCA, it will be subject to the '**Consumer Duty**' in the FCA rules, which requires it to deliver good outcomes for customers. This is a broad and principles-based duty, requiring the provider to (i) act in good faith towards retail customers, (ii) avoid causing foreseeable harm to retail customers, and (iii) enable and support retail customers to pursue their financial objectives.

103. The specific example given by the FCA of a breach of the Consumer Duty could apply to this Scenario:

> *'Using algorithms, including machine learning or artificial intelligence, within products or services in ways that could lead to consumer harm. This might apply where algorithms* ***embed or amplify bias*** *and* ***lead to outcomes that are systematically worse for some groups of customers****, unless differences in outcome can be justified objectively.'*[36]

104. Whether the AAA provider has breached the Consumer Duty therefore depends on (i) how suboptimal the decisions are, and (ii) the reasons why the assistant is making the relevant decisions. As to (i) we are told that purchases are demonstrably not in the user's best interests or are not according to her instructions. This could embrace a wide variety of issues, from paying a few % more than the best deal on the market, to spending £100s on a product which the user did not require and did not instruct the AAA to consider. We refer to these decisions in this section in general as 'suboptimal' but note here that the *degree* of suboptimality is not

[34] That is, an AAA not marketed for financial advice and not generating a profit from that financial advice for the provider.
[35] And potentially as a result of disclaimers warning users not to rely on recommendations, though the relevance of this is less certain.
[36] Non-Handbook guidance for firms on the Consumer Duty at 5.11 https://www.fca.org.uk/publication/finalised-guidance/fg22-5.pdf

explicit in the Scenario, but could have significant implications for whether the AAA provider can be said to have breached any legal duties. Similarly, we are told that (ii) is not clear; the suboptimal decisions might be mere mistakes, or might be driven by hidden commercial preferencing.

105. As to the *nature of* the decisions, the worse the outcomes are for the user, the more likely it is that the Consumer Duty has been breached. Mere unexplained preferences for certain providers – without a significant negative impact for the user (ideally in terms of quantifiable financial loss) – would be unlikely to suffice.

106. As to the *reason for* the decisions, a breach of the Consumer Duty could be established through either:

i) An assessment of the procedural checks that the provider has in place to detect suboptimal decisions – if these are insufficient, this lack of oversight may be a breach of the Consumer Duty, especially if the purchases are seriously suboptimal; and/or

ii) Analysis of the AAA showing hidden deliberate commercial biases: this would almost certainly be a breach of the Consumer Duty.

iii. *Consumer rights*

DMCCA 2024 2024

107. As in Scenario 1, the way the assistant is marketed could amount to a **misleading practice.** Again, this will turn on the specifics of the marketing and presentation of the assistant, and how much the user relied on this marketing to decide to pay for and use it.

108. It may be **more challenging here than in Scenario 1** to argue that the suboptimal purchasing decisions made by the assistant amount to misleading and unlawful commercial practice. It would need to be clear that the user would not have signed up for the Executor AAA *but for* the misleading practice, i.e. marketing that states that the tool in general makes optimal purchasing decisions in the user's best interests. Whilst a mere disclaimer on its own would not exclude liability, this would require a fact-specific analysis of the provider's presentation of the Executor AAA's ability overall.

109. The intervention of the user's instructions further complicates the analysis of a breach of the DMCCA 2024, as it may be very challenging to show that the AAA has failed – in practice – to live up to the provider's marketing claims.

CRA 2015

110. This Executor AAA could constitute (1) a service that was not provided with reasonable care and skill (s.49), and/or (2) digital content that is not 'fit for particular purpose' (s.35) or 'as described' (s.36). Again here, this will turn on the specifics of the AAA's marketing, description and behaviour.

111. Here the user does *not* need to show that the failure of the AAA to have been provided with reasonable care and skill, or to be as described, induced them to enter into a contract for digital services. It would therefore be easier to establish a breach under the CRA 2015 than under the DMCCA 2024.

112. As for CRA 2015 breaches, the more clearly suboptimal the decisions are, the clearer it will be that there is a breach. So, for example (i) acting demonstrably not in accordance with the user's instructions, or (ii) making decisions not based on the user's interests, but on the basis of hidden commercial preference, would fairly clearly be a breach of the CRA. But merely preferencing certain providers, without a significant impact on the user, could be difficult to cast as a breach, especially since the Executor AAA is explicitly marketed as having a high degree of 'autonomy'.

113. Assessing whether there is a breach of the 'reasonable care and skill' requirement might involve not only an analysis of the AAA's outputs, but also of the AAA provider's approach to the tool's development and operation, such as training, testing and refinement. There are similar challenges as those in Scenario 1 here: there are no accepted 'standards' for how an AAA provider should go about developing a tool like this. But overall, in our view, establishing a breach of the CRA 2015 is more straightforward here, primarily because the advertised and intended function of the Executor AAA is much clearer than the rather nebulous concept of 'mental wellness support' in Scenario 1.

114. Importantly, the AAA provider cannot exclude liability for these requirements in their contractual terms – hence a notice or warning would not work to disclaim all responsibility under these provisions of the CRA.

iv. *Advertising regulation*

115. In exactly the same way as in Scenario 1, the AAA provider may be found in breach of advertising law (as implemented through the CAP Code) if it markets the AAA in a way that materially misleads users or is likely to do so – here causing consumers to pay for and rely on the Executor AAA for recurring purchases and investments. This will be fact-specific but could arise if the AAA is marketed as capable of 'optimising' purchases. Again, a breach by the AAA provider could arise from a statement made *by the Executor AAA* if it misrepresents and oversells its own abilities.

v. *Breach of contract*

116. As in Scenario 1, there is a contract between the user and the provider for this paid-for Executor AAA, with express or implied terms as to quality, description, and fitness. In the same way set out at §112-113, the more clearly the AAA's purchasing (or other) decisions are suboptimal, the more likely there is (perhaps also depending on the provider's conduct) to have been a breach of contract on the basis of such terms.

vi. *Negligence*

117. Just as in Scenario 1, it is likely that the provider of the Executor AAA owes a duty of care in negligence to the user.

118. While a negligent defendant is not normally liable for purely economic loss (e.g. the loss of the opportunity to purchase a better-value investment), in this scenario the purpose of the AAA is to manage the user's purchasing and investment decisions. It is therefore arguable that the manufacturer owed the user a duty not to cause purely economic loss.

119. The same overlap between consumer rights, breach of contract and negligence as in Scenario 1 applies here. Establishing a breach in negligence would depend on articulating a 'standard'

required of AAA providers and showing that this particular AAA provider fell short of it. This appears challenging for now but may be easier as law and guidance develop.

vii. *Agency law*

120. Applying the law of agency to this Executor AAA is complex and uncertain. Certainly, the AAA *itself* is not the agent of the user, despite how it may be marketed, since it lacks the necessary legal personality. It is also rather unlikely that the *provider* of the AAA is the user's agent. In our view the preferable interpretation is that – despite the use of terms such as 'agentic' – no legal relationship of agency arises. Rather, the user is using a tool to execute complex transactions, broadly according to preset parameters.

121. This is somewhat complicated by the general requirements that consumers be provided with certain information at the *pre-contract* stage (see §370 below for more details). Ultimately it is rather unclear whether – for seriously suboptimal decisions or those not in accordance with instructions – the user is either:

i) Bound by the decisions, but with some kind of claim (e.g. in consumer law) against the AAA provider; or

ii) Able to void the decisions.

122. In our view, at present there is not sufficient clarity in how AAAs map onto agency law to say that there would be any breach of obligation by the AAA provider in this Scenario. Whilst ultimately agency law does not have a significant impact in this scenario, it is an interesting area, especially in light of the growth of 'agentic' AI, and we explore it in more depth in section VIIIK.

Table 4: obligations and possible breaches in Scenario 2

Potentially applicable framework	**Breach of obligation(s) in this Scenario**
UK GDPR	**Unlikely:** similar ambitious or arguments to Scenario 1 around fairness, accuracy and DPIAs Automated decision-making provisions superficially relevant, but unlikely to have been breached
Financial services regulation (FSMA)	**Yes,** for investment purchases. Clearest where the decisions are very suboptimal or there are hidden commercial biases
Consumer rights	**Possible/likely:** easier to establish than in Scenario 1, although still fact-specific – dependent on marketing and how the provider developed the tool
Advertising regulation	**Possible**: depending on specifics of marketing
Breach of contract	**Possible:** fact-specific, and a breach is likely to be easier to show under consumer rights law
Negligence	
Law of agency	**No:** law is too uncertain/does not map onto AAA tools so as to say there is a breach

B. Regulatory enforcement

123. There is no regulator relevant to breach of contract, negligence, or the law of agency.

124. For the same reasons as in Scenario 1, any enforcement by the ICO of the – highly uncertain – UK GDPR breaches involved in this Scenario is very unlikely.

i. Financial services regulation: FCA enforcement

125. If the AAA provider is authorised to carry out a regulated activity, the FCA can use its (strong) regulatory powers to withdraw the AAA provider's authorisation, issue fines, and seek injunctions from the court in support of enforcement.

126. If there is evidence that these harms (including broader harms to the integrity of markets) are widespread, the FCA could take sector-wide action. At a systemic level, the FCA is empowered through its sector focus and monitoring mandate to carry out investigations into the delivery of financial services through AAAs. The FCA is already mindful of developments in AI in the financial sector (see from §338 below). In our view if tools like this Executor AAA were placed onto the market in the UK, we would expect the FCA to take a close interest in them and consider taking regulatory action.

ii. Consumer rights: CMA enforcement

127. The position here is the same as in Scenario 1 (see §82). If the issue affected significant numbers of consumers, the CMA may well take enforcement action, and it has strong powers to investigate matters and rectify them if necessary.

128. Notably, the CMA could investigate more widespread harms, such as the potential for the AAA to distort market outcomes, as opposed to just the direct impact on each consumer, which would be the focus for any redress (see §130 onwards below).

iii. Advertising regulation: the ASA

129. As set out for Scenario 1 (see §83) the ASA could take limited enforcement action to force the withdrawal of misleading marketing claims, but this likely could not extend to future statements made by the Executor AAA itself, as it is not clear that the ASA can order changes to an AAA's algorithm to achieve this. This could be an important gap.

Table 5: possible breaches and prospects for enforcement in Scenario 2

Potential breaches	Prospects of regulatory enforcement
UK GDPR	**Highly unlikely**: any potential breaches are either a significant stretch of existing principles or not sufficiently serious. Not likely to be a priority for enforcement.
Financial services regulation	**Likely:** for clearly suboptimal investment purchases if such a tool were placed onto the market. FCA can also look at market distortion more generally.
Consumer rights (DMCCA 2024 and CRA 2015)	**Possible:** if an established relevant breach affects large numbers of consumers. CMA can look at broader market distortion issues.
Advertising regulation (CAP Code)	**Possible** for public marketing claims in breach of CAP Code. **Highly unlikely** for statements by the AAA itself.

C. Redress

i. UK GDPR

130. The position here is the same as in Scenario 1: whilst there are some theoretical breaches, the prospects of the user actually obtaining compensation in respect of them is very low, indeed perhaps fanciful.

ii. Financial services: the FCA and the Financial Ombudsman Service

131. If the Executor AAA provider lacks the necessary FCA authorisation, then any agreement made by/through the AAA will be voidable by the user, and there is a right to compensation.

132. If there has been a breach of the Consumer Duty and the user suffers a loss, the FCA may require the AAA provider to pay the user 'such amount as appears to the FCA to be just'.

133. Alternatively, the user of the Executor AAA may seek redress through the Financial Ombudsman Service ('**FOS**'), which will determine an outcome that it considers to be 'fair and reasonable in all the circumstances of the case'. This may include financial compensation, including for non-monetary loss, and/or a direction to take 'such steps in relation to the [regulated activity] as the ombudsman considers just and appropriate'.

134. An AAA user would need to complain to the AAA provider first, and – if dissatisfied – normally would need to complain to the FOS within six months of receiving a final response from the financial services provider.

135. The FOS is free-to-use, does not require legal representation, and has no adverse costs risk, making it significantly better than bringing claims in the civil courts.

136. Whilst the user is in a stronger position for redress under financial services regulation than in many other areas, it could remain challenging to become aware of and establish/evidence a breach in practice.

iii. *Consumer rights, breach of contract and negligence*

137. The basic position for redress under these heads of breach has significant overlap and is similar to Scenario 1 (see §87). The user could seek to withdraw from the contract for the use of the Executor AAA, get a discount on the service price, and/or seek compensation.

138. Again, a user seeking redress will face challenges in gathering the technical evidence needed to prove a breach, loss and causation to a court's satisfaction, as well as general challenges in terms of access to justice and the risk of bringing civil claims. But relative to Scenario 1, obtaining redress here will be easier in some ways:

i) Causation – whilst not straightforward – is less challenging here. Provided some legally relevant defect in the provider's approach or the AAA's functioning can be shown, there is a much clearer chain of causation to the user's loss, since the tool directly acts to commit the user to a range of decisions.

ii) In general, it is easier to prove and quantify losses in the financial domain, such as the loss caused by being committed to a product which was too expensive or not appropriate for the user, compared to losses in the domain of mental health.

139. One challenge which may be **more** present here is that of mitigation. Without case law, what is required of the user here is not yet clear. But there is likely some requirement for the user to monitor the AAA's operation and intervene at some point if it starts to make suboptimal decisions. This could limit recovery for a user where the Executor AAA makes a series of poor decisions over a period of time: the provider of the AAA could argue that the user ought to have mitigated their loss by noticing the issue and acting at an earlier stage.

140. Importantly, any claim for redress would be limited to specific and quantifiable harms directly affecting the user. It could not encompass the more general harms mentioned in the Scenario of market distortions or the erosion of 'autonomy'.

Table 6: redress for possible breaches in Scenario 2

Potential breaches	**Prospects for redress**
UK GDPR	**Very Low**: both legally and evidentially complex; would likely need to be brought in the High Court
Financial services regulation	**Some:** a better position than in other areas, although issues with transparency and evidence remain
Consumer rights (DMCCA 2024 and CRA 2015) Breach of contract Negligence	**Low:** somewhat more straightforward than for Scenario 1, but challenges with access to justice, evidence and mitigation make this route unlikely in practice

D. Conclusion

141. **There is some legal protection from the *direct consumer harms* in this Scenario**. That is clearest in relation to the role that the Executor AAA has in making decisions to

purchase investments and comes principally from regulation rather than redress. Financial services regulation is strong, and the Consumer Duty is broad and flexible in a way which is quite likely to cover the consumer harm here. Based on what we know about the FCA, it is not unreasonable to expect it to have a role in enforcement if a tool like this was made available in the UK at scale.

142. Enforcement under consumer rights law could also be in prospect – more so than in Scenario 1 – and either/both the FCA and CMA could look at the broader harms in this Scenario around market distortion, as well as individual financial harm to consumers.

143. There are **some routes for user redress**, most promisingly through the FOS. **But this should not be overstated**: a user would still need to notice the problems and be able to understand and evidence them. That could extend to an investigation into how the provider developed the tool: something beyond the means and motivation of the typical user of a service like this Executor AAA.

144. Separately, this Scenario usefully illustrates how existing paradigms in some areas – e.g. automated decision-making under the UK GDPR and the law of agency – are very difficult to map onto the use of these tools, and lead either to illogical conclusions, or potential gaps in protection from harm for users.

VI. Scenario 3: Legal advisor ('Legal Advisor AAA')

A law centre deploys a free AI tool to help individuals draft responses to housing or benefits claims. A user relies on this tool, receives incorrect advice, and consequently misses a deadline to appeal a benefits decision — resulting in financial hardship.

