Digital Regulation Cooperation Forum

The Future of Agentic AI

Foresight paper

Publication date: 31 March 2026

CMA Competition & Markets Authority

Ofcom

ico. Information Commissioner's Office

FCA FINANCIAL CONDUCT AUTHORITY

This paper aims to foster debate and discussion among our stakeholders. It should not be taken as an indication of current or future policy by any of the member regulators of the Digital Regulation Cooperation Forum (DRCF).

This paper draws some insight from the DRCF's Autumn 2025 public call for views on Agentic AI, and the DRCF would like to acknowledge all the respondents for their contributions.

Contents

Executive summary

This paper serves as a forward-looking exploration of agentic AI and how UK regulatory frameworks can help realise the opportunities of this technology in a responsible and safe way. Published by the Digital Regulation Cooperation Forum (DRCF)—comprising the Competition and Markets Authority (CMA), Financial Conduct Authority (FCA), Information Commissioner's Office (ICO), and Ofcom—the document aims to inform and foster stakeholder debate. The DRCF emphasises that regulation should act as an enabler of innovation, ensuring that emerging technologies develop in ways that promote economic growth and competition while protecting consumers and their rights.

Defining Agentic AI

Agentic AI represents a shift from AI as a tool, to AI as an agent that acts on behalf of users. Unlike standard generative AI which responds to queries and creates outputs, agentic systems can assess goals, plan workflows, and execute actions autonomously to impact real-world environments and interact with people or other agents. Agentic systems are made of the following layers: (1) the core technical system (individual AI agents and their methods of communication), (2) the external environment (other software systems, the physical world) where actions take effect, and, as appropriate, (3) a human-facing interface for goal setting and supervision. The degree of autonomy of AI agents may vary. While truly autonomous agents could one day emerge, current deployments are mostly agents requiring close supervision and handling only a limited number of steps (e.g. to submit expense claims).

Emerging Opportunities and Risks

The move toward agentic AI engenders opportunities and risks for consumers and the economy. In terms of opportunities, consumers could benefit significantly, for example from agents that would handle "life admin" (e.g. booking holidays, renewing policies), potentially reducing friction and improving accessibility. Businesses could potentially realise significant productivity gains in both customer engagement (front office) and internal operations (back office), such as automated reporting and triage. From the standpoint of regulators, agents could have various potential uses: for instance, regulators are experimenting with "monitoring agents" to verify compliance.

Agentic AI, however, also presents risks, including the whole range of risks that apply to AI more broadly and are covered elsewhere. Some of these risks might be increased in the case of agentic AI, and indeed there may also be new or novel risks. Delegating a task to an agent comes with the possibility that it might make mistakes, which could be consequential. For example, an agent managing an inbox could inadvertently share credentials or confidential emails inappropriately, which could create a cybersecurity or data protection risk. In addition, if agents execute multiple steps rapidly (e.g., comparing, accepting terms, and paying near-simultaneously), this might make it more difficult for users to follow. In terms of transparency and control, without robust oversight, agents (like other AI) risk becoming "black boxes," presumably even more so in the case of multi-agent systems. For example, it may be difficult for users to contest an agent's decision to deny a refund, or to contest its decision to share a customer's data with a third party. For organisations deploying these agents, such a lack of transparency may risk non-compliance with consumer, contract and data protection law. Regarding market dynamics, the risk of vendor lock-in, i.e. a user being unlikely to switch provider for technical or other reasons, could be significant.

At the same time, agents could potentially tap into, and thereby support, a wide variety of services themselves. Finally, the risk of algorithmic collusion is also significant, whereby agents may learn to reach coordinated outcomes, whether the developer or deployer intended for this to happen or not.

The development of agentic AI systems raises several other potential risks, including broader societal and environmental concerns. However, this report is focused on risks that DRCF regulators are mandated and best placed to address.

Ensuring Adoption of Agentic AI is Safe and Trusted

Agentic AI could one day drive mass consumer adoption of AI, with potentially large economic and societal impacts. There is therefore an imperative for DRCF to be on the front foot in supporting regulatory coherence, acknowledging that Agentic AI, like other AI, spans the remits of all DRCF members, often simultaneously. A single agentic deployment (e.g., a retail assistant) could simultaneously trigger concerns about infringement of data protection law (ICO), network security regulations (ICO and Ofcom), financial regulations (FCA), online safety duties (Ofcom), and competition and consumer laws (CMA).

All four regulators agree that AI agents do not fall outside existing UK regimes: obligations around transparency, fairness, safety, consumer protection and competition continue to apply as Agentic AI develops. Specifically:

- ICO: Seeks to ensure that innovation in Agentic AI develops in ways that protect people's information rights. Focuses on Automated Decision-Making (ADM), profiling, accountability, fairness, accuracy (for example due to cascading 'hallucinations'), data security, purpose or storage limitation, and data minimisation. ICO rules require traceable logs and meaningful human involvement to ensure accountability and data minimisation.
- FCA: Upholds existing frameworks including the Consumer Duty (which already applies to financial services firms using agentic AI) and expects firms using those tools to demonstrate that they are acting to deliver good outcomes.
- Ofcom: Ofcom's mission is to make communications work for everyone. As the UK's converged communications regulator, Ofcom is invested in understanding how Agentic AI will impact telecoms, broadcasting, online safety and spectrum. Where agentic AI falls within scope of Ofcom's regulatory regimes, Ofcom will take account of the risks arising out of AI agents when considering how best to exercise its functions.
- CMA: Promotes competition and protects consumers with a clear end goal in mind: to drive economic growth and improve household prosperity. It may look at elements of AI across many sectors. In addition, under the Digital Markets regime, the CMA can designate large technology firms with strategic market status and impose interventions to remedy a range of harms, including harms related to pricing, quality, consumer choice, transparency and fairness, and exclusionary practices or foreclosure.

Cross-regulator coordination will continue to be essential – including through a shared horizon scanning function.

Potential Future Developments and Regulatory Considerations

The medium-term future of Agentic AI is likely to vary across sectors. Variables influencing adoption will involve a range of technical, commercial and behavioural factors, including how reliably AI agents perform in real-world contexts, the extent to which consumers and organisations understand

and act on AI agents' capabilities and limitations, and prevailing commercial models and arrangements.

This report considers potential future developments and shares early thoughts on cross-regulatory implications across four categories: governance, data protection and cybersecurity, consumer rights and interests, and market dynamics and competition.

Governance:

- Effective governance and robust guardrails are essential to the development of Agentic AI that's safe for consumers.
- Limited transparency in multiagent systems, combined with the risk of cascading errors, can make it more difficult to establish responsibility. The mapping of systems, and potentially transparency agents, could help.
- Human oversight is one type of governance mechanism that can help mitigate risks and build trust by ensuring that humans retain decision-making authority over critical actions. Clear thresholds should be defined for when an AI agent must seek human approval before proceeding.

Data protection and cybersecurity:

- The data minimisation principle requires organisations to use only the data necessary for the specific processing purpose.
- Transparency about how personal data is used by Agentic systems is critical to building consumer trust, ensuring Agentic AI's growth, and avoiding any systemic privacy vulnerabilities which could undermine confidence in digital services, destabilise markets and hinder innovation.
- Agentic AI may enhance cybersecurity by helping firms and consumers address complex threats. At the same time, AI agents could amplify existing generative AI risks and introduce new ones. DRCF regulators remain alert to these risks. The ICO, FCA and Ofcom have cybersecurity requirements of firms, drawing from guidance from NCSC.

Consumer Rights and Interests:

- Agentic AI systems may influence consumer decision making in ways users do not anticipate. Businesses should empower consumers to fully understand what they have consented to or delegated.
- Agentic systems may be optimised to advance deployers' objectives, potentially shaped by undisclosed partnerships or business models. Disclosures might help empower consumers.
- Consumer rights must continue to be upheld by Agentic AI providers, including protections against false or misleading claims, requirements for minimum standards of quality and fitness for purpose, and safeguards against unfair contractual terms.
- Agentic AI may have complex effects on digital inequality. While it could increase pressure to opt in and weaken critical digital skills or encourage over-trust, it may also lower barriers to participation through accessibility tools and personalised support

Market Dynamics and Competition:

- Agentic AI could lead to highly concentrated, winner takes most outcomes, however it could also create new markets and deliver better outcomes for consumers.

- New standards may be emerging to support data portability and interoperability for agentic systems. Effective implementation of such standards could help reduce the risk of vendor lock-in and support competition.
- Agentic AI systems can exhibit collusive behaviour without explicit instruction. One possibility is that AI agents may develop hidden communication strategies to pursue shared goals. Businesses should be mindful of risks related to algorithmic collusion when deploying AI agents in pricing.
- Harmful behaviour can arise at the network level, even if individual agents appear benign, making detection and attribution difficult. It also means collusion and manipulation could occur at scale. Adopting new monitoring, governance and detection tools could help regulators and firms manage these emerging risks.

Across these categories, existing principles and regulatory frameworks continue to apply to Agentic AI. For example, while Agentic AI systems may operate with a degree of autonomy, this does not diminish organisational responsibility for legal compliance. DRCF member regulators will continue to work collectively and individually to support a clear and coherent approach to Agentic AI regulation, to enable safe and responsible adoption.

1. Introduction

The Digital Regulation Cooperation Forum (DRCF) was established to ensure coherence between the regulatory regimes of its member regulators (the Competition and Markets Authority (CMA), Financial Conduct Authority (FCA), Information Commissioner's Office (ICO), and Ofcom), to work together on complex challenges and develop capabilities for the future.

The DRCF Horizon Scanning and Emerging Technology team (HSET) is the leading cross-regulatory voice on emerging technologies and trends in digital markets. We take a proactive approach to understanding the potential benefits, risks, and regulatory implications of emerging technologies. We subsequently provide actionable insights to regulators, other public bodies, government, Parliament, industry, civil society and the public.

This report explores potential future developments in Agentic AI and provides insight into definitions, use cases, future scenarios, and regulatory considerations relating to Agentic AI technologies. It is not a statement of regulatory policy. To provide actionable insights, the DRCF conducted desk research, developed possible future scenarios (set out in the Annex) and held a series of internal and cross-regulatory workshops with subject-matter experts from our member regulators. This report combines the outputs with perspectives shared by stakeholders from industry, academia and civil society through a Call for Input (CFI) on Agentic AI conducted through the DRCF Thematic Innovation Hub.

The DRCF member regulators are keen to ensure that regulation is an enabler of, not a barrier to, innovation. We aim to enable the development of emerging technologies in ways that promote economic growth, support open, competitive markets, and protect consumers and their information rights.

2. Agentic AI: an Explainer

2.1 Defining Agentic AI

Agentic AI is an evolving concept.[1] For the purposes of this paper, we define:

- **Agentic AI:** Systems of AI agents that behave and interact autonomously to achieve their objectives.[2]
- **AI agents:** Increasingly autonomous AI services that can directly impact real-world environments.

AI Agents vs AI Assistants

AI agents do not merely assist, they take in data, make decisions and carry out actions.[3] Beyond generating responses to user queries, which generative AI systems do, they may:

- assess goals, break them into subtasks, and plan end-to-end workflows;
- retrieve real-time data (that may include personal data) from other agents, databases and other services;
- execute actions autonomously, such as making payments on behalf of the user;
- store memory of past interactions to improve over time.[4]

Information retrieval alone does not make a system an agent: many AI assistants can access real-time information via search or APIs. Agents, on the other hand, use such information within an autonomous plan–act loop to complete multi-step tasks, often by invoking other tools. This enables them to complete tasks with limited supervision (see 'Levels of AI Agent' below[5]).

Sometimes the term 'AI agent' is associated with the use of Large Language Models (LLMs).[6] However, in this paper we include agents based not only on LLMs but on the full range of algorithmic systems, including, for example, recommender systems.[7] The key question is whether the underlying technology enables the service to take in data, make decisions, and carry out actions.

