REVISTA ELECTRÓNICA DE DIREITO

RED

A (des)proteção dos dados pessoais nos modelos de IA de finalidade geral. Do Regulamento da IA às propostas *omnibus*

The (un)protection of personal data in general-purpose AI models. From the AI regulation to the omnibus proposals

Graça Enes

Professora Associada da FDUP, Investigadora Integrada do CIJ (FDUP)

Investigadora Colaboradora do CLUP (FLUP), Investigadora Associada do IRDEIC (Université Toulouse-Capitole)

Rua dos Bragas, 223, 4050-123 Porto, Portugal

gef@direito.up.pt

https://orcid.org/0000-0002-9461-5705

Outubro de 2026

RESUMO: O desenvolvimento e funcionamento dos modelos de IA de finalidade geral, em especial os modelos de *GenAI*, estão associados à recolha e processamento massivo de dados, incluindo de dados pessoais. Vários problemas emergem nessa realidade, seja o âmbito de proteção e o fundamento para tal utilização, sejam os riscos de *leakage* ou "regurgitação" de informação pessoal. O regime normativo vigente, assente no Regulamento Geral de Proteção de Dados e no Regulamento da IA, apresenta limitações e suscita incertezas, mas é também criticado por regular excessivamente e restringir a inovação e a competitividade da UE. A redefinição do conceito de dados pessoais e a previsão do interesse legítimo como fundamento para a utilização dos dados pessoais, incluídas nas propostas "*Digital omnibus*", promovem a certeza jurídica, mas também favorecem os prestadores dos modelos em detrimento dos titulares de dados pessoais. A proteção dos dados pessoais deve integrar o sistema de governação de risco, sistema esse que deve ser contextual e iterativo, envolvendo todos os *stakeholders* ao longo de todo o ciclo de vida do modelo, com vista a prevenir, eliminar e mitigar os riscos. Entre os diversos sistemas propostos para responder às incertezas, dificuldades e limites das medidas de proteção de dados pessoais, além da "AIDP", salienta-se o "*HUDERIA model*". A proteção dos dados pessoais está subjacente às obrigações instituídas no Regulamento da IA para os prestadores dos modelos de IA de finalidade geral, incluindo os modelos com risco sistémico, e concretizadas sobretudo nos capítulos primeiro e terceiro do Código de Práticas, aprovado em 2025 e subscrito por alguns dos mais relevantes operadores da indústria. Ainda que extensas e exaustivas, parece praticamente incontornável a imensa assimetria de informação entre os prestadores e os titulares de dados pessoais e sobressaem os limites decorrentes do princípio da proporcionalidade e do *balancing test* dos interesses em presença.

PALAVRAS-CHAVE: Dados pessoais; Modelos de IA de finalidade geral; Risco; Sistema de governação de risco; RGPD; RIA; Código de Práticas para os modelos de IA de finalidade geral.

ABSTRACT: The development and operation of general-purpose AI models, especially GenAI models, are associated with the massive collection and processing of data, including personal data. Several problems emerge, whether regarding the scope of protection and the basis for such use, or the risks of leakage or "regurgitation" of personal information. The regulations in force, namely the General Data Protection Regulation and the AI Regulation, present limitations and raise uncertainties, but are also criticized for over-regulating and restricting innovation and competitiveness in the EU. The redefinition of the concept of personal data and the provision of legitimate interest as a basis for the use of personal data, included in the "Digital omnibus" proposals, promote legal certainty, but also favor model providers at the expense of personal data subjects. The protection of personal data must be part of the risk governance system to be implemented. This system must be contextual and iterative, involving all stakeholders throughout the lifecycle of the model, with a view to preventing, eliminating, and mitigating risks. Among the various systems proposed to address the uncertainties,

difficulties, and limitations of personal data protection measures, besides the “DPIA”, the "HUDERIA model" stands out. The protection of personal data underlies the obligations established in the AI Regulation for providers of general-purpose AI models, including models with systemic risk, and is particularly detailed in chapters one and three of the Code of Practice approved in 2025 and signed by some of the most relevant industry operators. Even if they are extensive and comprehensive, the immense asymmetry of information between providers and personal data subjects seems unescapable, and the limitations arising from the principle of proportionality and the balancing test of the interests at stake are particularly noteworthy.

KEY WORDS: Personal data; General Purpose AI models; Risk; Risk governance system; GDPR; AIA; Code of Practice for General-Purpose AI Models.

SUMÁRIO:

1. Introdução

2. Os modelos de IA de finalidade geral e a proteção dos dados pessoais entre o RIA e o RGPD

2.1. Insuficiências e incertezas do regime de proteção de dados pessoais nos modelos de IA de finalidade geral

2.2. O âmbito e o alcance de proteção dos "dados pessoais". Da interpretação à revisão do conceito

2.3. O fundamento para a recolha e o processamento de dados pessoais nos modelos de IA de finalidade geral. Clarificação ou inovação?

3. Medidas de proteção dos dados pessoais no desenvolvimento e funcionamento dos modelos de IA de finalidade geral. Incertezas, dificuldades e limites

4. A proteção de dados pessoais nas obrigações dos prestadores de modelos de IA de finalidade geral

4.1. Obrigações dos prestadores de modelos de IA de finalidade geral

4.2. Obrigações dos prestadores de modelos de finalidade geral

4.3. Obrigações dos prestadores de modelos de IA de finalidade geral com risco sistémico

4.4. O Código de Práticas para os Modelos de IA de finalidade geral

5. Conclusão

Bibliografia

Jurisprudência

1. Introdução

No presente, a progressiva e praticamente imperativa interação digital das pessoas, nas relações pessoais, de consumo, profissionais ou até com as autoridades públicas, constrói uma "pegada digital" sem que haja uma plena consciência do seu alcance. Não é também despicienda a disseminação indiscriminada de dados pessoais fruto de comportamentos deliberados dos respetivos titulares, ainda que inconscientes ou displicentes, em especial através da ubíqua presença nas redes sociais. Pode até admitir-se o desenvolvimento de um novo paradigma de (des)proteção dos dados pessoais, ou, pelo menos, a existência de preferências conflituantes, em que a vontade de aparecer vai a par com preocupações de reserva da privacidade[1]. A partir do acesso a dados pessoais e não pessoais, incluindo expostos por terceiros, partilhados em conteúdos "postados" ou disponíveis publicamente[2], os detentores de infraestruturas digitais, plataformas ou *apps*, agregam e cruzam esses dados e, desse modo, conseguem saber praticamente tudo acerca das pessoas, até mais do que os próprios sujeitos, pois os padrões de comportamento revelam hábitos, preferências ou condicionamentos psicológicos, até inconscientes. As ferramentas de inteligência artificial exponenciam o alcance das inferências e das potenciais utilizações de informação pessoal, incluindo sensível (v.g., obtida através de aparelhos de telemetria). Esse fenómeno, decorrente da incontornável assimetria da informação na sociedade digital, foi designado de "*inverse privacy*"[3]. As consequências mais conhecidas deste espaço virtual de "vigilância ubíqua"[4] são o *marketing* ou a propaganda direcionada[5]. Já a manipulação da identidade, a monitorização e o controlo emocional e cognitivo permanentes são cada vez mais reais e

[1] Os participantes num inquérito *online* realizado, em 2024, em oito Estados (Espanha, Coreia do Sul, Reino Unido, México, Alemanha, Índia, China e EUA), manifestaram genericamente uma baixa preocupação com a proteção de dados *online*, com a exceção da Espanha. STATISTA, *Statista Consumer Insights*, 2025, https://cdn.statcdn.com/Infographic/images/normal/31706.jpeg (02.06.2026).

[2] E logo aqui emergem questões controversas. O que são "dados pessoais que tenham sido manifestamente tornados públicos pelo seu titular" e que a al. e) do n.º 2 do artigo 9.º do RGPD exclui da proibição de tratamento de certas categorias especiais de dados pessoais prevista no n.º 1 do mesmo dispositivo? De acordo com o Tribunal de Justiça, para que sejam qualificados como tal impõe-se que o sujeito tenha deliberadamente agido no sentido de tornar esses dados acessíveis ao público em geral (Ac. do TJ de 4 de outubro de 2024, *Schrems v. Meta*, Proc. C-446/21, ECLI:EU:C:2024:834, n.º 77). Mais, segundo o TJ, só esses poderão ser licitamente processados, não resultando daí uma autorização para o operador tratar outros dados pessoais relativos a essa informação disponíveis em aplicações e em sítios Internet de terceiros, efetuar a respetiva agregação e análise, a fim de apresentar publicidade personalizada (n.ºs 80, 82 e 83); já o Ac. do TJ de 4 de julho de 2023, *Meta v. Bundeskartellamt*, Proc. C-252/21, ECLI:EU:C:2023:537, impôs para tal efeito que o titular tenha "manifestado expressamente a sua escolha prévia, eventualmente com base numa parametrização individual efetuada com pleno conhecimento de causa, de tornar os dados que lhe dizem respeito publicamente acessíveis a um número ilimitado de pessoas" (n.º 85), o que não sucede necessariamente com a simples ativação de botões "gosto" ou "partilhar" ou com os dados recolhidos através de "*cookies*".

[3] YURI GUREVICH et al., "Inverse Privacy", 2015, https://arxiv.org/abs/1510.03311 (02.06.2026). Nos EUA, em 2019, o Senador John Kennedy apresentou a proposta "*Own Your Own Data Act*". SEN. KENNEDY, J., 116th U.S. Congress, *Own Your Data Act, S. 806, 116th U.S. Congress (2019-2020)* introduced in 14th of March 2019, https://www.congress.gov/bill/116th-congress/senate-bill/806/text?format=txt (02.06.2026).

[4] SHOSHANA ZUBOFF, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York, PublicAffairs, 2019.

[5] Se estamos perante um novo paradigma, e na impossibilidade de escapar, pode equacionar-se se a solução será a monetização dos dados pessoais, que permitiria a todos (e não apenas a "*influencers*") beneficiar desta realidade. JARON LANIER defende essa possibilidade. *Jaron Lanier fixes the Internet*, New York, *New York Times*, The Privacy Project, 2019, (em especial, Episode 2), https://www.nytimes.com/interactive/2019/09/23/opinion/data-privacy-jaron-lanier.html (02.06.2026). Na Califórnia, foi debatida a possibilidade de os titulares partilharem dos benefícios económicos resultantes dos respetivos dados, através de "Data Dividends". JILL COWAN, "How Much Is Your Data Worth?", *The New York Times*, 25.03.2019, https://www.nytimes.com/2019/03/25/us/newsom-hertzberg-data-dividend.html (02.06.2026).

permitem formas de "*profilling*" avançadas, "*hypernudging*", pondo em causa os direitos individuais e a liberdade, autonomia e a dignidade humanas[6].

Para se compreender a relação dos modelos de IA de finalidade geral com os dados pessoais e os riscos para a sua proteção é indispensável ter em conta a natureza e o funcionamento desses modelos[7], em especial aqueles que têm maior impacto na atualidade, os LLM (*Large Language Models*) com capacidades multimodais de geração de conteúdos[8]. O desenvolvimento dos modelos de IA de finalidade geral depende do treino de *big data*[9], através da reutilização de *datasets* pré-existentes ou da recolha de dados[10] através de "*web crawling*" ou "*web scraping*"[11] da internet. Por essas vias foram e continuam a ser recolhidos e tratados dados não pessoais e dados pessoais[12], incluindo dados sensíveis, apesar da generalizada ausência de conhecimento dos seus titulares e do fundamento para tal ser discutível, como se verá adiante. Essa utilização de dados pessoais, tal como a paralela utilização de conteúdos

[6] Uma síntese desses riscos, pode encontrar-se em CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, T-PD(2025)3rev3, Council of Europe, Strasbourg, 12.05.2026, pp. 19-27, https://media.licdn.com/dms/document/media/v2/D4E1FAQHorwju8e1mrw/feedshare-document-sanitized-pdf/B4EZ4hHqZUKkBE-/0/1778672118494?e=1779530400&v=beta&t=N9oMFh0LPeNOmtNV-4hvWW6l3uuFmQdK9JCWNCrUFC4 (02.06.2026).

[7] A OCDE recorre ao ISO/IEC 22989 standard e define modelo como "*physical, mathematical or otherwise logical representation of a system, entity, phenomenon, process or data*". OECD, "Explanatory Memorandum on the updated OECD definition of an AI System", *OECD Artificial Intelligence Papers* March 2024, No. 8, OECD Publishing, 2024, p. 8, https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/03/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_3c815e51/623da898-en.pdf (02.06.2026). Anteriormente, a Organização tinha definido "AI model" como "*a computational representation of all or part of the external environment of an AI system – encompassing, for example, processes, objects, ideas, people and/or interactions that take place in that environment. Core characteristics include technical type, how the model is built (using expert knowledge, machine learning or both) and how the model is used (for what objectives and using what performance measures)*". OECD, "OECD Framework for the Classification of AI systems", *OECD Digital Economy Papers*, 22.02.2022, pp. 7 e 20, https://www.oecd.org/en/publications/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en.html (02.06.2026). São três as categorias de modelos discriminados: simbólicos, estatísticos e híbridos. Estes últimos combinam as duas modalidades, como sucede nos NLP (*Natural Language processing algorithms*) (p. 44). Para a OCDE, a finalidade geral significa que o modelo pode ser utilizado "em qualquer setor da economia". OECD, "OECD Framework for the Classification of AI systems", cit., p. 63. Os *standards* ISO podem encontrar-se em ISO/IEC, *Information technology — Artificial intelligence — Artificial intelligence concepts and terminology*, 2022; ISO/IEC, *Information technology — Artificial intelligence — AI system life cycle processes*, 2023, https://www.iso.org/search.html (02.06.2026). Mesmo que não versem especificamente sobre os modelos também os integra nas definições apresentadas.

[8] Uma descrição sumária destes, do seu funcionamento e capacidades, pode encontrar-se em ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, EDPB, março 2025, pp. 6-23, https://www.edpb.europa.eu/system/files/2025-04/ai-privacy-risks-and-mitigations-in-llms.pdf (02.06.2026).

[9] São vários os métodos de treino habitualmente utilizados. O considerando 97 do Regulamento da IA refere a aprendizagem autossupervisionada, não supervisionada ou por reforço.

[10] A Comissão adota uma noção lata de modelo e inclui no respetivo ciclo de vida o processamento dos dados pré-treino. COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, C(2025) 7719 final, Bruxelas, 19.11.2025, p. 8, https://digital-strategy.ec.europa.eu/en/library/guidelines-scope-obligations-providers-general-purpose-ai-models-under-ai-act (02.06.2026). Para se compreender o alcance do conceito de "ciclo de vida", veja-se também os *standards* ISO 22989:2022 e ISO 5338:2023, já referidos.

[11] Sobre a recolha de dados por "*web scraping*", UGO PAGALLO, JACOPO CIANI SCIOLLA, "Anatomy of web data scraping: ethics, standards, and the troubles of the law", *European Journal of Privacy Law & Technologies*, n.º 2, Nápoles, Suor Orsola Benincasa Università Editrice, 2023, pp. 1- 9, https://doi.org/10.57230/EJPLT232PS (02.06.2026).

[12] A OpenAI reconhece que recolhe e processa dados pessoais em várias dimensões, mesmo que afirme não recolher intencionalmente dados pessoais e procurar limitar essa recolha, excluindo fontes que agregam grandes quantidades de dados pessoais. Declara também que treina os modelos "para evitar responder a pedidos de informações privadas ou sensíveis sobre indivíduos" (tradução nossa). OPENAI, *How ChatGPT and our foundation models are developed*, https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed#h_6b40d4df3c (02.06.2026).

protegidos por direitos de autor, tem sido contestada[13], incluindo nos tribunais[14]. A proteção desses bens pode mesmo conflituar entre si[15]. O risco para a proteção de dados pessoais continua após a colocação no mercado dos modelos de IA de finalidade geral, em virtude da respetiva interação ambiental, da interação com o sistema de IA onde o modelo é integrado, ou da interação com o utilizador, como sucede com informação pessoal exposta nos *outputs*, em resposta a solicitações do utilizador, ou introduzida pelo utilizador nos respetivos *prompts*. As múltiplas finalidades e utilizações não determinadas condicionam a suficiência e a efetividade das medidas de identificação, avaliação e mitigação dos riscos efetuada na fase de desenvolvimento. Não pode ignorar-se ainda o risco de inexatidão associada ao problema das "alucinações" dos modelos de *GenAI*. Finalmente, como em outras ferramentas digitais, também são importantes os riscos decorrentes de atividade maliciosa ou "leak" acidental. O elevado volume de dados processado é um fator que eleva o risco de impacto negativo nos direitos individuais em caso de "*breach*" e que tem de ser considerado no processo de

[13] Em 31 de março de 2023, poucos meses depois do lançamento do ChatGPT, a *GARANTE PER LA PROTEZIONE DEI DATI PERSONALI* determinou a sua suspensão em Itália, com o fundamento na utilização ilegítima de dados pessoais, incluindo de menores. *Intelligenza artificiale: il Garante blocca ChatGPT. Raccolta illecita di dati personali. Assenza di sistemi per la verifica dell'età dei minori*, 31.03.2023, https://gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870847#english (02.06.2026). Ainda que a suspensão tenha durado menos de um mês, condenou a OpenAI ao pagamento de uma coima de €15 milhões e a realizar uma campanha de informação, com fundamento na violação do regime vigente de proteção de dados pessoais. Em 21 de março de 2025, a XVIII secção civil do Tribunal de Roma ordenou a suspensão da decisão. *GARANTE PER LA PROTEZIONE DEI DATI PERSONALI*, "Comunicato Stampa - ChatGPT, il Garante privacy chiude l'istruttoria. OpenAI dovrà realizzare una campagna informativa di sei mesi e pagare una sanzione di 15 milioni di euro", *Garante Per La Protezione Dei Dati Personali*, 20.12.2024, https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10085432 (02.06.2026). Em 10 de abril de 2025, a autoridade italiana aplicou também uma coima de €5 milhões à Luka Inc. relacionada com o *chatbot Replika* baseado num modelo de GenAI. *Comunicato Stampa - AI: Il Garante sanziona la società che gestisce il chatbot "Replika"*, 19.05.2025, https://gpdp.it/home/docweb/-/docweb-display/docweb/10132048 (02.06.2026).

[14] JONATHAM GILLHAM elencou litígios envolvendo a OpenAI, *Originality.ai*, December 7, 2025. Disponível em https://originality.ai/blog/openai-chatgpt-lawsuit-list (10.05.2025). Uma resenha de várias ações é apresentada em BAKERHOSTETLER, *Case Tracker: Artificial Intelligence, Copyrights and Class Actions*, 2026, https://www.bakerlaw.com/services/artificial-intelligence-ai/case-tracker-artificial-intelligence-copyrights-and-class-actions/ (02.06.2026), e em MCKOOL SMITH, "AI Infringement Case Updates: July 28, 2025", McKool Smith, 28.07.2025, https://www.mckoolsmith.com/newsroom-ailitigation-33 (02.06.2026). Uma análise sumária de alguns dos primeiros casos nos EUA encontra-se em CHRISTOPHER J. VALENTE et al., *Recent Trends in Generative Artificial Intelligence Litigation in the United States*, K&L Gates, 5.09.2023, https://www.klgates.com/Recent-Trends-in-Generative-Artificial-Intelligence-Litigation-in-the-United-States-9-5-2023 (02.06.2026). Em especial no que respeita à proteção de dados pessoais, nos EUA (Califórnia), a Class Action apresentada em 28 de junho de 2023 terminou com a desistência dos autores. Class Action Case P.M. v. OpenAI LP (3:23-cv-03199) District Court, N.D. California, https://www.courtlistener.com/docket/67535351/pm-v-openai-lp/ (02.06.2026). A Class Action Case 3:23-cv-04557-VC foi favorável à OpenAI, com uma "dismissal (…) with leave to Amend", por considerar que o pedido era excessivamente longo e assente em considerações de "general policy". United States District Court Northern California, Order Granting Motions to Dismiss at 1, Cousart v. OpenAI LP, No. 3:23-cv-4557 (N.D. Cal. May 24, 2024), ECF No. 78, https://storage.courtlistener.com/recap/gov.uscourts.cand.417880/gov.uscourts.cand.417880.78.0.pdf (02.06.2026). A Class Action Case 3:23-cv-03440-LB, apresentada contra a Alphabet/Google terminou com uma decisão semelhante. United States District Court Norhern California, Order Graning Motion to Dismiss with Leave to Amend and Denying Administrative Motion to relate without Prejudice, JL et al. v. ALPHABET INC., et al., No. Case 3:23-cv-03440-AMO, https://storage.courtlistener.com/recap/gov.uscourts.cand.415223/gov.uscourts.cand.415223.46.0.pdf (02.06.2026).

[15] Tal conflito foi suscitado pela OPENAI quando um tribunal nova-iorquino solicitou os *logs* do ChatGPT numa ação relacionada com a alegada violação de direitos de autor intentada pelo New York Times. THE NEW YORK TIMES, *Motion to vacate the Order*, 12.11.2025, https://fingfx.thomsonreuters.com/gfx/legaldocs/dwpkqdwwwpm/OPENAI%20NEW%20YORK%20TIMES%20COPYRIGHT%20LAWSUIT%20reconsider.pdf (02.06.2026). Southern District of New York, The New York Times Company v. Microsoft Corporation, 1:23-cv-11195, (S.D.N.Y.), https://www.courtlistener.com/docket/68117049/the-new-york-times-company-v-microsoft-corporation/ 02.06.2026).

governação[16]. Todos estes riscos impõem medidas de prevenção, eliminação e mitigação. Nos pontos seguintes, percorrer-se-á o regime jurídico vigente e em elaboração na UE para responder às exigências de proteção dos dados pessoais, cujas traves-mestras são o Regulamento Geral de Proteção de Dados (Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016, também RGPD) e o Regulamento da IA (Regulamento (EU) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, também RIA), mas em que é igualmente importante a *soft law* do Código de Práticas e Orientações aprovados em 2025, bem como as recentes propostas legislativas "*omnibus*". Considera-se ainda a Convenção-quadro sobre a IA, do Conselho da Europa, e o projeto de *Guidelines* elaborado em maio de 2026. Na secção 2, analisar-se-ão as insuficiências e incertezas para a proteção dos dados pessoais que decorrem do regime do RIA e do RGPD, incluindo no que respeita ao conceito de "dados pessoais" e o fundamento para a sua recolha e processamento no desenvolvimento e funcionamento dos modelos de IA de finalidade geral. Na secção 3, a atenção dirige-se para as medidas de proteção dos dados pessoais, em especial o trabalho desenvolvido pelo Comité Europeu de Proteção de Dados (também CEDP ou EDPB[17]) e pelo Comité Consultivo da Convenção 108 do Conselho da Europa. A secção 4 incidirá sobre o regime específico de obrigações previsto para os prestadores dos modelos de IA de finalidade geral, tendo em conta não apenas os atos jurídicos vinculativos (RIA e RGPD), mas igualmente as Orientações da Comissão Europeia e o Código de Práticas adotado em 2025 e que foi subscrito por algumas dezenas de prestadores.