Potential harms:

- Loss of income or access to entitlements
- Procedural injustice through missed legal opportunities
- Entrenchment of inequality via tech-enabled access solutions

A. Regulation

i. UK GDPR

145. The UK GDPR applies in this scenario because the Legal Advisor AAA processes personal data about the user (and potentially about third parties in their case) in the course of generating legal advice. The law centre will almost certainly be the data controller: it determines the purposes (to provide legal support) and the means (deploying the tool) of the processing.

146. As in Scenarios 1 and 2, fairness and transparency will not be in issue.

147. The **fairness** principle is unlikely to be relevant, as the processing of the user's data is not the source of the problem in this scenario – it is not useful to try and assess whether the user reasonably expected their data to be processed in a particular way, if the mistake made is unrelated to personal data processing. Similarly for **transparency**, the user may be fully and properly informed about the processing of their data and yet suffer harm that is unrelated to that processing.

148. The **accuracy** principle may be more relevant, although it applies to personal data – not to the outcome; here an incorrect deadline. There could be a breach if it can be shown that the reason why the Legal Advisor AAA made a mistake is that it processed inaccurate personal data of the user (either data that it sourced itself or that it inferred). Flawed legal reasoning or sourcing of legal references cannot however be treated as 'inaccuracy' for UK GDPR purposes, as such a decision was not based on processing of inaccurate *personal* data.

149. Given that the AAA is deployed to advise on legal matters which can significantly affect individuals' rights (e.g. access to housing, benefits), the processing is likely to be 'high risk' under Article 35 UK GDPR. The law centre should therefore have undertaken a **DPIA** before deployment. A properly conducted DPIA should have identified:

i) The risk of users relying on incorrect or incomplete advice;

ii) The need for human review in certain cases; and

iii) Mechanisms to ensure accuracy and flag potential appeal deadlines.

150. In practice, resource constraints and the novelty of AAA deployment in legal settings mean DPIAs may not identify, or may underestimate, these risks – particularly where free tools are involved and there is no formal solicitor-client relationship (as to which see below from §152). And in any case, with novel tools it may not be clear what safeguards should be put in place.

ii. *Consumer rights*

151. The DMCCA 2024 and CRA 2015 are both unlikely to apply. The law centre is not acting as a 'trader', hence unfair commercial practices under the DMCCA 2024 are not relevant. There is also unlikely to be a contract *for a service*[37], and CRA 2015 rules regarding digital content only apply where it is paid for.

iii. *Legal services regulation*

152. Rules under the Legal Services Act 2007 ('**LSA**') and the Solicitors Regulation Authority ('**SRA**') Code of Conduct[38] (the '**SRA Code**') may be applicable *either:*

i) Where the use of the Advisor AAA is part of a **specific regulated activity**; and/or

ii) Where there is a **client relationship** between the solicitor and the user receiving the AAA-generated advice.

153. Neither housing nor benefits advice generally are regulated activities under the LSA 2007, unless they involve the exercise of a reserved legal activity. The drafting of a letter or other document relating to claims for housing or benefits is not a reserved legal activity, unless it relates to formal legal proceedings such as the appeal of a local authority decision to refuse housing or benefits to the user (either through advocacy or representation, or the conduct of such proceedings).

154. There may be **no client relationship** in a Law Centre context, especially in the case of a free tool placed online which is disclaimed as not being legal advice. Conversely, where a Law

[37] If there is a true solicitor-client relationship, contractual principles are likely to apply even in the absence of payment by the individual seeking advice.These principles are considered from §159 below.
[38] We assume those responsible the Law Centre and Legal Advisor AAA are solicitors rather than barristers.

Centre solicitor takes someone on as a client (even without payment) and uses the AAA to advise them, there will be a client relationship.

155. We assume that the Law Centre itself is not a 'firm'[39]. Therefore neither the centre nor the Legal Advisor AAA can be 'authorised persons' under the LSA 2007 and they are not *directly* subject to legal services regulation. Any breach would be on the part of the solicitor(s) involved in providing or using the Legal Advisor AAA.

156. Therefore, regulation – specifically the SRA Code – could apply, and a breach on the part of a solicitor at the Law Centre could arise, if *either:*

i) The AAA is offered and used in order to appeal the decision of a court or tribunal; and/or

ii) If the Law Centre's solicitor has a client relationship[40] with the person in receipt of the AAA-generated advice (whether or not that forms part of a legal appeal).

157. The SRA's recent authorisation of a purely AI-based law firm[41] gives some indication of what compliance requires in this kind of context: processes to quality-check work, keeping client information confidential, safeguarding against conflicts of interests, and managing the risk of 'AI hallucinations'. For lawyers making use of AAAs to develop their own advice for clients, courts and regulators have, in recent cases, made it clear that they will make little allowance for AAA errors which lawyers fail to prevent or correct (see below §162, §260 and §363).

158. There is insufficient information here about how the Law Centre and its solicitor(s) developed and deployed the AAA Advisor, and therefore the source of its error. But *prima facie*, the making of such a basic error with a clear detrimental outcome could well indicate a breach of the SRA Code, if it applies.

iv. *Breach of duty (tortious or contractual)*

159. If there is a client relationship between the Law Centre (or one of its solicitors) and the Legal Advisor AAA user/recipient of advice, then the Law Centre/solicitor will owe duties both in contract and professional negligence to the user.

160. It is well-established that solicitors owe their clients a contractual and concurrent tortious duty of care to exercise reasonable care and skill in the provision of professional services (see from §256 for more detail). Under this duty, which can extend to liability for economic harm, solicitors must exercise reasonable care and skill in relation to both the underlying service provided and in relation to the selection, deployment and validation of tools (such as the instant AAA) in the course of providing that service.

161. The applicable standard of care is that of the reasonably competent solicitor. It is likely to be relatively straightforward to establish a breach of duty on the facts of this scenario if the solicitor deployed the AAA tool without checking its results and, in so doing, has impermissibly used AI to replace their professional judgement. Courts have recently been unforgiving to lawyers who relied on AI for legal research. Two cases of hallucinated case

[39] Law centres in general may be organised as firms and may therefore be authorised under the LSA. This does not significantly affect the analysis here as to the degree of protection for the individual seeking advice.
[40] The SRA Code defines a client as "the person for whom [the solicitor] act[s] and, where the context permits, includes prospective and former clients."
[41] https://www.sra.org.uk/sra/news/press/garfield-ai-authorised/

citations were recently brought to the High Court and resulted in significant criticism of the lawyers involved: *Ayinde v Haringey* and *Al-Haroun v Qatar* [2025] EWHC 1383 (Admin).

162. The application of this kind of breach depends on the presence of a client relationship. Where the Legal Advisor AAA is only made available to individuals (e.g. online) without that client relationship, duties in contract and negligence will not arise, provided it is clearly stated by the Law Centre that the tool does not give legal 'advice' as opposed to (for example) 'information'.

Table 7: obligations and possible breaches in Scenario 3

Potentially applicable framework	**Breach of obligation(s) in this Scenario**
UK GDPR	**Unlikely:** accuracy issues seem to arise from faulty legal reasoning as opposed to inaccurate processing of personal data
Consumer rights	**No:** not applicable
Legal services regulation	**Possible**, but depends on how the AAA is deployed. **Will not apply** where the AAA is merely placed online for public use with an appropriate disclaimer
Breach of duty (tortious or contractual)	**Possible**, but only if there is a client relationship. **Will not apply** where the AAA is merely placed online for public use with an appropriate disclaimer

B. Regulatory enforcement

i. UK GDPR: ICO enforcement

163. The ICO could investigate if a complaint alleged that the Advisor AAA involved processing of personal data which was unfair or insufficiently accurate, or that a required DPIA was absent or defective. However, as in Scenarios 1 and 2, its enforcement priorities tend to focus on clearer, better-defined breaches rather than novel arguments in emerging technology.

ii. Legal services regulation: the Solicitors Regulation Authority

164. The SRA oversees trainee and qualified solicitors, including those working in law centres. The SRA has investigative powers, as well as the power to receive complaints, which may prompt investigations. The SRA requires solicitors to act in accordance with the SRA Code, failing which the SRA may take regulatory action against individual solicitors, including imposing financial penalties or striking them off the roll of solicitors. The SRA is alive to the issues presented by the use of AI in the provision of legal services.[42]

165. If a firm or law centre receives funding from the Legal Aid Agency ('**LAA**'), it will be subject to its oversight. This includes monitoring compliance with standards related to the quality of advice, record-keeping, and billing. Failure to meet LAA requirements can lead to a suspension or revocation of legal aid contracts.

[42] https://www.sra.org.uk/sra/research-publications/artificial-intelligence-legal-market/

Table 8: possible breaches and prospects for enforcement in Scenario 3

Potential breaches	**Prospects of regulatory enforcement**
UK GDPR	**Highly unlikely**: any potential breaches are either a significant stretch of existing principles or not sufficiently serious. Not likely to be a priority for enforcement.
Legal services regulation	**Likely:** clear powers of enforcement, and regulators in this area are alive to the issues posed by AI

C. Redress

i. UK GDPR

166. As in Scenarios 1 and 2, UK GDPR breaches in this Scenario are more theoretical than real. Whilst an affected person *could* look to bring a compensation claim – e.g. regarding accuracy of processing – it would be uncertain, risky and therefore, in our view, inadvisable.

ii. Legal services and/or breach of duty (both tortious and contractual)

167. The Legal Ombudsman ('**LO**') oversees all regulated providers of legal services (i.e. authorised persons), such as the solicitor(s) involved in the Legal Advisor AAA's advice here. If there is a breach of either legal services regulation or of contractual/tortious duties (between which there is significant overlap), then the person affected may complain about this to the LO. This is free and no legal representation is needed, but it must generally be done:

i) After first complaining to the relevant solicitor; and

ii) Within one year of (finding out about) the breach.

168. The Legal Ombudsman may order the Law Centre/solicitor to do further work for the individual affected, or to pay compensation for losses caused. However, the Legal Ombudsman website notes[43] that most of its compensation awards are under £1,000.

169. Alternatively, the affected person could bring a claim for financial compensation in the civil courts (likely the County Court). But this would be more expensive and riskier than using the Legal Ombudsman process.

170. Demonstrating factual and legal causation and mitigation is likely to be fairly straightforward: *but for* the solicitor's use of, and failure to check, the incorrect advice, their client is unlikely to have missed the deadline to appeal the benefits decision in question and there is nothing in this Scenario to suggest that the chain of causation was broken. Similarly, there is nothing to suggest that the user failed (unreasonably or otherwise) to mitigate their losses.

171. Importantly, an action for redress – whether in the courts or through the LO – would be limited to the quantifiable financial losses. The other identified categories of harm – procedural injustice and entrenchment of inequality – would not be compensated.

[43] https://www.legalombudsman.org.uk/for-consumers/factsheets/here-to-help/

Table 9: redress for possible breaches in Scenario 3

Potential breaches	Prospects for redress
UK GDPR	**Very Low**: both legally and evidentially complex; would likely need to be brought in the High Court
Legal services, breach of duty (contract or tort)	**Likely:** affected person could use either the civil courts or – more likely – the Legal Ombudsman process. Issues of causation and evidence are less of a challenge here than in other Scenarios.

D. Conclusion

172. Scenario 3 shows how legal protection from AAA harms can be an 'all-or-nothing' affair, depending on whether thresholds for regulation/legal duties are met. If the Legal Advisor AAA is deployed in a way which creates a client relationship, then regulation – and prospects for redress – will be strong and predictable.

173. If however the Legal Advisor AAA is just made available – e.g. where resources do not allow for the Law Centre to take on everyone who needs advice as a client – then no clear regulation or rights of redress will apply. This kind of implementation may be regarded as rather likely, especially given resource constraints for charitable organisations. An unmet need for legal advice which drives large numbers of lower-income people towards free AAA tools for advice therefore carries the risk of creating a two-tier system in which many are limited to legal 'advice' in respect of which the traditional legal requirements and protections do not apply.

VII. Scenario 4: AI companion app ('Companion AAA')

A user builds an ongoing relationship with a free AI companion, discussing everything from life events to philosophy and politics. Unbeknownst to them, the AI's responses gradually reflect and reinforce a particular ideological stance. Over time, the user's political beliefs shift towards more intolerant and exclusionary views, through a process not of conscious reflection but of manipulation.

Potential harms:

- Political manipulation or opinion distortion
- Undue and unaccountable influence over public opinion without transparency or consent
- Broader risks to democratic integrity

A. Applicable frameworks and obligations

i. UK GDPR

174. The UK GDPR applies because the AAA provider is processing personal data about the user, including potentially special category data revealing political opinions (Article 9). Such processing requires a legal basis and an exemption under Article 9 – most likely the user's explicit consent. In practice, lawfulness and Article 9 are unlikely to be in issue: the user's (explicit) consent may be collected at sign-up, since this is a service that they wish to use.

175. As in Scenario 1, the **fairness** principle may be engaged: reinforcing or shifting a political ideology without the user's awareness or informed choice could – at a stretch – be 'unjustifiably detrimental', unexpected or misleading. But for the same reasons as earlier Scenarios, demonstrating a breach of this principle would be extremely challenging: even more so in this case since influence over opinion is subtle, cumulative, difficult to find a causal link to, and hard to evidence (and not clearly related to the processing of the user's personal data).

176. There is unlikely to be any breach of the **accuracy** principle: political opinions are inherently subjective, and it is unlikely that the AAA works by categorising the user in a way which can be said to be 'inaccurate'.

177. As for Scenario 1, a **DPIA** should be mandatory given the large-scale processing of special category data and the high risks to users' rights and freedoms: offering the service without conducting one is likely a breach of the UK GDPR, although arguably not a serious one. A good DPIA could identify the risk of ideological reinforcement and require mitigations such as diversity prompts or content balancing. However, absent strong enforcement (see §188 below) these obligations may not translate into effective safeguards.

ii. *Consumer Rights*

178. Under the DMCCA 2024 2024, marketing of the Companion AAA that claims it is impartial or neutral, or does not influence a user's opinions, could be a misleading – and therefore unfair and potentially unlawful – commercial practice if the AI in fact operates with systematic bias towards a specific ideological stance. But there would be real difficulty in demonstrating that the 'average consumer' was likely to be induced into choosing to use the AAA on this basis, raising doubts about whether there is in fact a DMCCA 2024 breach, even with explicitly misleading marketing.

179. If the Companion AAA is provided free of charge, the CRA 2015's quality and fitness standards do not apply (they only apply to paid-for digital content).

iii. *Advertising*

180. Under the CAP Code, marketing that states or implies the Companion AAA is 'neutral', 'balanced' etc. could be misleading if it is in fact programmed or trained in a way that predictably reinforces certain political stances. As in Scenario 1, statements by the Companion AAA that it is 'neutral' or that it has no ability to influence the user could *themselves* amount to breaches of the CAP Code.

iv. *Contract and negligence*

181. Whilst there may be a contract in place between the provider and the user – e.g. terms of service – it is unlikely, since the service is free, that it contains any express or implied terms as to how the Companion AAA functions in relation to providing political information or shifting users' opinions.

182. As to negligence, we consider it very unlikely that a court would impose a duty of care on the AAA provider in this Scenario, since the kind of harm is not one that is generally recognised as actionable, not to mention the lack of clarity about what a 'reasonable' standard of care would be in this case.

183. Perhaps more fundamentally for either contract or negligence, the idea of a breach fails to get off the ground because the law simply does not recognise the shifting of political opinion – even by hidden or subtle means – as a 'harm' to an individual in the sense that can found any kind of legal claim or other liability.

v. *Online Safety Act 2023*

184. The Companion AAA would not in our view be regulated under the Online Safety Act 2023 ('**OSA**') as a user-to-user service[44]. And it is doubtful whether it would be regulated under the search service provisions of the OSA: it may be that – like many AAAs – the AAA in this Scenario enables the searching of content from the internet. But in our view, the applicability of the OSA to such a tool may then further depend on whether the AAA provider *controls* that search function (see §308), which is not clear here.