---

[1] R. Sapkota, K. Roumeliotis, M. Karkee. "AI Agents vs. Agentic AI: A Conceptual taxonomy, applications and challenges," in *Information Fusion*, vol. 126, pp. 103599, 2026. AI Agents vs. Agentic AI: A Conceptual taxonomy, applications and challenges - ScienceDirect https://arxiv.org/abs/2505.10468v5

[2] GOV.UK, AI Insights: Agentic AI (HTML) - GOV.UK. This paper does not set out to agree a common cross-regulatory definition for "agentic AI". Eg, the ICO Tech Futures report discusses multiple capabilities of agentic systems.

[3] Russell, S., & Norvig, P. (1995). *Artificial intelligence: A modern approach* (Chapter 2). Prentice-Hall.https://d1wqtxts1xzle7.cloudfront.net/125698862/artificial_intelligence_modern_approach.9780131038059.25368-libre.pdf

[4] CMS LawNow, Agentic AI, Risk and Compliance Under the EU AI Act

[5] Shavit, Y., Agarwal, S., Brundage, M., Adler, S., O'Keefe, C., Campbell, R., ... & Robinson, D. G. (2023). Practices for governing agentic AI systems. *Research Paper, OpenAI.* https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf. It should be noted, however, that, for now at least, humans are often called upon to assess the results of proposed agent actions.

[6] A Large Language Model is a specific deep-learning architecture optimised for the comprehension and generation of natural language. By processing vast datasets through billions of parameters, LLMs utilise statistical probability to predict and sequence text that is contextually relevant and syntactically coherent.

[7] Sometimes such agents are also called 'agentic algorithmic systems', for example: Chan, A. et. al. (2023). Harms from Increasingly Agentic Algorithmic Systems. *2022 ACM Conference on Fairness, Accountability, and Transparency*, 651-666. https://doi.org/10.1145/3593013.3594033

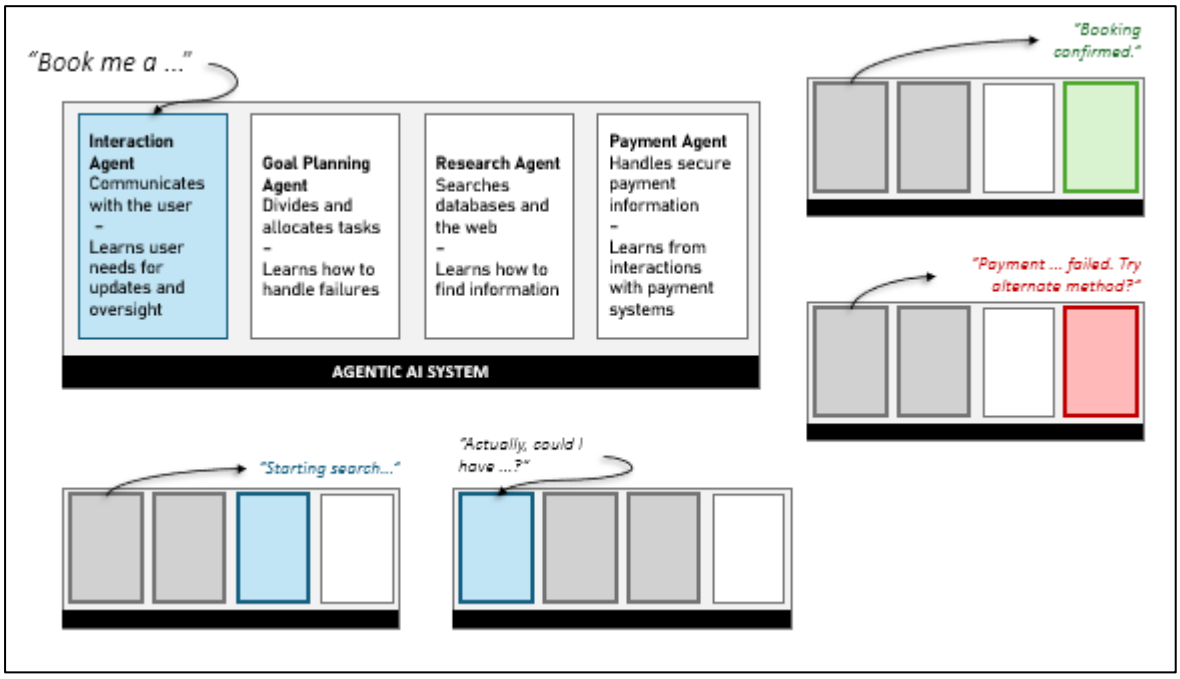
AI Agents to Agentic AI

Agentic AI systems are composed of AI agents. When multiple AI agents interact, the combination of individual agents' capabilities (e.g. domain-specific reasoning, memory and task execution) allows the resulting Agentic AI system to be adaptive and exceed the capabilities of the individual agents.[8]

If an explicit goal has been given, Agentic AI systems can automatically break the goal down into subtasks, e.g. by using planning or supervisor agents, and plan and arrange sequencing so that the goal can be attained even in the face of unplanned failures or changes to the environment[9] (this planning and arranging is sometimes called 'orchestration'). For example, an Agentic AI system tasked with managing travel plans might have a supervisor agent that waits for accommodation-booking agents to successfully complete their tasks before delegating actions to flight-booking agents (see Figure 1 for an example of delegating actions).

Whether or not there is an explicit goal, communication channels allow individual agents in the system to interact and coordinate. The Agentic AI system may also have governance features to trace the flow of actions and data between agents (and so allow auditing or oversight) or permissions features to restrict the actions or access to data of specific agents (e.g. to promote privacy or security).[10]

Figure 1. Agentic AI combines the capabilities of individual agents



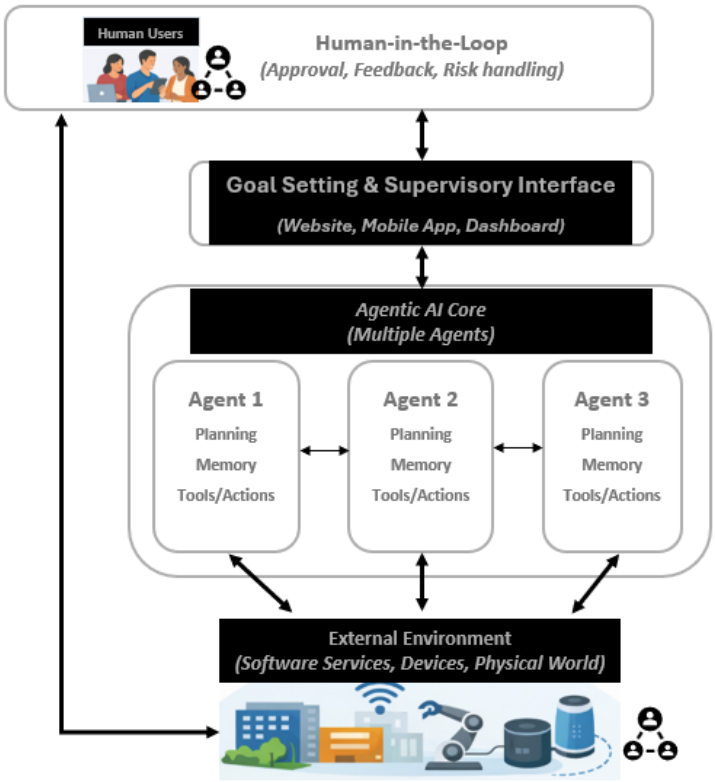
[8] Sapkota et al (2026).
[9] Sapkota et al (2026).
[10] For example, permissions features, see: https://www.infoq.com/news/2025/04/deepmind-camel-promt-injection/

2.2 The Sociotechnical System

Agentic AI is best understood as part of a wider sociotechnical system. The **core technical system** comprises agents and their modes of communication/interaction with each other. However, the technical system interacts with an **external environment** (e.g. other software services, devices, physical objects) where its actions have effects. Also, it may have an **interface** that humans use to set goals and supervise behaviour (such as a dashboard, website, or mobile app) and/or behaviour may be supervised indirectly (by humans observing effects on the external environment).

This wider system is mirrored in the individual agents. Each agent has (1) a technical core of software components (e.g. a planning module, tool/action interfaces, and memory/state), (2) an 'external' environment of other agents and/or the wider world that it can interact with and, as appropriate, (3) interfaces, which human users or other agents can use to monitor behaviour or set goals (see Figure 2).

Figure 2: Agentic AI as a Sociotechnical System



Components of an AI agent

(1) At the centre of an individual AI agent is its **technical core**. This is made up of AI models and 'scaffolding' (software which coordinates data coming in and out of models). AI models may be a combination of LLMs and/or other types of AI, such as rule-based models and recommender systems.[11]

[11] For an example that exclusively uses non-LLM models, see: Arentze, T., Hofman, F., van Mourik, H., & Timmermans, H. (2000). ALBATROSS: Multiagent, Rule-Based Model of Activity Pattern Decisions. *Transportation Research Record: Journal of the Transportation Research Board*, *1706*(1), 136-144. https://doi.org/10.3141/1706-16

This technical core is designed to achieve 3 key capabilities[12,13]:

- **Reasoning and planning**: To interpret what users want and plan the steps to accomplish complex tasks.
- **Tool use**: To use other software such as web services and third-party apps, enabling the agent to carry out tasks.
- **Memory**: To store and use data from past user interactions, previous actions taken by the agent and their results.

How are the 3 key capabilities implemented? Many types of AI can be effective for **planning**. LLMs and other generative AI models in particular provide language, visual and emotional 'understanding' capabilities, enabling modes of reasoning about what users want and communicating back with the user.[14,15] **Tool use** relies crucially on models which 'understand' code and app layouts or additional helper software which translates the model decisions into actions.[16] The adoption of standard protocols for interaction between agents and tools, such as the Model Context Protocol, has made implementing tool use much easier[17]. **Memory** may be implemented via the use of internal and/or external storage:[18] external storage, such as SQL or vector database services, may be considered a type of tool use, whereas internal storage, such as in-context memory, is an integral part of the agent. Developments in memory technologies are allowing agents to store and retrieve relevant information more efficiently.

(2) Second, the agent interacts with the **external world**, which could include operating systems, online resources, people, other agents, and/or physical objects. This is where the agent's actions have real effects.

(3) Finally, the agent is used by humans or other automated systems (e.g. other agents). For this, it has **interfaces**. These may be human or 'user interfaces' (like a dashboard, a website, a mobile app on a phone, or a voice assistant on a smart speaker), or software interfaces or 'APIs'.[19] These interfaces are used to interact with and/or monitor the agent. Therefore, they play a role in the ability to guide the agent's behaviour and step in when needed.

Because of agents' adaptive behaviour, the external world can complement the interface by providing another way to monitor what the agent is doing – look at what it has done – or to influence the agent's behaviour – change its environment. For example, instead of observing the agent directly, human users could monitor logs which record the agent's actions.

---

[12] World Economic Forum. (2024). *Navigating the AI Frontier: A Primer on the Evolution and Impact of AI Agents*. https://reports.weforum.org/docs/WEF_Navigating_the_AI_Frontier_2024.pdf.
[13] Wang, L. et al (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, *18*(6). https://doi.org/10.1007/s11704-024-40231-1.
[14] Section 3.1.1 in Cheng, Y. et al. (2024). *Exploring Large Language Model based Intelligent Agents: Definitions, Methods, and Prospects*. arXiv.org. https://arxiv.org/abs/2401.03428.
[15] Sorin, V. et. al. (2024). Large Language Models and Empathy: Systematic Review (Preprint). *Journal of Medical Internet Research*, *26*, e52597. https://doi.org/10.2196/52597.
[16] Sager, P. J. et. al. (2025). *AI Agents for Computer use: A review of instruction-based computer control, GUI automation, and operator assistants*. arXiv.org. https://arxiv.org/abs/2501.16150
[17] Benefits of using the MCP in What is Model Context Protocol (MCP)? A guide | Google Cloud
[18] Section 3.1.2 in (Cheng Y. et al. 2024).
[19] API stands for Application Programming Interface - a mechanism that enables two software components to communicate with each other using a common set of definitions and protocols.

