DRCF zur Zukunft agentischer KI-Systeme

Am 31. März 2026 veröffentlichte das DRCF (Digital Regulation Cooperation Forum) - bestehend aus der Competition and Markets Authority (CMA), der Financial Conduct Authority (FCA), dem Information Commissioner's Office (ICO) und Ofcom - ein Foresight Paper zur Zukunft **agentischer KI**. Es dient als vorausschauende Untersuchung von KI-Agenten und zeigt auf, wie regulatorische Rahmenbedingungen in Großbritannien dazu beitragen können, die Chancen dieser Technologie verantwortungsvoll zu nutzen. Das DRCF betont, dass Regulierung dabei ausdrücklich als Innovationstreiber und nicht als Innovationsbremse verstanden wird.

Das Foresight Paper umfasst die folgenden **zentralen Punkte**:

1. Definition von „Agentic AI"

Agentische KI bezeichnet Systeme, die aus mehreren KI-Agenten bestehen und selbstständig handeln, um vorgegebene Ziele zu erreichen. Im Unterschied zu klassischen generativen KI-Systemen, die auf Anfragen reagieren und Inhalte erzeugen, können agentische Systeme:

- Ziele analysieren, in Teilaufgaben aufteilen und Abläufe planen,
- Aktionen eigenständig ausführen (z. B. Zahlungen veranlassen, Formulare ausfüllen),
- in Echtzeit auf externe Daten und Dienste zugreifen,
- aus früheren Interaktionen lernen und sich anpassen.

Ein einzelner KI-Agent setzt sich aus einem technischen Kern (KI-Modelle und Steuerungssoftware), einer Schnittstelle zur Außenwelt (andere Systeme, Dienste, physische Umgebungen) sowie einer Nutzeroberfläche zusammen. Der Autonomiegrad variiert stark - von einfachen Werkzeugen bis hin zu weitgehend selbstständigen Akteure. Die meisten Anwendungen befinden sich derzeit auf den unteren bis mittleren Stufen des Autonomiegrades.

2. Chancen und Potenziale

Für Verbraucher:innen: Agentische KI könnte alltägliche administrative Aufgaben übernehmen - bspw. Rechnungen verwalten, Versicherungen wechseln, Urlaubsreisen buchen oder Tarife vergleichen. Dadurch entstünden erhebliche Zeitersparnisse im Alltag. Besonders profitieren könnten Personen mit Unterstützungsbedarf oder Sprachbarrieren.

Für Unternehmen: Im Unternehmensumfeld könnte agentische KI sowohl den Kundenkontakt (Front Office) als auch interne Abläufe (Back Office) transformieren. Das Foresight Paper führt in diesem Zusammenhang mehrere Studien an. Eine davon untersuchte den Einsatz eines generativen KI-Assistenten im Kundensupport und stellte Produktivitätssteigerungen von etwa 14–15 % fest. Dabei profitierten weniger erfahrene Mitarbeiter besonders stark von der Technologie.

Für Regulierungsbehörden: Auch Aufsichtsbehörden selbst können von agentischer KI profitieren - zur automatisierten Erkennung von Regelverletzungen (z. B. Preismanipulation, Ausschreibungsbetrug) oder zur Skalierung von Marktüberwachung. Die CMA setzt bereits KI-Agenten ein, um Verbraucherschäden wie „Drip Pricing“ aufzudecken.

3. Risiken und Herausforderungen

Das Foresight Paper identifiziert mehrere spezifische **Risikofelder**, die durch die erhöhte Autonomie agentischer Systeme betroffen sein könnten.

Ein Feld ist der *Datenschutz*, da die weitreichende Vernetzung der Agenten Prinzipien wie Datensparsamkeit und Transparenz beeinträchtigen könnten. Damit eng verknüpft ist das Problem der *Haftung*: In komplexen Mehrfach-Agenten-Systemen ist es schwierig zu bestimmen, wer für Fehler verantwortlich ist. Wenn daraus eine Kettenreaktion entsteht, können die Folgen besonders schwerwiegend sein. Agentische Systeme können darüber hinaus Entscheidungen beeinflussen, die Verbraucher:innen nicht vollständig überblicken. *Intransparenz* und sogenannte *„Behavioural Steering“* können das informierte Einverständnis unterhöhlen. Daneben könnte das Feld der *Cybersicherheit* betroffen sein. Die Vernetzung agentischer Systeme vergrößert die Angriffsoberfläche. Prompt-Injection-Angriffe und der Missbrauch weitreichender Zugriffsrechte stellen konkrete Bedrohungen dar. Weiterhin können Agentenbasierte KI-Systeme *kollusives Verhalten* zeigen, ohne dass sie dazu explizit angewiesen wurden. Unternehmen müssen daher vorsichtig sein: Wenn KI-Agenten die Preise festlegen, besteht die Gefahr von automatischen Preisabsprachen. Zudem kann die tiefe Integration agentischer Systeme in Workflows die *Abhängigkeit* von einzelnen Anbietern erhöhen. Zuletzt könnte eine *digitale Ungleichheit* entstehen. Während agentische KI Barrieren abbauen kann, besteht gleichzeitig das Risiko einer zunehmenden Ausgrenzung für Menschen mit geringer Medienkompetenz.