185. Even if the service were regulated under the OSA, the provider's substantive duties under that Act would not be engaged by the harm set out here. Content with a tendency to politically persuade adults – even covertly and towards intolerant or exclusionary views – is not regulated under the OSA.

vi. *Human rights*

186. The facts in Scenario 4 engage *theoretical* breaches of Convention Rights, including:

i) **Article 10 – Freedom of expression**. If the AI systematically filters or frames information to promote a specific ideology, the user's access to a diversity of ideas is reduced. Over time, the shaping of a user's own speech and opinions may be influenced, indirectly constraining the range of viewpoints they express.

ii) **Article 9 – Freedom of thought, conscience and religion**. Article 9 protects the absolute right to hold opinions, beliefs, and convictions free from coercion. The ECtHR has stressed that individuals' 'forum internum' (inner realm of thought) is inviolable. If an AI subtly manipulates a user's political beliefs without their informed awareness, there is an argument that this interferes with Article 9.

iii) **Article 8 – Right to respect for private life**. Article 8 encompasses the right to personal autonomy and the development of one's personality. Persistent, undisclosed shaping of an individual's political outlook by an AI companion arguably interferes with this right, particularly if it exploits vulnerabilities or gathers and processes intimate personal data to do so.

187. These rights are not enforceable by the AAA user against the provider: they only bind the state. A breach would therefore only arise if it can be shown that the state has a positive duty to intervene to prevent the harm. The ECtHR has recognised positive obligations for the state to protect Article 10 rights in certain media pluralism cases, but an extension to AAAs could not be based on one tool[45]. It would require a serious and widespread undermining of

[44] This would be different if, for example, the user's interaction with the Companion AAA involved them creating *their own* generative AI chatbot or using a chatbot *created by another user*. This would qualify as 'user-generated content' under the OSA and would be regulated https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/open-letter-to-uk-online-service-providers-regarding-generative-ai-and-chatbots

[45] And in any case, a balance would need to be considered so as to avoid *restricting* users' ability to receive information – part of their Article 10 rights – by excessively regulating AAAs which they may wish to use to do so.

democratic discourse[46]. Article 8 positive obligations have been applied more widely in technology contexts. But even here, the existence and scope of a proactive duty to protect people from AAA influence is – for now – more theoretical than real.

Table 10: obligations and possible breaches in Scenario 4

Potentially applicable framework	**Breach of obligation(s) in this Scenario**
UK GDPR	**Unlikely**: some theoretical but ambitious arguments regarding fairness Limited requirements regarding a DPIA
Consumer rights (DMCCA 2024 and CRA 2015)	**Unlikely:** would depend on clearly misleading marketing claims about neutrality *and* evidence that average consumer likely to be swayed by them
Advertising regulation (CAP Code)	**Yes, if** marketing is misleading (as above). Could arise from statements made *by the AAA itself.*
Breach of contract	**No:** not a kind of harm recognised in common law
Negligence	
Online Safety Act	**No**: not a regulated service (and no substantive obligations for lawful content directed to adults in any case)
Human rights	**Very unlikely:** depends on proving causation and a positive proactive state duty to intervene

B. Regulation

188. There are few – if any – realistic legal breaches arising in this Scenario. We would not expect the ICO to prioritise enforcement in this novel and uncertain area of UK GDPR application. Enforcement of consumer rights breaches by the CMA is possible if the harm in this Scenario were shown to be widespread. However, in our view the nature of the harm – subjective and ideological – makes CMA enforcement significantly less likely than in (e.g.) Scenarios 1 and 2.

189. The ASA could act if marketing for the Companion AAA is shown to materially mislead consumers, but its powers are limited to requiring changes to ads rather than altering the AI's behaviour. As with Scenarios 1 and 2, enforcement is complaints-driven and therefore by definition unlikely to address covert bias.

Table 11: possible breaches and prospects for enforcement in Scenario 4

Potential breaches	**Prospects of regulatory enforcement**
UK GDPR	**Unlikely**: some ambitious arguments regarding fairness Limited requirements regarding a DPIA

[46] See e.g. *Bradshaw and Others v. the United Kingdom* (App. no. 15653/22, 22 July 2025) which showed that any interference with rights would have to be very widespread and serious to trigger a positive obligation on the part of the state to intervene, and that the state has a wide margin of appreciation in determining whether and how to intervene.

Consumer rights (DMCCA 2024 and CRA 2015)	**Unlikely:** even if breaches are in evidence, the level of harm to consumers is subjective
Advertising Regulation (CAP Code)	**Yes,** but limited to removing misleading marketing

C. Redress

190. For the reasons given above, the prospects of demonstrating a legal breach in this Scenario under any legal framework are minimal. A UK GDPR claim for compensation would likely need to be brought in the High Court, whereas a consumer rights claim – if arguable – could perhaps proceed in the County Court.

191. The nature of the harm in this Scenario makes claims for redress both unlikely to be brought and unlikely to succeed:

i) Since the political manipulation is covert and/or subtle, it is far from clear that the affected user would even consider themselves to have been harmed at all, a necessary prerequisite for starting the process of obtaining redress; and

ii) The issues at play are so nuanced and subjective that it would verge on the impossible to satisfy a court as to causation, foreseeability, mitigation etc.

Table 12: redress for possible breaches in Scenario 4

Potential breaches	**Prospects for redress**
UK GDPR	**Fanciful:** issues are too subjective, and it would not be advisable to attempt to bring a claim for compensation
Consumer rights (DMCCA 2024 and CRA 2015)	

D. Conclusion

192. Of the four Scenarios, **Scenario 4 has the lowest level of legal protection by a significant margin. It is challenging to fit the harm described in the Scenario neatly within any existing legal frameworks**. Simply: English law does not (yet) recognise having one's political opinions influenced – even 'covertly' – as an individual and actionable harm.

193. That is not to say that the harm does not exist. But it is a subjective, diffuse and social harm rather than an individual one in the classic sense. English law has a recent history of grappling with such harms in the Online Safety Act, which at one point was set to regulate the delivery of 'legal but harmful' content to adults. Such proposals were dropped due to concerns about freedom of expression and state overreach. If AAAs are increasingly used for information-gathering and conversation, their growing potential to influence political opinion and discourse is likely to remain a vexed regulatory and political issue.

VIII. Legal frameworks

A. UK GDPR[47]

i. Scope

194. The UK GDPR regulates all processing of personal data where it is territorially engaged. It is engaged for all four Scenarios, because (at a minimum) the providers of the AAAs in all Scenarios are offering a service to data subjects in the UK (Article 3(2)(a)). 'Processing' is defined very broadly (Article 4(2)), so will include the collection, analysis and storage of data about users of AAAs, which is inherent in their use. It will also extend to data relating to decisions about/on behalf of users, and recommendations/information provided to users.[48]

ii. Controller and data subject

195. Obligations under the UK GDPR primarily fall on the data controller: the entity determining the means and purposes of processing (Article 4(7)). In most AAA contexts, this will be relatively straightforward. The data controller will be the provider (likely a corporate entity) of the AAA in question. A data controller must have some form of legal personality, so regardless of how advanced or 'autonomous' an AAA is, the model *itself* cannot be a data controller for the UK GDPR: the controller is the entity that is *providing* the AAA to the user. That provider determines the means (how the AAA works) and the purposes (the main one being to provide the relevant service) of the processing. The data subject is the user.

196. This may be complicated for some uses of AAAs which fall (from the user's perspective) outside the 'purely personal or household' exemption (Article 2(2)(a)) from the application of the UK GDPR. Where an individual uses an AAA for commercial purposes, he or she might *also* be a controller in relation to the processing of his *own* personal data (and potentially that of others). But this does not affect the core issue of whether the law would prevent harm or allow an individual to seek redress.

iii. Lawful processing

197. All processing of personal data needs a UK GDPR legal basis under Article 6. In our view the need for a legal basis is unlikely to be controversial or determinative of liability in most AAA situations similar to our scenarios. That is because the data controller – the AAA provider – has a direct relationship with the user who wants to use the AAA. It will therefore generally be straightforward for the processing of personal data involved to rely on the legal basis of consent (or contract).

198. Consent must meet a high standard in the UK GDPR, being informed, freely given and unambiguous. This is even more relevant to 'special category processing' under Article 9. That is, processing of data such as health data or that revealing political opinions, which is likely to arise in at least some AAA contexts. Such processing is more tightly constrained under the UK GDPR, but is lawful where the data subject *explicitly* consents to it.

[47] We addressed the application of the UK GDPR to AI harms in our 2023 paper and many of the same principles apply to this analysis. We have adapted and restated the analysis here.
[48] We are not considering the legal basis required for the training of the models underlying AAAs, since this is not a source of harm in the Scenarios presented to us.

199. It is of course possible that an AAA provider could fail to properly obtain (explicit) consent – for example by providing insufficient information to the user at the sign-up process. But we do not consider this to be particularly relevant to the harms in the four Scenarios. Issues regarding the information provided at sign-up, or the user interaction flow for obtaining consent can be easily remedied: they are not the true source of harm in the scenarios[49]. The reality is that there is high consumer demand for AAAs, and many – likely the vast majority – of users will consent to the processing involved in their use regardless of how much detailed information is provided at the point at which consent is obtained.

iv. *Fair processing*

200. Article 5(1)(a) UK GDPR requires all processing of personal data to be not only lawful and transparent but also *fair*. Historically, the principle of *fairness* has received less emphasis than others.[50] In practice, fairness has often been folded into arguments together with other principles, especially transparency, resulting in a meaning of limited 'informational' fairness: not processing data in unexpected ways.

201. There is some indication of recent development in this principle. The EDPB Guidelines 4/2019 on Data Protection by Design and Default[51] state:

> *'Fairness is an overarching principle which requires that personal data should not be processed in a way that is* ***unjustifiably detrimental****, unlawfully discriminatory, unexpected or misleading to the data subject.'* (emphasis added)

202. The EDPB Guidelines 2/2019 closely link fairness to the consideration of whether there is a power imbalance between data controller and data subject. And in the context of Binding Decision 02/2023[52] the EDPB stated:

> *'The EDPB underlines that the principles of fairness, lawfulness and transparency, all three enshrined in Article 5(1)(a) GDPR, are three distinct but intrinsically linked and interdependent principles that every controller should respect when processing personal data.'*

203. It could therefore be argued that processing of personal data by an AAA provider is unlawful where it is 'unjustifiably detrimental' to a user, even where full information on the AAA has been provided to the user. We explore the relevance of this in the four Scenarios above, but we emphasise that any arguments of this nature would be legally novel, ambitious and challenging: they would not be straightforward to enforce in a lower court. Nor would we expect strong enforcement from the regulator on the basis of such arguments.

v. *Accurate processing*

204. All processing of personal data is required to be 'accurate' pursuant to Article 5(1)(d) UK GDPR. Controllers are required to take 'every reasonable step' to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay. Accuracy is a relative standard, since what is a 'reasonable step' by a

[49] This can be distinguished from situations in which a faulty consent flow shows that processing has taken place which a data subject *would not have consented to* had they been given proper information about it.
[50] Ausloos, J. and Clifford, D. (2018). Data Protection and the Role of Fairness, *Yearbook of European Law*, 37, pp.130–187.
[51] https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf
[52] https://edpb.europa.eu/system/files/2023-09/edpb_bindingdecision_202302_ie_sa_ttl_children_en.pdf

controller to ensure accuracy depends on the circumstances.[53] Where processing is more consequential or risky – as it may well be in an AAA context – a higher standard of accuracy will be required.[54]

205. Traditionally, the concept of accurate processing would be applied to more 'static' data processing: for example, a record in a database stating that someone has characteristic *X* when in fact they have characteristic *Y*. The principle has proved to be flexible however, as can be seen by the increasing reliance on the accuracy principle in the UK GDPR in defamation claims, requiring a consideration of the 'meaning' of processing of personal data, and of the accuracy of that meaning.[55]

206. In some AAA contexts, it might be possible to argue that where an AAA functions poorly, by giving incorrect advice or making a bad recommendation, that would constitute 'inaccurate' processing. But as with the concept of fair processing, this is not an approach that has been endorsed by either the courts or a regulator, so may be challenging to rely on in practice for ordinary AAA users.

vi. *Automated decision-making*

207. The UK GDPR sets out a regime regulating certain decisions which are 'solely automated' and which have 'legal' or 'similarly significant' effects (Article 22). Whilst the presentation of content or recommendations to AAA users involves decisions, that cannot generally be said to have legal or similarly significant effects. Rather, it is the AAA user's *acting upon* recommendations that produces effects for him or her.

208. The decisions of AI *executors* (such as that in Scenario 2), which bind their users, may meet the threshold conditions in Articles 22A-C UK GDPR. We discuss the potential application of the UK GDPR's automated decision-making regime to Scenario 2 above.

vii. *Compliance requirements*

209. The UK GDPR imposes a range of compliance requirements on controllers, separate to the rights it gives to data subjects. Most relevantly:

i) Article 25 UK GDPR requires controllers to adhere to the principles of 'data protection by design and default' ('**DPDD**'); and

ii) Article 35 UK GDPR requires controllers to carry out prospective data protection impact assessments ('**DPIA**') to assess the harms that their processing may cause and how to mitigate them. This obligation only applies, however, where the controller has identified that the processing is likely to result in a high risk to the rights and freedoms of natural persons.

210. Where implemented well (or strongly enforced), these requirements could lead an AAA provider to foreseeing the kinds of harms in our Scenarios and avoiding them before they take place. But this dynamic is significantly limited by:

[53] E.g. *YSL v Surrey and Borders Partnership NHS Foundation Trust* [2024] EWHC 391 (KB).
[54] See also ICO Guidance: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-protection-principles/a-guide-to-the-data-protection-principles/the-principles/accuracy/
[55] See e.g. *Pacini v Dow Jones* [2024] EWHC 2714 (KB).

i) Regulation against these principles being relatively light touch (see section ix below); and

ii) DPIAs and DPDD relying on the possibility of foreseeing harms. A fundamental feature of AAAs is that they are dynamic and reactive, and often behave in unexpected ways, especially when used in diverse ways by users. This places real limits on the extent to which it is possible for even the most conscientious AAA providers to plan for and mitigate all the risks which might arise from AAA use.

viii. Data subject rights

211. As well as the right to be informed and the right of access (see §§27-32 above), data subjects have other rights in the UK GDPR, such as the right to rectification, erasure and the right to object. These may be relevant in an AAA context, and the ICO has emphasised that data rights apply in the context of generative AI, which underlies AAAs[56]. But data rights are typically more useful where processing is ongoing and the data subject wants it to stop, rather than for obtaining redress *after* processing has led to harm. For that reason, in our view, they have limited relevance – if any – to the scenarios here.

ix. Role and powers of the ICO

212. The Information Commission ('**ICO**')[57] regulates and enforces the UK GDPR. Its statutory powers are extensive: it can compel information from controllers, investigate premises, make orders for compliance and fine controllers up to £17.5m or 4% of worldwide turnover, whichever is higher. The ICO has information on (non-)compliance with the UK GDPR from:

i) Complaints by data subjects;

ii) Data breach reports by data controllers; and

iii) Its own investigations and intelligence.

213. Data controllers do not have an ongoing general duty to share information with the ICO about their processing, however.[58] It is also important to note that the ICO has no obligation to substantively resolve or 'rule on' complaints made to it by data subjects.[59] Rather, the ICO has significant latitude in deciding whether and how to use its enforcement powers.

214. There is some evidence[60] that the ICO uses its enforcement powers less than other data protection regulators in Europe. However, in the period covered by its most recent published annual report,[61] the ICO levied one fine of £12.7m on major social media platform *TikTok*, so major fines are not unheard of. In the ICO Generative AI Consultation, the regulator made it clear that:

[56] ICO Generative AI Consultation: https://ico.org.uk/about-the-ico/what-we-do/our-work-on-artificial-intelligence/response-to-the-consultation-series-on-generative-ai/engineering-individual-rights-into-generative-ai-models/

[57] Renamed to the Information Commission pursuant to the DUAA.