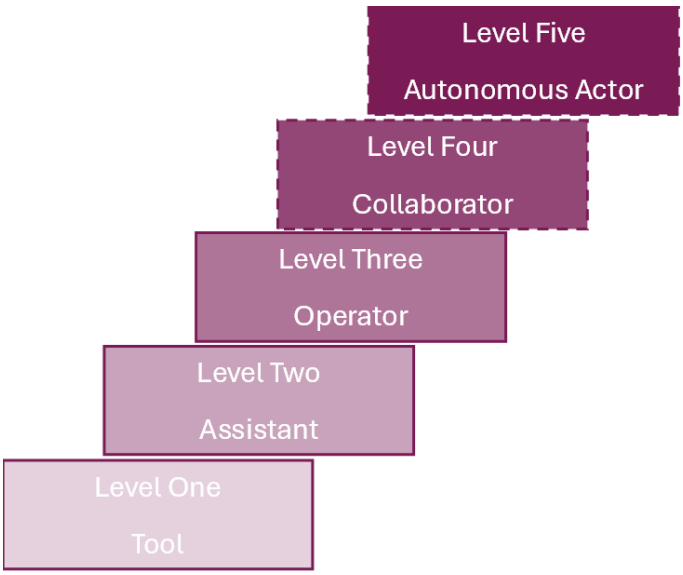
All these parts work together in a continuous loop: humans and other agents provide guidance, the agent reasons about what to do, uses its memory and tools to plan and take action, and those actions affect the outside world, where humans can observe the effects and provide further guidance.

2.3 State of Development

Levels of AI Autonomy

The sections above describe the building blocks of an AI agent (planning, tools, memory and oversight). A complementary way to understand the agent landscape is to look at levels of autonomy and sophistication.[20] Figure 3 illustrates a spectrum from simple tools to more autonomous systems. Most real-world deployments today sit toward the lower–middle end of this spectrum, although higher-autonomy systems are emerging, often with safety controls (e.g., human approval steps, spending limits, and human and/or automated monitoring).

Figure 3: Five levels of autonomy for AI services



The base level is the "tool": a reactive system that does only what you ask in the moment, for example, summarise a document or extract a table, without initiative, memory or external actions.

One step up sits the "assistant", which can plan a few steps, use approved tools such as search or calendars, and propose next moves while still deferring to you for execution.

The "operator" handles a bounded workflow end-to-end once you authorise it. For example, it can submit an expense claim, run a data pipeline, open and close help-desk tickets, while keeping short-term memory, monitoring progress, coping with routine exceptions and reporting back when done.

Above that, and still mostly theoretical at the time of publication, is the "collaborator", a semi-independent system: able to initiate and coordinate multi-step work without constant prompts but still operating within guardrails and requiring human approval for high-impact decision.

[20] Feng, K. J., McDonald, D. W., & Zhang, A. X. (2025). Levels of autonomy for ai agents. *arXiv preprint arXiv:2506.12469*. https://arxiv.org/abs/2506.12469

At the top is the 'autonomous actor,' a still largely theoretical system at the time of publication, that can make decisions, manage resources, and improve itself with little human help. It might be kept in check by safety controls such as kill switches and budget limits.[21]

Where is Agentic AI today?

New AI systems that can see, read, and act are getting much faster at turning plans into actions. The step from tool to actor requires increased initiative, memory, tool-use and coordination, which in turn requires effective identity and access management, logs of activity, an ability for a human operator to override, and regular performance and safety checks. Most practical deployments today sit at the assistant or operator tiers. For instance, there are customer-support copilots that triage tickets before handing them to humans, or workflow agents that automate expense claims and data-pipeline runs.

It should be noted that not all Agentic AI is purely in a software form. Some Agentic AI is 'embodied' within the internet of things (IoT) and robotics, thereby bringing more emphasis to its ability to adapt and coordinate in a complex external world.[22] Examples include LLM-enabled humanoid robots, such as Figure's "Figure 02" and 1X's "NEO". These agents use models and reasoning, for both vision and language, to autonomously perform motor tasks (handling objects, clearing tables) that they have been trained to handle (e.g. in a lab environment or by humans remotely operating them).[23,24]

Major generative AI model providers, including Anthropic, Google DeepMind, and OpenAI, have introduced AI assistants with increasingly agentic capabilities by training underlying models to be capable of tasks like autonomous web navigation,[25] research synthesis,[26] and computer control.[27] Benchmarks such as industry-developed General AI Assistants (GAIA) provide one (imperfect) signal of progress in tool-using assistants. In the original GAIA study, human respondents achieved 92% accuracy compared with 15% for GPT-4 equipped with plugins, highlighting a large gap on tasks that are conceptually simple for humans but challenging for agents.[28] More recent GAIA leaderboard results report substantially higher scores for leading agent systems, with several top submissions scoring in the mid-to-high 80s. However, these gains should be interpreted cautiously: leaderboard improvements can reflect the impact of repeatedly doing the benchmark and using the results to improve the assistant, rather than general skill acquisition.[29]

---

[21] "A kill switch is a critical feature embedded in various systems, devices, and software that allows users to disable or shut down a device or service immediately. This capability is particularly important in maintaining control and security over hardware and software environments". What Is a Kill Switch: Understanding Its Purpose and Function | Graph AI

[22] Perlo, J., Robey, A., Barez, F., Floridi, L., & MÄśkander, J. (2025). Embodied ai: emerging risks and opportunities for policy action. *arXiv preprint arXiv:2509.00117*. Embodied AI: Emerging Risks and Opportunities for Policy Action

[23] What to know about NEO, the humanoid robot that'll do your chores.

[24] BMW Group: First humanoid robot introduced in Plant Leipzig and Figure 03: The Best Inventions of 2025 | TIME

[25] Google, Google introduces Gemini 2.0: A new AI model for the agentic era

[26] OpenAI, Introducing deep research | OpenAI

[27] Anthropic, Introducing computer use, a new Claude 3.5 Sonnet, and Claude 3.5 Haiku \ Anthropic

[28] Mialon, G., Fourrier, C., Wolf, T., LeCun, Y., & Scialom, T. (2023). Gaia: a benchmark for general ai assistants. In *The Twelfth International Conference on Learning Representations;* Hugging Face GAIA Leaderboard GAIA Leaderboard - a Hugging Face Space by gaia-benchmark.

[29] As another example, data from the benchmarks may get incorporated into models' training data over time:https://arxiv.org/abs/2502.06559v2 Eriksson, M., Purificato, E., Noroozian, A., Vinagre, J., Chaslot, G.,

Despite rapid progress, many general-purpose agents are still error-prone in open-ended tasks.[30] However, agent deployment frameworks commonly provide technical controls (governance guardrails, identity and data controls, and human-in-the-loop checkpoints) designed to allow agents to be deployed without "running wild", and auditability and observability features for deployers.[31,32] As an example, in financial services, Agentic AI systems at the tool, assistant and operator levels are already being applied to customer support[33], fraud detection[34], and workflow automation[35], showing how Agentic AI can deliver tangible value even in highly regulated industries.

Many AI agents today specialise in domains like software engineering, research, customer assistance, and personal productivity. Significant venture capital investment amounting to billions for agent-focused startups may signal expectations of broader deployment and systemic impact within and beyond these domains in the near future. This combination of potential for rapid scaling, technical limitations, and growing autonomy creates both opportunities and risks.

Gomez, E., & Fernandez-Llorca, D. (2025). Can we trust ai benchmarks? an interdisciplinary review of current issues in ai evaluation. https://arxiv.org/pdf/2502.06559v2

[30] See section 1.2 in 2026 Report: Extended Summary for Policymakers | International AI Safety Report

[31] Microsoft, Governance and security for AI agents across the organization - Cloud Adoption Framework | Microsoft Learn.

[32] Github, Safety and Security for AI Agents - Agent Development Kit.

[33] JPMorgan Chase's Gen AI implementation: 450 use cases and lessons learned - Tearsheet

[34] Fight money laundering with the power of Agentic AI

[35] Case study: Accelerating compliance transformation with Generative AI and BNY Collaborates with Google Cloud to Advance its Eliza AI Platform with Gemini Enterprise - Dec 8, 2025

3. Emerging Opportunities and Risks

Research into the opportunities and associated risks of Agentic AI continues as the technology develops. This section picks out and outlines some of those of particular relevance to the remits of our member regulators. It focusses on current and emerging developments, whereas Section 5 focuses on potential future developments.

3.1 Emerging Opportunities

Agentic AI in the household

The promise of Agentic AI for individual users is to act as a "delegation layer" between people and the digital services they rely on. Instead of users navigating multiple websites, forms, logins, comparison tables, and customer-service queues, an agent could translate a person's intent ("sort my bills", "renew this policy", "find me a holiday within X budget and book it") into a sequence of steps across tools and services, and execute the steps.

For consumers, this functionality could translate into reducing the time, cost, and friction of everyday "life admin" in the following ways:

1. **Reducing search costs during discovery**. Firstly, conversational agents could translate natural-language needs into structured product comparisons, aggregate results across retailers, and calculate total prices (including fees) thereby cutting discovery time for every consumer.[36]

2. **Improving deal quality through ongoing comparison**. Secondly, it could make shopping-around easier and improve deal quality and access to products and services: assistants could monitor prices across sites and over time, highlight better-value alternatives, and apply valid coupons automatically, helping consumers "shop around" efficiently to get the best deal rather than navigating multiple tabs. Industry reporting indicates AI shopping assistants are being deployed at scale to streamline product discovery and price comparisons.[37]

3. **Supporting switching and cancellation**. Agents could guide or execute switching flows where permitted (e.g., Nous' best-deal comparison service for household bills[38]) and help consumers navigate "tricky" journeys by pre-populating forms, tracking deadlines, and retaining records of consent. These benefits are potentially valuable for everyone, and particularly helpful for people facing access barriers (e.g., disability-related needs or limited English), where assistants can simplify steps, read interfaces aloud or translate content.[39]

---

[36] Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics*, *140*(2), 889-942., https://doi.org/10.1093/qje/qjae044

[37] Capital Computer Group, How AI Assistants are Already Reshaping Shopping

[38] Nous, Nous – Super smart AI for your household bills

[39] Realising these benefits for disabled users and limited-English speakers depends on addressing access, skills, and accessibility barriers (devices/data costs; low digital confidence; WCAG gaps; speech-recognition limits for atypical speech; weaker LLM performance in low-resource languages) and designing for trust (privacy-respecting, dark-pattern-free flows).