2. Os modelos de IA de finalidade geral e a proteção dos dados pessoais entre o RIA e o RGPD

2.1. Insuficiências e incertezas do regime de proteção de dados pessoais nos modelos de IA de finalidade geral

A regulação é tecnologicamente neutral e a proteção de dados pessoais conferida pelo RGPD aplica-se qualquer que seja a tecnologia usada para o respetivo processamento. O RIA não prevê nada em especial sobre dados pessoais no que respeita aos modelos de IA, mas o considerando 10 esclarece que esse ato não visa afetar a aplicação do direito da União já em vigor para o tratamento de dados pessoais, incluindo as funções e as competências das autoridades de supervisão independentes responsáveis pelo controlo do cumprimento desses instrumentos. Assim, os titulares de dados pessoais continuam a usufruir de todos os direitos e garantias que lhes são conferidos pelo direito da União. O Regulamento não afeta as obrigações dos prestadores e utilizadores profissionais, enquanto responsáveis pelo tratamento de dados pessoais ou subcontratantes, decorrentes do direito da União ou do

[16] ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., p. 48.
[17] Sigla correspondente à designação em inglês: European Data Protection Board.

direito nacional em matéria de proteção de dados pessoais, na conceção, no desenvolvimento ou no funcionamento de ferramentas de IA, seja no que respeita ao fundamento para o tratamento de dados pessoais, seja no que respeita aos direitos dos titulares, incluindo no que respeita a decisões individuais automatizadas ou à definição de perfis[18].

Essa premissa e essas ressalvas genéricas implicam a prevalência dos normativos de proteção dos dados pessoais. Daí decorre, em primeiro lugar, o respeito pelo artigo 25.º, n.º 1 do RGPD[19], que consagra o princípio "*data protection by design*"[20]. Este princípio deve orientar o desenvolvimento dos modelos de IA de finalidade geral, não sendo admissível a justificação liminar de que não é tecnicamente possível[21], embora aquela disposição submeta a adequação das medidas de mitigação a uma ponderação onde se incluem também os custos da aplicação das técnicas mais avançadas. As medidas devem ser ser "*taylored*" aos modelos e às respetivas fases, bem como aos riscos identificados em cada uma delas[22]. Entre as medidas a implementar, podem enunciar-se a anonimização e a cifragem, a utilização de tecnologias que permitam a introdução de algoritmos nos dados e o treino dos sistemas de IA sem a transmissão entre as partes ou a cópia dos próprios dados em bruto ou estruturados. Na verdade, é irrecusável a existência de uma contradição de princípio entre o treino de grandes quantidades de dados e o contínuo processamento em ecossistemas multimodais dinâmicos o princípio da minimização e do tratamento para finalidades específicas em que assenta o regime do RGPD[23]. Numa perspetiva pragmática, estão comprovadas as dificuldades e ambiguidades no regime instituído em relação à aplicação e ao exercício efetivo dos direitos dos titulares de

---

[18] Também o Comité Consultivo da Convenção do Conselho da Europa para a Proteção dos Indivíduos em relação ao Processamento Automático de Dados Pessoais (também Comité da Convenção 108) recorda os princípios consagrados nesta Convenção que vincula os EM da UE: "*fairness, transparency, accountability, purpose limitation, effective remedies and safeguard for data subjects´ rights*". CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 7. As multifacetadas implicações nos velhos e novos riscos, incluindo os resultantes de alucinações e manipulação sintética de identidade, são percorridas nas pp. 19-27. Mesmo não vinculando juridicamente a UE, que não é Parte, é relevante no ordenamento da UE, na medida em que todos os EM são Partes.

[19] Com a epígrafe "*Proteção de dados desde a conceção e por defeito*", esta disposição prevê o seguinte:
"*1. Tendo em conta as técnicas mais avançadas, os custos da sua aplicação, e a natureza, o âmbito, o contexto e as finalidades do tratamento dos dados, bem como os riscos decorrentes do tratamento para os direitos e liberdades das pessoas singulares, cuja probabilidade e gravidade podem ser variáveis, o responsável pelo tratamento aplica, tanto no momento de definição dos meios de tratamento como no momento do próprio tratamento, as medidas técnicas e organizativas adequadas, como a pseudonimização, destinadas a aplicar com eficácia os princípios da proteção de dados, tais como a minimização, e a incluir as garantias necessárias no tratamento, de uma forma que este cumpra os requisitos do presente regulamento e proteja os direitos dos titulares dos dados.*
2. O responsável pelo tratamento aplica medidas técnicas e organizativas para assegurar que, por defeito, só sejam tratados os dados pessoais que forem necessários para cada finalidade específica do tratamento. Essa obrigação aplica-se à quantidade de dados pessoais recolhidos, à extensão do seu tratamento, ao seu prazo de conservação e à sua acessibilidade. Em especial, essas medidas asseguram que, por defeito, os dados pessoais não sejam disponibilizados sem intervenção humana a um número indeterminado de pessoas singulares.
3. Pode ser utilizado como elemento para demonstrar o cumprimento das obrigações estabelecidas nos n.ºs 1 e 2 do presente artigo, um procedimento de certificação aprovado nos termos do artigo 42.º."

[20] Este princípio é também referido pelo Comité Consultivo da Convenção 108. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 11-12 e 29.

[21] EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, 23 May 2024, p. 5, https://www.edpb.europa.eu/our-work-tools/our-documents/other/report-work-undertaken-chatgpt-taskforce_en (02.06.2026). Este relatório incluiu mesmo em anexo um modelo de questionário para a avaliação dos modelos.

[22] Uma exposição desses riscos, das vias através das quais ocorre o tratamento de dados nos diferentes tipos de modelos e fases do ciclo de vida e das medidas de mitigação adequadas é apresentada em ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 24-43.

[23] Artigo 5.º, n.º 1, als. b) e c) do RGPD.

dados pessoais, bem como a desadequação das medidas de proteção e de governação atualmente previstas perante a complexidade, a autonomia e dinâmica dos ecossistemas associados aos modelos de IA[24].

É compreensível que a massiva necessidade de dados destes desenvolvimentos tecnológicos não pudesse ser prevista pelo legislador quando aprovou o RGPD. As dificuldades de previsão também se explicam pelas dificuldades de compreensão do funcionamento e do impacto dos modelos de IA. Até no RIA, originalmente, apenas se abrangiam os sistemas de IA e foi já durante o procedimento legislativo que foram introduzidas modificações a contemplar os modelos de IA. Ainda assim, os modelos de IA de finalidade geral continuam ausentes em diversas disposições relevantes. As múltiplas referências à salvaguarda e proteção dos dados pessoais têm lugar a propósito dos sistemas de IA, ou seja, para além dos modelos, sendo certo, contudo, que as obrigações que impendem sobre o prestador de um sistema que incorporou um modelo podem estar dependentes do respeito dessas obrigações pelo prestador do modelo. Como podem os prestadores dos sistemas de IA de risco elevado respeitar o RGPD quando integram modelos onde se podem encontrar os problemas e para os quais não há nenhuma previsão que proporcione certeza jurídica? As obrigações de informação aos operadores a jusante na cadeia, instituídas no RIA, podem não ser suficientes. São várias as questões em aberto[25]. Pode defender-se a interpretação extensiva ou até a analogia para ultrapassar as ausências de previsão, mas a incerteza impera em várias questões fundamentais, nomeadamente no que respeita ao fundamento para o processamento de dados pessoais no desenvolvimento dos modelos. Os problemas e o potencial conflito do desenvolvimento e funcionamento dos modelos de IA de finalidade geral com o regime vigente de tutela dos dados pessoais são expressamente reconhecidos no Parecer 28/2024 do Comité Europeu de Proteção de Dados[26], onde são avançados esclarecimentos e orientações[27]. Têm-se sucedido as controvérsias em que o TJ é chamado a interpretar o sentido e o alcance de conceitos tão centrais como o próprio conceito de "dados pessoais"[28]. Estes problemas, a par com o objetivo de promover a inovação e a competitividade europeia estão subjacentes à

[24] CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 19 e 25-27.

[25] GEORGIOS FERETZAKIS et al., "GDPR and Large Language Models: Technical and Legal Obstacles", *Future Internet 2025*, 17(4), Basel, MDPI, p. 151, https://doi.org/10.3390/fi17040151 (02.06.2026).

[26] A incerteza sobre alguns dos problemas que os modelos de IA colocam no que tange aos dados pessoais, conduziu a Autoridade irlandesa de proteção de dados a requerer ao CEDP um parecer sobre a utilização de dados pessoais pelos modelos de IA. O Parecer 28/2024 debruça-se sobre a questão da anonimização/pseudonimização, sobre o fundamento para o processamento de dados pessoais e as consequências do uso ilícito do tratamento de dados pessoais. Não dissipa todas as dúvidas e ao preconizar sobretudo a realização de avaliações casuísticas não elimina a incerteza. CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, 17 de dezembro de 2024, https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf (02.06.2026). A adoção das propostas "*omnibus*" poderá trazer alguma certeza, mas várias dificuldades permanecem. Sobre o tema e, em especial, sobre aquele parecer, FEDERICA VELLI, "Processing Personal Data in the Context of AI Models: EDPB's Opinion 28/2024", *European Papers*, Vol. 9, No. 3, 2024, pp. 1497-1500. doi: 10.15166/2499-8249/822, https://www.europeanpapers.eu/system/files/pdf_version/EP_EF_2024_H_004_Federica_Velli_00822.pdf (02.06.2026).

[27] Mais recentemente, o EDPB publicou um estudo mais exaustivo sobre os riscos e medidas de mitigação para proteção de dados pessoais nos LLM, o qual, no entanto, não constitui uma posição oficial do organismo. ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit.

[28] Ac. do TJ de 4 de setembro de 2025, *AEPD/CUR (Notion de données à caractère personnel)*, Proc. C-413/23 P, ECLI:EU:C:2025:645.

reorientação político-legislativa proposta pela Comissão Europeia no pacote "*omnibus*" para o domínio digital. As propostas da Comissão Europeia "*Digital omnibus on AI*" e "*Digital omnibus*" preveem modificações com impacto significativo no regime de tutela dos dados pessoais em relação aos modelos de IA de finalidade geral, seja na definição do conceito de "dados pessoais", seja no fundamento para o respetivo processamento por modelos de IA[29].

2.2. O âmbito e o alcance da proteção dos "dados pessoais". Da interpretação à revisão do conceito

Se a utilização de dados no desenvolvimento do modelo de IA abarcar apenas dados não pessoais ou dados pessoais anonimizados, esta fase não coloca qualquer problema, pois as condições e restrições do RGPD não se aplicam. Dados anonimizados são os dados que não respeitam a uma pessoa identificada ou identificável, tendo em conta todos os meios razoavelmente prováveis de serem utilizados pelo controlador ou outra pessoa[30]. Por outro lado, segundo o considerando 16 desse ato, "[o]s dados pessoais que tenham sido pseudonimizados, que possam ser atribuídos a uma pessoa singular mediante a utilização de informações suplementares, deverão ser considerados informações sobre uma pessoa singular identificável", i.e., são dados pessoais nos termos do artigo 4.º, al. 1) do RGPD. Ainda assim, a pseudonimização é considerada pelo RGPD como uma das medidas de proteção e mitigação dos riscos[31] e uma das salvaguardas de proteção quando o tratamento (lícito) dos dados pessoais é realizado sem o consentimento do sujeito para finalidades distintas (lícitas) daquelas para que foram recolhidos[32]. Mas, se permanecem como dados pessoais, será sempre

[29] Ambas as propostas foram apresentadas pela Comissão no mesmo dia, em 19.11.2025. Proposta de Regulamento do PE e do Conselho que altera os Regulamentos (UE) 2024/1689 e (UE) 2018/1139 no respeitante à simplificação da aplicação das regras harmonizadas em matéria de inteligência artificial (*Omnibus Digital sobre a IA*) [COM/2025/836/FINAL], https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A52025PC0837 (02.06.2026). Esta proposta foi objeto de acordo político entre o Parlamento Europeu e o Conselho em 7.05.2026. PE, *AI Act: deal on simplification measures, ban on "nudifier" apps*, 7.05.2026, https://www.europarl.europa.eu/news/en/press-room/20260427IPR42011/ai-act-deal-on-simplification-measures-ban-on-nudifier-apps (02.06.2026). O seu resultado pode ser visto como uma vitória da indústria, em que a simplificação constitui em grande medida desregulamentação, em especial no que respeita aos sistemas abrangidos pelo "Regulamento Máquinas" (Regulamento (UE) 2023/1230 do Parlamento Europeu e do Conselho, de 14 de junho de 2023, JO L 165 de 29.06.2023, pp. 1-102), que deixarão de estar sujeitos ao regime do RIA para ficarem submetidos apenas ao regime setorial decorrente do outro regulamento. MATHEW NEUMAN, "Deal over EU's AI omnibus package represents political victory for business", *in MLex*, Newsletter May 7, 2026, https://www.mlex.com/mlex/artificial-intelligence/articles/2474987?nl_pk=94e4dd0b-c9bf-4432-85b4-95a382f04fa2&utm_source=newsletter&utm_medium=email&utm_campaign=mlex/artificial-intelligence&utm_content=2026-05-08&read_main=1&nlsidx=0&nlaidx=0 (02.06.2026). Se a duplicação da regulamentação para a "IA industrial" era criticável, também é necessário garantir que o "Regulamento Máquinas" não se traduz numa redução da proteção e segurança. PHILIP HACKER, *Comentário no LinkedIn*, 8.05.2026, https://www.linkedin.com/feed/?highlightedUpdateUrn=urn%3Ali%3Aactivity%3A7458099365094674432&highlightedUpdateType=TOPIC_TRENDING_CONVERSATION_IN_YOUR_NETWORK&origin=inapp&showCommentBox=true (02.06.2026). A Proposta de Regulamento do Parlamento Europeu e do Conselho que altera os Regulamentos (UE) 2016/679, (UE) 2018/1724, (UE) 2018/1725 e (UE) 2023/2854 e as Diretivas 2002/58/CE, (UE) 2022/2555 e (UE) 2022/2557 no respeitante à simplificação do quadro legislativo digital e que revoga os Regulamentos (UE) 2018/1807, (UE) 2019/1150, (UE) 2022/868 e a Diretiva (UE) 2019/1024 (*Omnibus Digital*) [COM/2025/837 final] ainda continua em discussão, https://eur-lex.europa.eu/legal-content/pt/TXT/?uri=COM%3A2025%3A837%3AFIN (02.06.2026).

[30] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 13.

[31] Artigo 25.º, n.º 1.

[32] Artigo 6.º, n.º 4, al. e).

indispensável que o fundamento para o processamento de dados pseudonimizados no treino de modelos de IA seja lícito, de acordo com o artigo 6.º, n.º 1 do RGPD, e sejam respeitadas as condições estabelecidas nos números seguintes desse artigo.

E se os dados pessoais pseudonimizados são transmitidos a um terceiro sem as informações suplementares que permitem a identificação da pessoa, devem ainda qualificar-se como dados pessoais ou não será de os incluir nesta categoria e devem ser considerados dados anonimizados? Em 4 de setembro de 2025, o TJ[33] entendeu (como já tinha entendido o TG no acórdão recorrido[34]) que dados pessoais pseudonimizados mudam de natureza e se convertem em dados anónimos quando são transmitidos a outra entidade sem serem acompanhados das informações que permitem a identificação do titular de consentimento, desde que a entidade recetora não disponha de "outros meios de identificação"[35] que "apresentem uma probabilidade razoável de ser utilizados pelo responsável pelo tratamento ou por outra pessoa para identificar direta ou indiretamente a pessoa singular"[36], tendo em conta "todos os fatores objetivos, como os custos e o tempo necessário para a identificação (…) as tecnologias disponíveis à data do tratamento dos dados e a evolução tecnológica". O facto de uma outra entidade, numa transmissão posterior, dispor de meios que venham a permitir essa identificação[37] não foi considerado relevante, ainda que nesse caso os dados readquiram a natureza de "dados pessoais"[38]. Esta orientação era recusada pela recorrente (a Autoridade Europeia de Proteção de Dados)[39]. A natureza de dados pessoais dos dados pseudonimizados não depende da existência objetiva das informações complementares que permitem a identificação pessoal, mas passa a depender da capacidade subjetiva, baseada na disponibilidade de meios técnicos, para conseguir fazer a re-identificação. Este acórdão reflete-se na proposta "*Digital omnibus*". O artigo 3.º, n.º 1 da proposta altera o artigo 4.º do RGPD e retira da categoria de "dados pessoais" para uma entidade os dados pseudonimizados "se essa entidade não conseguir identificar a pessoa singular a quem dizem respeito, tendo em conta os meios que apresentem uma probabilidade razoável de ser utilizados por essa entidade. Essas informações não se tornam pessoais para essa entidade pelo simples facto de um potencial destinatário subsequente dispor de meios que apresentem uma probabilidade razoável de ser utilizados para identificar a pessoa singular a quem as informações dizem respeito"[40]. Esta modificação, associada à proposta previsão expressa atinente ao fundamento do processamento, ao abrigo do interesse legítimo, promove os interesses das empresas que desenvolvem os modelos e fragiliza a posição dos titulares de dados pessoais.

[33] Ac. do TJ de 4.09.2025, *AEPD/CUR (Notion de données à caractère personnel)*, Proc. C-413/23 P, ECLI:EU:C:2025:645, n.º 64. Ainda que não fosse o RGPD que estava em causa, mas o Regulamento 2018/1725, o conceito de "dados pessoais" é idêntico.
[34] Ac. do TG de 26.04.2023, *CUR/AEPD*, Proc. T-557/20, EU:T:2023:219.
[35] Proc. C-413/23 P, cit., n.º 77.
[36] Proc. C-413/23 P, cit., n.º 79.
[37] A OpenAI já reconheceu essa capacidade, ainda que refira só o fazer quando "exigido por lei". OPENAI, *Política de Privacidade*, 6.02.2026, https://openai.com/pt-PT/policies/row-privacy-policy/ (02.06.2026).
[38] Proc. C-413/23 P, cit., n.º 85.
[39] Proc. C-413/23 P, cit., n.ºs 64 a 66.
[40] A proposta acrescenta o artigo 41a, que confere à Comissão, com a colaboração do CEDP, a tarefa de especificar os meios e os critérios através dos quais a pseudonimização altera a natureza dos dados.

2.3. O fundamento para a recolha e o processamento de dados pessoais nos modelos de IA de finalidade geral. Clarificação ou inovação?

O considerando 63 declara que o RIA "não deverá ser entendido como um fundamento jurídico para o tratamento de dados pessoais, inclusive de categorias especiais de dados pessoais, se for caso disso, salvo disposição específica em contrário no presente regulamento". Essa possibilidade ocorre em ambientes de testagem ou na adoção de medidas de gestão de risco para prevenir ou mitigar enviesamentos. Já o considerando 69 declara que "[o] direito à privacidade e à proteção de dados pessoais tem de ser garantido ao longo de todo o ciclo de vida do sistema de IA", ou seja, desde a conceção e o desenvolvimento até à sua retirada de utilização. Não se encontra qualquer referência expressa aos modelos de IA, sucedendo o mesmo com as disposições normativas relativas ao processamento de dados pessoais. Não obstante, estando os modelos de IA de finalidade geral previstos no Capítulo V do Regulamento, é-lhe aplicável o artigo 2.º, n.º 7 do Regulamento, que dispõe que "[o] direito da União em matéria de proteção de dados pessoais, privacidade e confidencialidade das comunicações aplica-se aos dados pessoais tratados em virtude dos direitos e obrigações estabelecidos no presente regulamento". Assim, continua esse dispositivo, o RIA não afeta o disposto no Regulamento (UE) 2016/679, no Regulamento (UE) 2018/1725, nem na Diretiva 2002/58/CE ou na Diretiva (UE) 2016/680, com a ressalva dos artigos 10.º, n.º 5, e do artigo 59.º. São precisamente estas disposições que preveem fundamentos especiais para o processamento de dados pessoais.

O artigo 10.º, n.º 5 do RIA diz respeito aos sistemas de IA de risco elevado e à respetiva governação e gestão de dados. A fim de detetar e corrigir vieses suscetíveis de afetar a saúde e a segurança das pessoas, de ter repercussões negativas nos direitos fundamentais ou de resultar em discriminações proibidas pelo direito da União, os prestadores são autorizados a tratar categorias especiais de dados pessoais, nomeadamente, os dados previstos no artigo 9.º, n.º 1 do RGPD. Tal utilização fica sujeita a uma condição geral de adoção de "garantias adequadas dos direitos e liberdades fundamentais das pessoas singulares" e às condições estabelecidas nas alíneas desse artigo 10.º, n.º 5. No entanto, essa disposição não refere os modelos de IA. Por sua vez, o artigo 59.º do RIA legitima o tratamento de dados pessoais recolhidos para outras finalidades para efeitos de desenvolvimento no ambiente de testagem da regulamentação da IA de certos sistemas de IA "para salvaguarda de um interesse público substancial" nos domínios da segurança pública e saúde pública, no ambiente, na energia, na mobilidade, infraestruturas críticas e transportes, na administração pública e serviços públicos. Tal como a anterior, esta disposição refere-se aos sistemas de IA e nada diz em relação aos modelos.