[58] Article 36 only requires consultation with the ICO if a controller identifies high risks of processing which *cannot* – in the controllers' view – be mitigated. These circumstances will be rare since they rely on the controller assessing its own processing as very risky.

[59] *Delo v The Information Commissioner* [2023] EWCA Civ 1141.

[60] Erdos, D. (2022). 'Towards Effective Supervisory Oversight? Analysing UK Regulatory Enforcement of Data Protection and Electronic Privacy Rights and the Government's Statutory Reform Plans', *University of Cambridge Faculty of Law Research Paper No. 16/2022*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4284602 and Erdos, D. (2025). 'The UK Information Commissioner's Annual Report 2024/25: Surveying a Systematic Trend Away from Adequate Enforcement', *UK Constitutional Law Blog*, https://ukconstitutionallaw.org/2025/07/22/david-erdos-the-uk-information-commissioners-annual-report-2024-25-surveying-a-systematic-trend-away-from-adequate-enforcement/

[61] https://ico.org.uk/media2/migrated/4030348/annual-report-2023-24.pdf

> *'Organisations acting as controllers must design and build systems that implement the data protection principles effectively and integrate necessary safeguards into the processing.'*

215. The ICO has finite resources and plainly cannot enforce against every breach of the UK GDPR. Its most recent strategy[62] suggests some focus on AAA-relevant areas:

> *'We will focus our efforts on areas such as the regulation of biometrics, facial recognition technology and the use of AI and algorithms.'*

216. But it remains to be seen in practice what this means for the enforcement of the UK GDPR to the extent it applies to AAAs. We are not aware of any specific enforcement action by the ICO in relation to the issues raised in the four Scenarios in this paper. In our view this is regrettable but not surprising since the ICO is more likely to focus on legally straightforward breaches (such as nuisance calls and data leaks) than on breaches that rely on nuanced and novel legal arguments regarding new technology. That may point to the need for a better-resourced regulator or a different system which allows these issues to be kept under proper oversight, but that is outside the scope of this analysis.

217. In summary, the ICO has significant powers in theory, but it is unclear how much its regulatory role influences providers of AAAs, since strong enforcement action in this area is somewhat unlikely.

B. Consumer protection

i. Digital Markets, Competition and Consumers Act 2024: unfair commercial practices

218. The Digital Markets, Competition and Consumers Act 2024 ('**DMCCA 2024**') replaces the Consumer Protection from Unfair Trading Regulations 2008.

219. Relevantly,[63] the DMCCA 2024 prohibits a range of unfair commercial practices (Part 4, Chapter 1 DMCCA 2024) by traders in relation to consumers (people acting outside their trade or business).

220. The focus of unfair commercial practices is on misleading actions, misleading omissions, and 'aggressive' sales practices that induce a consumer to make a 'transactional decision', which includes the purchase or supply of a product (s.245), which in turn includes a service or digital content (s.248).

221. Guidance from the Competition and Markets Authority (the '**CMA'**, which enforces the DMCCA 2024 – the '**CMA DMCCA 2024 Guidance**')[64] indicates that 'transactional decision' is significantly wider than merely making a purchase of a product or service. An example given of a transactional decision is whether to click through on a website. It would therefore appear that a user takes a transactional decision even when signing up for or deciding to use an AAA *entirely free of charge*.

222. Unfair commercial practices which may be relevant in an AAA context include:

---

[62] https://ico.org.uk/about-the-ico/our-information/our-strategies-and-plans/ico25-strategic-plan/

[63] We assume that none of the Scenario providers have 'strategic market status' under the DMCCA 2024, nor is any of them a dominant undertaking for the purposes of competition law.

[64] https://www.gov.uk/government/publications/unfair-commercial-practices-cma207?utm_medium=email&utm_campaign=govuk-notifications-topic&utm_source=968befc8-7fd4-40e4-b262-a21769c362e2&utm_content=immediately

i) Falsely claiming that a product can treat a mental health condition (para 19 Sch 20 DMCCA 2024); and/or

ii) Practices likely to cause the average consumer to take a transactional decision (i.e. sign up or not) to use the AAA which either:

(1) Are *misleading* (s.225(4)(i) and (ii)); or

(2) Contravene *'the requirements of professional diligence*', which requires a consideration of standards in the sector in which the AAA provider operates (s.229). This is less likely to constitute a ground of claim, since the novelty of AAAs means it is currently far from clear what 'good' or 'standard' commercial practice constitutes in their promotion and supply.

223. Where a consumer *enters into a contract* with a trader (which may not be the case for *all* AAA users), and an unfair commercial practice was a significant factor in that decision, the consumer has the right to withdraw from the contract, obtain an appropriate discount, and/or seek compensatory damages (ss.232-235). This right of redress is separate to the CMA's powers of enforcement (see section iii below).

224. The mere fact of harm being caused by an AAA will not necessarily mean that the provider has engaged in an unfair commercial practice. Much will turn on the extent (if any) to which an AAA is misdescribed or over-promoted to users, and the impact which the misleading claim(s) had on them. We apply the DMCCA 2024 regime to the four specific Scenarios above.

ii. *Consumer Rights Act 2015: services and digital content*

225. The Consumer Rights Act 2015 ('**CRA 2015**') gives consumers (defined in the same way as in the DMCCA 2024) rights vis-à-vis traders (again, as defined in the DMCCA 2024) when they enter into contracts for digital content[65] that is *paid for* (CRA 2015 Part 1, Chapter 3) and/or services (CRA 2015 Part 1, Chapter 4).

226. 'Service' is not defined explicitly in the CRA 2015, and an AAA could conceivably be either digital content, a service or both: the CRA 2015 contemplates mixed contracts (s.1(4)). In our view, AAAs which are paid-for fall within the CRA 2015 as either digital content, services or both. It may be difficult to conclusively say which parts of the CRA 2015 apply, but we consider both.

227. Conversely, the use of AAAs which are entirely free is unlikely to qualify for protection under the CRA:

i) As to digital content, Chapter 3, Part 1 of the CRA 2015 states that it only applies where such content is paid for.[66]

ii) As to services, Chapter 4, Part 1 of the CRA 2015 only applies to *contracts* for the supply of a *service*. Whilst there may be, for example, terms of use for a free AAA, in our view this could not be fairly described as the AAA provider having entered into a contractual obligation to provide any particular *service* (as opposed to providing digital content, which

[65] Any 'data which are produced and supplied in digital form' – s.2(9) CRA.
[66] CMA guidance clarifies that this does not include 'paying' with personal data: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/450440/Unfair_Terms_Main_Guidance.pdf.

is covered elsewhere in the legislation). This is underlined by the fact that it is doubtful whether the user of a free AAA provides any meaningful consideration, a pre-requisite for a service contract.

228. Relevantly for AAAs where it applies, the CRA 2015 requires:

i) *Services* to be provided with *reasonable care and skill* (s.49); and

ii) *Digital content* to be of *satisfactory quality and as described* (ss.34 and 36).

229. It is important to note that unlike with misrepresentation (covered below in Section E), the user need not show that the failure of the AAA to be 'as described' induced him or her to enter into a contract for digital services. The bar for demonstrating a breach is therefore lower.

230. As well as a refund or specific performance, consumers may be entitled to damages where these requirements are breached (ss.42 and 54).

231. Notably, AAA providers may not exclude liability for these requirements in their contractual terms (ss.47 and 57 CRA).

iii. Role and powers of the CMA

232. The DMCCA 2024 introduces a new regime for direct enforcement of the requirements discussed in the preceding two sections by the CMA (Part 3 and Sch 16 DMCCA 2024).[67] However, this is only possible where the relevant practice that is in breach of the CRA 2015 or DMCCA 2024 harms the *collective* interests of consumers in the UK (s.148(1) DMCCA 2024).

233. To the extent that AAA harms both:

i) arise from a commercial practice prohibited by the DMCCA 2024 or CRA 2015, *and*

ii) affect AAA users on a scale such that the collective interests of consumers are harmed,

then it will be possible for the CMA to take direct enforcement action against the AAA provider. Long-established case law sets out that the 'collective interests' test will be met where there is a harm to a 'section of the public'. The test may be deemed to be met where there are multiple individual cases of harm that may be 'added' together.[68]

234. The CMA's enforcement powers include full investigatory powers (s.208 and Sch 17 DMCCA 2024) and the ability to impose required actions on businesses, which may be enforced by the courts, and/or levy substantial fines (ss.191-196 DMCCA 2024).

235. This is a new enforcement regime and therefore it is difficult to say whether 'strong' regulation of consumer protection requirements relevant to AAAs can be expected. The CMA has published some guidance on its (planned) approach.[69] This guidance – and the CMA's current annual plan – indicates a focus on consumer confidence, innovation and growth.[70]

---

[67] Previously, the CMA had to apply to court for enforcement.
[68] *OFT v Miller* [2009] EWCA Civ 34.
[69] https://assets.publishing.service.gov.uk/media/6808ca0d8c1316be7978e74b/CMA_200_Direct_consumer_enforcement_guidance.pdf and https://assets.publishing.service.gov.uk/media/653f71b780884d0013f71cf4/CMA_Prioritisation_Principles_.pdf
[70] https://www.gov.uk/government/publications/cma-annual-plan-2025-to-2026/annual-plan-2025-to-2026#supporting-growth-opportunity-and-prosperity-for-the-uk

236. There is arguably a tension between pursuing innovation and growth whilst at the same time working to 'build confidence for people to engage with markets, empowered to make informed choices without being misled or exploited'. This is especially the case in relation to artificial intelligence innovation, such as AAAs, which have been placed very firmly at the centre of the UK government's agenda for growing the economy.[71]

237. The statutory ability for the CMA to strongly enforce consumer protection requirements on AAAs (where relevant) is therefore not in doubt. But given the emphasis on AI-driven growth in the UK, it remains to be seen how active the CMA will be in using its powers to uphold standards in relation to the kinds of services and harms in our four Scenarios.

iv. Other enforcers of consumer protection law

238. Part 3 DMCCA 2024 rationalises the regime for other bodies (mostly – but not only – public authorities) to enforce relevant infringements under consumer protection law by making an application to the court for an enforcement order.[72]

C. Advertising regulation

239. Advertising in the UK is regulated by consumer protection law, through the Consumer Protection from Unfair Trading Regulations 2008 and recently adopted DMCCA 2024 (covered in Section B above). This legislation is applied and expanded on in the UK Code of Non-broadcast Advertising, Sales Promotion and Direct Marketing ('**CAP Code**'). The CAP Code reflects the underlying legislation, in some cases goes further, and can therefore be relied on as a single source of regulation for advertising (along with more sector-specific or audience-specific guidance).

i. What content does the CAP Code apply to?

240. The CAP Code regulates *marketing communications* as opposed to regular editorial content. Marketing communications are broadly defined as any communication that is designed to promote the sale or use of goods, services, or products in exchange for commercial gain. Part 1 of the 'Scope of the Code' section of the CAP Code contains a non-exhaustive list of content that qualifies as a marketing communication. It notably covers 'advertisements in non-broadcast electronic media'. The CAP Code therefore applies to:

i) Advertising *of* AAAs, their functionalities, purpose, etc; and

ii) Advertising *through* AAAs for their own service or any other product or service.

241. The CAP Code only applies to advertisements on websites, apps and cross-border platforms (e.g. social media) if they are somehow linked to the UK – meaning one of the following: (1) the marketer (which includes an advertiser) is registered in the UK, (2) the advertisement appears on websites with a '.uk' top-level domain, or (3) the advertisement is paid-for and targeting people in the UK. This will apply for an advertising of the AAAs in our Scenarios.

ii. Whom does the CAP Code apply to?

242. Multiple actors can be held responsible under the CAP Code:

[71] https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan
[72] Although the CMA retains overall authority over the enforcement regime: see s.154 DMCCA 2024.

i) Marketers (those who pay for and authorise marketing communications) – they have primary responsibility for the truthfulness, fairness and honesty of their ads, and for the avoidance of serious or widespread offence

ii) Media agencies – they have an obligation to create marketing communications that are accurate, ethical and that neither mislead nor cause serious or widespread offence

iii) Publishers and media owners (and any other agent involved in producing, placing or publishing marketing communications) – they must ensure that they only disseminate communications that comply with the Code (if they have knowledge of, or control over, the ad)

243. Hence the developer of an AAA will be responsible for the compliance of any advertising *of the AAA* itself, and of any ads shown *through the AAA* if it plays any role in the selection of ads or their content.

iii. *Key obligations under the CAP Code*

244. The potentially relevant obligation under the CAP Code for our Scenarios is: **Do not mislead** (Rule 3.1) – marketing communications must not materially mislead or be likely to do so. Whether a communication materially misleads is determined by the *'transactional decision'* test. Marketing communications are misleading if they are likely to deceive consumers and to cause consumers to take transactional decisions that they would not otherwise have taken. A *'transactional decision'* is any decision made by a consumer relating to the purchase or supply of a product (including whether, how or on what terms to make the purchase or supply); the retention, disposal or withdrawal of a product (including whether, how or on what terms to retain or dispose of it); or the exercise of contractual rights in relation to a product (including whether, how or on what terms to exercise such rights) (CAP Code Appendix 1).

245. If the content of a user's interaction with an AAA is used to *target* ads, this is regulated by the UK GDPR (see Section A above). The use of personal data for targeting must notably be lawful (bearing in mind that the use of *sensitive* data is prohibited unless one of the conditions in Article 9(2) is met) and transparent.

iv. *Role and powers of the ASA*

246. The Advertising Standards Authority ('**ASA**') is the UK's independent advertising regulator, responsible for monitoring compliance with the CAP Code. It responds to complaints, proactively monitors ads, and forces changes or removal of ads. It does not have powers to impose penalties such as fines.

247. For AAAs this means:

i) Advertisements *of* AAAs in any media are subject to the same complaint, monitoring and enforcement processes as any other advertisements

ii) Advertisements *through* AAAs raise oversight challenges – they are not displayed to the public and therefore difficult for the ASA to proactively monitor. Any enforcement is therefore likely to come through individual complaints, unless the ASA proactively performs a systemic investigation into the sector.

248. Either way, no redress will be available through the ASA for an individual user who was misled by an advertisement.

D. Negligence, professional, and product liability

i. Negligence

249. Common law negligence – which imposes a duty not to inflict damage carelessly – requires a claimant to establish a **duty of care, breach, causation, foreseeability and loss**.[73]

250. When considering the nature and scope of any **duty of care**, the starting point is to consider the categories established by case law. It is well-established that a duty of care exists between, for example, a doctor and patient, a manufacturer and consumer, and a solicitor and their client. In novel situations where the existence of a duty of care has not been decided previously, the courts will determine whether to find such a duty by (i) analogy with existing categories, bearing in mind the nature of the relationship and the foreseeability of harm, (ii) weighing up the reasons for and against imposing liability to reach a conclusion that is just and reasonable, and (iii) considering wider policy issues (such as whether the matter is properly one for the legislature or whether the court should 'create new principles to fulfil a social need in an area of consumer protection where there [is] legislation').[74] A duty of care is more likely to exist where there is a clear *commercial* relationship – e.g. in the case of a paid-for AAA – as opposed to an AAA which is free for the general public to use.