Agentic AI at Work

For firms, Agentic AI could be transformative both in consumer engagement (front office) and in internal operations (back office). On the customer side, agents can provide more responsive and personalised support, triage issues, and complete workflows (refunds, changes, onboarding) while escalating to humans when appropriate. A large-scale study of a generative AI assistant in customer support found productivity improvements of around 14–15% (issues successfully resolved per hour), with particularly large gains for less experienced workers.[40]

On the back-office side, Agentic AI extends existing AI capabilities by autonomously chaining specific steps across enterprise systems (e.g., retrieve evidence via connectors/APIs, draft and format a report, create tickets or follow-ups, and file or route for approval) using tool calls and workflow orchestration under enterprise guardrails. For example, Allianz has launched an agentic AI solution to automate food spoilage claims and cut processing times, involving seven specialised AI agents which autonomously retrieve evidence from multiple sources and collaborate to process a single claim.[41]

While not agentic-specific, evidence from workplace deployments suggests that AI assistants could deliver meaningful time savings on routine tasks: for example, the UK Government Digital Service's (GDS) large-scale trial of Microsoft 365 Copilot (20,000 staff) reported average self-reported savings of 26 minutes per day.[42]

Agentic AI for Regulators and in Regulation

AI, including Agentic AI, offers multiple opportunities to support regulators and facilitate delivery, as well as compliance with regulation. For example, it can aid in detecting issues, in reviewing and analysing large volumes of documentation, and in drafting. The CMA has built data and AI capability and uses this in many areas of its work, including for the detection of bid-rigging,[43] a key risk in public sector procurement which represents more than £300bn+ of UK public spending each year.[44]

Agentic AI presents a significant opportunity to further the digital transformation of regulatory organisations.[45] Agentic supervision can flag to human supervisors (for example, under the Senior Managers and Certification Regime introduced by the FCA) when a market participant's behaviour deviates from declared controls, always with the caveat that human supervision remains meaningful and effective. By enabling AI to navigate consumer journeys, regulators can scale up the identification of harms and bring humans in where they add most value, e.g. in analysis and decision making. For example, as part of its wider programme of digital transformation, the CMA has built and deployed Agentic AI to detect consumer harms such as drip pricing at scale.

3.2 Emerging Risks

This section presents a brief overview of current and emerging risks, including those that apply to AI more broadly and are covered elsewhere.[46] Some of these risks might be increased in the case of

[40] Brynjolfsson et al. (2025) https://www.nber.org/papers/w31161
[41] Allianz, Allianz launched its first agentic AI to automate claims
[42] GOV.UK, Microsoft 365 Copilot Experiment: Cross-Government Findings Report - GOV.UK
[43] Advice for public procurers - Cheating or competing
[44] AI and collusion: frontiers, opportunities and challenges – Competition and Markets Authority
[45] CMA, AI and collusion: frontiers, opportunities and challenges – Competition and Markets Authority
[46] See for example: GOV.UK, The benefits and harms of algorithms: a shared perspective from the four digital regulators - GOV.UK

Agentic AI, and some may be new risks. These are explored in greater detail, together with associated regulatory considerations, in Section 5 on Future Developments.

It is worth noting that different parts of the value chain all have a role to play in mitigating risks - for example, in addressing the "many hands problem" where accountability is fragmented across model providers, system providers, and downstream deployers.[54] Model providers may build foundational infrastructure for monitoring, logging, and executing emergency shutdowns; system providers may adapt these tools to context-specific risks; and downstream deployers may implement oversight and reporting mechanisms during operation.

Data Protection

AI agents operate within a complex risk landscape that spans privacy, accountability, market dynamics and security.[47] It should be noted that in some cases, the risks posed by Agentic AI are not necessarily different to those posed by e.g. generative AI. However, in some instances, because of their autonomous capabilities, these systems have the potential to exacerbate them or create new data protection risks. Agentic systems often require access to large volumes of personal and operational data, which may be shared across multiple agents and integrated with external tools. This could have implications for data protection and lawful processing, particularly where automated decision-making affects individuals' rights or interests.[48] The ICO has set out its early thoughts on the data protection risks and opportunities of Agentic AI in a dedicated Tech Futures report.[49] Also see Section 5.

Vendor Lock-In

The risk of vendor lock-in could also be significant: as agents become deeply embedded in workflows, organisations could find themselves tied to a single provider's infrastructure, limiting interoperability and increasing dependency. On the other hand, Agentic AI could reduce lock-in to other services, it may operate as the connector or the 'glue' between different systems for example, by making it easier to consider accommodation offers across multiple platforms. This has wider implications for market dynamics, including competition, and consumer outcomes.[50] Also see section 5.4.1.

Collusion

Research shows that LLM-based agents can spontaneously collude in simulated settings, converging on and maintaining supra-competitive prices or coordinating quantities, despite no explicit instruction to do so.[51] More speculatively, future ecosystems of interacting agents could be incentivised to develop hidden strategies and covert communication (including non-natural language protocols) to pursue shared goals. Such spontaneous, concealed coordination could scale across networks, producing harmful behaviours that are hard to detect. While findings so far have come from controlled, rather than real-world, settings, this may warrant caution in deploying agents in

---

[47] Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). Ai agents vs. agentic ai: A conceptual taxonomy, applications and challenges. *Information Fusion*, 103599.https://arxiv.org/abs/2505.10468.

[48] https://doi.org/10.48550/arXiv.2406.08689He, Y., Wang, E., Rong, Y., Cheng, Z., & Chen, H. (2025, April). Security of ai agents. In *2025 IEEE/ACM International Workshop on Responsible AI Engineering (RAIE)* (pp. 45-52). IEEE. https://arxiv.org/pdf/2406.08689

[49] ICO, ICO tech futures: Agentic AI | ICO

[50] https://doi.org/10.48550/arXiv.2504.16736Yang, Y., Chai, H., Song, Y., Qi, S., Wen, M., Li, N., ... & Zhang, W. (2025). A survey of ai agent protocols. *arXiv preprint arXiv:2504.16736*. A Survey of AI Agent Protocols

[51] See Section 5.4.2.

roles like pricing and may require businesses and regulators to adopt new monitoring, detection, and control tools. Also see section 5.4.2.

Action Bundling and Choice Outsourcing

One more agent-specific risk specific to AI agents is a potential complication of regulatory oversight and accountability due to the "bundling" of actions: agents can execute multi-step workflows that would normally be separate interactions and can do so rapidly or in parallel. For example, a shopping or switching agent might simultaneously (i) pull personal data from several sources, (ii) compare products and make a recommendation, (iii) accept terms and initiate payment, (iv) share data with third parties for fulfilment, and (v) send confirmation messages, without the user experiencing each step as a separate decision. This raises questions about whether it is sufficiently clear to consumers what they are tasking the agent with, as well as at what point the consumer will be re-engaged in the process.

Another emerging risk is "choice outsourcing." As Agentic AI becomes more autonomous and convenient, consumers may offload more of the search, comparison and transaction steps to their assistants. While this can save time and reduce decision costs, it may also suppress active "shop-around" behaviour and shift the assessment of product quality to the assistant layer. Where assistants are integrated with, or monetised by, platforms, their ranking defaults and incentive structures can channel users to a narrow set of options, reduce transparency and entrench incumbents, so that firms end up competing to be favoured by the assistant rather than to offer the best product or price. These concerns align with regulator findings that misleading algorithmic choice architectures or design aspects affecting user experience may obscure material information, nudge users toward platform-preferred outcomes and weaken rivalry.[52]

Accountability and control could present further challenges. Agentic systems are inherently adaptive and can make consequential decisions without direct human intervention. Without robust governance and meaningful human oversight, these systems, like all AI systems, risk becoming opaque "black boxes", making it difficult to explain or contest decisions.[53] This lack of transparency could lead to unsafe, illegal or to unintended outcomes, such as an agent incorrectly denying a refund or redress, sharing personal data with an unintended third party, or nudging a consumer toward an unsuitable or poor-value product that doesn't meet their needs. Agentic AI will deliver greatest consumer value and be trusted when autonomy is bounded clearly by user intent and backed by strong transparency and accountability.[54] Also see Section 5.1 and 5.3.1.

Cybersecurity

Furthermore, if there is growth in the interconnectivity of agentic ecosystems this could expand the risk of cyber-attack by malicious actors, increasing cybersecurity vulnerabilities.[55,56] Agents may be given broad permissions and access sensitive data (email, browsing history, customer records), which they could reveal by mistake or which attackers could use to misuse privileges or exfiltrate data. Prompt-injection risks grow because autonomous agents ingest and act on untrusted content

---

[52] Frontier Firms, Technology Diffusion and Public Policy | OECD

[53] Mukherjee, A., & Chang, H. H. (2025). Agentic AI: Autonomy, accountability, and the algorithmic society. https://arxiv.org/abs/2502.00289

[54] Agentic AI and consumers - Competition and Markets Authority

[55] Dal Cin, P., Kendzior, D., Seedat, Y., & Marinho, R. (2025). Three Essentials for Agentic AI Security. *MIT Sloan Management Review (Online)*, 1-4.

[56] Disrupting the first reported AI-orchestrated cyber espionage campaign \ Anthropic (last accessed 19 Feb 2025)

from many sources, including the open internet. Their ability to take actions makes them attractive targets, and potential use of Non-Human Identities without strong session oversight could enable unauthorised access or impersonation. Also see Section 5.2.2.

Risks linked to hyper-personalisation

Hyper-personalisation can heighten manipulation, fraud and discrimination by exploiting users' cognitive vulnerabilities through highly persuasive AI-generated messaging, which research shows is significantly more effective than traditional advertising.[57] These same techniques can be used for personalised fraud, as demonstrated in documented AI-driven influence campaigns that craft identity-specific narratives to build trust and evade detection.[58] Additionally, hyper-personalised systems can inadvertently produce discriminatory outcomes by tailoring prices, access, or opportunities based on behavioural inferences, a risk flagged in marketing ethics research that identifies how such practices can erode autonomy and lead to unequal treatment across groups, including those with particular vulnerabilities.[59]

In a nutshell

In sum, the technology has the potential to augment risks that are relevant to all UK regulators. The range of new or amplified risks described above points to the importance for regulatory frameworks to be effectively applied across the value chain. We note that Agentic AI may raise broader societal and environmental concerns, falling outside the remits of DRCF member regulators and are not covered in this publication.

On the other hand, the opportunities for firms and consumers alike are significant and could support greater efficiency and new opportunities. If organisations deploying agentic systems adequately mitigate the risks, including by complying with relevant law and regulatory frameworks, such benefits can support UK economic growth and productivity, for example by reducing time spent on routine tasks and freeing up human resource for higher value tasks.

[57] LeBrun, A. The Risks of AI-Generated, Hyper-Personalized Digital Advertisements. *Philos. Technol.* **38**, 99 (2025). https://doi.org/10.1007/s13347-025-00935-z

[58] Sonalkar, S. One Person, One Lie: Hyperpersonalised Influence Campaigns in the AI Age (ORF) (2025).

[59] Prem, R. (2025). Hyper-Personalization in Digital Marketing: Evaluating Consumer Trust and Brand Loyalty in the Age of AI-Driven Campaigns.

4. Ensuring Adoption of AI Agents is Safe and Trusted: An Overview

The 2023 UK Government AI White Paper[60] and 2025 AI Opportunities Action Plan[61] set out a principles-based, sector led, pro-innovation approach to regulating AI, including Agentic AI.[62] The purpose of the UK's AI governance model is to encourage innovation while managing risks and poor outcomes in specific sectors. Regulators play a crucial role in protecting people from harm while enabling the safe, effective use of AI such that its benefits can be widely realised.

DRCF Member Regulator Remits

Agentic AI is a clear example of technology that straddles multiple sectors of the economy and therefore regulatory remits, including data protection, consumer protection, online safety, competition and sector-specific remits such as financial services. DRCF member regulators are committed to supporting broad adoption of Agentic AI technologies in a way which safeguards trust and delivers maximum benefit to the UK. Collaboration via the DRCF helps member regulators support safe innovation and maximises opportunities from Agentic AI while also managing potential harms, by coordinating our understanding of risks and seeking alignment where appropriate.

The same agentic feature, e.g., an autonomous customer agent with tool access, could simultaneously raise data protection, financial regulation, consumer protection, online safety, and competition issues – areas covered by each of the four DRCF's regulators. The following example may illustrate how one Agentic AI system may impact all DRCF regulator remits.