Sendo essas disposições normas especiais, que afastam o regime geral previsto no RGPD, a respetiva interpretação deve ser restritiva, o que justifica as dúvidas sobre a licitude da recolha e processamento de dados pessoais nos modelos de IA de finalidade geral.

Desde logo, sendo uma tecnologia com finalidades gerais, como pode o fundamento para o tratamento de dados pessoais no seu treino e funcionamento ser o consentimento do titular quando o RGPD refere que o consentimento deve ser "para uma ou mais finalidades específicas"[41]? Se os dados pessoais estiverem disponíveis publicamente, será lícita tal utilização? O CEDP, no Parecer 28/2024, recorda que o facto de os dados serem acessíveis publicamente, não significa que sejam públicos, o que é relevante no quadro da isenção conferida pelo artigo 9.º, n.º 2, al. e) do RGPD à proibição do tratamento de certas categorias de dados pessoais[42]. Como declarou o TJ, é necessário que o sujeito tenha deliberadamente agido no sentido de tornar esses dados acessíveis ao público em geral[43].

O consentimento como fundamento não soluciona as dúvidas suscitadas pelo processamento dos dados pessoais não anonimizados integrados em *datasets* ou dos dados obtidos por "*web crawling*" e "*web scraping*"[44], dos dados de terceiros introduzidos pelo utilizador ou obtidos através da interação ambiental já durante o funcionamento do modelo.

O tratamento dos dados pessoais integrados no modelo pelos utilizadores deliberadamente pode ser necessário para cumprir a solicitação do utilizador. Neste caso, a execução contratual será fundamento válido para o processamento desses dados. Porém, há um problema quando esses dados pessoais respeitam a outras pessoas. Com a previsão expressa avançada pela Comissão na proposta "*digital omnibus*", o fundamento do interesse legítimo que tem sido invocado pelos prestadores será válido, mas não deixa de suscitar questões, nomeadamente a distribuição da responsabilidade pelos danos causados ao titular dos dados. Deverá ser do utilizador, exceto quando o prestador/*deployer* pudesse saber que os dados introduzidos não eram do utilizador e que este não tinha fundamento válido para tal utilização? Deverá ser distribuída entre o prestador/*deployer* do modelo e o utilizador que introduziu esses dados? A previsão desta situação nos termos e condições de utilização pode não ser suficiente para acautelar este problema.

Os prestadores dos modelos têm invocado o interesse legítimo como fundamento para o processamento de dados pessoais nos modelos de IA de finalidade geral, segundo o artigo 6.º, n.º 1, al. f) do RGPD[45]. Na ausência de previsão expressa e em resposta às dúvidas da autoridade irlandesa de proteção de dados (Data Protection Commission), o Parecer 28/2024 do CEDP[46] recorda que não há prevalência entre os possíveis fundamentos e admite o

[41] Artigo 6.º, n.º 1, al. a).
[42] EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, cit., p. 7.
[43] Proc. C-252/21, cit., § 77.
[44] No "*web crawling*", obtêm-se URLs; no "*web scraping*", extraem-se dados daqueles. A aplicação do regime previsto no artigo 14.º do RGPD a esta forma de recolha de dados levanta dúvidas. EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, cit., p. 8.
[45] OPENAI, *How ChatGPT and our foundation models are developed*, https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed (02.06.2026).
[46] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 20.

fundamento do interesse legítimo quando se debruça sobre as condições de aplicação do regime dos artigos 13.º e 14.º do RGPD.

O artigo 3.º, n.º 15 da proposta "*Digital omnibus*" adita ao RGPD o artigo 88c, com a epígrafe "Processamento no contexto do desenvolvimento e funcionamento da IA". Segundo o enunciado desta nova disposição, "[s]empre que o tratamento de dados pessoais seja necessário para efeito dos interesses do responsável pelo tratamento no contexto do desenvolvimento e funcionamento de um (…) modelo de IA, esse tratamento pode ser prosseguido para efeitos dos interesses legítimos, na aceção do artigo 6.º, n.º 1, alínea f), do Regulamento (UE) 2016/679, quando aplicável, exceto se outra legislação da União ou nacional exigir explicitamente o consentimento e se prevalecerem sobre esses interesses os interesses ou os direitos e liberdades fundamentais do titular dos dados que exijam a proteção dos dados pessoais, em especial se o titular dos dados for uma criança"[47].

Com esta solução, ultrapassa-se alguma incerteza resultante da ausência de previsão de fundamento para a recolha e processamento de dados nos modelos de IA e acolhe-se a posição defendida pelos prestadores. Porém, a aplicação deste fundamento está sujeita a requisitos.

As orientações apresentadas pelo CEDP nas *Guidelines 1/2024*[48] e igualmente referidas no Parecer 28/2024 sobre o significado e alcance do "interesse legítimo" continuam relevantes. São três os critérios a respeitar do interesse legítimo – a licitude do interesse a prosseguir com o tratamento dos dados, a necessidade do tratamento e a ponderação dos interesses em presença.

Assim, nos modelos de IA de finalidade geral, o interesse do prestador é o interesse (habitualmente económico-comercial e técnico) associado ao melhor desenvolvimento da sua atividade e ao aperfeiçoamento do modelo. Obviamente, o treino dos modelos para assegurar o seu potencial, melhorar a sua eficiência e exatidão, mas igualmente a correção de vieses, promover a segurança, prevenir a fraude, etc., são exemplos de interesses legítimos. Imprescindível é que o interesse seja claro, preciso e efetivo (não ambíguo, vago ou hipotético). Pode admitir-se um interesse social da inovação tecnológica[49] e do desenvolvimento de uma tecnologia que depois poderá ser aplicada a numerosas finalidades, inclusive para beneficiar os titulares dos dados pessoais utilizados no acesso a serviços diversos. Além, evidentemente, do interesse político-económico de promoção da

[47] Um dos exemplos é a proibição de processamento de dados sem consentimento prevista no artigo 5.º, n.º 2, em especial as alíneas b), c) e d) do Regulamento dos Mercados Digitais para os controladores de acesso ("*gatekeepers*"). Ainda que não faça referência direta ao processamento para treino de modelos de IA, impede a ligação, combinação e cruzamento de dados pessoais provenientes do serviço essencial de plataforma com outros serviços do prestador, o que inclui a atividade de treinamento de modelos de IA.

[48] EDPB, *Guidelines 1/2024 on processing of personal da.ta based on Article 6(1)(f) GDPR*, Version 1.0, 8 October 2024, https://www.edpb.europa.eu/system/files/2024-10/edpb_guidelines_202401_legitimateinterest_en.pdf (02.06.2026).

[49] COMISSÃO EUROPEIA, *Commission Staff Working Document Accompanying the documents Proposal for a Regulation of the European Parliament and of the Council Amending Regulations (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 and Directives 2002/58/EC, (EU) 2022/2555 and (EU) 2022/2557 as regards the simplification of the digital legislative framework, and repealing Regulations (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868, and Directive (EU) 2019/1024 (Digital Omnibus) Amending Regulations (EU) 2024/1689 and (EU) 2018/1139 as regards the simplification of the implementation of harmonised rules on artificial intelligence (Digital Omnibus on AI)*, SWD(2025) 836 final, Brussels, 19.11.2025, p. 135.

competitividade europeia. No entanto, estes não parecem ser interesses suficientemente precisos.

A necessidade do tratamento afere-se pela inexistência de outros meios alternativos menos intrusivos para alcançar os objetivos. Deve ainda limitar-se o tratamento ao estritamente necessário[50].

Já os interesses dos titulares dos dados[51] são, em primeira linha, o interesse na autodeterminação e na manutenção do controlo sobre os seus próprios dados pessoais [52], mas podem ser também outros interesses pessoais, sociais ou financeiros[53], além da proteção dos direitos fundamentais de que são também titulares[54]. Todos estes interesses não obscurecem os interesses diretamente tutelados pelo regime de proteção dos dados pessoais e a ponderação dos interesses deve ser exigente[55].

Segundo o CEDP, o interesse legítimo do operador deve ter em consideração as expectativas razoáveis do indivíduo. A "razoabilidade" das expectativas baseia-se na relação mantida com o prestador e nas medidas de mitigação adotadas. Continuam a ser pertinentes as *Guidelines 1/24*, reiteradas no Parecer 28/2024. A orientação preconizada enuncia diversos elementos relevantes, sempre a avaliar à luz das circunstâncias específicas. Assim, não basta a inclusão dessa possibilidade na política de privacidade do prestador e as obrigações de transparência aí preconizadas mantêm relevância[56]. Os dados estarem ou não disponíveis para o público e quem os publicitou (pois a sua disponibilidade pública não se confunde com a disponibilização

[50] Além das orientações do CEDP, o TJ já se pronunciou diversas vezes sobre esta questão. Por todos, Proc. C-252/21, cit., n.ºs 108 e 109.
[51] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 24.
[52] É controversa a natureza jurídica da titularidade dos dados pessoais, questionando-se se é semelhante à propriedade. Não é também claro se, quando transmitidos, os dados têm a natureza de bens ou serviços. GRAÇA ENES, "Building a European Data Space", in INÊS QUADROS E MIGUEL GORJÃO-HENRIQUES (eds.), *Revisiting the Liberal Constitutional Features of the European Union*, Cham, Springer, 247-285, pp. 251-252. https://doi.org/10.1007/978-3-031-95717-8_11. A nosso ver, independentemente dessa natureza, os direitos substantivos associados aos dados pessoais devem ser considerados como direitos de personalidade, oponíveis universalmente a qualquer perturbação. Sobre a questão, em especial em relação aos dados biométricos, os mais pessoalíssimos de todos e qualificados pelo RGPD como categoria especial de dados, ESRA DEMIR, "The protection of human biodata: Is there any role for data ownership?", *Computer Law & Security Review*, Volume 51, Amsterdam, Elsevier, *2023*, pp. 1-13, pp. 4-6, https://doi.org/10.1016/j.clsr.2023.105905 (02.06.2026).
[53] Na esfera do titular, podem coexistir interesses conflituantes com a reserva subjacente ao regime de proteção dos dados pessoais, nomeadamente interesses económico-financeiros (por exemplo, quando um modelo de IA é utilizado pelo titular dos dados para gerar proveitos), benefícios pessoais (por exemplo, quando um modelo de IA é utilizado para melhorar a acessibilidade a determinados serviços), ou interesses socioeconómicos (por exemplo, quando um modelo de IA permite o acesso a melhores cuidados de saúde ou facilita o exercício de um direito fundamental, como o acesso à educação). CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 24.
[54] Além dos "clássicos" direitos fundamentais, entre os direitos enunciados pelo CEDP deve salientar-se o direito à integridade física e mental. *Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR*, cit., p. 13. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 31-32.
[55] Um dos exemplos fornecidos é a criação de um *chatbot* para assistência virtual ao cliente. Não deveria aqui ser ponderada a existência de alternativas para assegurar o objetivo de assistência do cliente, nomeadamente um assistente humano. É questionável se o principal beneficiário desta ferramenta é o consumidor, particularmente se em comparação com a assistência humana.
[56] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., pp. 26-27. Como é preconizado no considerando 47 do RGPD e foi considerado pelo TJ. Por todos, Proc. C-252/21, cit.

pública pelo titular[57]) é um dos elementos a considerar, bem como as potenciais utilizações do modelo, se há uma relação direta com o prestador, que teria permitido informar o titular dos dados sobre esse processamento e, assim, fundar as expectativas dessa utilização, etc. Neste particular ponto, na verdade, quanto menor a relação entre o titular dos dados e o prestador do modelo maior a liberdade deste, o que é perverso e promove a realização da tarefa de recolha através de intermediários (*brokers*) a montante[58].

Com a proposta "*Digital omnibus*", o interesse legítimo passa a ser inequivocamente fundamento para a recolha e processamento de dados, mas essa utilização está sujeita à adoção de salvaguardas e de medidas técnicas e organizativas destinadas a proteger os direitos do titular dos dados (*v.g.*, a cifragem ou pseudonimização, como previsto no artigo 6.º, n.º 4 do RGPD), incluindo a minimização (através da seleção de fontes e curadoria no treino e testagem do modelo), medidas para prevenir a divulgação de dados, medidas de reforço da transparência para com os titulares dos dados e, ainda, a concessão ao titular de um direito incondicional a objetar ao processamento dos seus dados pessoais[59]. Este direito de objeção incondicional (*opt-out*) é significativo, pois não fica sujeito à alegação de "motivos relacionados com a sua situação particular", previstos na faculdade de objeção vigente no RGPD[60], estando em linha com a sugestão do CEDP no Parecer 28/2024[61].

Por sua vez, o artigo 1.º, n.º 7, al. b) da proposta "*Digital omnibus on AI*" revoga o artigo 10.º, n.º 5 do RIA e o artigo 1.º, n.º 5 adita o artigo 4.º-A. A proposta da Comissão eliminava a expressão "estritamente" e representava um aligeiramento das condições para legitimar o processamento de categorias especiais de dados pessoais (previstos no artigo 9.º, n.º 1 do RGPD) com a finalidade de monitorizar, detetar e corrigir vieses nos sistemas de IA de risco elevado. Essa condição acabou por ser mantida na posição do PE e na posição do Conselho[62]. O n.º 2 do novo artigo 4.º-A propõe colmatar a lacuna da inclusão dos modelos de IA e prevê que o regime previsto no n.º 1 para os sistemas de IA de risco elevado pode ser aplicado aos modelos "sempre que tal for necessário e proporcionado, se o tratamento ocorrer para os fins

[57] Como reconhece o CEDP no *Report of the work undertaken by the ChatGPT Taskforce*, cit. Segundo o TJ, é necessário que tenha havido essa intenção de tornar os dados acessíveis ao público em geral, através de uma ação explícita e afirmativa (Proc. C-252/21, cit., n.º 77).

[58] O Comité Consultivo da Convenção 108 dá nota da menor transparência e da dificuldade em controlar a licitude do tratamento dos dados. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 17. Tal não pode ser causa de exclusão da responsabilidade do prestador do modelo.

[59] A OpenAI anuncia que "[o]utilizador pode facilmente escolher se o seu conteúdo pode ser utilizado para melhorar e treinar os nossos modelos", através de um formulário no *Privacy Portal*. OPENAI, *Política de Privacidade*, cit.

[60] Artigo 21.º.

[61] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 29.

[62] A posição aprovada pelo PE, em 26.03.2026, propôs a manutenção dessa condição. PE, *Alterações aprovadas pelo Parlamento Europeu, em 26 de março de 2026, sobre a proposta de regulamento do Parlamento Europeu e do Conselho que altera os Regulamentos (UE) 2024/1689 e (UE) 2018/1139 no respeitante à simplificação da aplicação das regras harmonizadas em matéria de inteligência artificial (pacote Omnibus Digital sobre a IA)*, https://www.europarl.europa.eu/doceo/document/TA-10-2026-0098_PT.html (02.06.2026). Na posição adotada em 13 de março de 2026, o Conselho acompanhou o PE. COUNCIL OF THE EU, *Council agrees position to streamline rules on Artificial Intelligence* (Press Release), 13.03.2026, https://www.consilium.europa.eu/en/press/press-releases/2026/03/13/council-agrees-position-to-streamline-rules-on-artificial-intelligence/ (02.06.2026). O acordo político alcançado entre o PE e o Conselho, em 7.05.2026, mantém essa condição para legitimar o tratamento de categorias especiais de dados. PE, *AI Act: deal on simplification measures, ban on "nudifier" apps*, cit.

nele previstos e desde que estejam reunidas as condições estabelecidas ao abrigo das garantias previstas no presente número". Os requisitos de "necessidade" e de "proporcionalidade" que se acrescentam ao requisito da necessidade enunciado no n.º 1, além da redundância, não afastam a incerteza.

A alteração do artigo 9.º do RGPD, também incluída na proposta "*Digital omnibus*", alarga a exceção que permite o processamento de certas categorias especiais de dados pessoais, nomeadamente das categorias especiais de dados previstos no n.º 1 dessa disposição ("dados pessoais que revelem a origem racial ou étnica, as opiniões políticas, as convicções religiosas ou filosóficas, ou a filiação sindical, bem como o tratamento de dados genéticos, dados biométricos para identificar uma pessoa de forma inequívoca, dados relativos à saúde ou dados relativos à vida sexual ou orientação sexual de uma pessoa"). O artigo 3.º, n.º 3 da proposta da Comissão acrescenta a alínea k) ao artigo 9.º, n.º 2, isentando da proibição absoluta de processamento daquelas categorias de dados pessoais o processamento operado no contexto do desenvolvimento e funcionamento de modelos e sistemas de IA, embora submetido ao respeito das condições técnicas e organizacionais indicadas no novo n.º 5 incluído nessa mesma proposta (medidas adequadas para evitar a recolha e o tratamento dessas categorias especiais de dados pessoais e a sua remoção, caso sejam detetados). Segundo a Comissão[63], a exclusão da proibição não visa atividades de processamento intencional dessas categorias de dados, não resultando daí a possibilidade da sua utilização na produção de inferências ou de "*outputs*", nem a sua transmissão a terceiros[64]. Visa, sim, acautelar as situações em que ocorre um tratamento "residual" desses dados, apesar das medidas referidas tendentes a prevenir a sua recolha ou a eliminá-los, logo que detetados, em virtude da sua inclusão em grandes *datasets.* Cumpridas as obrigações de diligência, esta cláusula salvaguarda a responsabilidade do prestador. Deve assinalar-se, contudo, que nessa mesma disposição se prevê que se a remoção "exigir um esforço desproporcionado", a remoção não será exigível e pode ser substituída, "sem demora indevida", por medidas de proteção desses dados contra a sua utilização para a produção de resultados, divulgação ou disponibilização a terceiros, ou seja, a sua inclusão nos "*outputs*" produzidos com o modelo. A remoção da proibição de princípio associada à incerteza resultante dos conceitos indeterminados utilizados nas condições impostas facilita a utilização desses dados pessoais. Em especial, a condição de que a eliminação não constitua um "esforço desproporcionado" responde essencialmente aos objetivos da indústria e se "contribui diretamente para o objetivo de apoiar a inovação e desenvolver uma IA europeia de confiança e não discriminatória"[65] reduz o nível de proteção dos direitos dos indivíduos[66]. Afasta-se da perspetiva do Comité Consultivo da Convenção 108

[63] COMISSÃO EUROPEIA , *Commission Staff Working Document* (…), SWD(2025) 836 final, cit., pp. 39-40.
[64] O Comité Consultivo da Convenção 108 recorda que a implementação do quadro de gestão de risco que propõe nas respetivas Orientações não legitima *de per se* o tratamento de dados pessoais e não elimina necessariamente os potenciais problemas. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 8.
[65] COMISSÃO EUROPEIA, *Commission Staff Working Document* (…), SWD(2025) 836 final, cit., p. 40 (tradução nossa).
[66] A crítica de que o pacote "*omnibus*" favorece os interesses das "*Big Tech*" em detrimento dos direitos fundamentais, em especial a proteção dos dados pessoais, foi feito por diversas organizações. NOYB, "Digital Omnibus: EU Commission wants to wreck core GDPR principles", *NOYB*, 19.11.2025, https://noyb.eu/en/digital-

do Conselho da Europa, segundo a qual se as medidas de mitigação se revelarem insuficientes para assegurar o cumprimento das obrigações da Convenção, não deve iniciar-se o processamento de dados pessoais, ou este deve ser suspenso, interrompido, redefinido ou abandonado[67].

Com estas propostas também se ultrapassa a limitação da exceção contida no artigo 59.º do Regulamento da IA.

É verdade que com o pacote de propostas "*omnibus*" se sistematiza de modo mais claro a relação entre o RGPD e o RIA[68]. O reconhecimento liminar do interesse legítimo como fundamento para o processamento de dados pessoais no desenvolvimento dos modelos de IA de finalidade geral e a isenção (mesmo que condicionada a objetivos de proteção dos indivíduos) da proibição de processamento das categorias especiais de dados, dispensando o requisito do consentimento do seu titular, clarificam dúvidas e colmatam lacunas do regime vigente. Contudo, as modificações propostas não vão no sentido de assegurar primacialmente a proteção e a soberania dos indivíduos sobre os seus dados pessoais. Se preocupações de promoção da inovação tecnológica das empresas europeias estarão subjacentes às propostas[69], as modificações correspondem sobretudo aos interesses da indústria e favorecem os grandes prestadores de modelos de IA de finalidade geral presentes no mercado. Como se verá de seguida, podem fragilizar a tutela dos titulares de dados pessoais e não eliminam todos os problemas, nomeadamente em relação à operacionalização dos direitos dos titulares dos dados pessoais que são salvaguardados, incluindo a objeção incondicional ao processamento de dados pessoais. Se o CEDP propôs um direito de objeção ou um direito à eliminação dos dados em termos mais favoráveis aos previstos nos artigos 21.º e 17.º do RGPD e a proposta "Digital *omnibus*" vai no mesmo sentido, permanecem algumas incertezas[70].

São múltiplas as questões e dificuldades que se colocam e vão permanecer: na verificação do fundamento do interesse legítimo para o tratamento de dados pessoais no treinamento dos

---

omnibus-eu-commission-wants-wreck-core-gdpr-principles (02.06.2026); CORPORATE EUROPE OBSERVATORY, "Article by article, how Big Tech shaped the EU's roll-back of digital rights", *Corporate Europe Observatory*, 18.01.2026, https://corporateeurope.org/en/2026/01/article-article-how-big-tech-shaped-eus-roll-back-digital-rights (02.06.2026); ITXASO DOMÍNGUEZ DE OLAZÁBAL, "Europe is dismantling its digital rights from within", *EDRI*, 2025, https://edri.org/our-work/europe-is-dismantling-its-digital-rights-from-within/ (02.06.2026).