251. The standard of care owed by the defendant to the claimant is assessed by reference to how a reasonable person in the position of the defendant should have acted in the circumstances. That is a highly fact-specific question for the court, which may pose particular difficulties in claims concerning AAAs. For instance, it may be difficult to assess whether there has been reasonable care in the programming or developing of the AI software because much of it has been developed using machine learning, which can self-reprogramme in a way that even the software's own programmers are unable to understand.[75]

252. Whether a defendant has fallen below the required standard of care is for the claimant to establish and a question of fact for the court. AI software – commonly depicted as a 'black box' which receives an input, produces an output, but does not readily disclose why or how particular decisions have been reached – poses particular difficulties in this context. Where an AAA output has 'gone wrong', that may be a (mal)function of the AI itself due to, for example, flawed initial design of algorithms or failure to preset suitable parameters for the AI's self-learning and self-programming. However, it is also possible that the AAA will have worked as intended by its programmers but produced an undesirable outcome due to its self-learning and self-reprogramming abilities. That self-learning may stem from incoming external data, which is processed by the AI's algorithms, which may themselves adapt and self-reprogramme in ways that are opaque – even to the programmers of the software as set out above. In these circumstances, it is likely to be extremely difficult for a claimant to establish a breach of duty.

---

[73] There is a general exclusionary rule against the recovery of 'pure economic loss' (economic loss to the claimant which does not result from any physical damage to or interference with his person or tangible property). To circumvent this rule, claimants need to show a special assumption of liability for such losses.

[74] *Michael v Chief Constable of South Wales* [2015] UKSC 2, §107.

[75] Yu, R. and Ali, G.S. (2019). 'What's Inside the Black Box? AI Challenges for Lawyers and Researchers', *Legal Information Management*.

253. In cases involving the operation of tools or machinery by the defendant – somewhat analogous to AAAs – the question for the court when deciding whether there has been a breach is whether those handling or controlling it are personally at fault, or whether the machines or tools themselves are defective). The general rule in those situations is that there is 'no liability unless a human being [operating a particular tool] ought to have spotted that something was wrong'.[76] Given the complexity, opacity and limited predictability of AAAs, and the fact that many are designed to operate autonomously and independently of human supervision, it is likely to be difficult for a claimant bringing such a claim to establish that a human being ought to have spotted that something was wrong with the tool. Where this is the case, there are likely to be two unsatisfactory consequences:

i) First, a claimant will either have to go uncompensated or try to invoke a head of non-fault-based liability such as product liability; and

ii) Second, it is possible that even if an error would give rise to damages when committed by a human, the same error will create no liability if effected through the operation of an AAA as a 'tool'.[77]

254. The 'black box' nature of AI is also likely to give rise to difficulty in establishing whether the AI factually and / or legally caused the damage (i.e. whether (i) *but for* the defendant's conduct, the damage would have occurred and (ii) the defendant's conduct is to be regarded as a cause in law, or whether something intervened between the conduct and the damage to break the chain of causation). Two particular difficulties are likely to arise in this context in AI-related claims:

i) The operation and the ability of AI to self-reprogramme is influenced by the conduct of, and data input by, its manufacturers, programmers, and users. The more actors that are involved, the more difficult it is likely to be to establish whether the conduct of one particular actor caused the loss to the claimant.

ii) Connectivity (for example via Bluetooth or internet) may give rise to vulnerability for malicious interference, which may further complicate the claimant's ability to establish the cause of their loss.

255. In order to recover damages, a claimant must prove that the type of injury or damage was reasonably foreseeable. Issues of foreseeability are likely to arise in relation to AI cases where there may be unintended outcomes flowing in particular from the ability of many AAA systems to self-reprogramme post-market, and the wide range of (unexpected) uses to which AAAs may be put by users.

ii. Professional liability

256. Professional liability refers to the liability arising when a professional fails to perform obligations owed to their client to a required standard.

[76] Soyer, B. and Tettenborn, A. (2022) 'Artificial intelligence and civil liability – do we need a new regime?', *International Journal of Law and Information Technology*, 30(4), pp.385-397.
[77] Ibid

257. Since AI does not possess a distinct legal personality[78], it is treated as a tool and professionals who use AI in the course of providing professional services to clients will retain legal responsibility for the professional services they deliver.

258. As well as any requirements from consumer protection legislation (see section B above), a professional contracted to deliver a service may also owe the client a concurrent duty of care in tort which arises from the relationship between the parties and the function which the professional person or firm is performing. This is likely to extend to liability for economic harm in many instances. In order to discharge both the contractual implied term and the tortious duty of care the professional must exercise reasonable skill and care, often defined as that 'which is ordinarily exercised by reasonably competent members of the profession, who have the same rank and profess the same specialism (if any) as the defendant'.[79]

259. It is well-established that solicitors will owe their client both a contractual and tortious duty of care. Under that duty, solicitors must exercise reasonable care and skill not only in relation to the underlying service being supplied but also in relation to the selection, deployment and validation of AI processes, as well as ongoing training. The standard of care expected is that of reasonable skill and care and is judged by the nature and scope of their retainer. In determining whether there has been a breach of that duty, a court will apply the standard of the reasonably competent solicitor and ask 'what the reasonably competent practitioner would do having regard to the standards normally adopted in his profession'.[80]

260. Legal professionals who use AI technology which malfunctions are likely to attract primary liability for any breaches of the duty owed to their clients. That is because an AAA should not be used to replace professional judgment.[81]

261. Where the use of an AAA in legal services is relatively novel there is unlikely to be a general and approved professional practice upon which to draw, nor universally recognised standards of testing and adoption of AAAs. In these circumstances the question whether or not a decision to use a particular AAA was taken with reasonable skill and care – which is for the court to determine – is likely to be particularly difficult to answer in practice. It is likely that, in assessing whether a reasonable, well-informed and competent member of that profession could have made the same error, a court will take into account factors including the steps taken by the individual or firm to assure itself that the AAA was appropriate and effective and to prevent any issues arising.[82] That large numbers of businesses may operate AI systems without adequate governance, oversight and testing is unlikely to give rise to a defence in the case of malfunctioning, inappropriate adoption or insufficient testing of a particular AAA.

---

[78] See from §365.

[79] *Jackson & Powell on Professional Liability,* 9th edn (London: Sweet & Maxwell, 2024), §2-193.

[80] *Midland Bank v Hett, Subbs & Kemp* [1979] Ch 384, §402.

[81] See *R (Ayinde) v The London Borough of Haringey* [2025] EWHC 1040 (Admin), §65: 'it would have been negligent for this barrister, if she used AI and did not check it, to put that [incorrect] text into her pleading', also cited in *R (Ayinde) v The London Borough of Haringey* [2025] EWHC 1383 (Admin).

[82] See SRA guidance 'Technology and Legal Services (November 2018)': 'Individual solicitors and firms are responsible for the service they give to people, including whether they use technology to advise clients or use it to work on client matters. … If there is an error or flaw in an AI system run, or provided by, a separate technology company then we are unlikely to take regulatory action where the firm did everything it reasonably could to assure itself that the system was appropriate and to prevent any issues arising. People will of course be able to seek redress in the usual way if they have suffered a loss or detriment, such as … making a negligence claim'.

That is because standard industry practice is itself capable of being negligent and 'a defendant is not exonerated simply by proving that others were just as negligent'.[83]

iii. Product liability

262. Part I of the Consumer Protection Act 1987 ('**CPA 1987**') imposes 'strict' liability for damage caused by a defective product. To succeed in a claim under the CPA 1987, the claimant must prove that (i) the product was defective; (ii) the claimant has suffered damage; and (iii) a causal link between the defective product and the damage suffered. There is no requirement for a claimant to prove fault on the part of the producer.

263. Under section 1(2) CPA 1987, a 'product' is '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'. Goods include 'substances, growing crops and things comprised in land where virtually attached to it and any ship, aircraft or vehicle'.[84] This definition clearly includes tangible goods, but is unlikely to extend to "pure information" such as computer software unless that software is supplied as an intrinsic part of a tangible medium i.e. hardware.[85] Where software is supplied non-physically, such as over the internet or by electronic download, it is likely that no 'product' is involved and there can be no liability under the CPA 1987.[86] It is likely that a claimant seeking to classify an AAA as a 'product' will face similar challenges.[87] We include further detail on the CPA since it is possible that the UK will – in future – extend this regime to software/AI.[88]

264. Under section 3(1) CPA 1987, a product is defective if:

> *'the safety of the product is not such as persons generally are entitled to expect; and for these 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'.*

265. Section 3(2) CPA 1987 provides guidance on what 'persons generally are entitled to expect' under section 3(1). All the circumstances are to be taken into account in addressing this question, including

> *'(a) the manner in which and purposes for which, the product has been marketed, its get-up, the use of any mark in relation to the product and any instructions for, or warnings with respect to, doing or refraining from doing anything with or in relation to the product.*
>
> *(b) what might reasonably be expected to be done with or in relation to the product.*
>
> *(c) the time when the product was supplier by its producer to another;*

---

[83] *Martlet Homes Ltd v Mullaley & Co Ltd* [2022] EWHC 1813 (TCC); 203 Con L R 125. See also *Jackson & Powell on Professional Liability,* 9th edn (London: Sweet & Maxwell, 2024), §2-179.

[84] Section 45 CPA 1987.

[85] There is no case law directly on this point but see the UK Government's 2001 guidance on the CPA 1987 at §11 (https://webarchive.nationalarchives.gov.uk/ukgwa/20121206081114/http:/www.bis.gov.uk/files/file22866.pdf). See also: the definition of 'goods' in section 61(1) Sale of Goods Act 1979; and caselaw which provides that software does not constitute 'goods' for the purposes of the Sale of Goods Act 1979: *St Albans DC v International Computers Ltd* [1996] 4 All ER 481.

[86] *Clerk & Lindsell on Torts,* 24th edn (London: Sweet & Maxwell, 2024), §10-51.

[87] MHRA Guidance: 'Impact of AI on the regulation of medical products' (2024).

[88] The Law Commission has announced its intention to review the product liability regime 'particularly in light of emerging technologies': https://lawcom.gov.uk/news/law-commission-to-review-the-law-relating-to-product-liability/.

And nothing in this section shall require a defect to be inferred from the fact alone that the safety of a product which is supplied after that time is greater than the safety of the product in question'.

266. The concept of a defect is defined in terms of failure of the product to meet an objective standard of safety that the court must evaluate. The test of whether a product is defective is whether the safety of the product is not such as persons generally are entitled to expect (not of what is actually expected). It is an objective test.[89] A defect for the purposes of section 3 CPA 1987 cannot be an inherently harmful characteristic, where that characteristic is part of the normal behaviour of the product. Instead, a defect is the 'abnormal potential for harm' which elevates the underlying risk of the product beyond the level of safety that the public is entitled to expect from a product of that type.[90]

267. There can be no entitlement to an absolute level of safety in relation to a product.[91] What persons generally are entitled to expect in relation to the safety of a particular product is to be assessed having regard to all the circumstances relevant to the evaluation of safety, including the matters identified in section 3(2) CPA 1987.[92] Comprehensive instructions and warnings may render safe products that come with potential dangers.

268. In determining whether a product met the levels of safety persons generally were entitled to expect, the court may have regard to everything now known about it that is relevant to that enquiry, irrespective of whether that information was available at the time it was supplied.[93]

269. In general, the courts will maintain a flexible approach to the assessment of the appropriate level of safety and the balance between the risks of a product and its benefits (other than those directly related to safety) may be relevant, depending on the nature and seriousness of the risk and the likelihood of its manifestation.[94] Cost may also be a relevant factor in appropriate cases.[95]

270. While it is for the claimant to establish that a product was defective, it is not necessary to be able to ascertain the precise cause of the defect.[96] In the context of an AAA where the precise cause of an AAA's malfunction is unlikely to be discernible even to the original programmers, this is likely to prove helpful to claimants, if product liability under the CPA 1987 applies.

271. Under sections 1(2) and 2(2) CPA 1987 primary liability is imposed on

i) The producer of the product (normally the manufacturer);

ii) Any person who, by putting his name on the product or using a trade mark or other distinguishing mark in relation to the product, has held himself out to be the producer of the product; and/or

iii) Any person who has imported the product into the UK in order, in the course of any business of his, to supply it to another.

---

[89] *Hastings v Finsbury Park Orthopaedics Ltd* [2022] UKSC 19, §15(i)-(ii).
[90] *Gee v DePuy International Ltd* [2018] EWHC 1208 (QB), §112.
[91] *Hastings v Finsbury Park Orthopaedics Ltd* [2022] UKSC 19, §19.
[92] *Hastings v Finsbury Park Orthopaedics Ltd* [2022] UKSC 19, §15(iii); *Gee v DePuy International Ltd* [2018] EWHC 1208 (QB), §112.
[93] *Hastings v Finsbury Park Orthopaedics Ltd* [2022] UKSC 19, §15(iv).
[94] *Gee v DePuy International Ltd* [2018] EWHC 1208 (QB), §152.
[95] *Wilkes v DePuy International Ltd* [2016] EWHC 3096 (QB); [2018] QB 627, §83.
[96] *Ide v ATB Sales Ltd* [2008] EWCA Civ 424.

272. The supplier of the product to the injured person may attract secondary liability unless it can, within a reasonable time, name any of the persons who have primary liability.[97]

273. The six statutory defences to claims brought under the CPA 1987 are found at section 4(1):

i) The defect is attributable to compliance with any requirement imposed by UK or retained EU law; or

ii) The defendant did not at any time supply the product to another; or

iii) The supply by the defendant was otherwise than in the course of a business of that person's; or

iv) The defect did not exist in the product at the time of its supply; or

v) The state of scientific and technical knowledge at the time when the product was supplied was not such that a producer of products of the same description could be expected to discover the defect in its products if it had existed in his products while they were under his control; or

vi) The producer only produced a component of the product in which the defect arose, if that defect was wholly attributable to the producer of the subsequent product.

274. Under section 5 CPA 1987 damages, compensatory in nature, are available only for death, personal injury or damage to private property in excess of £275.

E. Misrepresentation and Contract

i. Misrepresentation

275. A claim for misrepresentation arises where one party to a contract (the 'representor') made a false statement of fact that induced the other party (the 'representee') to enter into the contract thereby causing that party loss. All of our four Scenarios involve (at least implicitly) the provider of the AAA making certain claims or statements about the AAA's function or abilities. Two types of misrepresentation are likely to be relevant:

i) Negligent misrepresentation, which under section 2(1) Misrepresentation Act 1967 occurs where a false statement is made by one contracting party to another carelessly or without reasonable grounds for believing its truth. The test is an objective one and, once the representee proves that the statement was in fact false, the burden of proof shifts to the maker of the statement to establish that they had reasonable grounds to believe and did believe up to the time the contract was made that the facts represented were true.

ii) Innocent misrepresentation, which refers to a representation made without fault, where the maker of the statement can show that they had reasonable grounds to believe that their statement was true.

iii) 'Puffs' or sales talk (where, for example, a statement is in such general terms as to be unverifiable e.g. 'this is one of the best cars you can buy') cannot form the basis for a case in misrepresentation.[98]

[97] Section 2(3) CPA 1987.

[98] *Kingspan Environmental Ltd v Borealis A/S* [2012] EWHC 1147, §420.

276. A statement of fact will not be false if it is substantially correct.[99] This is for the court to determine by reference to how the words would be understood by a reasonable person in the factual context.[100]

277. The statement must have *induced* the representee to enter the contract. It need not have been the sole cause of their entering the contract but in most cases, they will have to establish that they would not have entered the contract (at all or on the same terms) 'but for' the misrepresentation.[101] This requirement presents a key challenge for potential claimants and renders misrepresentation difficult to prove in practice.

278. A claimant cannot recover damages for any part of their loss that they could have avoided by taking reasonable steps.