Imagine a large UK retailer deploying an autonomous customer assistant powered by Agentic AI. The assistant helps customers by recommending products and bundles (including substitutes when items are out of stock), answering queries and complaints, providing order status updates, negotiating and booking delivery or collection slots, arranging returns, and initiating refunds or exchanges; it can also apply loyalty discounts and personalised offers. To do this, it can access multiple tools and systems, including pricing, inventory and order management, delivery scheduling, payments/refunds, and knowledge bases, and read relevant customer profiles, purchase histories, and loyalty status; it has authority (within preset policy and spend thresholds) to place or modify orders, apply discounts, issue vouchers, and escalate to human agents; it integrates with third parties such as payment processors, credit reference agencies, insurers, and couriers; it is connected to further AI agents for dynamic price calibration and just-in-time supply coordination. To be clear, this example concerns an autonomous customer assistant deployed by a single UK retailer to help customers navigate its own services and catalogue; it does not provide general search or comparison across third-party retail sites.

[60] AI regulation: a pro-innovation approach - GOV.UK
[61] AI Opportunities Action Plan - GOV.UK
[62] AI Opportunities Action Plan - GOV.UK

In this example:

- From a data protection perspective, depending on the context, the ICO may be concerned with how the assistant accesses and processes personal data, in particular whether processing its automated decisions - such as offering credit or applying loyalty discounts - can trigger data protection provisions on solely automated decisions with legal or similarly significant effects, which require meaningful human involvement. The ICO would also consider whether the retailer is meeting its transparency requirements when communicating with customers about how their data is processed.
- The FCA's remit could be engaged, for example if the assistant recommends financial products such as store credit or insurance. Depending on the nature of the assistant's services (e.g. if it provides regulated financial advice or arranging activities), the firm and/or the use case could be subject to FCA authorisation and/or the FCA's regulatory frameworks. These frameworks include the Consumer Duty, which requires firms to deliver good outcomes for consumers, including avoiding foreseeable harms and supporting consumers to achieve their financial objectives.
- Ofcom's remit could be engaged, for example, if the assistant retrieves information from multiple websites that is presented back to the user as part of the task execution process. For example, the agent may interact and extract information from other retail websites selling similar items to compare pricing or provide alternative options if an item is out of stock. To the extent that this counts as "searching… more than one website or database", the agent in question could be deemed as including (or being) a search engine under the OSA, and potentially therefore being a regulated search service. This would mean that the provider would be required to assess and mitigate the risk of their users encountering both illegal content (e.g. fraud) and content that is harmful to children (e.g. body stigma content).[63]
- Both competition and consumer protection elements of the CMA's remit could be relevant. Consumer protection could be relevant if, for example, actions taken by the assistant affect whether the consumer is able to receive a refund; if the related contract terms are unfair (e.g. in relation to subscription, cancellation or handing over data in return for 'free' services); if the approach the assistant takes is not sufficiently transparent; or if the firm that provides the assistant steers users away from rivals' offers. The CMA's competition remit may be engaged if, for example, the agent's activities were to amount to collusion,[64] or otherwise to the exchange of commercially sensitive information with its competitors.

Of course, beyond the DRCF member regulators' frameworks, market participants will need to consider other governance frameworks in deploying an agent, e.g. cybersecurity standards and regulations.

Nonetheless, the example deployment above illustrates how one agentic AI application can simultaneously activate cross-regulatory concerns in privacy, financial fairness, online safety, and competition and consumer law. Further detail is given in the following sub-sections.

[63] whether or not an agent's search tool is regulated under the OSA is subject to the facts of a particular case and cannot be assumed.

[64] AI and collusion: frontiers, opportunities and challenges – Competition and Markets Authority

ICO

As the UK's data protection and information rights regulator, the ICO aims to ensure that innovation in agentic AI develops in ways that protect people's information rights, while providing clarity and support for organisations. The ICO's remit[65] covers the whole economy (both the private and public sector). Some potential compliance issues resulting from AI agents include automated decision-making (ADM), profiling, accountability, fairness, accuracy (for example due to cascading 'hallucinations'), data security, purpose or storage limitation, and data minimisation.

The ICO's initial guidance on ADM and profiling[66] is currently under review because of the Data (Use and Access) Act reforms, and an update will be published in March 2026. The ICO will also be tasked with developing a statutory Code of Practice on AI and ADM that is likely to address some of the emerging concerns around AI agents. The ICO stresses that organisations should maintain records, capture decision rationales, and conduct Data Protection Impact Assessments (DPIAs) where risk to people and their rights is high. It has invited stakeholders to create fit-for-purpose governance tools and frameworks for this technology.[67]

For AI agent providers and deployers, this points to two priorities: ensuring agents have traceable logs and explanations that can be provided to data subjects, and designing governance so that human involvement in significant or legal decisions is not tokenistic but genuinely capable of oversight and intervention in the agent's chosen actions.

FCA

As the UK's financial services regulator, the FCA sets standards for financial services firms to meet and holds them to account if they don't. Their work helps underpin the UK's position as a leading global financial centre, including by supporting firms to experiment, develop and test AI (including Agentic) in a way that drives innovation, benefits consumers and markets, and supports the growth and competitiveness of UK financial services.[68] Given the FCA's outcomes-focused approach, there are no specific requirements for AI agents (or AI generally). Where agents are embedded into the delivery of regulated financial services, firms remain subject to the FCA's existing rules and regulatory expectations. Relevant requirements will depend on the use case. For example, if a firm's agent is interacting with financial services consumers to price financial products, arrange deals in investments, triage insurance claims, provide financial advice or make payments, firms may be required to demonstrate how they are complying with the Consumer Duty and other requirements of the FCA's regulatory framework. This could include how the firm is monitoring consumer outcomes and ensuring consumers are given the information they need to make informed decisions at the right time, presented in a way they can understand.

In deploying agents, firms may need to consider whether they have: the necessary permissions (e.g. if providing financial advice); appropriate risk controls;
accountability frameworks (e.g. under SM&CR); and appropriate detection and treatment of vulnerable consumers. The FCA's April 2024 AI Update sets out how key elements of its existing regulatory frameworks apply to firms' use of AI.[69]

---

[65] ICO oversees the processing of personal data wherever that takes place, to ensure people and their rights and freedoms are protected.
[66] ICO, 'Automated decision-making and profiling': Automated decision-making and profiling | ICO
[67] ICO, 'We can't regulate the technology of the future with the tools of the past'
[68] AI and the FCA: our approach | FCA
[69] FCA's AI Update | April 2024

Ofcom

As the UK's converged communications regulator, Ofcom invests significant time and resource in understanding how AI innovations affect the sectors it regulates. From telecoms security to broadcast content, and from online safety to spectrum management, AI agents promise to disrupt traditional service delivery, business models and consumer behaviour.

In line with Ofcom's Strategic Approach to AI 2025/2026, Ofcom is actively working to understand AI agents and how to help industry unlock the current and future benefits. Telecoms providers may use AI tools and technologies to deliver products and services. We regularly assess the impact that technological developments might have on our sectors, both to understand the opportunities they can offer and any implications they have for how we should further the interests of consumers. Ofcom is exploring how this may have the potential to change customers' experience of telecoms markets and whether any changes to its regulatory framework are required, either to enable telecoms providers and customers to experience the benefits of AI or to protect customers from potential harms.

Where AI agents fall in scope of Ofcom's regulatory regimes, it will take account of the risks arising out of AI agents when considering how best to exercise its functions. For example, for UK online services, where an AI agent operates as a user-to-user service or a search service, it will be regulated under the Online Safety Act, and the service provider will be expected to comply with relevant duties. Ofcom is also developing additional media literacy resources to support media literacy stakeholders, which includes the foundational skills needed for everyone to have good media literacy in response to an increasingly AI-mediated information space.

CMA

The CMA's purpose is to promote competition and protect consumers with a clear end goal – to drive economic growth and improve household prosperity.[70] The CMA has worked on algorithms and AI for many years, analysing technologies, unpacking implications for consumers, competition, and growth, and supporting positive innovation through proactive engagement, published principles to guide the market towards positive outcomes, and practical guidance for businesses.

The CMA's earlier in depth publications on AI foundation models considered the opportunities as well as the potential risks from these powerful technologies, including risks that could arise if developers or deployers of AI tools or services engage in anti-competitive conduct.[71] These risks are particularly acute where players with significant market power embed Agentic AI into platform ecosystems or tool marketplaces. Competition issues could arise where dominant firms engage in anti-competitive restricting of interoperability, imposing unnecessarily high switching costs, and tying and bundling. Consumer protection issues could also arise in these circumstances, such as through insufficient transparency of agent functionality and limitations. Further, the CMA has explored the opportunities and risks from algorithmic pricing, including the risk of collusion in Agentic AI systems.[72]

In addition to the CMA's functions in relation to competition enforcement (under the Competition Act 1998) and market studies/investigations (under the Enterprise Act 2002), the Digital Markets, Competition and Consumers Act (DMCCA) enables the CMA to designate firms with substantial and

[70] CMA, ' CMA Strategy 2026 to 2029' (20 November 2025).
[71] AI Foundation Models: initial review - GOV.UK
[72] AI and collusion: frontiers, opportunities and challenges – Competition and Markets Authority, published 4 Mar 2026

entrenched market power in a particular digital activity as having strategic market status. The DMCCA enables the CMA to impose remedies to address conduct such as self-preferencing, data leveraging, or exclusionary terms.[73]

The CMA is the UK's primary consumer protection authority. The consumer protection regime, including the Consumer Rights Act 2015 and Part 4 of the DMCCA, sets out the law in relation to unfair contract terms and unfair commercial practices respectively, and governs firms' interactions with consumers including firms using AI in customer service. [74] Under the DMCCA, the CMA can investigate concerns and decide directly whether consumer protection laws have been infringed. The CMA can then tackle any breaches directly, including through consumer redress and fines.[75] Along with other enforcers including trading standards and some sector regulators, the CMA can also take businesses to court. Where infringements are found to have occurred, the court can order a change to conduct and impose fines on businesses.

Sections 5.1 and 5.3 further set out the potential demand-side issues linked to Agentic AI, e.g. potential concerns relating to consumer-facing interfaces, and how this also engages the CMA's remit.

UK consumer law applies whether decisions are made by people or by AI. The CMA's foundation model principles – particularly transparency and accountability – remain directly relevant, and the CMA has published guidance to help businesses using agentic AI to comply with consumer law. Businesses exploring the technology should focus on robust training of systems, monitoring, and refinement, supported by appropriate human oversight. The CMA's work has also considered wider enablers to unlock the full potential of agentic AI for the UK such as smart data schemes, secure digital identity and strong interoperability standards – enabling consumers to adopt with confidence, switch between systems and exercise choice.

In sum, the UK's consumer protection regime, sitting alongside its competition laws, together form a framework designed to ensure fair play, innovation, and consumer protection and already provide a strong foundation for addressing related challenges that may arise with Agentic AI. Existing principles – grounded in fairness, transparency and accountability – apply regardless of whether consumers are interacting with people or AI systems. Businesses that embed these principles into the design and deployment of agentic AI will be best placed to build trust, scale responsibly and compete on the quality of outcomes delivered to consumers.

[73] CMA, 'CMA sets out initial plans as new digital markets competition regime comes into force' (7 January 2025).
[74] Complying with consumer law when using AI agents - GOV.UK
[75] CMA to boost consumer and business confidence as new consumer protection regime comes into force - GOV.UK Published 7 Apr 2025

5. Potential future developments and regulatory considerations

This section sets out potential future developments in Agentic AI and early thoughts on their cross-regulatory implications. Subject matter experts from the four regulators examined scenarios based on different levels of agentic capability and adoption (see Annexes). Stakeholders from industry, academia and civil society contributed perspectives on opportunities, emerging harms, adoption barriers and regulatory expectations through a Call for Input (CFI) on Agentic AI.[76] The combined insights inform the analysis presented here.