[67] CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 9 e 29.

[68] Reconhece-se que o RGPD é a "pedra angular" da legislação digital da União. COMISSÃO EUROPEIA, *Commission Staff Working Document* (…), SWD(2025) 836 final, cit., p. 33.

[69] Há uma coincidência significativa entre as posições apresentadas pela indústria e as propostas "*omnibus*" para o domínio digital. COMISSÃO EUROPEIA, *Call for evidence and public consultation on the Apply AI Strategy*, 2025, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/14625-Apply-AI-Strategy-strengthening-the-AI-continent_en (02.06.2026); COMISSÃO EUROPEIA, *Commission Staff Working Document* (…), SWD(2025) 836 final, cit.

[70] Recorde-se, no entanto, que o TJ limitou o alcance territorial do "direito ao esquecimento" ao espaço da UE (Ac. do TJ de 24 de setembro de 2019, *Google (Territorial scope of de-referencing)*, Proc. C-507/17, ECLI:EU:C:2019:772, n.ºs 69 e 73), o que desde logo levanta um problema quando o processamento de dados pessoais, nomeadamente a exposição de informação pessoal nos *outputs*, ocorre fora do território da UE. A preocupação com a transferência de dados transfronteiriça é objeto de uma atenção particular pelo Comité Consultivo da Convenção 108, que salienta a sua especial complexidade. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 34-35.

modelos, na "regurgitação" de informação sobre pessoas, no direito à correção ou ao apagamento de dados pessoais que foram integrados no modelo ou constam nos seus *outputs*[71], na responsabilidade pelo processamento ilícito de dados pessoais, em especial quando os modelos são objeto de *finetuning* a jusante da colocação no mercado[72] ou quando os *inputs* e a informação colocada pelos utilizadores nos *prompts* retroalimentam o modelo, como sucede em regra nos modelos não *pro*[73]. O problema torna-se mais complexo em relação à proteção de dados pessoais de terceiros introduzidos pelo utilizador ou obtidos da interação ambiental com a realidade física, com outros dispositivos ou com fontes de informação. A proteção dos dados pessoais é um problema de governação de risco ao longo de toda a cadeia de valor baseada em modelos de IA de finalidade geral[74]. Nos pontos seguintes percorrem-se as medidas de proteção e as obrigações dos prestadores destinadas a responder a essas questões e dificuldades.

3. Medidas de proteção dos dados pessoais no desenvolvimento e funcionamento dos modelos de IA de finalidade geral. Incertezas, dificuldades e limites

O Parecer 28/2024 do CEDP constitui uma bússola útil na elaboração do sistema de governação de risco. Coloca a ênfase no princípio da lealdade, que relaciona com o princípio da transparência. Não obstante, ainda que apresente os critérios orientadores, verifica-se uma profusão de conceitos indeterminados, a carecer de preenchimento[75], alguns até exigindo uma

[71] No que concerne à alegação de que os modelos não "desaprendem", ou não será possível apagar memórias registadas ou inferidas, a tecnologia não deve ser antropomorfizada, embora as dificuldades e os custos possam ser significativos, como sucederá se for necessário o retreinamento do modelo. Pode deduzir-se da política de privacidade da OpenAI que tal é possível, apesar da autonomia do modelo. A evolução tecnológica também parece ser rápida neste domínio, pois se a política de privacidade da OpenAI, em dezembro de 2023 (*Europe privacy policy of OpenAI*, Version 15 December 2023, effective by 15 February 2024, § 6), apontava maiores dificuldades na correção do que na eliminação dos dados, a política atual, de 6.02.2026, não indicia esse problema. OPENAI, *Política de Privacidade*, cit..; EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, cit., p. 9.

[72] Consultative Committee Of The Convention For The Protection Of Individuals With Regard To Automatic Processing Of Personal Data (Convention 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 16.

[73] A OpenAI expressamente informa que utiliza os *prompts* no treino dos modelos não pagos, prevendo a possibilidade de oposição. OPENAI, *How your data is used to improve model performance*, https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance (02.06.2026).

[74] Consultative Committee Of The Convention For The Protection Of Individuals With Regard To Automatic Processing Of Personal Data (Convention 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 7.

[75] Afirma-se que os dados pessoais não devem ser tratados "por métodos desleais, ou por engano, ou de uma forma que seja "injustificadamente prejudicial, ilegalmente discriminatória, inesperada ou enganadora para o titular dos dados" (...) [t]endo em conta a complexidade das tecnologias envolvidas, as informações sobre o tratamento de dados pessoais no âmbito dos modelos de IA devem, por conseguinte, ser fornecidas de forma acessível, compreensível e fácil de utilizar". CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., pp. 21.

ponderação com outros bens jurídicos[76]. O Comité Consultivo da Convenção 108[77] aponta ainda o conflito entre os diversos princípios aplicáveis, nomeadamente entre a transparência ou a minimização e a segurança ou a não discriminação, indicando como princípios orientadores os princípios da proporcionalidade, necessidade, lealdade e proteção dos direitos fundamentais. O Comité considera que os riscos para os dados pessoais e para a privacidade decorrentes dos LLM não são adequadamente protegidos pelos meios já instituídos, nem por medidas *ad-hoc*[78], sendo indispensável uma metodologia específica para o sistema de governação de risco, propondo a metodologia "HUDERIA" e o modelo "COBRA".

As medidas de transparência impostas no RIA aos modelos de finalidade geral, em especial aos que têm risco sistémico, são complementares das obrigações de informação previstas nos artigos 13.º e 14.º do RGPD, as quais são consideradas pelo CEPD relevantes, mas não suficientes. No primeiro, preveem-se as obrigações quando os dados são recolhidos junto do titular e, no segundo, quando são recolhidos através de terceiros[79]. Sendo a maioria dos dados utilizados obtidos por "*web scrawling*" e "*web scraping*", a recolha dos dados pessoais para o desenvolvimento dos modelos cai sobretudo no âmbito do artigo 14.º. Mas, como cumprir as obrigações contidas nesse dispositivo? Será de aplicar o n.º 5 do artigo 14.º, dispensando-se a comunicação ao titular dos dados pessoais, pois tal traduzir-se-ia num esforço impossível ou desproporcionado[80].

Perante a realidade, a primeira e mais efetiva linha de proteção dos dados pessoais é a literacia digital dos respetivos titulares que permita que sejam verdadeiros agentes conscientes e livres do mundo digital[81]. O RIA, embora reconhecendo a sua importância no considerando 56 do Preâmbulo, onde menciona "as competências, os conhecimentos e a compreensão" das

[76] Parece admitir-se a existência de formas de tratamento "justificadamente" prejudiciais ou "legalmente" discriminatórias. Há ainda que considerar outros interesses (*v.g*. o combate à fraude, a cibersegurança, a exatidão ou a restrição dos *outputs*?) em relação aos quais é necessário fazer um segundo *balancing test*.

[77] CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 27-28.

[78] CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 9-10. Entre as medidas propostas estão mecanismos de autenticação, compartimentação de ambientes sensíveis, monitorização, teste adversariais e "*red-teaming*" ou a encriptação (pp. 31-32).

[79] A questão da aplicabilidade destes dispositivos já se levantou para os dados inferidos, pelo facto de não serem recolhidos. SANDRA WACHTER, BRENT MITTELSTADT, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI", *Columbia Business Law Review*, n.º 2, Nova Iorque, Columbia School of Law, 2019, pp. 1-135, https://ssrn.com/abstract=3248829 (02.06.2026).

[80] O CEDP considera que pode aplicar-se a isenção do artigo 14.º, porque é impossível informar cada titular dos dados obtidos por "*web scraping*", desde que se cumpram os requisitos previstos nessa disposição. *Report of the work undertaken by the ChatGPT Taskforce*, cit., p. 8.

[81] A Convenção-Quadro sobre a IA do Conselho da Europa (CoE) salienta, no seu artigo 20.º, a importância da promoção da literacia digital dos cidadãos ("*Each Party shall encourage and promote adequate digital literacy and digital skills for all segments of the population, including specific expert skills for those responsible for the identification, assessment, prevention and mitigation of risks posed by artificial intelligence systems*"). As orientações elaboradas pelo Comité Consultivo da Convenção 108 colocam entre os *stakeholders* relevantes as organizações da sociedade civil e os utilizadores finais. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 10-11. Também o CEDP recomenda, entre as medidas de transparência a adotar pelos prestadores dos modelos, a realização de campanhas mediáticas e outras medidas de comunicação para "ajudar a superar a assimetria da informação e permitir que os titulares dos dados obtenham uma melhor compreensão do tratamento envolvido na fase de desenvolvimento". *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 31.

pessoas afetadas, apenas contém uma disposição normativa a esse propósito, o artigo 4.º. Aí se institui uma obrigação de promover a formação dos colaboradores dos prestadores e responsáveis pela implantação. Este dispositivo não menciona os modelos, mas apenas os sistemas, facto que se deverá ao momento tardio em que os modelos foram incluídos no âmbito do RIA. A insuficiência da letra desse dispositivo pode resolver-se com uma interpretação extensiva, pois a inclusão dos modelos estará no espírito da norma. Ainda que se considere uma verdadeira lacuna, esta pode ser colmatada através da aplicação analógica do regime previsto para os sistemas de IA. A proposta "Digital Omnibus on AI" (artigo 1.º, n.º 4) altera esta disposição e elimina a obrigação para os prestadores e responsáveis pela implantação, convertendo-a num compromisso dos EM e da Comissão Europeia de incentivar esses operadores a "adotar medidas para garantir que o seu pessoal e outras pessoas envolvidas na operação e utilização de sistemas de IA em seu nome dispõem de um nível suficiente de literacia no domínio da IA". Os modelos continuam ausentes da previsão, o que agora é incompreensível. Também incompreensível é a completa ausência de preocupação com a literacia das pessoas, afastando-se do compromisso assumido pela UE e pelos EM na Convenção- Quadro do CoE[82] e como é indispensável à governação "WoS" ("Whole of Society") que se impõe num domínio com os riscos e o impacto social da IA.

Continuam a ser importantes as medidas de mitigação gerais e específicas indicadas pelo CEDP, sejam medidas que facilitem o exercício dos direitos dos titulares dos dados, tais como mecanismos de reclamação ágeis[83], sejam medidas técnicas de curadoria e governação de dados[84]. Além da pseudonimização e "mascaramento", com substituição de dados verdadeiros por dados falsos, por exemplo de nomes ou endereços[85], o CEDP refere, na fase de desenvolvimento do modelo, a exclusão de certas fontes (contendo informação de pessoas vulneráveis ou informação inapropriada, por exemplo), nomeadamente através do reconhecimento de mecanismos que exprimem a recusa do *web scraping*, tais como o

[82] O Conselho da UE aprovou a Convenção, em 21.04.2026, através da Decisão (UE) 2026/1080 do Conselho, de 21 de abril de 2026, JO L, 2026/1080, 13.5.2026, concluindo, assim, o processo de vinculação da União.

[83] A OpenAI disponibiliza um "Portal de Privacidade". Os direitos dos titulares de dados pessoais podem ser exercidos na conta ou por email. Os direitos descritos que podem ser exercidos diretamente pelo titular na respetiva conta são o acesso, a correção e a atualização da informação, a eliminação das interações, a eliminação da conta e a objeção à utilização "*do seu conteúdo para ajudar a treinar os nossos modelos*". É ainda possível ao titular solicitar o descarregamento dos seus dados pessoais e a remoção desses dados das respostas do ChatGPT, embora sujeita a aprovação, caso em que a OPENAI assume o seguinte compromisso: "trabalharemos para impedir que essas informações voltem a aparecer nas respostas do ChatGPT", https://privacy.openai.com/policies (02.06.2026). Se aparentemente se asseguram os direitos dos titulares, subsistem questões: o que constitui "o seu conteúdo"? o conteúdo produzido pelo sujeito? todo o conteúdo respeitante ao sujeito, independentemente da sua origem? distingue-se da "informação" do sujeito, também aí referida? Segundo o CEDP, o "content" inclui o "*input of data subjects when interacting with LLMs such as ChatGPT as well as file uploads and the user feedback regarding the quality of the data output*" (*Report of the work undertaken by the ChatGPT Taskforce*, cit., p. 7). A remoção dos dados pessoais de um sujeito das respostas do ChatGPT é admitida como possível, mas a OpenAI não se compromete com esse resultado, sujeitando-o a aprovação e assumindo apenas uma obrigação de comportamento, agir no sentido de "prevenir" essa regurgitação nos *outputs* futuros. Fica a dúvida sobre se a dificuldade é técnica ou se a aprovação fica dependente de uma avaliação dos fundamentos e ponderação de interesses.

[84] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., pp. 28-30.

[85] Esta medida pode ter um efeito negativo na qualidade dos dados, especialmente relevante nos modelos de finalidade geral, pois, como a respetiva utilização não está definida à partida, essa falsificação pode conduzir a futuras inferências imprecisas.

robots.txt ou o *ai.txt* files[86], mas igualmente através da utilização de filtros de *input*[87], da autenticação federada, que isola a informação pessoal, ou de outras técnicas de proteção (*v.g.* o ruído estatístico). Também se apontam a inviabilização da memorização e da "regurgitação" aquando da utilização, através de "*output filters*" que recusem responder ou replicar informação pessoal, mesmo que solicitada[88], a eliminação dos "*logs*" da interação, ou a colocação das marcas de conteúdo gerado por IA, bem como a supervisão humana e até a "desaprendizagem" dos modelos[89].

Se o CEDP considera que a utilização em larga escala dessas técnicas permite mitigar os riscos, todas são complexas e onerosas e permanecem as dúvidas sobre o seu sucesso, especialmente com o desenvolvimento das técnicas de extração (v.g. *membership inference attacks, adversarial attacks,* ou *model inversion*)[90]. De qualquer modo, impõe-se um sistema de governação de risco com uma abordagem preventiva, nomeadamente através da condução de testes adversariais, simulação de cenários ou "*red-teaming*"[91]. Algumas medidas necessárias para a identificação e avaliação dos riscos podem ser contrárias a técnicas de proteção de dados acima indicadas, como a eliminação dos "*logs*", pois estes podem ser importantes para a recolha de indícios dos riscos[92].

A Comissão admite a dificuldade de implementação das medidas de mitigação dos riscos e nas orientações de 18 de julho de 2025 dispensa os prestadores de modelos de IA de finalidade geral colocados no mercado antes de 2 de agosto de 2025 de "retreinarem" ou promoverem a "desaprendizagem" dos modelos "se tal não for possível para ações realizadas anteriormente, se algumas das informações sobre os dados de treino não estiverem disponíveis ou se a sua recuperação implicar encargos desproporcionados para o prestador"[93]. Essas situações devem ser publicitadas no resumo pormenorizado previsto no artigo 53.º, n.º 1, al. d) do RIA.

---

[86] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 30.

[87] EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, 23 May 2024, cit., p. 7. O possível problema com esta última medida reside no facto dos modelos de finalidade geral também serem aptos para utilizações relacionadas com essas informações sensíveis (v.g., no domínio da saúde), caso em que o uso desses filtros terá um potencial impacto negativo na exatidão do modelo e, no limite, poderá mesmo impossibilitar a utilização para essa finalidade.

[88] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 30.

[89] Desde 2019 que essa possibilidade é estudada. A. GINART et al., "Making AI forget you: data deletion in machine learning", in H. WALLACH et al., *Proc. 33rd Conference on Neural Information Processing Systems*, Vancouver, NeurIPS, 2019, https://arxiv.org/pdf/1907.05012 (02.06.2026). Em especial em relação aos LLM, SIJIA LIU et al., "Rethinking machine unlearning for large language models*", Nature Machine Intelligence*, volume 7, Londres, Nature Portfolio, 2025, pp. 181–194, https://www.nature.com/articles/s42256-025-00985-0 (02.06.2026).

[90] MICHAEL VEALE, REUBEN BINNS & LILIAN EDWARDS, "Algorithms that remember: model inversion attacks and data protection law", *Philosophical Transactions of the Royal Society A*, Vol 376, Ed. 2133, Londres, Royal. Society Publishing, 2018, http://dx.doi.org/10.1098/rsta.2018.0083 (02.06.2026). Se essas técnicas forem suscetíveis de ser razoavelmente utilizadas impedem a classificação de um modelo como um modelo anónimo.

[91] ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 50-51. Kim Martineau, *What is red teaming for generative AI?*, 11.04.2024, https://research.ibm.com/blog/what-is-red-teaming-gen-AI (02.06.2026).

[92] ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., p. 51.

[93] COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, C(2025) 7719 final, Bruxelas, 19.11.2025, pp. 30-31, https://digital-strategy.ec.europa.eu/en/library/guidelines-scope-obligations-providers-general-purpose-ai-models-under-ai-act (02.06.2026).

O CEDP recorda também medidas de mitigação dos riscos para os direitos fundamentais previstas no RGPD e que propõe sejam incluídas na política de governação de risco[94], nomeadamente a AIDP (Avaliação de Impacto nos Dados Pessoais) ou "DPIA" (*Data protection impact assessment*)[95]. Aponta ainda um conjunto de métodos e elementos a considerar. Estes respeitam às características intrínsecas dos modelos, incluindo a sua documentação, ao contexto da respetiva utilização, distinguindo a acessibilidade do grande público da utilização apenas no interior de uma organização e tendo em conta a formação dos utilizadores. É ainda considerado relevante o custo da eventual extração dos dados[96]. A implementação do DPIA, mesmo que não se preencham os requisitos que a tornam obrigatória, segundo o artigo 35.º do RGPD, é recomendável como medida de mitigação preventiva de potenciais riscos[97].

A perspetiva alargada de proteção dos dados pessoais encontra-se igualmente nas orientações propostas do Comité Consultivo do Conselho da Europa (também Comité Convenção 108), com uma referência especial aos Agentes de IA. As orientações do Comité Consultivo, em que os Agentes de IA são objeto de uma atenção especial, fundamentam-se nos princípios estabelecidos na Convenção 108 e no seu Protocolo n. 223 (também Convenção 108+), na CEDH (particularmente, o artigo 8.º) e na Convenção-Quadro sobre a IA e têm como objetivo clarificar a aplicação da Convenção 108+ no contexto do desenvolvimento da IA baseada em LLM[98]. Propõe uma "metodologia estruturada" para "Identificar, Avaliar, Mitigar e Monitorizar" (também "IAMM") os riscos em ambiente real e com vários grupos de interessados[99]. Além de

---

[94] *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*, 2024, cit., pp. 14-19.

[95] Previsto no artigo 35.º do RGPD. EDPB, ARTICLE 29 DATA PROTECTION WORKING GROUP, *Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679*, 2017, http://ec.europa.eu/newsroom/document.cfm?doc_id=44137 (02.06.2026).

[96] Segundo o CEDP, os elementos a considerar na avaliação, nomeadamente pelas autoridades de proteção de dados, são diversos e abrangem as condições da utilização dos dados de treino (v.g. a agregação, randomização, etc.), a testagem estruturada, entre outros, contra "(i) inferência de atributos e associações; (ii) exfiltração; (iii) regurgitação de dados de treino; (iv) inversão de modelos; ou (v) ataques de reconstrução". A documentação deve incluir, entre outras, informações relacionadas com a DPIA, bem como "as medidas específicas para cada fonte de conjuntos de dados de formação, incluindo os URL das fontes pertinentes e descrições das medidas tomadas (ou já tomadas por terceiros fornecedores de conjuntos de dados)", "as medidas técnicas e organizativas adotadas em todas as fases ao longo do ciclo de vida do modelo, que contribuíram para, ou verificaram, a falta de dados pessoais no modelo", "a documentação que demonstra a resistência teórica do modelo de IA às técnicas de reidentificação, bem como os controlos concebidos para limitar ou avaliar o êxito e o impacto dos principais ataques (regurgitação, ataques de inferência de filiação, exfiltração, etc.)." *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 19.

[97] ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., p. 83.

[98] CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, pp. 7-8. Uma exemplificação das medidas adequadas para os principais riscos e operadores responsáveis pode encontrar-se em ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 67-73.

[99] Outras metodologias têm sido apresentadas. CLAUDIO NOVELLI *et al*. propõem uma metodologia baseada no modelo delineado pelo IPPC (Painel Intergovernamental para as Alterações Climáticas). "AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act", in *Digital Society*, Volume 3, n.º 13, Londres, Springer Nature 2024, https://link.springer.com/article/10.1007/s44206-024-00095-1 (21.05.2026). LAURA WEIDINGER *et al*. salientam a desadequação das metodologias "tecnológicas" habitualmente utilizadas a uma realidade que é intrinsecamente socio-tecnológica, em que a sua funcionalidade está dependente não apenas dos componentes tecnológicos, mas de ações humanas indeterminadas. Propõem uma abordagem com três planos: a avaliação das capacidades, a avaliação da interação humana e a avaliação do impacto sistémico. *Sociotechnical Safety Evaluation of Generative AI Systems*, 2023, https://arxiv.org/pdf/2310.11986 (29.05.2026). ISABEL BARBERÀ, nos casos de uso que apresenta, salienta a importância da colaboração *multistakeholder*, incluindo especialistas em ética e direitos fundamentais e representantes da sociedade civil e dos utilizadores finais. *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., p. 82.

estruturada, deve ser iterativa, proporcionada e contextual[100]. Cabe ao controlador dos dados pessoais a principal responsável pela prevenção e minimização dos riscos, devendo ser articulada com as obrigações inscritas na Convenção 108 e com outras metodologias de proteção dos direitos humanos[101] (v.g., "*HUDERIA Methodology*" e "*COBRA Model*"[102]). Tem quatro fases[103]: (1) Identificação dos riscos para a proteção de dados e para a privacidade nas sucessivas fases do ciclo de vida; (2) Avaliação a dois níveis, o nível do modelo e o nível do sistema em que este é implantado, com atenção às ações e responsabilidades dos diferentes *stakeholders* e com o objetivo de dotá-los com ferramentas adequadas aos riscos nas diferentes fases, sejam relacionados com o treino de dados, a memorização, a inferência, as alucinações, a integração do modelo no sistema e a respetiva adaptação a jusante ou a interação com o utilizador; (3) Mitigação dos riscos através de medidas técnicas e organizacionais "adequadas à natureza, probabilidade e severidade dos riscos identificados"[104], tais como controlos organizativos, medidas de privacidade tecnológica, medidas de avaliação, monitorização e salvaguarda operacional; (4) Monitorização dos riscos residuais e potenciais ao longo do ciclo de vida através da avaliação contínua, supervisão, auditoria e revisão. Além de medidas tecnológicas automáticas e de métricas quantitativas, impõe-se a supervisão humana qualitativa[105].