4. Regulatorischer Rahmen: Bestehende Regeln gelten weiterhin

Alle vier DRCF-Mitgliedsbehörden betonen: Agentische KI fällt nicht außerhalb bestehender Regelwerke. Bestehende Pflichten zur Transparenz, Fairness, Datenschutz, Verbraucherschutz und Wettbewerb gelten unverändert. Ein einzelner Einsatz agentischer KI kann gleichzeitig den Regulierungsbereich aller vier Behörden berühren:

- **ICO (Datenschutz)**: Fokus auf automatisierte Entscheidungen, Datenminimierung, Transparenz und menschliche Aufsicht bei rechtlich bedeutsamen Entscheidungen.
- **FCA (Finanzmarktregulierung)**: Unternehmen müssen nachweisen, dass sie im Einsatz agentischer KI die Consumer Duty erfüllen und gute Ergebnisse für Verbraucher:innen erzielen.
- **Ofcom (Kommunikationsregulierung)**: Prüfung, ob KI-Agenten als Suchdienste im Sinne des Online Safety Act gelten und damit bestimmten Sicherheitspflichten unterliegen.
- **CMA (Wettbewerb und Verbraucherschutz)**: Beobachtung von Marktmacht, algorithmischer Kollusion und Verbraucherrechten bei KI-gestützten Dienstleistungen.

5. Schlüsselprinzipien für einen sicheren Einsatz

Das Foresight Paper leitet aus den vier Regulierungsbereichen (Governance, Datenschutz, Verbraucherrechte, Marktdynamik) konkrete Empfehlungen ab:

- **Menschliche Aufsicht („Human in the Loop“)**: Klare Schwellenwerte definieren, ab wann ein KI-Agent menschliche Genehmigung benötigt - insbesondere bei Entscheidungen mit rechtlicher oder finanzieller Tragweite.
- **Beobachtbarkeit und Erklärbarkeit**: Systeme müssen nachvollziehbare Protokolle ihrer Aktionen führen, damit Fehler identifiziert und Verantwortlichkeiten geklärt werden können.
- **Datenschutz by Design**: Personenbezogene Daten dürfen nur im notwendigen Umfang verarbeitet werden („Datenminimierung“).
- **Transparenz für Verbraucher:innen**: Unternehmen müssen klar kommunizieren, welche Aufgaben sie an KI-Agenten delegieren und welche Konsequenzen das für die Nutzer:innen hat.

- **Interoperabilität und Datenportabilität**: Offene Standards (z. B. Model Context Protocol) können Vendor Lock-in reduzieren und den Wettbewerb stärken.
- **Wachsamkeit gegenüber algorithmischer Kollusion**: Unternehmen, die KI-Agenten in der Preisgestaltung einsetzen, müssen aktiv überwachen, ob kollusive Verhaltensweisen entstehen.

6. Ausblick und nächste Schritte

Die zukünftige Entwicklung agentischer KI wird je nach Branche unterschiedlich verlaufen. Maßgebliche Faktoren sind die Verlässlichkeit der Systeme in der Praxis, das Verständnis von Nutzer:innen und Organisationen sowie die vorherrschenden Geschäftsmodelle. Das DRCF plant für 2026/27 weitere Untersuchungen zu zukünftigen Nutzeroberflächen, Consumer Robotics und dem allgemeinen Verbrauchererlebnis im KI-Zeitalter.

Die einzelnen Behörden kündigen konkrete Folgemaßnahmen an: Der ICO veröffentlicht aktualisierte Leitlinien zu automatisierten Entscheidungen und erarbeitet einen Praxiskodex für KI. Die FCA setzt ihr „Supercharged Sandbox“-Programm fort und evaluiert den KI-Einsatz im Retailbanking. Ofcom prüft die Auswirkungen agentischer KI auf Telekommunikationsmärkte und die Online-Sicherheit. Die CMA veröffentlicht weiterhin Leitfäden für Unternehmen und setzt KI intern zur Erkennung wettbewerbswidriger Verhaltensweisen ein.