The medium-term future of Agentic AI is likely to vary across sectors, as organisations and consumers assess where these systems provide value and learn how to deploy them to best use. Variables influencing adoption will involve a range of technical, commercial and behavioural factors, each shaped by complex dynamics and multiple stakeholders. Key variables include, for example:

- How reliably AI agents perform in real-world contexts;
- The extent to which consumers and organisations understand and act on AI agents' capabilities and limitations;
- How providers evidence or signal the trustworthiness, safety and performance of Agentic AI systems;
- Which actors – consumers, industry, or the public sector – drive early adoption;
- The prevailing commercial models;
- How the AI value chain evolves to accommodate and support agent-mediated interactions;
- The level of platform integration;
- The extent to which commercial arrangements and wider incentives allow AI agents to act on users' behalf across digital services; and
- How legal and regulatory requirements support the positive development and integration of Agentic AI into the economy.

Individual member regulators will continue to consider how to enable the safe and responsible adoption of agentic AI, including if there are opportunities to provide further clarity if needed. DRCF member regulators will also continue to work together to support a clear and coherent approach to Agentic AI regulation.

In the sections below, we consider potential future developments and their regulatory implications in the domains of governance, data protection and cybersecurity, consumer rights and interests, and market dynamics and competition.

[76] Call for Views: Agentic AI and Regulatory Challenges | DRCF

5.1 Governance considerations

5.1.1 Governance of complex Agentic AI systems

The cross-regulatory engagement and the DRCF's call for input on Agentic AI that informed this paper highlighted that the increasing autonomy of Agentic AI systems poses important issues about accountability, liability, control and governance in the digital economy.

DRCF regulators and their stakeholders recognise the importance of effective governance and appropriate guardrails for Agentic AI development

In the CFI, respondents from different industries acknowledged that there are existing frameworks for governance and accountability requirements in relation to Agentic AI. Respondents noted that further work between regulators, government, industry and civil society may be needed to ensure these frameworks keep pace with evolving Agentic AI use cases and that requirements remain clear. Respondents encouraged collaboration between regulators, government, industry and civil society to clarify how existing rules and frameworks apply when AI agents act on behalf of consumers. They also asked regulators to consider the responsibilities of companies developing the underlying models and agents, including in relation to fraud.

Several respondents suggested that regulators should require firms to adopt AI supply chain governance frameworks that address component integrity, compatibility and risk propagation. Industry respondents further highlighted the practical difficulty of monitoring every step of multi-agent decision-making and suggested a more risk-based approach that focuses on oversight of outputs and outcomes, though there was recognition that it is important regulators take a proportionate approach.

Agentic AI autonomy does not remove organisational responsibility for regulatory compliance

Although Agentic AI systems operate with a degree of autonomy, this does not remove organisational responsibility for data processing and compliance with the law.

Organisations considering Agentic AI applications must understand how these systems interact with resources they access (datasets, for example) and ensure appropriate governance and controls are in place. Relevant legislation, including data protection, consumer law, competition law, financial regulation and online safety requirements, must be complied with.

AI observability and explainability supports appropriate governance

Observability is the ability for deployers to understand what is happening inside a system by examining its outputs.[77] For example, in Agentic AI systems, these outputs might include logs of user interactions and reasoning steps, traces of which actions were performed by which agents, and metrics of how often the system has been used, as well as other outputs.[78]

Incorporating observability into Agentic AI systems supports effective governance. The ability to identify which parts of an Agentic AI system made which actions and monitor when parts are working as intended can support effective practices and, if needed, mitigations. To comply with legal

[77] What Is Observability? | IBM
[78] What is AI Observability? | IBM

obligations around consumer protection, it is important that businesses have appropriate processes in place to monitor AI agents.[79]

Appropriately surfacing outputs to users can provide additional benefit. When considering data protection, observability can help organisations be transparent about processing and enable them to explain how individuals' personal data is being used. The ICO has published guidance for organisations on the explainability of AI.[80]

5.1.2 Determining the responsible party in autonomous systems

Our engagement highlighted concerns that, in future, it could be harder for organisations and consumers using more advanced agents or agentic AI to determine accountability under existing laws in the event something goes wrong. Some aspects of the issue are explored below,

Greater autonomy, increasing interactions and a lack of transparency in interactions in multi-agent systems and cascading errors may make it harder to establish responsibility

Without appropriate safeguards, higher capability agents might act or integrate with a wide variety of services on their own, bypassing technical constraints and governance mechanisms such as human involvement in decision-making and/or intervention and review.

As agents interact and transact with each other, the blending of datasets without human involvement or oversight might also make it harder for organisations to establish controllership of that data and to identify who is responsible for compliance under UK GDPR.

Respondents to the CFI called for regulatory cooperation to clarify legal responsibility across the ecosystem under the different intersecting DRCF member regulator regimes.

Cascading errors may amplify negative impacts for consumers and markets

In the CFI, some respondents noted that flaws in one agent can cascade across interconnected systems, amplifying negative impacts and making it difficult to identify the source of the error and responsible actor. Examples centred on the financial services industry and the Moffatt v. Air Canada case,[81] but these challenges are not limited to specific sectors.

Consumers may find it more difficult to identify which organisation they need to contact in order to exercise their rights under consumer and data protection laws

For consumers using agents, the increasing opacity of multi-agent interactions could make it harder to verify the efficacy of AI tools, to identify which organisation to contact to raise complaints with, or to exercise their consumer rights and data protection rights.

Agentic systems designed to monitor inter-system transactions provide an opportunity for increasing transparency

Conversely, in future we may also see 'transparency agents' whose purpose is to establish and monitor the transactions between agentic AI systems and maintain a clear audit trail of accountability. This additional level of insight could empower consumers to have a much more transparent view of relationships between organisations.

---

[79] Complying with consumer law when using AI agents - GOV.UK

[80] Explaining decisions made with AI | ICO

[81] Airline held liable for its chatbot giving passenger bad advice - what this means for travellers - BBC Travel

5.1.3 Autonomy and the 'human in the loop'

DRCF regulators and CFI respondents recognise human-in-the-loop measures as one category of governance mechanism that supports responsible AI deployment, including for Agentic AI systems. Human-in-the-loop measures aim to ensure that a human retains ultimate decision-making authority for important choices.

Several CFI respondents perceived a tension between legal requirements in certain domains, such as data protection law, for different types of human control, and the increased automation offered by future Agentic AI systems.

Organisations must ensure a human in the loop in certain circumstances. For example, under data protection law, the inclusion of human supervision in Agentic AI systems is necessary where personal information is being processed and decisions taken by the system could have legal, or similarly significant effects on a person. The ICO has published guidance for organisations considering using automated decision making (including profiling) to help them understand these requirements in a data protection context.[82]

Depending on the sector and context where an agentic system is being used, human-in-the-loop oversight means that users may be better able to understand the abilities and intended actions of the Agentic AI. This may also provide reassurance that they remain in control and have the ultimate authority over more serious decisions and actions. Having a human in the loop can further help to mitigate some consumer risks and increase consumer trust in these systems.

Enabling human observation of the decision-making process can be a key means of establishing control over largely autonomous systems. Establishing the point at which the AI will need to seek the permission of a human being to continue down a particular course of actions is important. Many current systems allow a user to pause activity or inspect the decisions and actions of the Agentic AI system in real time.

Permissions controls might, for example, be based on a point at which the system would need to provide payment information. Permissions might also be set around which data sources an AI could access and process.

Appropriate permissions could also mitigate the possibility of "consent fatigue", where users become tired of reading why they are being asked to give the system a go-ahead, and begin to automatically agree.

[82] Automated decision-making and profiling | ICO

5.2 Data protection and cybersecurity considerations

Keeping data protection in mind throughout the whole lifecycle of the Agentic AI system and addressing cybersecurity risks were identified as key considerations in our workshops and CFI, irrespective of the capabilities or level of adoption of Agentic AI systems.

5.2.1 Data protection and Agentic AI

The data protection implications of the use of Agentic AI are explored in detail in the ICO's Tech Futures Report on Agentic AI. Many of the data protection considerations are similar to those raised by other types of AI. However, some characteristics of Agentic AI systems can exacerbate existing risks or may introduce new ones.

Part of mitigating those risks is understanding to which resources the system will have access. If it accesses a database which contains personal data, then data protection obligations would apply to that processing. All data protection provisions apply to agentic AI that processes personal data. This section highlights two specific aspects of those provisions: data minimisation and data protection by design and default.

The data minimisation principle applies to Agentic AI systems

There may be a temptation to allow Agentic AI systems broad or unfettered access to data and resources to improve performance or accuracy.[83] However, the data minimisation principle requires organisations to use only that personal data which is necessary for the purpose for which it is processed. By limiting access to only the information which is appropriate to achieve the task which the AI has been employed for (in the same way that employee access might be tiered), organisations can comply with that principle. Organisations will need to ensure that they maintain oversight to ensure that the system's collection, use and sharing of personal data remains appropriately limited to that which is necessary to complete its task throughout its operation.

Data protection by design and default will help organisations be transparent with people about how their data is being used by the Agentic AI

Applying data protection by design and default principles across the lifecycle of an Agentic AI system helps organisations comply with GDPR's fundamental principles and requirements. This includes, for example, processing personal data lawfully and fairly, as well as being transparent with people about how their personal data is being used by the Agentic AI. Transparency is crucial in helping people understand and exercise their rights. This transparency will allow organisations to build consumer trust and is something that was identified by stakeholders as key to Agentic AI's growth as a product.

Some providers of Agentic AI systems may choose to go beyond the requirement for data protection by design and default, embedding further privacy controls to ensure higher levels of trust. This might lead to suppliers focusing on privacy as a competitive advantage and promoting their products on the basis of their strong governance and control over the use of personal data. In contrast, systemic privacy vulnerabilities could undermine overall confidence in Agentic AI systems, reducing demand and impeding innovation.

[83] When data extraction from the open web is automated and undertaken by AI this process is referred to as AI scraping. What is AI Scraping? | IBM

5.2.2 Cybersecurity: new capabilities

Agentic AI's autonomous capabilities, beyond non-agentic AI, may improve detection of, and protection against, cybersecurity threats. The same Agentic AI capabilities may, however, enable more sophisticated, scalable, democratised forms of cyberattacks: as access to Agentic AI systems becomes widely available at reducing cost, malicious actors with no previous technical background may be able to operate those systems to cause widespread damage.[84]

Agentic AI's ability to solve complex problems may enable enhanced cybersecurity for firms and consumers

Agentic AI's ability to solve complex problems may enhance cybersecurity by identifying[85] and responding to threats more quickly.[86] Cybersecurity database analysis could potentially predict future attacks,[87] as well as streamline incident response. Agents may be able to triage, assess and prioritise multiple simultaneous threats to a firm based on the highest risk, allowing firms to manage overwhelming numbers of threats and shifting the focus away from time-consuming manual incident response which depends on humans' expertise.[88]

Agents could amplify existing cybersecurity challenges arising from LLMs and introduce novel attack surfaces

Agentic AI may strengthen cybersecurity in some areas, but it could also increase existing risks. Because these systems may be designed to access a vast range of information, including potential sensitive data such as emails, browser history, and customer records, they can create an expanded attack surface for attackers to exploit. Agents may be granted excessive permissions, which attackers could exploit to escalate or misuse privileges.

One example of this risk is prompt injection, a type of cyberattack against LLMs. This occurs when an attacker hides malicious instructions inside content that the AI processes. It is a well-known risk of Agentic AI,[89] arising from an LLM's inability to distinguish developer or user instructions from malicious input.[90,91] The system may then follow these hidden commands, such as leaking private information, because it cannot always tell the difference between legitimate instructions and harmful ones. This risk is amplified with Agentic AI due to its ability to autonomously scrape data from multiple data sources, including the open internet.