As orientações do CEDP e do Comité Consultivo da Convenção 108 coincidem na adequação à finalidade do modelo, na operacionalização iterativa ao longo de todo o ciclo de vida do modelo, ou seja, para lá da implantação num sistema de IA, considerando os impactos da sua utilização. Os objetivos do quadro de governação de risco podem ser a eliminação ou a mitigação do risco até que este se torne aceitável, caso a eliminação não seja possível ou não se mostre adequada após a ponderação dos interesses. Ainda assim, o objetivo principal das medidas indicadas não parece coincidir, pois no documento do CEDP, embora se pretenda reduzir o impacto negativo nos interesses dos titulares dos dados, sobressai uma lógica instrumental em ordem a

[100] Os potenciais usos e o contexto da utilização, seja o contexto demográfico, seja o contexto cultural, linguístico, empresarial, etc., são importantes. ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., p. 50.

[101] CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 6 e 8.

[102] A "*HUDERIA Methodology*" e o "*HUDERIA Model*" foram aprovados pelo Comité de Ministros do Conselho da Europa, em 25.02.2025, e em 25.02.2026, respetivamente. COUNCIL OF EUROPE, *HUDERIA Methodology and Model. Methodology for assessing the risks and impacts of artificial intelligence systems from the perspective of human rights, democracy and the rule of law. Resource model for context-based risk analysis (COBRA)*, abril 2026, https://rm.coe.int/prems-002726-gbr-2006-huderia-texte-web-a4/48802ba7b1 (02.06.2026). A severidade dos riscos tem quatro níveis: o "dano mínimo ou moderado", o dano sério, o dano crítico e o dano catastrófico.

[103] CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 9-10. Indicações mais exaustivas das várias fases de governação do risco para os dados pessoais podem encontrar-se em ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 48-79. A autora apresenta alguns casos de uso, tais como um Chatbot de assistência a clientes, uma ferramenta de monitorização e apoio educacional a estudantes, um assistente de organização de viagens ou de agenda (pp. 80-96).

[104] Uma matriz exemplificativa contextual dessa avaliação encontra-se em ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 58-65. Essa interseção entre probabilidade e severidade determina a classificação do risco entre baixo, médio, elevado ou muito elevado. Serão os riscos classificados nestas duas categorias devem ter prioridade no quadro de gestão e de mitigação.

[105] ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 76 e 77. Aa autora qualifica a abordagem que propõe como híbrida.

preservar os interesses do prestador do modelo ao passo que nas "*Draft Guidelines*" é mais evidente a prioridade da proteção dos dados pessoais[106].

As medidas devem ser adaptadas às circunstâncias e às finalidades do modelo e devem ter em consideração o dinamismo e a autonomia dos modelos, o que atenuará o peso dos interesses do titular dos dados no indispensável *balancing test* com os interesses do prestador do modelo[107]. Efetivamente, a mitigação dos riscos é também um exercício de ponderação de interesses[108] e, como tal, inclui a consideração dos interesses do prestador[109], nomeadamente o interesse do desenvolvimento e exatidão do modelo[110] ou a liberdade de expressão e informação[111]. A prevalência dos interesses do prestador ou do responsável pela implantação pode até conduzir à opção pela não adoção de medidas de mitigação, se o risco for considerado aceitável. O facto de os dados pessoais serem recolhidos de *dataset* mistos ou através de *webscraping* e estarem integrados ("*embedded*") nos parâmetros dos modelos como representações matemáticas em conjunto com outros dados, dificulta a sua eliminação ou correção sem afetar o desempenho do modelo ou os direitos de outrem. O TJ, no Proc. C-434/16, considerou que o direito à eliminação está condicionado ao cumprimento da finalidade (v.g. a avaliação definitiva do titular)[112]. Ora, o processamento de dados pessoais no desenvolvimento de um modelo de IA de finalidade geral destina-se ao treinamento do modelo. Após a conclusão deste está cumprida a finalidade da utilização dos dados pessoais recolhidos para o desenvolvimento do modelo, pelo que a sua eliminação se impõe. Tal não exclui as dificuldades no que respeita à sua efetivação, em virtude das características dos modelos. Estando os dados integrados no modelo em parâmetros matemáticos abstratos que integram outros dados, tal pode obrigar a uma eliminação de dados agregados. Como se asseguram os direitos de terceiros, sejam outros dados integrados nesses parâmetros, sejam os direitos de autor respeitantes ao *software*[113]?

Já na fase posterior de funcionamento do modelo, o facto de a informação pessoal incluída nos *outputs* não constituir a reprodução de dados recolhidos na fase de desenvolvimento ou até

---

[106] Consultative Committee Of The Convention For The Protection Of Individuals With Regard To Automatic Processing Of Personal Data (Convention 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 9.

[107] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 28.

[108] Isabel Barberà, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., p. 66.

[109] Gianclaudio Malgieri admitiu como inevitável a assimetria entre os consumidores e titulares de dados pessoais e os controladores de dados, considerando a proteção da propriedade intelectual e a liberdade de empresa. "Trade Secrets v Personal Data: a possible solution for balancing rights", *International Data Privacy Law*, Volume 6, Issue 2, Oxford, Oxford University Press,2016, pp. 102–116, https://doi.org/10.1093/idpl/ipv030 (02.06.2026).

[110] Norbert Nolte e Christoph Werkmeister contrapuseram o artigo 16.º da CDF (Liberdade de Empresa) e o interesse legítimo do desenvolvimento técnico como fundamento para recusar a eliminação de dados enquanto for necessário para esse desenvolvimento. "DS-GVO Art. 17 Recht auf Löschung ("Recht auf Vergessenwerden")", Peter Gola (ed.), *Datenschutz-Grundverordnung VO (EU) 2016/679*, 1.ª Ed., Munique, CHBeck, 2017, *apud* Sandra Wachter, Brent Mittelstadt, "A Right to Reasonable Inferences: re-thinking data protection law in the age of big data and AI", cit., pp. 39-40.

[111] Em especial nos modelos de GenAI, que geram conteúdos de texto ou gráficos, não será possível invocar o exercício da liberdade de expressão ou de informação, previsto no n.º 3, al. c) do artigo 17 e que exclui a aplicação dos n.ºs 1 e 2, que preveem o direito ao apagamento?

[112] Ac. de 20 de dezembro de 2017, *Nowak*, Proc. C-434/16), ECLI:EU:C:2017:994. E o direito à correção é condicionado, não permitindo, obviamente, "corrigir" a posteriori respostas erradas (n.ºs 52 e 55).

[113] Considerando 63 do RGPD; EDPB Article 29 Data Protection Working Party, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, cit., p. 17; Proc. C- 487/21, cit., n.º 43.

posteriormente, através dos *prompts* introduzidos pelo utilizador ou da interação ambiental (*v.g.*, o recurso a fontes de informação *online* ou disponibilizadas na utilização, sejam documentos ou dados telemétricos, incluindo biométricos), mas constituir informação inferida autonomamente pelo modelo, uma mimetização ou um simulacro de dados pessoais[114], não exclui a sujeição ao regime do RGPD[115], ainda que se possa questionar a aplicação dos artigos 13.º e 14.º do RGPD, pois as referidas disposições referem-se a dados recolhidos diretamente ("*first party*" ou "*third party*"). Acompanhando o entendimento do TJ de que também a informação inferida, quando respeite a um indivíduo, se inclui no conceito de dados pessoais, o Parecer 28/2024 do CEDP entende que essas obrigações se aplicam, embora admita a aplicação da dispensa de comunicação prevista no n.º 5 do artigo 14.º quando tal seja demasiado difícil ou oneroso, como sucede quando não há uma relação com o titular dos dados. Por outro lado, recomenda que as medidas de transparência sejam voluntariamente reforçadas, através de informação suplementar sobre os critérios que presidiram à recolha e à escolha dos *datasets*, da implementação de campanhas de informação nos *media*, disponibilização de e-mail para contacto, através de *FAQ* (*frequent asked questions*), do uso de suportes gráficos para a apresentação dos modelos ("*model cards*"[116]), ou de relatórios anuais de transparência[117].

A complexidade manifesta-se em outras dimensões, nomeadamente se a informação pessoal é não apenas reproduzida ("regurgitada") nos *outputs* em resposta a *prompts* dos utilizadores, como é também integrada no modelo a partir dos *prompts* dos utilizadores ou através dos *inputs* recolhidos de fontes de informação com as quais o modelo interage e é utilizada e replicada posteriormente por este?[118] A solução preventiva passará, em primeiro lugar, pela

---

[114] É o caso dos modelos treinados com a voz de alguém e que reproduzem essa voz. Se por um lado, estes modelos não podem ser considerados anónimos, por outro lado, a manipulação da identidade ou a desinformação associada a estas ferramentas que mimetizam indivíduos coloca outras preocupações para além da proteção dos dados pessoais. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 16.

[115] O TJ declarou que informação inferida, subjetiva, sob a forma de opiniões ou de apreciações, constitui "dados pessoais" quando respeitem a uma pessoa (Ac. de 20 de dezembro de 2017, *Nowak*, Proc. C-434/16), cit., n.ºs 34, 35, 42. No Ac. de 4 de maio de 2023, *Österreichische Datenschutzbehörde e CRIF*, Proc. C-487/21, EU:C:2023:369, n.º 26, o TJ declarou que o conceito de dados pessoais "não abrange apenas os dados recolhidos e conservados pelo responsável pelo tratamento, mas inclui igualmente todas as informações resultantes de um tratamento de dados pessoais que digam respeito a uma pessoa identificada ou identificável" e que o elenco do artigo 4.º, n.º 2 do RGPD não é exaustivo (n.º 27). JACOB LEON KRÖGER, "Recognizing Information Inferred about Individuals as Personal Data", Chapter adapted from: KRÖGER, J. L., *Rogue Apps, Hidden Web Tracking and Ubiquitous Sensors*, Doctoral dissertation, Technische Universität Berlin, 2022, pp. 222–225, https://doi.org/10.14279/depositonce-16043, https://ssrn.com/abstract=4349200 or http://dx.doi.org/10.2139/ssrn.4349200 (02.06.2026). SANDRA WACHTER, BRENT MITTELSTADT, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI", cit.. O TJ considerou dignos de proteção, enquanto dados sensíveis, os dados que indiretamente, "no termo de uma operação intelectual de dedução ou de aproximação", revelam dados sensíveis, como a orientação sexual. Ac. do TJ de 1.08.2022, *Vyriausioji tarnybinės etikos komisija*, Proc. C-184/20, ECLI:EU:C:2022:601, n.º 123.

[116] Sobre os "*model cards*", veja-se HUGGING FACE, "Model Cards", 2024, https://huggingface.co/docs/hub/en/model-cards (02.06.2026).

[117] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 29.

[118] No Proc. C-340/21, o TJ considerou que "o responsável pelo tratamento não pode ficar isento da sua obrigação de reparar o dano sofrido por uma pessoa, nos termos do artigo 82.º, n.º 1 e 2 [do RGPD], (…), pelo simples facto de esse dano resultar de uma divulgação ou acesso não autorizados a dados pessoais por "terceiro[s]", na aceção do artigo 4.º, ponto 10, do referido regulamento, devendo então o referido responsável provar que não é de modo nenhum responsável pelo evento que deu origem aos danos (n.º 74). Por outro lado, "uma divulgação ou um acesso não autorizados a dados pessoais por "terceiro[s]", na aceção do artigo 4.º, ponto 10 [do RGPD), (…), não são suficientes, por si só, para se considerar que as medidas técnicas e organizativas aplicadas pelo

definição de uma política de uso que restrinja tal possibilidade[119] ou, o que será mais efetivo, por filtros para eliminar ou impedir o tratamento desses dados[120], embora tal possa limitar o potencial dos modelos[121]. A OpenAI declara que já implementou medidas para filtrar o que qualifica como "conteúdo sensível"[122].

Além do direito à eliminação dos respetivos dados pessoais dos parâmetros do modelo, os titulares dos dados pessoais têm o direito de impedir o processamento e a divulgação em *outputs* futuros, bem como têm o direito à correção dos *outputs*[123]. Não obstante, subsistem ambiguidades sobre o alcance de alguns dos direitos[124], nomeadamente o direito à correção[125]. O TJ, no Proc. C-434/16, em relação a um exame escrito, além das respostas do próprio titular (mesmo que sejam replicação de conteúdos da autoria de terceiros)[126], considerou que constituem dados pessoais do examinando qualquer tipo de informações, tanto objetivas como subjetivas, sob a forma de opiniões ou de apreciações do corretor, na condição de "dizerem respeito" à pessoa em causa, pelo conteúdo, finalidade ou efeito[127]. Por outro lado, em especial nos modelos de *GenAI*, uma outra questão importante relaciona-se com a (im)possibilidade de assegurar a exatidão dos *outputs*[128], pois o objetivo daqueles não é tanto produzir resultados fiéis aos dados de um sujeito, mas inferir predições probabilísticas de *tokens* que constituem palavras ou representam outras realidades (v.g. imagens ou sons), o que pode constituir uma dificuldade para assegurar a correção da informação pessoal. O produto apresentado aos utilizadores constitui um encadeamento das palavras que mais provavelmente se seguem ou simulacros de outras realidades. As "alucinações" em que os

responsável pelo tratamento em causa não eram "adequadas", na aceção dos artigos 24.º e 32.º" (n.º 39). A aferição dessa adequação foi remetida para uma avaliação casuística pelos órgãos judiciais das medidas técnicas e organizativas aplicadas pelo responsável pelo tratamento pelos órgãos jurisdicionais nacionais, tendo em conta os riscos associados ao tratamento em causa e apreciando se a natureza, o teor e a aplicação dessas medidas são adequados a esses riscos (n.º 47). Ac. do TJ de 14 de dezembro de 2023, *Natsionalna agentsia za prihodite*, Proc. C-340/21, ECLI:EU:C:2023:986.

[119] Essa é uma das medidas de mitigação para tornar o risco aceitável inscritas no Capítulo 3 do Código de Práticas (Medida 4.2). COMISSÃO EUROPEIA, *Code of Practice for General-Purpose AI Models. Safety and Security Chapter*, cit., p. 16.

[120] O Comité Consultivo da Convenção 108 refere ainda mecanismos de monitorização e auditoria, bem como a supervisão humana. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 29.

[121] EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, cit., pp. 7-8.

[122] "[P]ersonally identifiable information (PII), sexual content, hate-speech, and similar content". OPENAI, *GPT-4 technical report*, Dec. 2023, Chapter 6, https://arxiv.org/abs/2303.08774 (02.06.2026).

[123] EDPB ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, 17/EN WP251 rev.01, pp. 17-18, https://ec.europa.eu/newsroom/article29/items/612053/en (02.06.2026).

[124] O WP29 considerou, nas suas orientações, que o direito à portabilidade não é aplicável aos dados inferidos. EDPB ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on the right to data portability*, 16/EN WP 242 rev.01, 5.04.2017, p. 10, https://ec.europa.eu/newsroom/article29/items/611233/en (02.06.2026).

[125] BART CUSTERS, HELENA VRABEC, "Tell me something new: data subject rights applied to inferred data and profiles", *Computer Law & Security Review*, Volume 52, Amsterdam, Elsevier, 2024, https://doi.org/10.1016/j.clsr.2024.105956 (02.06.2026). Os autores apresentam uma análise da possível aplicação dos vários direitos consagrados no RGPD às diversas categorias de dados, incluindo dados inferidos.

[126] Como pode suceder com as respostas num exame com consulta (n.º 40).

[127] N.ºs 34, 35 e 42.

[128] Como previsto no artigo 5.º, n.º 1, al. d) do RGPD.

modelos incorrem agravam esse risco de inexatidão[129]. A lealdade e a transparência obrigam a esclarecer o usuário sobre essa potencial inexatidão dos resultados apresentados[130].

Quanto ao direito de acesso, sendo os dados utilizados no treino dos modelos como representações matemáticas integradas nos parâmetros dos modelos, como se assegura o acesso nas condições previstas no artigo 15.º, n.º 3 do RGPD (uma cópia em formato eletrónico de uso corrente ou outro formato solicitado pelo titular[131])? No plano substantivo, a proposta "*Digital omnibus*", no seu artigo 3.º, n.º 4, propõe a alteração do artigo 12.º, n.º 5 do RGPD, especificando e expandindo o conceito de "pedidos manifestamente infundados ou excessivos" através do conceito do "abuso de direito" ("o titular utiliza de forma abusiva os direitos conferidos pelo presente regulamento para outros fins que não a proteção dos seus dados"), e que permite a exigência de pagamento ou até a recusa de acesso. O TJ[132], recentemente, acolheu essa orientação sobre os limites do direito de acesso por parte dos titulares de dados pessoais. Independentemente da validade das razões, há um potencial "*chilling effect*" sobre os titulares dos dados pessoais.

Há que sublinhar que o ónus da prova pertence ao controlador dos dados, ou seja, no que respeita à fase de desenvolvimento, ao prestador do modelo[133], que deverá comprovar que utilizou todas as medidas possíveis[134]. Na verdade, não há uma coincidência necessária entre os responsáveis relevantes no RGPD, nomeadamente o controlador e o processador de dados, e o prestador ou o responsável pela implantação dos sistemas de IA no RIA[135]. Em princípio, cada prestador é responsável pelos dados de que é controlador, como reconhece o considerando 109 do RIA[136], mas a dificuldade está em determinar a existência e imputação desse controlo. Com a introdução ou manipulação de dados a jusante, com o *finetuning/taylorization* do modelo ou com a sua implantação em sistemas de IA utilizados por outros operadores, pode haver partilha de responsabilidade e/ou controlo conjunto, nos termos do artigo 26.º do RGPD[137], pois a decisão sobre a finalidade tem lugar posteriormente e outro operador pode ser também controlador de dados, ainda que o prestador que desenvolveu o modelo possa continuar a ser o processador de dados. A complexidade exponencia-se quando

[129] LONG OUYANG, *Training language models to follow instructions with human feedback*, 2022, https://arxiv.org/pdf/2203.02155 (02.06.2026).
[130] EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, cit., pp. 8-9. As ferramentas atualmente disponíveis salvaguardam essa possibilidade de inexatidão.
[131] Segundo o TJ, deve assegurar-se o acesso de "forma concisa, transparente, inteligível e de fácil acesso, utilizando uma linguagem clara e simples, sendo que as informações devem ser prestadas por escrito ou por outros meios, incluindo, se for caso disso, por meios eletrónicos, a menos que o interessado solicite que a informação seja prestada oralmente". Proc. C-487/21, cit., n.º 38.
[132] Ac. do TJ de 19 de março de 2026, *Brillen Rottler*, Proc. C-526/24, ECLI:EU:C:2026:216. As Conclusões do Advogado-Geral, de 18.09.2025, orientavam-se no mesmo sentido. Não se tratando de um pedido repetido, previsto no RGPD, o Tribunal aplicou o instituto do abuso do direito a esse primeiro pedido de acesso e admitiu que mesmo um primeiro pedido pode ser um pedido excessivo "qualitativamente" (n.º 34). Jurisprudência anterior do Tribunal já se tinha pronunciado sobre o conceito de "pedidos excessivos", mas em situações de pedidos repetidos. Ac. do TJ de 9 de janeiro de 2025, *Österreichische Datenschutzbehörde (Pedidos excessivos)*, C-416/23, ECLI:EU:C:2025:3. Também o princípio da proporcionalidade já tinha sido considerado na ponderação dos interesses do titular dos dados pessoais em face dos interesses do controlador. Ac. do TJ de 21 de dezembro de 2023, *Krankenversicherung Nordrhein*, C-667/21, EU:C:2023:1022, n.º 54.
[133] Proc. C-340/21, cit. n.º 55; EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, cit., p. 7.
[134] Proc. C-340/21, cit.
[135] ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 43-47.
[136] O prestador a jusante ficará obrigado a complementar a documentação técnica já existente com informações sobre as alterações.
[137] EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, cit., p. 14.

o desenvolvimento do modelo continua com a interação posterior, particularmente quando a interação é multimodal[138] e o tratamento de dados é contínuo, e o prestador do modelo utiliza os dados introduzidos a jusante para as suas próprias finalidades de treino e melhoria do modelo, caso em que os riscos que se colocam a jusante se vão projetar a montante. Neste caso, o prestador do modelo assume igualmente o papel de controlador desses dados e deve incluir esses riscos na sua política de governação de risco[139]. Assim sucede com os principais modelos de *GenAI* disponibilizados ao público em geral[140]. São fundamentais as obrigações de informação impostas aos prestadores dos modelos em relação aos operadores a jusante, como preconizado no RIA e concretizado no Código de Práticas, mas não suficientes e há que determinar a distribuição das responsabilidades. Em abstrato, dir-se-á que as responsabilidades do prestador do modelo incidem sobre a infraestrutura e a respetiva manutenção, as condições do treino e das funcionalidades do modelo; já sobre o responsável pela implantação/utilizador profissional, quando não seja ele próprio um prestador (o que pode acontecer se for responsável por um *finetuning* posterior), é responsável pelas condições de uso e implantação e pode ser responsável pela proteção de dados de que seja controlador[141]. Nas situações em que o modelo é utilizado em ambiente fechado do operador a jusante, sem que o respetivo prestador defina a utilização e tenha acesso aos dados aí integrados, o prestador não é controlador de dados ou pode sê-lo limitadamente (quando possa utilizar alguns dados por razões técnicas ou de segurança) e a sua responsabilidade nesta fase do ciclo de vida do modelo corresponde às obrigações que decorrem do RIA. Quando o processamento tenha lugar na infraestrutura do prestador do modelo, mas sob as instruções do operador a jusante, o prestador será processador, segundo o RGPD. Se o operador a jusante modificar significativamente o modelo, pode mesmo assumir o papel de prestador, para efeitos do RIA, e será o controlador dos dados, para efeitos do RGPD[142].