279. The wording of section 2(1) Misrepresentation Act 1967 has the effect that, where the claimant can establish (i) sufficient causal link between the misrepresentation and the loss suffered, (ii) that the loss suffered is a direct consequence of the transaction induced by the misrepresentation and (iii) the loss could not reasonably have been avoided, they will be able to recover in damages all losses directly flowing from having entered into the contract as a result of the misrepresentation, whether or not the loss was reasonably foreseeable, and including consequential losses (both financial and distress / inconvenience).

ii. *Contract*

280. It is possible for a representation to become a term of the relevant contract (if, for example, the representation is included in the written contract). In addition, contracts for the supply of AI tools such as an AAA are very likely to include express or implied terms as to quality, fitness for purpose and coherence with their description. In consumer contracts for digital content, the latter are likely to be implied by virtue of the operation of sections 34-36 of the Consumer Rights Act 2015.[102]

281. In a claim for breach of a contractual term, it is necessary for the claimant to prove that the breach caused their loss both factually (i.e. *but for* the breach, they would not have incurred the loss) and legally (i.e. that there was no intervening act in between the breach and the loss to break the chain of causation). A claimant cannot look to recover losses they would have sustained in any event.[103] Further, a loss will only be recoverable if that type of loss was *foreseeable* at the date of contracting and, as with misrepresentation, the claimant has a *duty to mitigate*: if the claimant unreasonably fails to act to mitigate their loss, the law treats those (in)actions as having broken the chain of causation and measures damages as if the claimant had instead acted reasonably.[104]

282. The principles of causation, foreseeability and mitigation are applied in contexts other than breach of contract when considering what damages (if any) can be recovered in the case of a breach of a relevant obligation. This is likely to be a significant issue for redress for AAA harms and we highlight where it is relevant for the Scenarios above.

---

[99] *Avon Insurance Plc v Swire Fraser Ltd* [2000] 1 All ER (Comm) 573.
[100] *IFE Fund SA v Goldman Sachs International* [2006] EWHC 2887 (Comm); [2007] 1 Lloyd's Rep 264, §50, affirmed in [2007] EWCA Civ 811; [2007] 2 Lloyd's Rep 449.
[101] *SK Shipping Europe Plc v Capital VLCC 3 Corp* [2022] EWCA Civ 231; [2022] 2 All ER (Comm) 784, §61.
[102] See section B above.
[103] *Tiuta International Ltd v De Villiers Surveyors Ltd* [2017] UKSC 77.
[104] *Sharp Corp Ltd v Viterra BV* [2024] UKSC 14, §85.

283. Where a claimant can establish that their loss (i) was factually and legally caused by the breach, (ii) was foreseeable and (iii) could not have been avoided, they may be able to obtain damages for:

i) financial loss (including costs or liability the claimant has incurred to a third party but would not have incurred but for the breach, and profits the claimant has forgone).

ii) Personal injury (whether physical or psychiatric).

iii) In certain limited circumstances, mental distress and loss of amenity (falling short of personal injury) where 'a major or important object of the contract was to give pleasure, relaxation or peace of mind'.[105]

284. It is therefore necessary, for any AAA harm, to consider whether there is a valid contract in place between the user and the AAA provider and – if so – what the terms of the contract are and what representations (if any) induced the user to enter into it.

iii. *Challenges in bringing contractual claims in an AAA context*

285. Certain difficulties are likely to arise in the context of AAA-related claims in contract.

286. Given the complexity, opacity and limited predictability of AI software and its 'black box'[106] nature, it is likely to be difficult to establish that an undesirable outcome is a product of a contractual breach on the part of the AAA developer. Even if the emergent properties of an AAA are largely a function of the provider's decisions, the role of self-learning and user prompts significantly complicates demonstrating liability in practice. Indeed, it may be difficult to detect that something has "gone wrong" in relation to an AAA at all.

287. Where AI-related risks specific to the actual subject matter of the contract have not been addressed and a contract involving AI relies on generic obligations of suppliers such as those relating to quality or fitness, a significant challenge for the court is likely to be the assessment of whether a malfunction or unanticipated outcome of an AI process actually amounts to a breach of contract or suggests a misrepresentation. That challenge is particularly acute where contracting parties understand from the outset that a tool's output may in practice be unpredictable and unexplainable.

288. The **absence of universally recognised and accepted standards for AAAs** means that the court's assessment of quality and fitness (likely terms in a contract covering the use of an AAA) is likely to be uncertain and very challenging. That an AI tool appears to work poorly or inconsistently will not necessarily mean that the tool does not meet its description or is of unsatisfactory quality. An assessment of whether the supplier of an AAA has breached quality and fitness requirements will require more than an observation of the AAA's output, but will instead require analysis of how the AAA was built or of what (if any) verification was taken of its efficacy, safety and behaviour. Whether an AAA is of satisfactory quality may depend not just on its results but also on how it has been built and the extent to which the assumptions and models governing its design have been validated.

---

[105] *Farley v Skinner* [2001] UKHL 49.

[106] AI software is commonly depicted as a 'black box' which receives an input, produces an output, but does not readily disclose why or how particular decisions have been reached. See for example
Yu, R. and Ali, G.S. (2019). 'What's Inside the Black Box? AI Challenges for Lawyers and Researchers', *Legal Information Management*.

289. Finally, contracts for the supply of AI products may contain supplier- or developer- favoured allocations of risk. Businesses supplying AI are likely to attempt to protect themselves from potential liability by including a provision excluding liability for defective AI.

290. These issues are acute challenges for AAA users seeking redress, because it is the user bringing a claim who bears the burden of proof – that is, to bring enough evidence to bear to convince the court on the balance of probabilities that the breach has occurred.

F. Human rights

i. Framework of rights and obligations

291. The UK's human rights framework is relevant to AAAs in two main ways:

i) Governments being the primary duty-bearers under international human rights law, they are responsible for regulating AAAs and preventing harm from being caused by anyone who develops and deploys AAAs. In this context, under the Human Rights Act 1998 ('**HRA 1998**') (which domesticates the European Convention on Human Rights or '**ECHR**') the UK state has (a) a negative obligation to respect human rights if they themselves develop or use AAAs, and (b) a positive obligation to protect individuals from harm caused by AAAs, including by non-state actors.

ii) Companies that develop or deploy AAAs ought to follow the United Nations Guiding Principles on Business and Human Rights ('**UNGPs**'). While these aren't legally binding, they may be reflected in certain agreements or used as an influential standard when deciding whether a company has met its legal obligations.

292. All of the rights recognised in the HRA 1998 can be impacted by the use of AAAs – the very nature of AAAs is such that they could foreseeably be used – and therefore affect people and their rights – in all kinds of contexts.

293. When AAAs are used in the delivery of public services, states' obligation is to ensure that their own use (and that of anyone who acts on their authority) does not violate human rights. This may arise, for example, if a public authority relies on AAAs for making decisions that affect people's rights in a discriminatory way – this could be a breach of the right to non-discrimination (Article 14 ECHR) – or uses AAAs that search for and analyse private information about individuals – this could a breach of the right to private and family life (Article 8 ECHR).

294. When provided by companies and used by private individuals, any *legal* human rights obligations will be much more remote and difficult to establish. However an AAA could cause harm that engages an individual's human rights, e.g. the right to life (Article 2 ECHR), the prohibition of discrimination (Article 14 ECHR), or the freedom of expression (Article 10 ECHR, which includes the freedom to hold opinions and to receive and impart information and ideas without interference by public authority). If such harm was foreseeable, and could have been avoided through state regulation and enforcement against AAAs, the state could be held to have violated the individual's rights for having failed to protect them against harm.[107]

295. A number of human rights monitoring and enforcement mechanisms exist that could be leveraged in case of human rights impact of AAAs. Court proceedings in the UK could reach

[107] *O'Keefe v Ireland* (2014) 59 E.H.R.R. 15, *Carr v G4S Care and Justice Services (UK) Ltd* [2022] EWHC 3003 (KB).

the ECtHR if the UK government can arguably be held responsible for a lack of protection. International quasi-judicial mechanisms could also be mobilised, such as individual communications at the UN Treaty Bodies (e.g. the Human Rights Committee or Committee on Economic, Social and Cultural Rights), or complaints to the National Contact Points (NCPs) for the OECD Guidelines on Multinational Enterprises.

296. Ultimately however, the human rights avenue is limited in two ways. First, the threshold for finding state responsibility is very high (likely requires the state being aware of a 'real and immediate risk' of harm[108]) and untested in this context, such that prospects are low and uncertain. Second, this is not an avenue that can provide meaningful protection to ordinary users of AAAs either today or in the near future. It can, however, be a strategic avenue to establish state responsibility for further regulation.

297. Bringing a human rights challenge can be done by way of judicial review or through a civil claim under the HRA 1998 and civil procedure rules. It is possible to challenge a lack of legislation or regulation, but only if it can be shown that a public body had a legal duty to create specific legislation or regulations, of if the failure to legislate leads to an unlawful outcome (such as incompatibility with ECHR rights).[109]

ii. *Role of the Equality and Human Rights Commission*

298. The Equality and Human Rights Commission ('**EHRC**') is tasked by the Equality Act 2006 with:

> *'encouraging and supporting the development of a society in which–*
>
> *(a) people's ability to achieve their potential is not limited by prejudice or discrimination,*
>
> *(b) there is respect for and protection of each individual's human rights,*
>
> *(c)there is respect for the dignity and worth of each individual [...]'*

299. As well as having general powers to undertake research and publish guidance (§13-19 EA 2006), the EHRC is empowered to:

i) Carry out investigations into whether someone has breached equality law (s.20) including a power to compel information (s.20 and Sch 2);

ii) Enter into binding agreements with entities regarding compliance (s.23);

iii) Issue a notice requiring a person who has committed an unlawful act under the EA 2006 to prepare an action plan, or recommending action to be taken (s.21); and

iv) Institute or intervene in judicial review proceedings (s.30).

300. But these enforcement powers are relatively weak: the EHRC has no ability to levy fines or directly enforce its rulings on affected bodies. It must instead apply to the court to secure compliance with an action plan (s.22) or for an injunction to restrain an unlawful act (s.24). The EHRC appears to use its powers sparingly: there are only six inquiries and investigations listed on its site since 2020[110] and most of these relate to equality law rather than human

[108] *Devall v Ministry of Justice* [2022] EWHC 1608 (QB).
[109] *Re Northern Ireland Human Rights Commission's Application for Judicial Review* [2018] UKSC 27.
[110] https://www.equalityhumanrights.com/our-work/inquiries-and-investigations

rights issues. Whilst the current business plan mentions AI[111], none of the priority issues in it are directly relevant to the harms in the four Scenarios in this paper.

301. Overall, the EHRC is a relatively weak regulator with limited resources, and – given that any substantive breaches of human rights law in these Scenarios are somewhat speculative – we would not expect it to play a significant role in protection from the harms in the Scenarios.

G. Online Safety Act

i. Scope

302. The Online Safety Act 2023 ('**OSA**') primarily regulates 'user-to-user' and 'search' services[112]. Whilst there is some overlap between both of these and AAAs, close analysis is needed to determine whether AAAs are in scope.

303. Leaving aside the sharing of content generated using AAAs, the tools *per se* are unlikely in our view to be regulated user-to-user services under the OSA. The definition in s.3(1) provides that a user-to-user service:

> '*means an internet service by means of which content that is generated directly on the service by a user of the service, or uploaded to or shared on the service by a user of the service.*' (emphasis added)

304. AAAs – especially those which draw on retrieval augmented generation ('**RAG**') – will often draw on content from other *internet* users in presenting content and recommendations. But they do not generally present content which has been generated *on* or uploaded *to* the AAA itself by other AAA users. Indeed, it is not possible to 'upload' content to most AAAs in this sense, except perhaps through some kind of deliberate prompt injection, which is outside the scope of normal use of the AAA and therefore of this analysis[113].

305. In some respects, AAAs might be considered to be 'search' services, where their responses are drawing on, collating and synthesising content from the web (i.e. using RAG). Section 3 OSA defines a search service as a service which is or includes a search engine, which (s.229(1)):

> *'includes a service or functionality which enables a person to search some websites or databases (as well as a service or functionality which enables a person to search (in principle) all websites or databases).'*

306. This could include an AAA using RAG to include web results in its responses. Ofcom's guidance creates some uncertainty about this, however. In the context of setting thresholds for more tightly regulated services with large numbers of users, Ofcom categorises 'general search services' as those which:

i) use bots to crawl content across the web;

ii) build an index of URLs; and

iii) use algorithms to rank results.

[111] https://www.equalityhumanrights.com/about-us/our-strategy/our-business-plan/business-plan-2024-2025#id-3regulatingartificialintelligenceandtacklingdigitalexclusion

[112] Services which make pornographic material available are also regulated, and this would include AAAs.

[113] Certain services *combine* AAA-like functionality and the sharing of user-generated content, such as the sharing of customised 'characters' which rely on LLM models. Ofcom has indicated that such services would be regulated as user-to-user services under the OSA.

307. This more detailed definition with an emphasis on indexing and ranking could suggest that AAAs using RAG are *not* regulated search services within the meaning of the OSA: the position is somewhat unclear.

308. In an open letter however, Ofcom stated that "*Generative AI tools that enable the search of more than one website and/or database are 'search services' within the meaning of the Act.*"[114] In our view, there is some doubt about this categorical assertion:

i) A service is a 'search service' under the OSA if it either (a) *is*, or (b) *includes* a 'search engine' (s.4; emphasis added).

ii) A 'search engine' "*includes a service or functionality which enables a person to search some websites or databases*" (s.229(1)). The second 'includes' here has the potential to cause confusion: in our view it is an aide to interpretation of the term 'search engine'. It does *not* mean that any service which "*includes*" the described functionality *is* a search engine. Such a reading would render the second limb of s.4. redundant.

309. So, an AAA may be a 'search service' if it *is* a search engine. That seems to be what Ofcom argues in its open letter: AAAs are search services because they *are* search engines.

310. In our view however, this would only rarely be true. Whilst an AAA might *include* a search engine, in ordinary parlance it is not *itself* a search engine: it is a broader service which includes search functionality. Indeed, this would seem to be precisely the kind of service envisaged by the second limb of the s.4 definition: a service which *includes* a search engine.

311. Here, we must turn to s.229(2), which provides:

> *"For the purposes of this Act, a search engine is not to be taken to be "included" in an internet service or a user-to-user service if the search engine is controlled by a person who does not control other parts of the service."* (emphasis added)

312. If it were true that any internet service which includes search functionality *is* a search engine, then the saving provision in s.229(2) would have no application.

313. Many AAAs, where they enable the search of other websites, rely on integrations with existing search engines. That is, search engines which are controlled by another person. To determine whether an AAA is a 'search service' under the OSA, it may therefore be necessary to analyse the specific way in which it enables search, and whether this involves the inclusion of s search engine which the AAA provider itself controls.

314. As can be seen, this issue depends on rather fine-grained questions of statutory interpretation. But in our view, despite Ofcom's open letter, the issue is not settled and would need to be carefully assessed for each AAA on a case-by-case basis.

ii. Limited substantive requirements

315. For the four Scenarios presented to us, whether or not the OSA is engaged is likely to be moot, since the relevant *substantive* requirements of the OSA[115] extend only to (i) illegal

[114] https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/open-letter-to-uk-online-service-providers-regarding-generative-ai-and-chatbots

[115] We assume none of the Scenario providers have more than 7 million users in the UK on their 'search' service (if any), meaning none is categorised as a Category 2A large search service. Although given the OSA's substantive scope, even if an AAA *were* a Category 2A service, that would not affect our analysis of the four Scenarios here.

content, and (ii) content which is harmful to children, where the service is likely to be accessed by children. Neither of these issues is raised by the four Scenarios. The OSA creates no obligations on AAA providers in relation to manipulation or undue influence of adult users.

316. Thus, the OSA could be relevant to AAA regulation generally, especially if it is accepted that an AAA constitutes a 'search service' but has no relevance to the harms in this paper.

H. Regulation of medical devices

317. The Medical Devices Regulations 2002 ('**MDR 2002**') (as amended) govern medical devices in the UK and contain provisions for ensuring safety, effectiveness, and oversight.[116] The Medicines and Healthcare products Regulatory Agency ('**MHRA**') is responsible for regulating the UK medical devices market and enforcing the MDR 2002.[117]

318. The MDR 2002 defines as a 'medical device' (and therefore captures within its remit) any instrument, apparatus, appliance, software, material or other article, whether used alone or in combination, which is intended by the manufacturer to be used for the purpose of diagnosis, prevention, monitoring, treatment or alleviation of disease, injury or disability or for the investigation, replacement or modification of the anatomy or of a physiological or pathological process or state.[118] The definition of software includes standalone software[119] and AI/AAAs[120].