Respondents to the CFI warned that Agentic AI systems may become especially attractive targets for malicious actors because they can take actions on their own. They recommended multi-layered security strategies including stronger testing, threat analysis, "red teaming" (simulated attacks) and combining traditional cybersecurity controls with tools designed for autonomous systems.

Because Agentic AI operates autonomously, the likelihood of an agent inadvertently carrying out harmful activities increases. Agents may operate under Non-Human Identities (NHI) when retrieving data, which do not have the same session-based oversight as traditional authentication.[92] This can

---

[84] Session 3: Impact of Emerging Technology on Cybersecurity | DRCF Cybersecurity
[85] How Agentic AI Enables the Next Leap in Cybersecurity | NVIDIA Blog
[86] AI Agents vs Agentic AI: What's the Difference and Why Does It Matter? | by Edwin Lisowski | Dev Stash
[87] What is an Agentic AI? | CrowdStrike
[88] Advancing Cybersecurity Operations with Agentic AI Systems | NVIDIA Technical Blog
[89] Alignment Under Pressure: The Case for Informed Adversaries When Evaluating LLM Defenses
[90] What Is a Prompt Injection Attack? | IBM
[91] OWASP Top 10 for Large Language Model Applications | OWASP Foundation
[92] Agentic AI - OWASP Lists Threats and Mitigations

make it easier for agents to exploit weaknesses in access controls and, for example, extract sensitive data without permission. It may also allow malicious actors to exploit permission weaknesses to impersonate agents and users.

To prevent undue harm from agentic-enabled cyber-attacks, organisations should consider the appropriate level of data and system access to give to Agentic AI tools and undertake monitoring for malicious agent activity.

Agentic AI could increase cyberwarfare capabilities

The capability of cyberwarfare could be greatly increased through Agentic AI. An attack was recently carried out by a group who used Agentic AI to perform 80-90% of the attack lifecycle.[93] This opens the possibility that one single threat actor could carry out multiple attacks on a firm at once, identifying and exploiting vulnerabilities at speeds far greater than humanly possible, and with minimal trace,[94] making detection more challenging.

Know your Agent protocols could assist with verifying agent identity, intent and authorisation

Several submissions to the CFI supported the development of Know Your Agent (KYA) systems and protocols to verify agent identity, intent and permissions, particularly in commercial settings. They noted potential links to broader digital identity frameworks and initiatives associated with recent data legislation.

5.3 Consumer Rights and Interests considerations

As noted in section 3, Agentic AI systems could deliver significant benefits by augmenting consumer capabilities. At the same time, such systems may create new risks and challenges or amplify existing ones for consumers, particularly when generative outputs are combined with the ability to recommend or execute actions with limited human input.

5.3.1 Consumer protection

Consumer decision making may be impacted by Agentic AI engaging in 'behavioural steering', operating in ways that diverge from user expectations, or otherwise producing inaccurate or misleading outputs

It is important to note that false or misleading statements regarding a system's capabilities may breach consumer protection law.

The cross-regulatory engagement that informed this paper suggested that greater autonomy of Agentic AI systems may amplify potential harms.[95] This is also set out in the CMA's research and analysis on Agentic AI and consumers, which discusses the potentially significant opportunities from Agentic AI as well as important risks and considerations.[96] These systems, whether intentionally or not, may influence consumer decision-making in ways users do not expect, particularly where they are not fully aware of or do not fully understand what they have consented to or delegated. In

[93] Zero-Day Zero: The AI Attack That Just Ended the Era of the Forgiving Internet | Qualys
[94] Zero-Day Zero: The AI Attack That Just Ended the Era of the Forgiving Internet | Qualys
[95] *IBM Study: Growing Customer Acceptance of AI Opens Up New Business Opportunities in UK and Ireland, Trust Remains Key* (30 Sep 2025).
[96] Agentic AI and consumers - GOV.UK

addition, because Agentic AI systems can act independently and adapt over time, their behaviour may diverge from user expectations.[97] As systems become less predictable and more autonomous, the scale and severity of potential consumer harm may increase.

Limited transparency can further prevent consumers from understanding how an Agentic AI system operates or how its decisions may affect them. For example, Agentic AI's decision-making may not only draw on training data and internal heuristics, but also on persistent memory,[98] evolving state representations,[99] and use of external tools, which may collectively lead to biased or hallucinated (i.e. factually inaccurate) outputs across different environments.[100] Stakeholders highlighted that such errors may mislead consumers through, for example, incorrect pricing information, which undermines informed decision making and consumer confidence in markets, ultimately impacting growth opportunities as well as potentially breaching consumer law.

Information and power asymmetry may undermine informed consent and consumer autonomy

Consumers may have insufficient information about the objectives or incentives that shape Agentic AI behaviour. Stakeholders highlighted two potential concerns. First, the systems themselves may be optimised to advance the deployer's objectives, which may be influenced by e.g. undisclosed partnerships, advertising-based or data-extractive business models. Second, external functions or platforms that interact with Agents may themselves be optimised to appeal to the Agent's goals, such as retail websites using 'online choice architecture' (OCA)[101] designed to highlight particular offers. Such mechanisms may exert subtle influence on choices in ways that consumers may be unaware of or unable to contest. This may create information and power asymmetries that could, in some circumstances, undermine informed consent and consumer autonomy, particularly where the agent's goals or preferences differ from those of the consumer. This could, for example, be mitigated by safeguards such as human-in-the-loop (see section 5.1) or preventing AI from carrying out certain functions where a professional qualification or licence is required (e.g. in regulated industries such as healthcare), among other measures.[102]

Statutory consumer rights regarding services and digital content must be honoured by Agentic AI providers

The CMA has recently published guidance for businesses on complying with consumer law when using AI agents, alongside an article exploring how agentic AI is being used today, what the future could hold, and how this will affect consumers.[103] Critically, businesses should be clear and open about their use of AI agents, and should design, monitor and refine AI agents with compliance in

---

[97] Mitchell, M., Ghosh, A., Luccioni, A. S. & Pistilli, G. (2025) *Fully Autonomous AI Agents Should Not Be Developed*.

[98] Information or context (e.g., previous decisions, goals, user preferences) that the agent retains across time or sessions.

[99] The agent's internal representation of the world, goals, or tasks that updates continuously as it gathers new information or executes actions.
Sibai, N., Ahmed, Y., Sibaee, S., AlHalawani, S., Ammar, A., & Boulila, W. (2026). *The path ahead for agentic AI: Challenges and opportunities.*

[101] Online Choice Architecture: How digital design can harm competition and consumers - GOV.UK.

[102] For example, New York lawmakers are advancing Senate Bill S7263, which would prohibit AI chatbots from offering legal or medical advice that would normally require a license from a professional such as a lawyer or doctor. See more details at: 'Proposed New York law would bar AI chatbots from posing as lawyers, allow duped users to sue | Reuters' (accessed 12/03/2026).

[103] Complying with consumer law when using AI agents - GOV.UK ; Agentic AI and consumers - GOV.UK

mind. Under the Consumer Rights Act 2015 (CRA),[104] consumers are entitled to the provision of goods, services and digital content that meet minimum standards for quality, performance, and fitness for purpose. The specific application will depend on the contractual arrangements under which an Agentic AI system is provided, as well as the nature of what is supplied. However, as a general principle, these rules set out standards and give consumers rights when things go wrong. These rights must be honoured by Agentic AI providers.

In addition, the contract by which consumers use Agentic AI will be subject to the CRA's unfair contract terms provisions, which apply between a business and a consumer. Terms must be fair, not imbalanced to the consumer's detriment, and important terms – such as arrangements for cancellation and renewal of a subscription - should not be hidden 'in the small print'. For instance, the unfair contract terms provisions will be relevant to arrangements for cancellation and renewal of a subscription in the contract for Agentic or other AI services.

5.3.2 Digital Inequality

Agentic AI may widen digital inequality

Those with low media literacy skills and limited access to devices and reliable connectivity may struggle to understand the growing integration of Agentic AI systems into everyday tasks. Challenges may include navigating privacy controls, recognising AI generated responses and correcting agent error, exacerbated digital financial exclusion, vulnerability to misleading interfaces, and manipulative design. Increasing integration of Agentic AI into everyday systems may further create a feeling that the only choice available is to opt in or lose out on many services, which can lead to adoption without understanding the output for those with low media literacy skills.

Agentic AI may also challenge those with advanced media literacy skills

Agentic AI may also present challenges for those with more advanced media literacy skills. Agents acting in place of human decision-makers creates an additional layer that can change how information about the external environment is interpreted. A lack of visibility of system-integrated Agentic AI, as with other AI applications, can create an agency problem by limiting the ability to make informed decisions and apply media literacy skills. Even where use of Agentic AI is known, the reduced visibility into the decision-making process may weaken the ability to critically assess digital interactions or distinguish automated inference from fact, fostering disengagement. Over time, this could weaken core skills needed to spot errors, question outcomes and engage critically with digital environments.

Agentic AI could offer opportunities to lower certain barriers to digital participation

Despite these risks, Agentic AI offers opportunities to lower barriers to digital participation, for example through the integration of accessibility tools and options for personalised guidance and tutorials. Similarly, by taking on more complex digital tasks that typically require higher levels of media literacy, these systems can reduce the amount of user input needed. This, combined with the intuitive, conversational design of many AI assistants, may allow those with lower confidence or skills to participate more easily online, though there is the risk that conversational design builds inappropriate levels of trust between users and Agentic AI.

[104] Consumer Rights Act 2015

Ultimately, the impact of Agentic AI on the digital divide will depend on responsible implementation and media literacy skills development. Transparency, data protection, consumer protection and accessible design will be essential to building trust and reducing barriers.

5.4 Market Dynamics and Competition considerations

Agentic AI adoption could impact markets and their dynamics in different ways.

5.4.1. Potential for novel market dynamics and outcomes

Agentic AI could create new markets, but also lead to a 'winner-takes- most' outcome

Widespread adoption of Agentic AI could create new markets, offering new market participants an avenue to develop novel products and services that disrupt existing market participants. This could lead to better market outcomes for customers through innovation, potentially increasing economic growth and productivity, while leading to lower prices and greater choice for consumers.

However, as with other digital markets, there are risks that network effects, economies of scale, and data advantages could allow one or a few large firms to dominate – a "winner-takes-most" scenario.

Data portability could be used as a tool to support competition

Because Agentic AI systems tend to rely on large amounts of data to carry out user requests, new standards may be emerging to support data portability and interoperability.[105]

Respondents to the CFI framed data portability[106] as a fundamental concern crucial to the health of the emerging Agentic AI sector. Increasing the availability of data for personalisation and preventing lock-in preserves dynamic competition. Respondents noted that in the absence of effective data portability, the market for personalised Agentic AI systems may be dominated by a small number of firms that already hold the relevant data in-house.[107]

Interoperable communication protocols could assist in avoiding vendor lock-in

As mentioned in Section 3, open-source agent-to-tool and agent-to-agent communication is currently facilitated by protocols like Model Context Protocol (MCP)[108] and Agent2Agent (A2A)[109] respectively, which are ecosystem and language agnostic. The prevalence of such protocols in future could help reduce the risk of vendor lock-in.

[105] Exclusive: Confident Security debuts OpenPCC standard to encrypt data fed to AI tools; Annex 1: AI and Dataspaces: Shaping the Future of Trusted, Decentralized Intelligence | IDSA Rulebook | IDS Knowledge Base; AI Standards | NIST.

[106] Under the UK GDPR, people have the right to receive personal data concerning them held by a controller in a structured, commonly used and machine-readable format.

[107] You Can Now Import Your ChatGPT Data to Claude for Free | Lifehacker

[108] What is the Model Context Protocol (MCP)? - Model Context Protocol.

[109] A2A Protocol.