Os "Agentes de IA" (também "*AI Agents*") colocam desafios acrescidos, pois habitualmente o modelo com a arquitetura básica é objeto de modificação dinâmica significativa associada à interação multimodal e às finalidades do operador a jusante. O responsável por essa modificação é também ele prestador do modelo final e pode ser controlador de dados, bem como terceiros operadores com os quais há interação. Os diferentes papéis distribuem-se por múltiplos operadores, em virtude dessa (re)construção do modelo e da interação

[138] Os modelos multimodais também produzem *outputs* de diferentes tipos. LAURA WEIDINGER *et al*., *Sociotechnical Safety Evaluation of Generative AI Systems*, cit., p. 6, nt. 2.
[139] Tal pode dificultar a determinação da responsabilidade, mas é igualmente importante para as decisões de *design* do modelo e do sistema de governação de riscos. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 14 e 18.
[140] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., pp. 33-35.
[141] ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 43-44.
[142] São os chamados "off-the-shelf models". ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 44-45.

multimodal[143]. Para esta realidade especialmente complexa e indeterminada[144], a avaliação casuística parece a única solução, seja na determinação da existência e imputação do controlo, seja na verificação do cumprimento das obrigações[145]. No entanto, a certeza jurídica não fica assegurada com uma análise casuística ou com a indicação de que a atribuição das responsabilidades deve ser clara[146]. A adoção de acordos/protocolos é desejável, mas a assimetria entre os operadores não pode ser ignorada e é importante a definição de critérios[147]. A retirada pela Comissão Europeia da proposta de diretiva relativa à responsabilidade na IA deixa esta questão inteiramente no domínio dos ordenamentos dos EM.

Se os dados pessoais, ainda que indevidamente recolhidos, foram efetivamente anonimizados em algum momento do desenvolvimento do modelo ou do sistema a colocação no mercado ou em serviço o modelo deixará de suscitar problemas em face do RGPD, desde que não haja recolha/processamento posterior de dados pessoais. Independentemente das obrigações estabelecidas no RGPD, o Regulamento da IA impõe aos prestadores dos modelos de IA de finalidade geral um conjunto de obrigações específicas, as quais também têm relevância para a proteção dos dados pessoais.

4. A proteção de dados pessoais nas obrigações dos prestadores de modelos de IA de finalidade geral

4.1. Obrigações dos prestadores de modelos de IA de finalidade geral

Como explicita o respetivo artigo 1.º, n.º 2, al. e), o RIA estabelece "regras harmonizadas para a colocação no mercado de modelos de finalidade geral", regras essas que são específicas

[143] O Comité Consultivo da Convenção 108 sublinha que os "Agentes de IA" aumentam o âmbito, a persistência e sensibilidade do processamento de dados pessoais. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 31. ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., p. 46. Os riscos são ainda exponenciados quando os Agentes de IA são utilizados pelos colaboradores das organizações para os auxiliarem e até substituírem sem o conhecimento dos responsáveis, fenómeno designado de "*shadow AI*". DELINEA, *2026 Identity Security Report. Uncovering the Hidden Risks of the AI Race*, 2026, p. 19, https://delinea.com/hubfs/Delinea/whitepapers/delinea-report-identity-security-report-2026.pdf (02.06.2026)

[144] As múltiplas possibilidades são sistematizadas em ISABEL BARBERÀ, *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., p. 47. O Comité Consultivo da Convenção 108 refere a dificuldade para os titulares em determinar se os seus dados foram e como foram utilizados no treino ou se a informação permanece ou como são geradas as inferências e propõe medidas melhoradas de auditabilidade, de acesso para o exercício dos direitos dos titulares e a supervisão humana. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., p. 34.

[145] CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., p. 31. Obviamente, as autoridades nacionais devem impor medidas corretivas, especialmente importantes quando os modelos estão em desenvolvimento contínuo fruto da respetiva interação depois de implantados em sistemas de IA. Tal não deixará de enfrentar dificuldades com um quadro institucional de aplicação e fiscalização que diversas autoridades, nacionais e europeias.

[146] Como faz o Comité Consultivo da Convenção 108. CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, cit., pp. 34 e 35.

[147] ISABEL BARBERÀ apresenta um quadro com a exemplificação dos riscos mais significativos para os dados pessoais nas diversas modalidades e fases do ciclo de vida dos LLM e respetiva responsabilidade. *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, cit., pp. 53-55.

e aplicam-se também quando os modelos são integrados ou fazem parte integrante de um sistema de IA (considerando 97). As obrigações previstas aplicam-se a partir da colocação no mercado dos modelos de finalidade geral no território da UE, quer os prestadores estejam estabelecidos ou localizados na União ou num país terceiro (artigo 2.º, n.º 1, al. a))[148].

O RIA contem um capítulo dedicado aos modelos de finalidade geral, o capítulo V (artigos 51.º a 56.º)[149], onde estão previstas as obrigações específicas a cumprir. Esse capítulo não prejudica a aplicabilidade de obrigações decorrentes de outros capítulos do RIA, nomeadamente quando os modelos sejam integrados em sistemas de risco elevado ou caiam no âmbito das obrigações de transparência previstas no artigo 50.º[150]. Acresce que, como visto nos pontos anteriores e é explicitado pelo RIA, o RGPD é também aplicável. Efetivamente, há uma parcial coincidência e uma complementaridade entre as obrigações do RGPD e do RIA[151], como se pode ver por alguns exemplos. As obrigações de registo e comunicação de atividades de processamento de dados pessoais, impostas pelo artigo 30.º do RGPD, correspondem em parte às obrigações previstas no 55.º do RIA; já o artigo 33.º, n.º 5 do RGPD obriga a registar e comunicar as "*breaches*" de dados pessoais, ainda que tal não constitua necessariamente um incidente grave relevante para o RIA. Como já referido, a proposta "*Digital omnibus*" (artigo 3.º, n.º 8) limita a obrigação de comunicação aos incidentes de violação de dados pessoais de que resulte um elevado risco para os direitos e liberdades dos titulares, aproximando-o do regime de comunicação previsto pelo RIA para os prestadores de modelos de IA de finalidade geral com risco sistémico, que abrange apenas os "incidentes graves". Por outro lado, os prestadores dos modelos de IA sem risco sistémico estão igualmente obrigados a comunicar os incidentes com dados pessoais de acordo com o RGPD.

Como sucede em outras regulamentações de produtos técnicos, também no RIA a normalização é considerada como a via mais apropriada para assegurar o respeito pelas obrigações instituídas (considerando 121)[152]. Este processo de normalização apresenta especificidades significativas em relação aos objetivos habituais, pois, além da normalização técnica essencialmente vocacionada para a segurança e eficiência do funcionamento das ferramentas tecnológicas, propõe-se, pela primeira vez, o desenvolvimento de normas europeias harmonizadas (*standards*) dirigidas à proteção de bens intangíveis, nomeadamente os direitos fundamentais ou os valores e princípios das sociedades democráticas, ou seja, o

[148] Não pode ainda ignorar-se a articulação com outros normativos, nomeadamente o Regulamento (UE) 2022/2065 (Regulamento dos Serviços Digitais ou RSA). Também neste há uma preocupação central com os "riscos sistémicos" e o RIA considera que o cumprimento das obrigações previstas no RSA faz presumir o cumprimento das obrigações inscritas naquele, "a menos que surjam e sejam identificados nesses modelos riscos sistémicos significativos não abrangidos pelo Regulamento (UE) 2022/2065" (considerando 118).
[149] O Anexo XI contém a documentação técnica a apresentar pelos prestadores de modelos de IA de finalidade geral, relativa às obrigações previstas no artigo 53.º, n.º 1, alínea a). O Anexo XII indica a documentação técnica destinada aos prestadores a jusante que integrem o modelo de IA de finalidade geral no seu sistema de IA.
[150] Como se refere no considerando 133, os métodos e técnicas a utilizar para marcação e deteção de resultados gerados ou manipulados por IA "podem ser aplicados ao nível do modelo de IA, incluindo modelos de IA de finalidade geral que geram conteúdos, facilitando assim o cumprimento desta obrigação pelo prestador a jusante do sistema de IA".
[151] O CEDP considera que as medidas de mitigação que preconiza são salvaguardas adicionais às medidas exigidas para cumprir as regras do RGPD. *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., pp. 28-29.
[152] O artigo 40.º, n.º 3 incentiva o reforço da cooperação mundial e a governação multilateral em matéria de normalização.

alinhamento com os valores e princípios da União. A normalização deve ainda incidir sobre a dimensão da gestão, incluindo os processos de comunicação e documentação. Também o desempenho dos sistemas de IA no que respeita à utilização de recursos durante o seu ciclo de vida, como o consumo de energia e de outros recursos, é objeto de atenção no processo de normalização.

Os *standards* não são obrigatórios, mas, segundo o artigo 53.º, n.º 4 e 55.º, n.º 2, o cumprimento das normas harmonizadas ou dos produtos de normalização que tenham sido adotados pela Comissão e publicados no Jornal Oficial da UE confere uma presunção de conformidade com os requisitos impostos aos sistemas e modelos de IA abrangidos nessas normas[153]. Em 22 de maio de 2023, a Comissão Europeia solicitou aos organismos de normalização europeus CEN e CENELEC a elaboração de *standards* para a aplicação do RIA, mas não especificamente para os modelos de IA de finalidade geral[154]. Esta decisão foi substituída por outra, em 2025, que também não incluiu os modelos[155].

Cabe à Comissão Europeia, através de atos delegados, efetuar a especificação das metodologias de medição e cálculo, nomeadamente dos recursos computacionais utilizados para treinar o modelo, o tempo de treino e outros dados pertinentes e o consumo de energia do modelo, com vista a permitir documentação comparável e verificável[156]. Os Anexos XI e XII podem ser igualmente modificados pela Comissão, para garantir o acompanhamento da evolução tecnológica (artigo 53.º, n.º 6). O uso de conceitos vagos e indeterminados, associada à constante inovação tecnológica, acarreta incerteza. Dada a complexidade e o desenvolvimento recente da IA, também as definições e conceitos técnicos e jurídicos suscitam dúvidas e necessitam de clarificação complementar, incluindo através de exemplos. A *soft law*, nomeadamente as orientações que compete à Comissão adotar (artigo 96.º), tem esse intuito clarificador[157].

Até à aprovação de *standards*, o Regulamento confere um relevo primacial à subscrição pelos interessados de códigos de práticas, previstos no artigo 56.º. O considerando 117 qualifica os códigos de práticas como o "instrumento central" de *compliance* para os prestadores. Além de contribuírem para a elaboração de uma taxonomia do tipo e da natureza dos riscos sistémicos

[153] De acordo com o Regulamento (UE) n.º 1025/2012.
[154] A solicitação abrangeu 10 domínios: gestão de risco, governação e qualidade de dados, manutenção de registos, transparência, supervisão humana, exatidão, robustez, cibersegurança, gestão de qualidade, avaliação de conformidade.
[155] Implementing Decision C(2025) 3871 final, Brussels, 23.06.2025. O CEN e o CENELEC instituíram o Joint Technical Committee 21 (JTC 21), dedicado ao desenvolvimento de *standards* para a IA, incluindo para a aplicação do RIA. Este é um processo complexo e verifica-se uma demora significativa. Sobre o desenvolvimento do trabalho do JTC 21, *vide* https://jtc21.eu/about/ (02.06.2026). Em julho de 2025, o Corporate Europe Observatory apresentou uma queixa à Provedora de Justiça Europeia, acusando os organismos europeus de estandardização de falta de transparência, inclusão e responsabilização na condução do processo, nomeadamente a não publicitação dos membros do Comité Técnico Conjunto CEN-CENELEC para a IA (JTC 21), das atas e dos documentos. A Provedora abriu um inquérito e dirigiu à Comissão um conjunto de questões. Sobre o inquérito, veja-se EUROPEAN OMBUDSMAN, "How the European Commission ensures transparency, inclusiveness and accountability in the adoption of harmonised standards related to artificial intelligence. Case 1974/2025/MIK", 26.09.2025, https://www.ombudsman.europa.eu/en/opening-summary/en/212090#_ftn4 (02.06.2026).
[156] Artigo 53.º, n.º 5.
[157] Em 18 de julho de 2025, foram publicadas essas orientações. A versão portuguesa foi publicada em 19.11.2025. COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, cit.

a nível da União, incluindo das suas fontes, centram-se na avaliação específica dos riscos e nas medidas de mitigação (considerando 116), definindo compromissos e meios a utilizar pelos prestadores. A Comissão Europeia, através do Serviço Europeu para a IA (também AI Office), promoveu a elaboração do Código de Práticas para a IA de finalidade geral (também *General-Purpose AI Code of Practice*)[158], entre setembro de 2024 e abril de 2025. Foi publicado em 10 de julho de 2025. Para apoiar a elaboração deste, como facultado pelo artigo 56.º, n.º 3, o Serviço Europeu para a IA convidou representantes de organizações da sociedade civil, da indústria, do meio académico e peritos independentes, seguindo um modelo de *governance* "*WoS*", tendo em conta "as necessidades e os interesses de todas as partes interessadas, incluindo as pessoas afetadas" (n.º 4)[159]. A respetiva aprovação pela Comissão Europeia[160], em 1 de agosto de 2025, conferiu validade ao código em todo o território da UE como um instrumento voluntário para os prestadores demonstrarem o cumprimento do RIA. Nos termos dos artigos 53.º, n.º 4 e 55.º, n.º 2, o seu cumprimento (total ou parcial) demonstra o cumprimento pelos prestadores das obrigações previstas nos artigos 53.º, n.º 1 e 55.º, n.º 1 até que sejam adotadas normas harmonizadas[161].

O regime de cumprimento das obrigações deverá ainda adequar-se ao tipo de prestadores. Além da isenção conferida ao desenvolvimento e utilização de modelos para fins de investigação não profissional ou científica[162], deverá ter-se em conta "a dimensão do prestador e permitir formas simplificadas de conformidade para as PME, incluindo as empresas em fase de arranque, que não deverão representar um custo excessivo nem desencorajar a utilização de tais modelos", através de formas simplificadas de *compliance* (considerando 109 e Capítulo 3 do Código de Práticas).

A secção 5 do capítulo IX (artigos 88.º a 94.º) tem como epígrafe "Supervisão, investigação, execução e controlo no que respeita a prestadores de modelos de IA de finalidade geral" e atribui tais competências em exclusivo à Comissão Europeia[163], Cabe ao Serviço Europeu para a IA o seu desempenho de acordo com um modelo colaborativo[164], acompanhando a execução

[158] COMISSÃO EUROPEIA, *General-Purpose AI Code of Practice*, July 10, 2025. As suas três partes estão disponíveis em https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai (02.06.2026).
[159] Houve críticas ao Código de Práticas, em especial no que respeita aos direitos de autor, como foi comunicado à Vice-Presidente da Comissão Europeia por 15 organizações, em 27 de janeiro de 2025. AAVV, *Open Letter*, 27.01.2025, https://www.newsmediaeurope.eu/wp-content/uploads/2025/01/Right-holders-joint-letter-on-the-Second-Draft-Code-of-Practice-27012025.pdf (02.06.2026).
[160] COMISSÃO EUROPEIA, *Comission Opinion of 1.8.2025 on the assessment of the General-Purpose AI Code of Practice within the meaning of Article 56 of Regulation (EU) 2024/1689*, C(2025) 5361 final, Brussels, 1.08.2025, https://digital-strategy.ec.europa.eu/en/library/commission-opinion-assessment-general-purpose-ai-code-practice (02.06.2026). O Conselho da IA (também AI Board) também se pronunciou e aprovou o Código de Práticas. AI BOARD, *Conclusion of the Artificial Intelligence Board on the Assessment of the General-Purpose AI Code of Practice pursuant to Article 56 of Regulation 2024/1689 (Artificial Intelligence Act or "AI Act")*, https://digital-strategy.ec.europa.eu/en/policies/ai-board (02.06.2026).
[161] Os prestadores que não cumpram o código de práticas aprovado ou a norma europeia harmonizada "devem demonstrar meios de conformidade alternativos" (artigos 53.º, n.º 4 e 55.º, n.º 2).
[162] Ainda que se incentive o seu cumprimento voluntário.
[163] O artigo 53.º, n.º 3 prescreve a cooperação "na medida do necessário" dos prestadores dos modelos de IA de finalidade geral para o desempenho dessas competências.
[164] Se o Serviço Europeu para a IA se compromete com uma abordagem colaborativa, de cooperação e apoio aos prestadores desde o treino, em especial no caso dos modelos de IA de finalidade geral com risco sistémico, espera-se igualmente uma atitude proativa e colaborativa, incluindo nas fases de investigação e "*enforcement*" (*v.g.*, através de "diálogos estruturados"). COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, p. 29.

e o cumprimento das obrigações dos prestadores dos modelos[165]. Pode solicitar informações e a documentação das medidas, pode realizar avaliações e prescrever aos prestadores a adoção de medidas, incluindo "medidas de redução dos riscos em caso de riscos sistémicos identificados, bem como através da restrição da disponibilização no mercado, da retirada ou da recolha do modelo" (considerando 164)[166]. A Comissão Europeia é competente para aplicar coimas aos prestadores dos modelos de IA de finalidade geral pelo desrespeito das obrigações previstas no RIA, pelo incumprimento das medidas definidas, pela não prestação ou prestação incorreta de informações, bem como a não disponibilização do acesso ao modelo para a respetiva avaliação. As coimas devem ser proporcionadas e dissuasivas e devem ter em conta a subscrição dos compromissos dos códigos de práticas. O Tribunal de Justiça da União Europeia (TJUE) é competente para os recursos dessas decisões e tem plena jurisdição (artigo 101.º). O RIA admitiu o envolvimento das autoridades nacionais de fiscalização do mercado, competentes para fiscalizar os sistemas de IA de risco elevado, na medida em que o acesso a informação sobre um modelo de IA de finalidade geral se revele necessário para avaliar o cumprimento das obrigações do sistema de IA de risco elevado que integra esse modelo e o respetivo prestador é diferente do prestador do modelo (artigo 75.º, n.º 3)[167].

Segundo o artigo 113.º, al. b), as normas do capítulo V são aplicáveis desde 2 de agosto de 2025. Contudo, no que respeita aos modelos de IA de finalidade geral que tenham sido colocados no mercado antes dessa data, o artigo 111.º, n.º 3 dá aos prestadores um prazo até 2 de agosto de 2027 para adotarem as medidas necessárias para o cumprimento das respetivas obrigações. Nos termos do artigo 113.º, as disposições relativas ao exercício das competências de supervisão e fiscalização são aplicáveis a partir de 2 de agosto de 2026.

4.2. Obrigações dos prestadores de modelos de finalidade geral

A secção 2 do capítulo V (artigos 53.º e 54.º) contém as obrigações aplicáveis aos prestadores de todos os modelos de IA de finalidade geral. A Secção 3 (artigo 55.º) contém as obrigações que os prestadores de modelos de IA de finalidade geral com risco sistémico devem cumprir para além das anteriores.

As obrigações preconizadas pelo RIA para todos os prestadores de modelos de IA de finalidade geral reconduzem-se essencialmente a um regime de transparência, com obrigações de documentação e informação com atualização permanente, a fim de assegurar a auditabilidade, a explicabilidade e a responsabilidade dos prestadores, de modo a assegurar confiança e

[165] Tais competências e a "interação colaborativa" são objeto de clarificação nas orientações adotadas em 18 de julho de 2025. COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, cit., pp. 28-29.

[166] Deve ainda ser referido o papel de apoio do Painel Científico, nomeadamente com o alerta para riscos sistémicos (artigo 90.º).

[167] Quando o prestador é o mesmo, o Serviço Europeu para a IA tem as competências conferidas às autoridades nacionais de fiscalização do mercado (artigo 75.º, n.º 1). A proposta "*Digital omnibus on AI*" expande a competência de supervisão da Comissão Europeia/Serviço Europeu para a IA.

segurança. Essas obrigações são relevantes para a proteção dos dados pessoais, nomeadamente as informações sobre os dados utilizados no treino do modelo, ainda que o regime esteja permeado pela proporcionalidade e pela proteção dos interesses legítimos dos prestadores e da inovação, nomeadamente a propriedade intelectual e os segredos comerciais.

Essas obrigações estão inscritas sobretudo no artigo 53.º, n.º 1, als. a), b) e d)[168]. Uma parte dessas obrigações respeita a informação que deve ser facultada ao Serviço Europeu para a IA e às autoridades nacionais competentes que o solicitem, a fim de facilitar o exercício das respetivas competências de controlo e fiscalização. Assim, o prestador do modelo fica obrigado a elaborar e manter atualizada a documentação técnica do modelo, incluindo o seu processo de treino e de testagem e os resultados da sua avaliação (artigo 53.º, n.º 1, al. a)). A informação mínima que essa documentação tem de conter é elencada na Secção 1 do Anexo XI e, na verdade, é extensa e com elevado grau de pormenor. Têm especial relevância para a proteção de dados pessoais as obrigações elencadas no ponto 2, al. c) do Anexo[169]: "informações sobre os dados utilizados para treino, testagem e validação, se aplicável, incluindo o tipo e a proveniência dos dados e as metodologias de curadoria (por exemplo, limpeza, filtragem, etc.), o número de pontos de dados, o seu âmbito e as principais características; a forma como os dados foram obtidos e selecionados, bem como todas as outras medidas para detetar a inadequação das fontes de dados e dos métodos para detetar enviesamentos identificáveis, se aplicável". Ainda assim, a incerteza não é eliminada, em virtude da previsão da sua adaptação à "dimensão" e ao "perfil de risco" do modelo, de que resultará a determinação das concretas obrigações aplicáveis a cada modelo. O envolvimento de dados pessoais deve ser considerado um elemento que, evidentemente, eleva o perfil de risco do modelo.