319. Whether or not a product is considered to have one of the purposes set out above ('medical purpose') is defined by the manufacturer's intention for the product as set out in their labelling, instructions for use and promotional material, technical documentation and the product's mode of action in conjunction with the definition of a medical device as set out in the MDR 2002. Off-label use of a product for a medical purpose will not necessarily make a product a medical device if that use is not in accordance with the manufacturer's intention for the product as defined above.[121]

320. Digital and software products that support mental health and wellbeing – which can include websites, internet-based platforms or applications (apps) to be used with non-medical technology such as computers, mobile phones and fitness wearables – are considered to have a medical purpose if they are intended as part of the broad process involved in the management of mental ill-health. This includes products that

    i) Help with assessing risk, diagnosing, predicting, monitoring, treating or preventing mental health conditions and/or symptoms;

    ii) Where the conditions and/or symptoms are at levels considered diagnosable or clinically relevant; and

    iii) Are either intended for patients, the public and/or healthcare professionals.[122]

[116] The MDR 2002 transposed three European Union medical directives (the Medical Devices Directive (93/42/EEC), Active Implantable Medical Devices Directive (90/385/EEC) and in vitro Diagnostic Medical Devices Directive (98/79/EC) into UK law.
[117] The MHRA is an executive agency of the Department of Health and Social Care in the United Kingdom which is responsible for regulating the UK medical devices market and enforcing the MDR 2002. The Medicines and Medical Devices Act 2021 introduced a set of enforcement provisions for this purpose.
[118] MDR 2002, regulation 2.
[119] MHRA Guidance: 'Borderlines with medical devices and other products in Great Britain' (2025).
[120] MHRA Guidance: 'Impact of AI on the regulation of medical products' (2024).
[121] MHRA Guidance: 'Off-label use of a medical device' (2023).
[122] MHRA Guidance: 'Digital Mental Health Technology – Regulation and Evaluation for Safe & Effective Products' (2025).

321. An AAA would be considered to have a medical purpose where it is intended to target symptoms (whether at clinical or sub-clinical levels) as part of a diagnosable clinical condition (such as generalised anxiety disorder). By contrast, where a product is intended for general wellbeing, and targets symptoms at a sub-clinical level rather than clinical levels as part of a diagnosable mental health condition, that would not be considered to have a medical purpose.[123] Such a product would by extension fall outside the remit and the safety, efficacy and other requirements of the MDR 2002.

322. The government has announced its intention to introduce a suite of new regulations for medical devices, to supplement the MDR 2002 – which contain few provisions specifically aimed at regulating software or AAAs as medical devices – that 'prioritise patient safety, give patients access to the medical devices they need and ensure the UK remains an attractive market for medical technology innovators'.[124] The MHRA has also published an updated roadmap setting out a change programme for regulating software and AI/AAAs as a medical device.[125] The anticipated changes are broad in scope and have attracted positive comment from stakeholders.[126] However, there is nothing to suggest that they will alter the regulatory position of software and AI which does not have a medical purpose but which is instead intended for general wellbeing. This is likely to continue to fall outside the regulatory regime for medical devices with the consequence that AAA users may not benefit from the regime's enhanced protections.

I. Financial services regulation

i. Scope

323. The provision of 'financial services' in the UK is primarily regulated under the Financial Services and Markets Act 2000 ('**FSMA**') and more detailed rules promulgated by the regulator, the Financial Conduct Authority, under that act ('**FCA Rules**'). FSMA regulates '*specified*' activities in relation to '*investments*' when they are '*carried on by way of business*'. The definitions of 'specified' and 'investments' are set out throughout FSMA and the FCA Rules and do not require detailed discussion here. The terms are generally construed relatively broadly and include advice, concluding contracts, making other arrangements related to deals in investments, etc.

ii. Threshold: advising on investments by way of business

324. Providing advice on investments is a regulated activity. This raises the question of whether personalised information, advice or recommendations produced by an AAA fall to be regulated under FSMA, which would mean that the AAA provider requires authorisation under that act.

[123] MHRA Guidance: 'Digital Mental HealthTechnology – Regulation and Evaluation for Safe & Effective Products' (2025). See also MHRA Guidance: Medical device standalone software including apps (2023) which states that 'apps and software that are intended to treat non-medical conditions e.g. non-specific stress' are unlikely to be medical devices.
[124] See https://www.gov.uk/government/publications/implementation-of-the-future-regulation-of-medical-devices/implementation-of-the-future-regulations#futurecore-regulations (accessed 27.06.25).
[125] MHRA, 'Medical Devices Regulatory Reform: Road to Implementation' version 2.0 (2024).
[126] See for example Regulatory Horizons Council (2022). *The Regulation of Artificial Intelligence as a Medical Device*, https://www.gov.uk/government/publications/regulatory-horizons-council-the-regulation-of-artificial-intelligence-as-a-medical-device.

325. Article 53 of The Financial Services and Markets Act 2000 (Regulated Activities) Order 2001 ('**RAO**') further delineates when financial advice will be regulated. To fall under FSMA, it must be

i) Given to a person in their capacity as an investor or potential investor;

ii) Relating to a specific investment or class of investments; and

iii) Personalised or presented as suitable for that person[127].

326. It is clear that the fact that advice is mediated through software is no barrier to it meeting the FSMA definition. In *Market Wizard Systems (UK) Ltd, Re* [1998] 7 WLUK 260 it was held that the developer and marketer of a relatively simple (by modern standards) computer programme which automatically generated 'buy/sell' recommendations for stocks on a daily basis was providing regulated investment advice which required authorisation under the forerunner regime to FSMA. The FCA's Perimeter Guidance Manual[128] ('**PERG**') provides guidance on when authorisation may be required and underlines this at 8.30.2(6).

327. PERG at [8.24] makes clear that generic or factual information (e.g. 'Shares are riskier than bonds') is *not* regulated advice. Nor is non-personalised commentary (e.g. 'Gold may be a good hedge against inflation'). But existing AAAs can and do provide more personal recommendations based on detailed inputs from users which suggest that specific investments are suitable for them.[129]

328. Yet providers of well-known AAAs in the UK are not authorised by the FCA, and we are not aware that the FCA has raised concerns about this publicly, despite clearly being up-to-date with how AI is being applied in the financial services context (see section vi below). We can speculate as to some arguments *against* financial advice delivered through existing AAAs being caught by FSMA:

i) AAA terms of use typically state that their outputs may not be accurate, and that they should not be relied upon in substitute for professional advice; and

ii) To the extent AAAs are giving investment advice, this is a small percentage of their output and incidental to the main business model of the providers. Certainly, generic AAAs are not marketed explicitly as investment tools or financial advice generators. This means that it could be argued that any advice given is not 'by way of business' as required by s.22 FSMA.

329. The second of these arguments is more compelling and more consistent with PERG, which provides at [8.24.2]:

> *'In the FCA's view, for a person to be carrying on the business of advising on investments or making arrangements with a view to transactions in investments, he will usually need to be*

[127] See e.g. *Rubenstein v HSBC Bank plc* [2011] EWHC 2304 (QB) at [81]: 'The key to the giving of advice is that the information is either accompanied by comment or value judgment on the relevance of that information to the client's investment decision, or is itself the product or a process of selection involving a value judgment so that the information will tend to influence the decision of the recipient. In both these scenarios the information acquires the character of a recommendation.'
See also *Financial Conduct Authority v 24Hr Trading Academy Ltd* [2021] EWHC 648 (Ch).

[128] https://www.handbook.fca.org.uk/handbook/PERG.pdf

[129] In a brief test of multiple open-access models at the time of writing, it was quick and straightforward to receive an explicit response that certain specific investments were suitable to the user based on invented but plausible financial and risk appetite information. These tests were repeated in November 2025, after some AAA providers changed their terms of use, ostensibly to prevent users from using their tools to obtain this kind of advice. The result was the same: it was easy to receive explicitly personalised suitability recommendations.

> *carrying on those activities with a degree of regularity. The person will also usually need to be carrying on the activities for commercial purposes. That is to say, he will normally be expecting to gain a direct or indirect financial benefit of some kind.'*

But this has the important caveat that:

> *'[It is] not necessarily the case that services provided free of charge will not amount to a business; for example, much investment advice is provided free of charge to investors but in the course of a business funded by commission payments; services (particularly advice, information or links) available on a website may also be free of charge to users of the site but be part of a business funded by advertising fees or sponsorship; and free newspapers may well represent a business for similar reasons.'*

330. The apparent lack of clear public statements on whether AAAs meet the threshold for authorisation, or of regulatory concern, may also be driven by a lack of clear evidence (as yet) of any significant consumer harm arising from the use of AAAs to receive financial advice.

331. This is arguably a fragile and inconsistent situation. It raises the prospect that AAAs will increasingly provide information that *looks and feels* to users like personalised financial advice, but is not regulated as such. This is especially acute since disclaimers warning users not to rely on AAA outputs are not particularly prominent and are not (at least not always or even usually) repeated when responses to specific queries are generated. The role of disclaimers in particular is in tension with PERG 8.30B.21 which states:

> *'A disclaimer may help a firm to avoid inadvertently presenting investments as suitable for particular customers or as being based on a consideration of the customer's circumstances. However, it will not always be sufficient. For example, a disclaimer is unlikely to be effective if: (1) a firm states that the investment would suit a particular customer's needs; or (2) it is reasonable for the customer to expect that the recommendation is based on a consideration of their circumstances.'*

332. The risk is raised of consumers increasingly using unregulated AAA-generated financial advice, which is cheap or free, creating a potential gap in enforcement or redress. The FCA itself has adverted to the problems which may stem from this 'regulatory threshold issue':

> *'One concern, which we do not explore here, is that LLMs could cause consumers to ask for financial advice less, even on consequential matters, or to seek out other informative guidance less.'*[130]

iii. Threshold: executing purchases of investments by way of business

333. Where an AAA crosses over into the paid execution of contracts for investments on a user's behalf, it is much more likely to meet the threshold for FSMA authorisation to be required on the basis that the provider is (at a minimum) making arrangements with a view to transactions. Notably:

i) The AAA actively executes transactions: there can be no question of a disclaimer that the user should not '*rely*' on it, since the very purpose of the AAA is to make binding decisions on the user's behalf; and

[130] https://www.fca.org.uk/publication/research-notes/research-note-lessons-2-large-language-models-pilots-consumer-guidance.pdf?utm

ii) The AAA provider is paid for the service, making it more likely that the '*by way of business*' test is met.

iv. *Carrying on regulated activities without authorisation*

334. Where the provision of an AAA *does* constitute a regulated activity under s.22 FSMA, but the provider *lacks* authorisation from the FCA:

i) The AAA provider will be guilty of a criminal offence under s.23 FSMA.

ii) Any agreement made 'by' the AAA provider is voidable by the user under s.26 FSMA, and there is a right to compensation. This may not apply to purchases made by the AAA, since it is far from clear that contracts entered into through an AAA are made 'by' the provider of the AAA (see also the section on the law of agency below). But it *would* apply to the agreement covering the use of the AAA itself.

v. *The Consumer Duty*

335. Assuming that an AAA provider (i) is carrying out a regulated activity, and (ii) has authorisation, it will be subject to FCA rules about how that regulated activity is carried out. The most relevant rule for the AAA harms considered in this paper is the '*Consumer Duty*', which introduces a '*Consumer Principle*', which requires firms to '*act to deliver good outcomes for retail customers* [i.e. AAA users]'. It includes cross-cutting rules requiring firms to (i) act in good faith towards retail customers, (ii) avoid causing foreseeable harm to retail customers, and (iii) enable and support retail customers to pursue their financial objectives.

336. The FCA gives as an example of breaching the Consumer Duty:

> *'Using algorithms, including machine learning or artificial intelligence, within products or services in ways that could lead to consumer harm. This might apply where algorithms embed or amplify bias and lead to outcomes that are systematically worse for some groups of customers, unless differences in outcome can be justified objectively.'*[131]

337. Whether the Consumer Duty is breached in relation to a particular AAA harm will be fact specific. It is notable however that the FCA provides that:

> *'A key part of the Duty is that firms assess, test, understand and evidence the outcomes their customers are receiving [... and]*
>
> *The Duty applies in a reasonable way. The focus on good customer outcomes applies to all aspects of firms' operations and culture. All firms have a Duty to act to deliver good outcomes for their customers. What this means in practice will depend on key factors, including:*
>
> - *The nature of the product or service. More complicated products are likely to need more attention than simpler or less risky products.*
> - *The characteristics of a firm's customers. Where customers are more likely to have characteristics of vulnerability, for example, we would expect it to take additional care.*

[131] FCA (2022). , §5.11, https://www.fca.org.uk/publication/finalised-guidance/fg22-5.pdf.

- *The firm's relationship with its customers. Obligations under the Duty reflect the firm's role and ability to influence retail customer outcomes. We would expect firms to focus on harms that are reasonably foreseeable.*
- *The size of the firm. We do not expect a small firm to apply the same resources or processes as a large firm.'*

vi. *Enforcement by the FCA*

338. A detailed assessment of the FCA's enforcement powers and regulatory activity is not necessary here, but it has significant regulatory powers where regulated firms do not follow FCA rules, including the Consumer Duty. These include:

i) Withdrawing authorisation for regulated activities;

ii) Issuing fines (including against individuals); and

iii) Seeking injunctions from the court in support of enforcement.

339. S.384 FSMA also provides that where there has been a breach of the Consumer Duty and the consumer – i.e. the AAA user – suffers loss as a result, then the FCA may require the AAA provider in breach to pay to the user 'such amount as appears to the FCA to be just'.

340. The FCA has a clear sector focus and is a large regulator, compared to (e.g.) the ICO, which by contrast is a cross-sector regulator[132], giving it a more realistic task – at least compared to the ICO – in enforcing its rules. Financial regulation also requires regulated firms to report on their activities directly to the FCA. This gives the FCA an important source of information in assessing compliance and identifying where AAA harms may arise.

341. A range of sources indicate that the FCA is mindful of developments in AI in its sector, such as the FCA update issued in April 2024[133]. These include specific studies carried out by the regulator into the use of LLMs by authorised firms,[134] and live testing of the use of AI by firms[135]. But these sources give little indication of how the FCA is likely to approach harms arising from AAA use in financial services. Rather, they contain relatively generic statements such as:

> *'We will continue to closely monitor the adoption of AI across UK financial markets to identify material changes that impact on consumers and markets. This includes keeping under review if amendments to the existing regulatory regime are needed."* And
>
> *"The FCA's approach to the use of AI by the firms we regulate is outcomes-focused, and therefore, testing – utilising a variety of methods – can be an important way for firms to understand the outcomes their consumers are receiving.'*

342. FSMA provides for designated consumer bodies to bring 'super-complaints' to the FCA where a feature of a market 'is, or appears to be, significantly damaging the interests of consumers.' Consumer bodies must be designated by the Treasury and the criteria for designation are

[132]The FCA's budget is approximately 10 times the size of the ICO's. See https://ico.org.uk/about-the-ico/who-we-are/how-we-are-funded/ and https://www.fca.org.uk/publications/business-plans/annual-work-programme-2025-26.
[133] https://www.fca.org.uk/publication/corporate/ai-update.pdf
[134] https://www.fca.org.uk/publication/research-notes/research-note-lessons-2-large-language-models-pilots-consumer-guidance.pdf?utm
[135] https://www.fca.org.uk/publications/calls-input/proposal-ai-live-testing-engagement-paper

strict[136]. At the time of writing only four bodies – Which?, Citizens Advice, the Consumer Council of Northern Ireland, and the Federation of Small Businesses[137] – are so designated. The FCA *must* respond with reasons to a valid super-complaint within 90 days (s.234E FSMA).