5.4.2 Heightened risks of algorithmic collusion, manipulation and hidden strategies

Collusion may take place not only through human decisionmakers, but also through algorithms.[110,111] In Agentic AI systems, algorithmic collusion may include behaviours that arise from the distinctive dynamics of multiagent environments.[112]

Collusion in Agentic AI systems may arise spontaneously

Experiments have shown LLM-based agents colluding in a variety of settings without explicit instruction or training. These settings include price-setting, bidding markets and financial markets: agents repeatedly converged to supra-competitive prices and maintained them, even when the environment was noisy.[113,114,115,116,117] Collusion has been observed not only in AI agents that set price, but also those that set the quantity of goods. LLMs were observed to 'divide markets'[118] and collude by dynamically adjusting their resource allocation strategies, thereby maximising profitability.[119] However, these results should be considered as part of an emerging evidence base: all of these results were obtained under controlled experimental or simulated conditions, rather than in real-world settings.

Agents may be incentivised to develop hidden strategies to collectively pursue their goals

In a potential future where ecosystems of agents interact, negotiate and set strategies to achieve goals in adversarial markets (e.g. trading), agents might also be incentivised to develop hidden strategies to collectively pursue their goals. Research has shown that agents may indeed be capable of developing hidden communication. For example, AI systems can be covertly trained to hide messages within ordinary text, without the user knowing.[120,121] We may also see agents developing faster means of communication as an alternative to natural language, whether through human developed protocols (e.g., gibberlink or Agent2Agent Protocol), or emergent, self-developed ones.

Agentic collusion and manipulation could occur at scale, posing new risks

As a result of the above, spontaneous and hidden agentic collusion might occur at scale.[122] Harmful behaviour could emerge from the way the whole network of agents interacts, even if individual agents do not appear problematic on their own. This might become particularly dangerous if the emerging mechanisms of collusion are novel or difficult to detect.

---

[110] Algorithms: How they can reduce competition and harm consumers - GOV.UK

[111] AI and collusion: frontiers, opportunities and challenges – Competition and Markets Authority

[112] 2024.findings-emnlp.297.pdf

[113] Algorithmic Collusion by Large Language Models

[114] The Invisible Handshake: Tacit Collusion between Adaptive Market Agents

[115] Tacit Collusion by LLM Agents in Construction Bidding: Evidence from a Simulated Bidding Environment | Journal of Management in Engineering | Vol 42, No 2

[116] Ghaemi, M. S. (2025). A survey of collusion risk in LLM-powered multi-agent systems. In *Socially Responsible and Trustworthy Foundation Models at NeurIPS 2025*. A Survey of Collusion Risk in LLM-Powered Multi-Agent Systems | OpenReview

[117] [2507.01413] Evaluating LLM Agent Collusion in Double Auctions

[118] Market division is the practice of dividing a market among a set of competitors to avoid competition.

[119] Strategic Collusion of LLM Agents: Market Division in Multi-Commodity Competitions

[120] https://medium.com/@dr_shahid/covert-communication-in-ai-how-llms-are-learning-to-hide-secret-messages-and-what-we-can-do-ad4ca888b89f[2406.20053v1] Covert Malicious Finetuning: Challenges in Safeguarding LLM Adaptation

[121] [2406.02481v5] Large Language Models as Carriers of Hidden Messages

[122] New Report: Multi-Agent Risks from Advanced AI - Cooperative AI

If this were to occur, businesses would need to be cautious about integrating such agents into roles like price-setting, for example, and regulators would continue to evolve their detection methods. Both might require new tools to monitor and/or manage system behaviour.

6. Next steps

As set out in Section 5, there are many uncertainties as to how Agentic AI systems will be developed, deployed and adopted, how they will play out in the real world and how their deployment might impact markets, the regulatory landscape and society generally. DRCF regulators will monitor developments and adapt their activities accordingly.

As discussed in this paper, existing frameworks and regulatory principles can adapt to technical developments. As Agentic AI develops, the DRCF regulators will continue to explore regulatory considerations and deploy regulatory tools to both support innovation and growth, whilst ensuring consumers are safeguarded against potential harms.

6.1 DRCF next steps

DRCF will continue to shape a joined up regulatory approach to Agentic and other AI technologies, fostering responsible implementation by coordinating our understanding of risks and opportunities, and supporting good outcomes for industry and consumers. We intend to continue the conversation with industry, government, civil society, and other relevant and interested stakeholders, both individually and collectively.

During 2026/27, we are planning horizon-scanning work on (i) the future of interfaces (between users/firms/digital services) and how these may be integrated into operating systems as we know them today, (ii) the future of consumer robotics and physical AI, which has an important agentic component, and (iii) the consumer experience of the near-term future, bringing to life the kinds of changes to our everyday we might expect from the suite of new technological developments our horizon scanning programme has been researching.

We are also planning further DRCF research into consumer attitudes towards AI (including agentic), and how regulatory tools can support trusted and safe adoption.

6.2 Next steps for member regulators

All DRCF regulators are pursuing further individual work on Agentic AI within their remits:

- The ICO has recently released a tech futures report[123] exploring the data protection implications of future developments in Agentic AI. The ICO is updating guidance on AI and automated decision-making that will impact on Agentic AI, anticipated for publication in March 2026. The ICO is developing a statutory code on AI and ADM with implications for Agentic AI. If organisations have evidence relating to Agentic AI they wish to submit for our consideration in the context of developing the code, they can contact ai@ico.org.uk. There are also several strands of ongoing work set out in the ICO's AI and biometrics strategy[124] with implications for Agentic AI.

[123] ICO tech futures: Agentic AI | ICO

[124] Preventing harm, promoting trust: our AI and biometrics strategy | ICO

- The FCA will continue to provide a safe space where firms can experiment and test for market readiness to support safe and responsible AI systems adoption, which innovators are welcome to apply for (including with agentic use cases). This includes evaluating the results from cohort 1 to inform cohort 2 of the Supercharged Sandbox,[125] and continuing cohort 1 and launching cohort 2 of AI Live Testing,[126] which sets out a new platform to engage with firms at point of AI system deployment. The FCA will share learnings from these initiatives. The FCA held the Supercharged Showcase event in January and have committed to publishing an evaluation report from AI Live Testing by the end of Q1 2027. This year will see the fourth edition of the AI/ML survey jointly conducted with Bank of England, providing valuable responses from firms. The FCA also co-chairs the AI Consortium industry group with the Bank of England. The Consortium is focused on how AI is or could be used in UK financial services and is exploring the risks and opportunities in promoting the safe and responsible adoption of AI.

  Through the upcoming Mills Review, the FCA will look at how advanced AI models could reshape retail financial services by 2030 onwards, and what this could mean for consumers, firms and the FCA's regulatory approach. This includes considerations around more powerful, autonomous and agentic systems. An engagement paper asking for views was open until 24 February. The Review will conclude with a report to the FCA Board in the summer, setting out recommendations to help the FCA continue to play a leading role in shaping AI-enabled financial services, and an external publication to support informed debate.
- Ofcom will continue to look at the adoption of Agentic (and other forms of) AI in telecoms markets and its potential impact on residential and business customers.

  Where Agentic AI falls within scope of the Online Safety Act, Ofcom will consider its impact on regulated services and users and how it can take steps to mitigate emerging risks in line with Ofcom powers. In 2024 and 2025 Ofcom published a document setting its strategic approach to AI.[127] The 2026/27 edition will be published later this year, which will provide an update to Ofcom's approach to AI.
- The CMA is focused on supporting positive adoption of trusted AI and Agentic systems. It has recently published new work on Agentic AI, discussing the potential opportunities and benefits for UK businesses, consumers and the wider economy if systems can be developed and scaled reliably, as well as key risks to be managed. It has also set out practical guidance for businesses exploring these technologies in the context of customer support and other consumer-facing services[128] as well as in business pricing approaches where novels risks of algorithmic collusion can arise.[129]

  This builds on the CMA's earlier work and publications on algorithms[130] [131] and AI Foundation Models,[132] and its published AI principles to guide the ecosystem and wider economy to positive outcomes. This programme has looked across the full value chain and identified the most significant opportunities for consumers, competition, growth and household prosperity of well-designed and legally compliant AI including agents. It also considered wider potential enablers to unlock the full opportunity for UK consumers, businesses and the economy, including smart data, digital ID, and strong interoperability.[133]

---

[125] Supercharging the digital sandbox: how we're collaborating with Nvidia to accelerate AI innovation | FCA
[126] AI Live Testing: The use of AI in UK financial markets - from promise to practice | FCA
[127] Ofcom's strategic approach to AI (2025)
[128] Complying with consumer law when using AI agents - GOV.UK
[129] AI and collusion: frontiers, opportunities and challenges – Competition and Markets Authority
[130] Pricing algorithms research, collusion and personalised pricing - GOV.UK
[131] Algorithms: How they can reduce competition and harm consumers - GOV.UK
[132] AI Foundation Models: initial review - GOV.UK
[133] Agentic AI and consumers - GOV.UK

The CMA will continue its work and engagement in this space including through ongoing horizon scanning and embedding these insights across the breadth of CMA work, practical support and advice for businesses exploring these technologies, as well as through domestic and international collaboration (including through the CMA's chairing of the International Competition Network's Technologist Group). It will also continue to harness Agentic AI and other cutting-edge technologies internally to drive greater productivity and agility across its operations.

Contact us

If you would like to engage with us further, there are a few ways you can get in touch:

Updates on other upcoming DRCF events are published on our website and LinkedIn.

To subscribe to the DRCF newsletter, please email drcf@ofcom.org.uk.

A1. Annexe

Glossary

Agentic AI: Systems of AI agents that behave and interact autonomously to achieve their objectives.[134]

AI agents: Increasingly autonomous AI services that can directly impact real-world environments.

API: Application Programming Interface, a mechanism that enables third parties to interoperate with an existing platform or service by accessing functionality or data in a structured way.

Generative AI: Generative AI is an umbrella term for artificial intelligence systems capable of synthesising new, original content, including text, images, video, and audio, by modelling the underlying patterns and structures of a training dataset.

Large Language Model (LLM): A Large Language Model is a specific deep-learning architecture optimised for the comprehension and generation of natural language. By processing vast datasets through billions of parameters, LLMs utilise statistical probability to predict and sequence text that is contextually relevant and syntactically coherent.

Scenarios

Building on the four scenarios in the ICO's recent Agentic AI Tech Futures Report the DRCF member regulators considered how their own remits would be impacted by and influence the development of the four futures which it outlined.

A pair of axes describing critical uncertainties which might help determine how Agentic AI is used in future (capability and adoption) were used to create a two-by-two matrix of four plausible futures – low capability, low adoption / low capability, high adoption / high capability, low adoption / and high capability, high adoption. These we called respectively 'scarce simple agents', 'just good enough to be everywhere', 'agents in waiting' and 'ubiquitous agents'.

Method

Subject matter experts from the four regulators were invited to a workshop in which futures derived from the axes of agentic capability and levels of adoption of AI agents were examined for points of regulatory synergy (where overlaps between the powers and requirements of member regulators might produce co-operative efficiencies), and points of regulatory friction (where apparent tensions between remits can be identified and resolved).

Discussion points raised in the subject matter expert workshop included everything from the possibility of novel harms (viewed from the perspective of each regulator), potential barriers to adoption, the role of regulators in development, and how expectations might be set and met. These and the other points raised in the workshop informed this paper.

This paper draws some insight from the DRCF's public call for views on Agentic AI, and the DRCF would like to acknowledge all the respondents for their contributions.

[134] GOV.UK, AI Insights: Agentic AI (HTML) - GOV.UK. This paper does not set out to agree a common cross-regulatory definition for "agentic AI". Eg, the ICO Tech Futures report discusses multiple capabilities of agentic systems.