O RIA atribui aos prestadores de modelos de IA de finalidade geral "uma responsabilidade especial na cadeia de valor da IA" (considerando 101). A al. b) do n.º 1 do artigo 53.º prescreve obrigações de informação e documentação atualizadas a fornecer aos prestadores de sistemas de IA que pretendam integrar o modelo nesses sistemas, a fim de permitir àqueles uma boa compreensão das capacidades e limitações do modelo de IA de finalidade geral, seja para assegurar o devido funcionamento das suas ferramentas, seja para que possam cumprir as

[168] A alínea c) respeita à obrigação de implementação de uma política de governação de dados que respeite o direito europeu de proteção dos direitos de autor e direitos conexos, incluindo as técnicas para assegurar a reserva prevista no artigo 4.º, n.º 3 da Diretiva (UE) 2019/790.

[169] O ponto 1 do Anexo impõe uma descrição geral do modelo, através da informação das tarefas que o modelo se destina a desempenhar e o tipo e natureza dos sistemas de IA em que pode ser integrado (al. a)), das políticas de utilização aceitáveis (al. b)), da data de lançamento e os métodos de distribuição (al. c)), da arquitetura e o número de parâmetros (al. d)), da modalidade (por exemplo, texto, imagem) e formato das entradas e saídas (al. e), a licença (al. f)). Além das obrigações respeitantes aos dados, o ponto 2 do Anexo aponta outras informações que devem ser disponibilizadas e que pormenorizam os elementos referidos no ponto 1 e o processo de desenvolvimento do modelo, incluindo os meios técnicos (por exemplo, instruções de utilização, infraestruturas, ferramentas) necessários para que o modelo de IA de finalidade geral seja integrado nos sistemas de IA (al. a)), as especificações de conceção do modelo e do processo de treino, incluindo as metodologias e técnicas de treino, as principais opções de conceção, incluindo a fundamentação e os pressupostos assumidos; o que se pretende otimizar com o modelo e a importância dos diferentes parâmetros, se aplicável (al. b)); os recursos computacionais utilizados para treinar o modelo (por exemplo, o número de operações de vírgula flutuante), o tempo de treino e outros dados pertinentes relacionados com a formação (al. d)), o consumo de energia conhecido ou estimado do modelo ou, se tal for desconhecido, informação sobre os recursos computacionais utilizados (al. e)).

suas obrigações[170]. Salvaguarda-se a proteção dos direitos de propriedade intelectual e as informações comerciais de carácter confidencial ou segredos comerciais. O anexo XII apresenta o elenco mínimo de elementos a incluir nessa documentação[171], verificando-se que alguns são idênticos aos que devem ser disponibilizados ao Serviço Europeu para a IA e às autoridades nacionais competentes. Novamente, algumas dessas obrigações incidem sobre os dados, nomeadamente a prestação de informações sobre os dados utilizados para o treino, a testagem e a validação, se aplicável, incluindo o tipo e a proveniência dos dados e as metodologias de curadoria (al. c) do ponto 2 do Anexo).

Ficam isentos dessas obrigações os modelos de IA de finalidade geral lançados ao abrigo de licenças gratuitas e de fonte aberta (artigo 53.º, n.º 2), pois considera-se que "asseguram elevados níveis de transparência e abertura se os seus parâmetros, incluindo as ponderações, as informações sobre a arquitetura do modelo e as informações sobre a utilização do modelo, forem disponibilizados ao público", desse modo contribuindo "para a investigação e a inovação no mercado" e proporcionando "oportunidades de crescimento significativas para a economia da União" (considerando 102)[172]. Estes modelos lançados ao abrigo de licenças gratuitas e de fonte aberta suscitam problemas específicos relativamente ao respeito pela proteção dos dados pessoais. Exige-se que sejam "transparentes", ou seja, que seja disponibilizada publicamente informação com "clareza e especificidade" sobre as modalidades de *input* do modelo. Ora, estas podem comportar *datasets* que incluam dados pessoais, que assim podem tornar-se públicos e publicamente acessíveis. É absolutamente fundamental que no treinamento e funcionamento do modelo esteja salvaguardada a anonimização e/ou filtros de *input/output*. Ainda no que respeita aos dados pessoais, estes não podem ser moeda de troca para o acesso e utilização do modelo. Quando sejam exigidos, considera-se que há monetização e a isenção não se aplica, exceto quando o acesso e o processamento de dados pessoais sejam exclusiva e estritamente destinados a garantir a segurança do modelo e não ganhos financeiros ou

[170] Essa informação permitirá aos operadores a jusante confirmarem a diversidade dos dados utilizados, facilitando a adoção de medidas de mitigação para prevenir vieses. COMISSÃO EUROPEIA, *Comunicação da Comissão Nota explicativa e modelo para o resumo público dos conteúdos utilizados para o treino de modelos de IA de finalidade geral exigido pelo artigo 53.º, n.º 1, alínea d), do Regulamento (UE) 2024/1689 (Regulamento da Inteligência Artificial)*, C(2025) 8311 final, Bruxelas, 5.12.2025, https://digital-strategy.ec.europa.eu/en/library/explanatory-notice-and-template-public-summary-training-content-general-purpose-ai-models (02.06.2026).

[171] O ponto 1 do Anexo impõe também uma descrição geral do modelo, através da informação das tarefas que o modelo se destina a desempenhar e o tipo e a natureza dos sistemas de IA em que pode ser integrado (al. a)), as políticas de utilização aceitáveis (al. b)), a data de lançamento e os métodos de distribuição (al. c)), de que forma o modelo interage ou pode ser utilizado para interagir com hardware ou software que não faça parte do próprio modelo, quando aplicável (al. d)), as versões do software pertinente relacionadas com a utilização do modelo de IA de finalidade geral, se aplicável (al. e)), a arquitetura e o número de parâmetros (al. f)), a modalidade (por exemplo, texto, imagem) e o formato das entradas e saídas (al. g)), a licença para o modelo (al. h)). O ponto 2 do Anexo impõe alguns elementos técnicos mais específicos, incluindo relativos ao desenvolvimento do modelo, nomeadamente, os meios técnicos (por exemplo, instruções de utilização, infraestruturas, ferramentas) necessários para que o modelo de IA de finalidade geral seja integrado nos sistemas de IA (al. a)), a modalidade (por exemplo, texto, imagem, etc.) e o formato das entradas e saídas e a sua dimensão máxima (por exemplo, comprimento da janela de contexto, etc.) (al. b)), além das informações referidas sobre os dados previstas na al. c).

[172] COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, cit., pp. 22-26.

comerciais e sejam efetuados com respeito pelos normativos europeus de proteção de dados pessoais[173].

Essas isenções para os modelos lançados ao abrigo de licenças gratuitas e de fonte aberta não se aplicam quando se trate de modelos com risco sistémico.

Num setor dominado por um punhado de grandes empresas não europeias rodeado por uma imensidão de PME e *start-ups* inovadoras que a jusante alteram e aperfeiçoam os modelos de IA de finalidade geral, o legislador europeu está consciente de que a proteção da capacidade concorrencial destas últimas impõe uma adaptação do regime previsto, aligeirando as exigências previstas. Assim, as suas obrigações "deverão limitar-se a essa alteração ou aperfeiçoamento, por exemplo, complementando a documentação técnica já existente com informações sobre as alterações, nomeadamente novas fontes de dados de treino"[174], desse modo distribuindo as obrigações ao longo da cadeia de valor. Há ainda exceções para transações entre microempresas.

A obrigação de disponibilização desses elementos é, com certeza, relevante para a proteção de dados pessoais, nomeadamente a informação sobre os dados utilizados no treino, teste e validação do modelo, indicada na al. c) do ponto 2 do Anexo XI. Contudo, estando o acesso a esta informação previsto apenas para o Serviço Europeu para a IA e para as autoridades nacionais competentes, será mais diretamente relevante para os titulares dos dados o Resumo "suficientemente pormenorizado sobre os conteúdos utilizados para o treino do modelo de IA de finalidade geral", dada a obrigação da sua disponibilização pública imposta pelo artigo 53.º, n.º 1, al. d), incluindo por parte dos prestadores de modelos lançados ao abrigo de licenças gratuitas e de fonte aberta. Esse resumo deve ser disponibilizado igualmente pelos modelos que sejam um desenvolvimento (*finetuning*/*taylorização*) de um outro modelo.

A principal finalidade desse resumo é "facilitar às partes com interesses legítimos (…) o exercício e a aplicação dos seus direitos ao abrigo do direito da União" (considerando 107). Ainda que apenas refira expressamente os direitos de autor e sejam estes a principal preocupação, evidentemente os titulares de dados pessoais são "partes com interesses legítimos" para quem esse resumo é importante. A Nota Explicativa da Comissão Europeia esclarece que a disponibilização da informação do Resumo não exime do cumprimento das obrigações de informação que resultam do regime europeu de proteção de dados pessoais[175]. Foi disponibilizado um formulário para facilitar aos prestadores a elaboração do resumo[176].

---

[173] COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, cit., pp. 25-26. As orientações relativas à monetização não se aplicam às transações entre microempresas (p. 26).

[174] Considerando 109.

[175] COMISSÃO EUROPEIA, *Comunicação da Comissão Nota explicativa e modelo para o resumo público dos conteúdos utilizados para o treino de modelos de IA de finalidade geral exigido pelo artigo 53.º, n.º 1, alínea d), do Regulamento (UE) 2024/1689 (Regulamento da Inteligência Artificial)*, cit., p. 3.

[176] Esse formulário foi elaborado na sequência de uma consulta *multi-stakeholder* entre 30 de julho e 18 de setembro de 2024, e do *input* dos participantes na preparação do Código de Práticas. Foi igualmente discutido com o PE e o Conselho da IA. COMISSÃO EUROPEIA, *Comunicação da Comissão Nota explicativa e modelo para o resumo público dos conteúdos utilizados para o treino de modelos de IA de finalidade geral exigido pelo artigo 53.º, n.º 1, alínea d), do Regulamento (UE) 2024/1689 (Regulamento da Inteligência Artificial)*, cit..

Propugna-se um equilíbrio na informação a disponibilizar entre a amplitude e a precisão da informação. O resumo dos conteúdos utilizados para o treino do modelo de IA de finalidade geral deverá ser "suficientemente pormenorizado" e "abrangente no seu âmbito de aplicação", mas não "tecnicamente pormenorizado" (considerando 107). De acordo com a nota explicativa e o formulário disponibilizado pela Comissão Europeia, as informações a divulgar no Resumo sobre as fontes de informação e os dados utilizados no treino do modelo são numerosas. Incluem-se as línguas dos dados recolhidos, a indicação da última data de recolha dos dados e se o modelo é continuamente treinado com dados recolhidos após essa data ou dados dinâmicos. Também se devem ser publicitadas as modalidades (texto, imagem, áudio, vídeo, outra), a dimensão (*tokens* de texto; quantidade de imagens; horas de áudio e vídeo) e a natureza dos conteúdos utilizados no treino do modelo (v.g., textos de ficção, científicos, jornalísticos, etc.; fotografias, logos, etc.; gravações, audiolivros, programas de rádio; comunicações privadas em áudio; vídeos de música, filmes, programas de TV, jogos de vídeo, vídeos das redes sociais, etc.), bem como outras especificidades dos dados utilizados, nomeadamente regionais, nacionais ou demográficas. Muito relevante para a proteção de dados pessoais são as fontes dos dados recolhidos que devem ser indicadas no resumo, incluindo a descrição das respetivas modalidades, dimensão e conteúdos, sejam elas *datasets* disponíveis publicamente (v.g., em repositórios), *datasets* privados não disponíveis publicamente, obtidas diretamente dos titulares dos dados ou seus representantes ou não, com indicação da forma de obtenção (v.g., licenciamento comercial). Quando os dados sejam recolhidos através de *web crawling* ou *web scraping*, seja diretamente ou através de terceiros, deve ser indicado o período de recolha, as condições do *crawler* (v.g., respeito por *captchas*, robot.txt ou outros protocolos), as respetivas modalidades (texto, imagens, vídeo, etc.), enumerando os principais nomes de domínio utilizados (v.g., "*exemplo.com*"). Recorda-se que, ainda que estejam publicados, a reutilização de dados pessoais para treino de modelos pode não ser legítima[177]. Também deve ser informado se há utilização de dados (incluindo as respetivas modalidades) dos utilizadores obtidos através da interação com o modelo, sejam os *user inputs* e *prompts*, ou de interação com outros serviços ou produtos do prestador (v.g., redes sociais, plataformas e outros interfaces). Deve ser ainda incluída no resumo informação sobre a utilização de dados sintéticos gerados por modelos de IA e provenientes de outras fontes, incluindo *offline*, das quais deve ser incluída uma descrição narrativa. Finalmente, o resumo deve conter informação sobre as medidas postas em prática para assegurar o respeito pelas exceções e limitações ao *TDM* e a remoção de conteúdos ilegais. O formulário contém campos opcionais para a inclusão de outra informação relevante.

Apesar deste Resumo ser um instrumento importante para a salvaguarda dos direitos dos titulares de dados pessoais, é pertinente a dúvida sobre a sua adequação, nomeadamente a disponibilização de informação genérica sobre dados recolhidos através de *web scraping*. Mesmo quando se prevê que deva ser providenciada pelo Resumo uma descrição narrativa do

[177] EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, cit., p. 7; Proc. C-252/21, cit., § 77. O formulário da Comissão salvaguarda essa situação.

tipo de conteúdo e das fontes[178], é praticamente impossível aos indivíduos comuns controlarem efetivamente a presença de informação pessoal sua nos milhões de páginas da internet, incluindo categorias de dados especiais (v.g., imagens faciais). Resta o facto de a disponibilização desta informação não constituir a presunção ou sequer a demonstração do cumprimento das obrigações resultantes dos normativos de proteção dos dados pessoais, ainda que a previsão do fundamento do interesse legítimo na proposta da Comissão ampare o interesse dos prestadores dos modelos.

O CEDP vai mais longe e propõe a publicação nos relatórios anuais de transparência do *balancing test* que fundamenta o interesse legítimo do processamento de dados pessoais[179]. Um outro instrumento para tornar a informação mais acessível e compreensível são os "*model cards*"[180] a disponibilizar juntamente com os modelos.

4.3. Obrigações dos prestadores de modelos de IA de finalidade geral com risco sistémico

Às obrigações referidas no ponto anterior acrescem obrigações suplementares para os prestadores de modelos de IA de finalidade geral com risco sistémico, com o objetivo de prevenir e mitigar esse risco.

O Preâmbulo aponta a implementação de uma política de mitigação de riscos iterativa, com "processos de responsabilização e governação, executando o acompanhamento pós-comercialização, adotando medidas adequadas ao longo de todo o ciclo de vida do modelo e cooperando com os intervenientes pertinentes ao longo de toda a cadeia de valor da IA" (considerando 114). O quadro de obrigações especiais é denso e está inscrito na Secção terceira do Capítulo V, artigo 55.º. Essas obrigações têm como objetivo identificar, prevenir e mitigar os riscos sistémicos e ainda assegurar um nível adequado de cibersegurança e segurança física com medidas específicas que previnam riscos acidentais (lançamento, acesso ou fuga não autorizados de dados) e também a subtração do modelo e a sua "*weaponization*" maliciosa. São aplicáveis desde a sua construção e ao longo de todo o ciclo de vida.

As quatro alíneas do artigo 55.º, n.º 1 enunciam as obrigações dos prestadores. Em primeiro lugar, realizar avaliações internas ou externas independentes, *ex ante* (anteriores à sua colocação no mercado), incluindo testagem antagónica e respetiva documentação, a fim de identificar e mitigar os riscos sistémicos (al.a)). Devem implementar uma política preventiva e corretiva de gestão de riscos sistémicos e suas fontes, com avaliação e mitigação iterativa ao nível da União, incluindo processos de *governance* e *accountability* e acompanhamento ao

[178] São referidas as características geográficas, linguísticas ou demográficas, e o tipo de *websites* (v.g., notícias, *blogs*, redes sociais, websites, plataformas de jogo, de *streaming*, portais governamentais e websites culturais, educacionais, etc.).
[179] *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, cit., pp. 29-30. Refere ainda os selos de transparência.
[180] HUGGING FACE, "Model Cards", cit..

longo de todo o ciclo de vida do modelo e da cadeia de valor pós-comercialização, pois esses riscos podem advir do desenvolvimento, da colocação no mercado ou da utilização do modelo (al. b)). No caso de ocorrência de um incidente grave no desenvolvimento ou utilização do modelo, o prestador tem a obrigação de acompanhamento, documentação e comunicação de informações e eventuais medidas corretivas ao Serviço Europeu para a IA e às autoridades nacionais competentes (al. c)). Recorde-se a diferença em relação ao RGPD, em que não obrigações semelhantes para o controlador dos dados se aplica aos incidentes, em geral, prevendo a proposta "Digital *omnibus*" uma modificação que aproxima esse regime do previsto no artigo 55.º do RIA. Finalmente, a cibersegurança do modelo e da respetiva infraestrutura física devem ter um "nível adequado" (al.d)).

A Secção 2 do Anexo XI precisa as obrigações de informação especiais que os prestadores de modelos de IA de finalidade geral com risco sistémico devem cumprir. Inclui descrições pormenorizadas (1) das estratégias de avaliação, incluindo os critérios de avaliação, parâmetros, a metodologia de identificação de limitações e os resultados, (2) das medidas adotadas para efeitos de realização de testagens antagónicas internas e/ou externas (por exemplo, simulação de ataques), adaptações do modelo, incluindo alinhamento e ajustes, e, eventualmente, (3) da arquitetura do sistema, explicando de que forma os componentes de *software* se utilizam ou se alimentam mutuamente e como se integram no processamento global.

Mais uma vez, salienta-se a normalização, através da referência ao uso de protocolos e instrumentos normalizados que reflitam o estado da arte, como inscrito na al. a) do n.º 1 do artigo 55.º, e o reconhecimento da importância da elaboração de códigos de práticas para o cumprimento destas obrigações, em especial na ausência de norma harmonizada. Por outro lado, caso não respeite o *standard* aplicável ou o código de práticas, o prestador dos modelos de AI de finalidade geral com risco sistémico deve demostrar a utilização de meios de conformidade alternativos adequados para o cumprimento das obrigações impostas (artigo 55.º, n.º 2).

4.4. O Código de Práticas para os Modelos de IA de finalidade geral

Consciente das dificuldades do processo de normalização, o considerando 116 do RIA dispõe que o Serviço Europeu para a IA, em colaboração com as autoridades nacionais competentes, deverá incentivar e facilitar a elaboração de códigos de práticas para as obrigações dos prestadores de modelos de IA de finalidade geral e de modelos de IA de finalidade geral com riscos sistémicos. Parte-se da premissa de que só os modelos mais avançados terão riscos sistémicos e que estes serão poucos, ainda que estes sejam um dos principais focos de atenção, nomeadamente a sua avaliação e mitigação. Prevê-se mesmo a elaboração uma taxonomia do tipo e da respetiva natureza desses riscos, incluindo as suas fontes.

O artigo 56.º determina que os códigos de práticas incidam, no mínimo, sobre as obrigações previstas nos artigos 53.º a 55.º, indicando desde logo os seguintes elementos: a) Os meios para assegurar que as informações contidas no artigo 53.º, n.º 1, alíneas a) e b), sejam mantidas atualizadas à luz da evolução tecnológica e do mercado; b) O nível de pormenor adequado para o resumo dos conteúdos utilizados no treino; c) A identificação do tipo e da natureza dos riscos sistémicos ao nível da União, incluindo as respetivas fontes, se for caso disso; d) As medidas, procedimentos e modalidades de avaliação e gestão dos riscos sistémicos a nível da União, incluindo a respetiva documentação, que devem ser proporcionados em relação aos riscos, ter em conta a sua gravidade e probabilidade e ter em conta os desafios específicos da resposta a esses riscos à luz das possíveis formas como podem surgir e materializar-se ao longo da cadeia de valor da IA. O Código de Práticas não impõe mais obrigações, antes precisa e orienta a operacionalização das obrigações contidas no RIA.

No processo de elaboração do código de práticas, iniciado em 30 de setembro de 2024, foram envolvidos cerca de 1000 participantes, entre os quais os prestadores de modelos de IA de finalidade geral, além de organizações da sociedade civil e outras partes interessadas, bem como peritos, incluindo da academia. Esta abordagem colaborativa permite acompanhar o estado da arte e integra um conjunto diversificado de perspetivas. Com o contributo dos participantes, o documento foi evoluindo através de três *drafts*, preparados por um comité de 13 peritos independentes e submetidos a comentários escritos e discussão oral com todos os participantes. O Código de Práticas sobre IA de finalidade geral foi publicado em 10 de julho de 2025.

As vantagens apontadas são a certeza jurídica e a redução dos encargos. A adesão dos prestadores de modelos de IA de finalidade geral ao Código de Práticas, total ou parcial, constitui a demonstração de cumprimento das obrigações previstas no regulamento, na medida que essas normas abranjam essas obrigações. Os não subscritores deverão comprovar o cumprimento das obrigações através de outros meios adequados (artigos 53.º, n.º 4, e 55.º, n.º 2). Entre os subscritores encontram-se alguns dos prestadores mais relevantes de modelos de IA de finalidade geral, mas também há ausências notáveis (*v.g.*, a Meta) e adesões com restrições (a xAI apenas aderiu ao capítulo "*Safety and Security*").