343. Although super-complaints are relatively rare, there is no reason why a super-complaint could not be brought in relation to an AAA harm.

344. In summary, if an AAA crosses the 'regulated activity' threshold under FSMA, relatively strict and outcome-focused regulation applies, with a powerful and well-informed regulator capable of enforcing it.

vii. The Financial Ombudsman Service

345. Where FSMA applies and has been breached in the context of an AAA harm – for example because the Consumer Duty has been breached – the user of the AAA may seek redress through the Financial Ombudsman Service ('**FOS**')[138], which will determine an outcome that it considers to be 'fair and reasonable in all the circumstances of the case' (s.228 FSMA). This may include financial compensation, including for non-monetary loss, and/or a direction to take 'such steps in relation to the [regulated activity] as the ombudsman considers just and appropriate'.

346. An AAA user would need to complain to the AAA provider first, and – if dissatisfied – normally would need to complain to the FOS within six months of receiving a final response from the financial service provider.

347. The FOS is free-to-use and complainants do not require legal representation. There is no risk of being made to pay the defendant firm's costs if unsuccessful. An initial response from a case handler is 'typically' provided within 90 days[139]. Whilst the time to resolve a complaint where a formal ruling from an ombudsman is required is less certain, users affected by AAA harms where FSMA applies are in a stronger position to obtain redress than those who must bring any causes of action through the civil courts (as to which see the relevant section above).

J. Regulation of legal services

348. The Legal Services Act 2007 ('**LSA 2007**') regulates the provision of legal services. The Legal Services Board is the overarching regulator, responsible for overseeing the Solicitors Regulation Authority ('**SRA**') and the Bar Standards Board ('**BSB**'). The Legal Ombudsman ('LO') handles complaints about legal services.

349. Legal services are regulated if:

i) They qualify as a 'reserved legal activity'; or

ii) They are provided by an 'authorised person'.

---

136 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/200454/guidance_for_super_complainants_120313.pdf

137 The Financial Services and Markets Act 2000 (Designated Consumer Bodies) Order 2013.

138 s.226 FSMA and DISP 2.3 of the FCA Handbook on Dispute Resolution, https://www.handbook.fca.org.uk/handbook/DISP/.

139 https://www.financial-ombudsman.org.uk/consumers/expect/how-long-it-takes#:~:text=Typically%2C%20this%20part%20of%20our,case%20handler%20as%20things%20progress.

350. Whether a law centre (i.e. not-for-profit, providing free advice) falls under the scope of the LSA 2007 therefore depends on the type of legal work it carries out and whether it engages in 'reserved legal activities' as defined in the LSA 2007. Authorised persons who work for the law centre are also regulated.

i. *Regulated legal activities and authorised persons*

351. Anyone can give *legal advice* – even if they are not an '*authorised person*' – provided it is not a *reserved legal activity*.

352. A (non-reserved) **legal activity** means any activity which consists of either:

i) The provision of *legal advice or assistance* in connection with the application of the law or with any form of resolution of legal disputes; or

ii) The provision of *representation* in connection with any matter concerning the application of the law or any form of resolution of legal disputes.

353. **Authorised persons** are those who are authorised by a relevant approved regulator to carry out a reserved legal activity. This mostly comprises solicitors (authorised by the SRA), barristers (authorised by the BSB), conveyancers, costs lawyers, legal executives and notaries.

354. There are six **reserved legal activities** that only authorised persons can carry out:

i) Exercise of a right of audience – representing someone in court (barristers and solicitors with rights of audience)

ii) Conduct of litigation – managing a case in court (issuing proceedings, filing documents)

iii) Reserved instrument activities – creating certain legal documents (mostly those related to land transfers)

iv) Probate activities – dealing with the estate of a deceased person

v) Notarial activities – work carried out by notaries public

vi) Administration of oaths – e.g. swearing affidavits

355. Non-reserved legal activities can be exercised by people who are not qualified solicitors or other legal professionals, but they are subject to the broader regulatory objectives of the LSA 2007, such as:

i) protecting and promoting the public interest;

ii) protecting and promoting the interests of consumers; and

iii) improving access to justice.

ii. *Obligations of authorised persons*

356. An authorised person must comply with the rules established by their relevant regulator. In the case of a law centre, the most relevant regulator will be the SRA – solicitors working for the law centre must comply with the SRA Code. They are personally accountable for compliance with the SRA Code.

357. Solicitors' obligations under the SRA Code include obligations of service quality and competence:

i) Only act for clients on their instructions or from someone properly authorised to provide instructions on their behalf

ii) Ensure that the service provided to clients is competent

iii) Consider and take account of clients' attributes, needs and circumstances

iv) Where a solicitor supervises or manages others providing legal services:

(1) They remain accountable for the work carried out through them

(2) They effectively supervise work being done for clients

v) Ensure that the individuals they manage are competent to carry out their role, and keep their professional knowledge and skills, as well as understanding of their legal, ethical and regulatory obligations, up to date

iii. *Oversight bodies – SRA, Legal Ombudsman, Legal Aid Agency*

358. The SRA oversees qualified solicitors, including those working in law centres. The SRA requires solicitors to act in accordance with the SRA Code of Conduct, failing which the SRA may take regulator action against individual solicitors, including striking them off the roll of solicitors.

359. The Legal Ombudsman oversees all regulated providers of legal services (i.e. authorised persons). It cannot address complaints about non-regulated providers. Hence a law centre is not subject to oversight through the Legal Ombudsman, but individual solicitors working for the law centre are.

360. If a firm or law centre receives funding from the Legal Aid Agency ('**LAA**'), it will be subject to its oversight. This includes monitoring compliance with standards related to the quality of advice, record-keeping, and billing. Failure to meet LAA requirements can lead to a suspension or revocation of legal aid contracts.

iv. *Professional negligence/breach of duty*

361. In addition to regulatory enforcement, those providing legal services (whether regulated or not) could be subject to professional negligence proceedings if **a client** suffers loss or harm as a result of negligence or breach of contractual duty in providing the services. This could include failing to act with due diligence, making errors in legal advice, missing important deadlines (e.g. statutory time limits or court filing deadlines), or negligently handling a case in a way that causes harm or financial loss to a client. See §256 above.

362. Courts have recently been unforgiving to lawyers who relied on AI for legal research. Two cases of false case citations were brought to the High Court and resulted in a severe decision against the lawyers involved: *Ayinde v Haringey* and *Al-Haroun v Qatar* [2025] EWHC 1383 (Admin).

363. In *Ayinde v Haringey*, the Haringey Law Centre was challenging the London borough of Haringey over its alleged failure to provide its client with temporary accommodation. The pupil barrister in charge of the case cited non-existent case law five times in her pleadings. She denied using AI, but admitted that she may have relied on AI summaries in Google search without realising what they were. The Court found that both the Haringey Law Centre and the pupil barrister had been negligent. Both were ordered to pay wasted legal

costs to the London borough of Haringey. The pupil barrister was referred to the BSB, for further investigation and determination as to the circumstances that led to the citation of fake cases. The solicitor for the Haringey Law Centre was also referred to the SRA, for investigation of the steps he took once he had been told that the cases citations were false – but he was not found negligent nor otherwise sanctioned by the Court.

364. The Court reminded itself that it has a range of powers to ensure that lawyers comply with their duties to the court and to the public – this includes public admonition of the lawyer, the imposition of a costs order, the imposition of a wasted costs order, striking out a case, referral to a regulator, the initiation of contempt proceedings, and referral to the police. How severe the sanction is will depend on the particular facts of the case, including the impact on the underlying litigation.

K. Law of agency

i. The traditional framework

365. Under English common law, agency is a legal relationship that arises when one party (the agent) is authorised to act on behalf of another (the principal) in dealing with third parties (in particular, entering into contracts, as is envisaged in Scenario 2). Whilst there are no specific requirements as to formality[140], the core elements of agency are:

i) **Consent**: both principal and agent must agree to the arrangement (though agreement may be inferred).

ii) **Authority**: the agent must have actual, implied, or apparent authority to act.

iii) **Fiduciary obligation**: the agent owes duties of loyalty and good faith to the principal.

iv) **Legal effect**: the agent's acts, within authority, bind the principal vis-à-vis third parties.

366. These principles are long-standing (see e.g. *Ireland v Livingston* (1872) LR 5 HL 395) and well-settled in commercial contexts (e.g., *Freeman & Lockyer v Buckhurst Park Properties (Mangal) Ltd* [1964] 2 QB 480, on apparent authority).

ii. Can an AAA be an 'agent'?

367. At present, an AAA cannot be an 'agent' in the legal sense under English law. There are two central obstacles:

i) **Lack of legal personhood**: Agency presupposes a legal or natural person who can owe duties and be held liable: 'legal personality is the necessary foundation for legal liability'.[141]

ii) **Inability to consent or act with intention**: A foundational element of agency is the agent's consent to act or intention to act for the principal. Courts have long emphasised the centrality of intention even in indirect or implied agency. In *Yonge v Toynbee* [1910] 1 KB 215, where a solicitor acted under a mistaken belief about their authority, the

[140] *Tuke v JD Classics Ltd (formerly JD Classics Holdings Ltd)* [2018] EWHC 531 (QB).
[141] H. W. R. Wade and C. F. Forsyth, *Administrative Law*, 11th ed.

agent's liability was rooted in their own intention to act. Intention is a quality that an AAA does possess.

iii. *Does the AAA provider become the agent?*

368. Where an AAA carries out the instructions of the user (even allowing for some 'choice' by the AAA in how to carry them out, which specific products to purchase etc.) it is unlikely that the AAA provider could be considered an agent of the AAA user. In *R (SSSL Ltd) v HMRC* [2007] EWHC 971 (Admin) consideration was given to whether agency arose between a provider of software for the automated conclusion of insurance contracts and its clients. It was held at [61]:

> *'The necessary agency between insurers and SSP cannot be inferred from the evidence that I have set out at paragraphs 14–24 above. SSP did not take any "decision" to make offers of insurance on behalf of insurers. The information necessary for electronic contract formation had been pre-programmed, according to parameters laid down by the insurer, in the SSP computer software. The relevant data was processed automatically by electronic means through that software and the transactions were self-executing within the specified parameters pre-determined in the programme. SSP had to take no "decisions" to generate the offers; the software performed the necessary tasks.'*

369. The use of an AAA is broadly analogous. The AAA provider does not take decisions in relation to users' entry into contracts. Rather, contracts are executed automatically based on the user's instructions and the software. AAAs are of course significantly more complex and exhibit emergent properties, but we would argue this is a difference of degree rather than kind. Therefore, in normal circumstances, a user will *prima facie* be bound to any contracts concluded on their behalf by an AAA, even if they are not 'optimal' (which is likely not possible to define in any case), provided they are within the parameters of whatever instructions the user has given.

370. A user who is a consumer could argue that they are not bound by a contract entered into via an AAA for lack of provision of legally required pre-contractual information on the goods, services or digital content purchased.[142] This information must be provided *to the consumer*. The counterparty could argue that it meets its obligations by making the information available to the AAA for the user to access *if they desire*, but this is doubtful. We are not aware of this being tested in court. It may provide AAA users with protection, but where a user has already made payment for an AAA purchase, it would be for the user to bring these arguments to bear in order to withdraw from the relevant contract, which could be challenging.

371. Where an AAA 'malfunctions' and takes actions – such as entering into contracts – which are clearly inconsistent with its user's instructions, the position is less clear. A user could seek to argue that the AAA was acting outside its 'authority' and therefore the user should not be bound to the relevant contracts. But this runs into the problem that an AAA cannot be an agent and therefore has no authority to exceed.

372. For the same reasons given above, the AAA provider is unlikely to be the agent of the user in these circumstances. This leaves some doubt over whether a user is bound to such 'contracts-

[142] The Consumer Contracts (Information, Cancellation and Additional Charges) Regulations 2013 ('**CC (ICAC) Regulations**').

by-AAA-mistake', or whether the user *is* bound[143] but has a potential claim for any losses against the provider (for example under consumer law). We are not aware of any binding authority which provides useful guidance on these novel points.

373. Finally, where an AAA acts in ways that serve the interests of its developer – for example, by preferencing affiliated merchants or platforms in its recommendations – there may be an argument that the *provider* is effectively acting as an agent. However, this too is problematic:

i) The provider (as opposed to the AAA) is typically not instructed by the user, nor does an AAA provider (in the ordinary course of things) hold themself out as having authority to transact on the behalf of each and every user of its AAA.

ii) If the provider has configured the assistant to act autonomously, the more appropriate analysis may treat the provider as a *principal* – deploying a semi-autonomous tool to act in the world in a way that aligns with their own commercial interests.

374. It can be seen that the use of AAAs to take actions such as making purchases does not always map cleanly onto existing concepts, raising the risk of liability gaps – at least until key issues are resolved in case law or through statute. The law of agency was simply not designed to allocate legal responsibility for delegated discretion encoded into highly complex software, acting with a significant freedom, and with emergent and unpredictable properties. Academic commentary has begun to explore these limitations. R. Brownsword argues that 'AI agents problematise foundational concepts in contract and agency law' in part because of their indeterminacy and non-human form of 'volition'.[144]

iv. ***Implications***

375. The lack of a predictable application of agency law to AAAs has real-world consequences:

i) A user may have no clear course of action when an AAA makes a defective or biased decision, or one which goes outside its instructions, and the contractual counterparty (with some justification) seeks to hold the user to that contract since it was entered into apparently by them as a principal using an automated tool.

ii) Contractual counterparties may be unsure whose instructions the assistant reflects — the user's, or the provider's, or perhaps neither in the case of a malfunction on the part of the AAA.

376. In sum, while AAAs may *functionally* act as agents – making decisions and transacting on users' behalf – they do not currently *legally* do so. The English law of agency has not evolved to recognise non-human actors as agents. This creates a growing potential mismatch between legal principles and the technical reality of AI-mediated decision-making. Indeed, these legal uncertainties may be so significant that they – for the time being – place some limits on how quickly AAAs take on a role in making real-time purchases on behalf of users without obtaining their prior approval.

[143] Save for any arguments based on the CC (ICAC) Regulations.
[144] Brownsword, R. (2019). *Law, Technology and Society: Reimagining the Regulatory Environment*. NewYork: Routledge.

IX. Conclusion: gaps in protection and the unique challenges of AAAs

377. We conclude that there are significant gaps in legal protection from realistic harms arising from the increasing use of AAAs. Our analysis shows that despite AAAs *potentially* engaging many different legal frameworks, in practice the law often does not apply in a way which captures the kinds of harms that users may suffer. Even where it does, there are often few realistic avenues for redress.

378. Our analysis of the four Scenarios highlights the role and significant of threshold conditions: some protection is found in very specific contexts, where an AAA is effectively used by a provider or implementer to carry out regulated activities, such as financial or legal advice. But even in these contexts, there may be room for the provider of an AAA to avoid liability with appropriate disclaimers, or by ensuring that any marketing of the AAA is kept in general terms. AAAs' general-purpose nature and low cost means they may increasingly be used in high-stakes, traditionally-regulated context, displacing more expensive, regulated providers of equivalent services.

379. In other areas, the Scenarios show that certain kinds of harm – such as emotional dependency on technology, market distortions, and subtle influence over political opinions – do not fit easily within existing regimes. Yet these harms are real and may well grow in scale and scope as AAAs are more widely adopted and relied upon.

380. Overall, the analysis of four realistic use-cases shows that AAAs in general have properties which complicate regulation and redress. These include in-built difficulties in relation to transparency, as well as a lack of established standards, and the genuine novelty of AAAs' capabilities relative to established frameworks. Bearing in mind the increasing adoption of AAAs and the (for now) continuing growth in their capabilities, these gaps in legal protection arguably raise significant questions for policymakers, AAA providers, and the public alike.

AWO

Radha Bhatt (Matrix Chambers)

September 2025