O Código é uma "ponte" até à adoção de *standards*[181] e não tem vocação totalizante, pois remete em vários aspetos para atos complementares, nomeadamente as Orientações da Comissão. O Código propõe-se ser "*future-proof*", apto a uma atualização rápida, segundo uma lógica colaborativa, com remissão para bases de dados dinâmicas de incidentes[182], a partilha de boas práticas e a participação de todos os interessados. Sempre que haja modificações relevantes na informação disponibilizada, nomeadamente em novas versões dos

[181] Al. g) do Preâmbulo do terceiro *Draft*. COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, cit., p. 28.

[182] À semelhança da aviação, por exemplo. Sobre os exemplos de resposta regulatória e de *governance* de inovações tecnológicas com grande impacto societal, como a aviação ou a energia nuclear, AKASH R. WASIL et al., "Affirmative safety: An approach to risk management for high-risk AI", 24.04.2024, https://arxiv.org/pdf/2406.15371 (02.06.2026).

modelos, a documentação deve ser atualizada[183] e a informação a disponibilizar ao Serviço Europeu para a IA, autoridades nacionais competentes e prestadores a jusante deve ser a mais atual[184]. O Serviço Europeu para a IA promoverá a sua revisão a cada dois anos, que será submetida a nova adesão voluntária dos prestadores[185].

Os objetivos gerais do Código de Práticas acompanham os do RIA: melhorar o funcionamento do mercado interno, promover uma IA humana e de confiança, que assegure um elevado nível de proteção da saúde, segurança e direitos fundamentais, democracia, estado de direito e do ambiente, bem como apoiar a inovação. Os objetivos específicos indicados são dois[186]: o primeiro é servir de guia para demonstrar a conformidade com as obrigações dos artigos 53.º e 55.º do RIA. O regulamento distingue este efeito da presunção de cumprimento, que prevê para o cumprimento dos *standards*[187], mas, efetivamente, não há uma diferença prática, pois o não cumprimento de qualquer deles acarreta idêntica consequência, a obrigação de demonstração do cumprimento das obrigações do RIA através de "meios de conformidade alternativos adequados para avaliação pela Comissão" (artigos 53.º, n.º 4, e 55.º, n.º 2). O segundo objetivo é assegurar que os prestadores de modelos de IA de finalidade geral cumprem as respetivas obrigações e permitir ao Serviço Europeu para a IA avaliar esse cumprimento pelos prestadores que subscrevam o Código. Para a Comissão Europeia, a adesão ao Código "é uma forma simples de demonstrar o cumprimento (…) [e] a Comissão centrará as suas atividades de execução no controlo da respetiva observância do código de práticas"[188], pelo que o que é esperado dos prestadores é o empenho na transparência das medidas preconizadas no Código. Já os prestadores que não aderirem ao Código terão não apenas de comunicar os meios alternativos adotados, mas igualmente explicar como é que esses meios asseguram o cumprimento das obrigações impostas pelo RIA, incluindo através de uma avaliação comparativa com as medidas do Código[189]. A relevância do Código de Práticas

[183] Com obrigação de manutenção da informação anterior por um período de 10 anos após a colocação no mercado do modelo.
[184] Medida 1.2. do Capítulo "Transparência". COMISSÃO EUROPEIA, *Code of Practice for General-Purpose AI Models. Transparency Chapter*, cit., pp. 5-6.
[185] O terceiro draft do Código assumiu a respetiva precaridade, em virtude da rápida evolução deste domínio, desde logo considerando o desenvolvimento dos designados "*Agentic AI models*".
[186] COMISSÃO EUROPEIA, *Code of Practice for General-Purpose AI Models. Transparency Chapter*, cit., p. 3 (Copyright Chapter, p. 2; Safety and Security Chapter, p. 2).
[187] COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, cit., p. 28.
[188] COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, cit., p. 27.
[189] COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, cit., p. 27. A Comissão alerta imediatamente para a possibilidade de sujeitar os prestadores que não adiram ao Código de Práticas a mais pedidos de informação e a avaliações ao longo do ciclo de vida do modelo, aumentando os respetivos encargos administrativos e operacionais. Em relação aos prestadores que já tenham colocado no mercado os modelos antes de 2 de agosto de 2025, a Comissão reconhece que constitui um desafio conseguirem cumprir as respetivas obrigações até 2 de agosto de 2027, pelo que assume o compromisso de apoio por parte do Serviço Europeu para a IA (p. 30). Para os prestadores que colocarem no mercado modelos de IA de finalidade geral após 2 de agosto de 2025 e aderirem ao Código, não exigirá o cumprimento rigoroso de todas as obrigações durante o primeiro ano e também apoiará os prestadores, até porque as disposições relativas a essa competência de execução só produzirão efeito a partir de 2 de agosto de 2026 (p. 31). COMISSÃO EUROPEIA, *Questions and answers on the code of practice for General-Purpose AI*, https://digital-strategy.ec.europa.eu/en/faqs/questions-and-answers-code-practice-general-purpose-ai (02.06.2026).

também é visível no domínio das sanções, pois o artigo 101.º, n.º 1, § 2, determina que a Comissão deve ter em conta a adesão ao código de práticas na fixação do respetivo montante.

Não foram cumpridos plenamente todos objetivos inscritos no artigo 56.º. Desde logo, não foram adotados os indicadores-chave preconizados no n.º 4. O Código também não estabelece nenhum protocolo de boas práticas para a prestação de esclarecimentos ou resposta a reclamações de titulares de dados pessoais que sejam objeto de um processamento indevido, o que teria sido útil.

O Código assenta, em geral, em alguns princípios fundamentais: alinhamento com os princípios e valores da UE; alinhamento com as orientações, *standards* e métricas internacionais[190]; proporcionalidade das medidas aos riscos. No que respeita, em especial, ao terceiro capítulo do Código, "*Safety and Protection*", são assumidos mais alguns princípios.

No que respeita à proporcionalidade, encontramos as suas três dimensões (necessidade, adequação e proporcionalidade em sentido estrito). Assim, o artigo 56.º, n.º 2, al. b) refere que o Código deve determinar do "nível de pormenor adequado para o resumo dos conteúdos utilizados no treino", entre os quais se encontram com elevada certeza dados pessoais. No compromisso de transparência instituído no Código de Práticas (Capítulo 1), impõe-se a disponibilização de informação porque tal é necessário e na medida do necessário para o desempenho das competências do Serviço Europeu para a IA, das autoridades nacionais de fiscalização do mercado e/ou dos prestadores a jusante. É igualmente a proporcionalidade que é invocada quando o considerando b) assinala a limitação da responsabilidade dos prestadores a jusante de um modelo que é produto de uma modificação ou de um *fine-tuning* de um modelo de IA de finalidade geral. A subscrição do Código por parte destes prestadores consubstancia a assunção dos compromissos apenas no que respeita a essa modificação ou *fine-tuning*, não ocorrendo uma transferência do risco e das obrigações de cumprimento ao longo da cadeia. Por outro lado, se o critério é apropriado, pode não ser suficiente, nomeadamente se a modificação vem a retroalimentar o modelo original. A clarificação foi atribuída às Orientações da Comissão[191].

Já no que respeita aos modelos de IA de finalidade geral com risco sistémico, as medidas respeitantes a procedimentos e modalidades de avaliação e gestão dos riscos sistémicos a nível da União e respetiva documentação "devem ser proporcionados em relação aos riscos, ter em conta a sua gravidade e probabilidade (artigo 56.º, n.º 2, al. d)).

A proporcionalidade também está subjacente à consideração do tamanho dos modelos e na especial consideração das PME e *start-ups* (artigo 56.º, n.º 5 e considerando h) do Capítulo 3 do Código de Práticas[192]).

[190] COMISSÃO EUROPEIA, *Code of Practice for General-Purpose AI Models. Safety and Security Chapter*, cit., p. 3.
[191] Tal como previsto no artigo 96.º, n.º 1 do RIA. COMISSÃO EUROPEIA, *Code of Practice for General-Purpose AI Models. Transparency Chapter*, cit., p. 4. Estas orientações não incidem sobre questões de responsabilidade.
[192] Com informações de reporte menos exigentes. COMISSÃO EUROPEIA, *Code of Practice for General-Purpose AI Models. Safety and Security Chapter*, cit., p. 19.

O Código institui 12 compromissos, divididos em duas categorias e em três capítulos. Os dois primeiros compromissos intitulam-se "*Transparência*" e "*Copyright*"[193] e são aplicáveis a todos os prestadores de modelos de IA de finalidade geral. Os restantes 10 compromissos integram o terceiro capítulo – "*Safety and Security*" - e são aplicáveis apenas aos prestadores de modelos de IA de finalidade geral com riscos sistémicos. Evidencia-se uma densidade crescente das exigências. São o primeiro e o terceiro capítulos que têm maior importância para a proteção dos dados pessoais. O primeiro assegura o acesso à informação indispensável para a monitorização do respeito do Direito e dos direitos dos titulares na recolha e processamento de dados pessoais. O terceiro assegura a adoção de medidas de avaliação, prevenção e mitigação dos riscos sistémicos, incluindo com impacto nos dados pessoais.

O compromisso do primeiro capítulo – "Transparência" –, designado "Documentação" concretiza as obrigações previstas nos artigos 53.º, n.º 1, als. a) e b), n.º 2 e n.º 7, e nos anexos XI e XII. Desdobra-se em três medidas de documentação e de garantia da qualidade, segurança e integridade da informação relevante (Medidas 1.1. a 1.3.) e inclui em anexo um formulário modelo, que permite sistematizar a informação da Medida 1.1. a prestar a cada um dos destinatários (Serviço Europeu para a IA, autoridades nacionais competentes e prestadores a jusante). Há um incentivo à transmissão "proativa" da informação a disponibilizar aos operadores a jusante. A disponibilização da informação destinada ao Serviço Europeu para a IA ou às autoridades nacionais competentes (através daquele) só se faz após uma solicitação fundamentada, indicando a base legal e a finalidade do pedido, e restringir-se-á exclusivamente aos elementos estritamente necessários ao desempenho das tarefas fiscalização e supervisão que incumbem a essas entidades no momento do pedido[194].

No que respeita aos dados utilizados, a informação será acessível ao Serviço Europeu para a IA, às autoridades nacionais competentes e alguma também aos prestadores a jusante. A informação a disponibilizar a todos esses destinatários inclui o tipo/modalidade de dados (texto, imagens, áudio, vídeo ou outros a especificar) e a proveniência (v.g. *web crawling*, *private non-publicly available datasets obtained from third parties*, *user data*, *publicly available datasets*, *synthetic data*). Já os restantes elementos apenas ficarão disponíveis para o Serviço Europeu para a IA e/ou as autoridades nacionais de fiscalização do mercado, incluindo a breve descrição dos métodos implementados na aquisição e processamento para detetar "origens indevidas" que conduzam a um processamento ilegítimo. Excecionam-se as metodologias de curadoria dos dados, as quais também devem ser disponibilizadas aos operadores a jusante[195].

Quanto aos riscos associados à utilização do modelo e à "regurgitação" de dados pessoais, o formulário devia ser mais claro. Por um lado, refere entre os usos para que pode ser utilizado o modelo a análise e visualização de dados sem qualquer ressalva relativamente aos limites e condicionalismos que se se impõem no que respeita aos dados pessoais; por outro lado, no

[193] Este não é objeto de atenção neste texto, dado que não respeita à proteção de dados pessoais.
[194] Model Documentation Form. Assim se cumprem os requisitos de necessidade e proporcionalidade impostos no artigo 88.º, n.º 2.
[195] O Model Documentation Form dá os seguintes exemplos: "*cleaning* (*e.g. filtering out irrelevant content such as advertisements*), *normalisation* (*e.g. tokenizing*), *augmentation* (*e.g. back-translation*)".

que tange às restrições e proibições de uso, a obrigação de informação não se refere às restrições ou proibições resultantes da lei, entre as quais se encontram obviamente as normas de proteção dos dados pessoais. Ora, como se viu, são vários os problemas e as dúvidas que subsistem relativamente aos requisitos relacionados com o fundamento do interesse legítimo para o processamento, pelo que seria importante que as medidas de mitigação e as restrições de uso fossem claras para os prestadores a jusante. A cláusula aberta prevista na Medida 1.2 pode responder a esta limite da Medida 1.1, mas não assegura a certeza desejável e coloca o ónus da diligência nos operadores a jusante. Com salvaguarda do respeito pela propriedade intelectual e segredos comerciais legalmente tutelados, a Medida 1.2. determina que os prestadores de modelos de IA de finalidade geral providenciarão, a solicitação dos prestadores a jusante, informação adicional que seja necessária para que estes tenham uma boa compreensão das capacidades e limitações do modelo que sejam relevantes para a respetiva integração nos seus sistemas de IA e possam cumprir as respetivas obrigações. Saúda-se a indicação de um prazo para a disponibilização da informação[196].

De modo distinto dos capítulos anteriores, o capítulo terceiro enuncia, no Preâmbulo, um conjunto de princípios. Além do já mencionado princípio da proporcionalidade das medidas de mitigação aos riscos, salientam-se o princípio da avaliação e mitigação dos riscos ao longo do ciclo de vida dos modelos, o princípio da avaliação e mitigação contextual, segundo o qual os prestadores dos modelos devem ter em consideração os sistemas de IA nos quais serão previsivelmente implantados. Estes princípios importam uma colaboração com os prestadores a jusante e outros *stakeholders*, incluindo representantes da sociedade civil e da academia, que se traduz na formulação de um outro princípio, o princípio da cooperação. Finalmente, sobressaem, enquanto princípios no domínio da proteção e segurança, por um lado, o princípio da precaução e, por outro lado, o princípio da inovação, que incentiva os prestadores a promover o avanço do estado da arte naquele domínio.

Os 10 compromissos incluídos no Capítulo 3 – "*Safety and Protection*" - destinam-se aos prestadores de modelos de IA de finalidade geral com risco sistémico, embora os prestadores de modelos sem risco sistémico também possam subscrever esses compromissos e adotar as respetivas medidas[197]. Estes compromissos desdobram-se em 32 medidas: a elaboração, implementação e notificação de um Quadro de Segurança (Medidas 1.1 a 1.3, Appendix 2), onde se sistematizam os processos e medidas que assegurem que os riscos são mitigados até serem "aceitáveis"; a identificação dos riscos, incluindo a sua natureza e fontes, e respetivos cenários (Medidas 2.1 e 2.2; Appendix 1); a análise dos riscos, incluindo com monitorização após a colocação no mercado (Medidas 3.1. a 3.5; Appendix 3); a determinação do nível aceitável dos riscos (Medidas 4.1 e 4.2); a adoção de medidas de mitigação para tornar os riscos aceitáveis, entre as quais se enuncia a "filtragem e limpeza dos dados de treino" e a "monitorização e filtragem dos *inputs* e *outputs* do modelo", a realização do *fine-tuning* do

[196] COMISSÃO EUROPEIA, *Code of Practice for General-Purpose AI Models. Transparency Chapter*, cit., p. 6.
[197] COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, cit., p. 28.

modelo para "recusar certos pedidos"[198] (Medida 5.1); adotar medidas de cibersegurança para assegurar que os riscos de divulgação, acesso e/ou roubo do modelo são aceitáveis (Medidas 6.1 e 6.2; Appendix 4); a elaboração do Relatório do Modelo para a documentação e o reporte ao Serviço Europeu para a IA da informação sobre o modelo (incluindo os dados e métodos de treino) e sobre os processos e medidas de avaliação e mitigação dos riscos sistémicos, antes da colocação no mercado, a sua atualização e notificação (Medidas 7.1 a 7.7); a determinação das responsabilidades pela política de gestão dos riscos e afetação dos recursos humanos, técnicos e financeiros apropriados para a sua implementação (Medidas 8.1 a 8.3), num quadro de uma cultura de risco saudável; a monitorização, documentação e comunicação de incidentes graves e medidas corretivas ao Serviço Europeu para a IA e, eventualmente, às autoridades nacionais competentes, ao longo do ciclo de vida do modelo (Medidas 9.1 a 9.4); e, a documentação das medidas preconizadas neste capítulo e o reforço da transparência através do Resumo pormenorizado do Quadro de Segurança e do Relatório do Modelo, a disponibilizar publicamente (Medidas 10.1 e 10.2). Foi ainda elaborado um glossário para clarificar alguns conceitos técnicos fundamentais.

Muitos *stakeholders* empresariais pronunciaram-se negativamente sobre o RIA[199], considerando as obrigações complexas, excessivamente onerosas e prejudiciais para a competitividade europeia, propugnando a sua simplificação[200]. Seguiu-se uma reação de sentido contrário de responsáveis de diversas organizações da sociedade civil e da academia[201].

5. Conclusão

Os modelos de IA de finalidade geral, em especial os LLM, não têm uma natureza estritamente tecnológica, mas, antes, socio-tecnológica, pois a sua funcionalidade envolve necessariamente componentes humanos e sociais (desde os dados utilizados no seu treino até aos *inputs* ambientais e *prompts* introduzidos na sua utilização) que se refletem nos riscos que deles decorrem, nomeadamente para os dados pessoais. Por isso, a identificação, a avaliação e a governação dos riscos que envolvem não pode reduzir-se a um quadro estritamente tecnológico e impõe-se o envolvimento de todos os interessados. Os seus riscos para a proteção dos dados pessoais estão bem documentados. Logo no seu desenvolvimento, suscita-

[198] No entanto, não parece que as preocupações subjacentes a estas medidas sejam a proteção de dados, pois, no que concerne às respostas, refere-se apenas a respetiva inutilidade. COMISSÃO EUROPEIA, *Code of Practice for General-Purpose AI Models. Safety and Security Chapter*, cit., p. 17.

[199] A carta aberta de 58 CEO de empresas europeias criticava a regulação e solicitava à Comissão Europeia o adiamento da aplicação das obrigações previstas no RIA. EU CHAMPIONS INITIATIVE, *Stop the Clock - Open Letter*, https://aichampions.eu/ (02.06.2026).

[200] PHILIP HACKER, ROBERT KILIAN, JANA COSTAS, *Simplifying European AI Regulation - An Evidence-based Study*, Gütersloh: Bertelsmann Stiftung, 2025, https://www.bertelsmann-stiftung.de/en/publications/publication/did/simplifying-european-ai-regulation-an-evidence-based-study (02.06.2026).

[201] AAVV, *Open Joint Letter against the Delaying and Reopening of the AI Act*, 9 July 2025, https://www.beuc.eu/sites/default/files/publications/BEUC-X-2025-066_Open_Joint_Letter_against_the_Delaying_and_Reopening_of_the_AI_Act.pdf (02.06.2026).

se a questão do fundamento para o tratamento de dados pessoais utilizados no treino dos modelos, quando não se possa enquadrar na execução contratual ou haja consentimento. A proposta *Digital Omnibus* ultrapassa a dúvida e reconhece o fundamento do interesse legítimo que tem sido invocada pelos prestadores, passo que favorece o interesse da indústria em detrimento dos titulares dos dados pessoais. No funcionamento dos modelos, além dos riscos associados a *leakages* acidentais ou maliciosos, os potenciais impactos nos direitos dos titulares de dados pessoais podem ser a "regurgitação" de informação pessoal e a recolha e tratamento de dados associada às interações ambientais ou aos *prompts* introduzidos pelos utilizadores das ferramentas onde os modelos estejam integrados. Além do recurso ao AIDP para a avaliação dos riscos para os dados pessoais, são vários os meios tecnológicos que têm sido apontados para responder a esses riscos – v.g., filtros -, para além da supervisão humana, mas as incertezas, dificuldades e limites permanecem. A implementação de um sistema de governação do risco contextual e iterativo ao longo de todo o ciclo de vida dos modelos é fundamental, com vista a prevenir, eliminar ou mitigar os riscos identificados. Entre os diversos modelos de governação de risco que têm sido sugeridos, salienta-se o HUDERIA, proposto no Conselho da Europa. O RIA prevê um conjunto de obrigações para os prestadores de modelos de IA de finalidade geral e obrigações acrescidas para os prestadores de modelos de IA de finalidade geral com risco sistémico, algumas com relevância para a proteção dos dados pessoais. Por um lado, essas obrigações traduzem-se sobretudo em obrigações de informação e transparência. Por outro lado, impõem a adoção de medidas de segurança e proteção. Na ausência de *standards* que as densifiquem, o Regulamento previu a aprovação de Códigos de Práticas elaborados de modo colaborativo e de adesão voluntária pelos interessados. O Código de Práticas, aprovado em 2025, constitui um documento abrangente e pormenorizado, embora não isento de críticas. Dois dos seus três capítulos têm relevância para a proteção dos dados pessoais, o primeiro ("*Transparency*"), aplicável a todos os modelos de IA de finalidade geral, e o terceiro ("*Safety and Protection*"), aplicável aos modelos de IA de finalidade geral com risco sistémico. Além de enunciarem princípios fundamentais, instituem compromissos e concretizam as medidas a implementar. Ainda que a sua adoção não constitua presunção do cumprimento das obrigações, isenta os prestadores da obrigação de demonstração de cumprimento por outros meios. Em suma, no quadro normativo de regulação dos modelos de IA a proteção dos dados pessoais é uma preocupação do legislador europeu. Não obstante, as soluções das propostas "*omnibus*" em resposta às incertezas e insuficiências do regime resultante do RGPD e do RIA e às críticas de regulação excessiva fragilizam a proteção dos dados pessoais.

Bibliografia

AAVV, *Open Letter*, 27.01.2025, https://www.newsmediaeurope.eu/wp-content/uploads/2025/01/Right-holders-joint-letter-on-the-Second-Draft-Code-of-Practice-27012025.pdf (02.06.2026)

AAVV, *Open Joint Letter against the Delaying and Reopening of the AI Act*, 9 July 2025, https://www.beuc.eu/sites/default/files/publications/BEUC-X-2025-066_Open_Joint_Letter_against_the_Delaying_and_Reopening_of_the_AI_Act.pdf (02.06.2026)

AI BOARD, *Conclusion of the Artificial Intelligence Board on the Assessment of the General-Purpose AI Code of* Practice pursuant to Article 56 of Regulation 2024/1689 (Artificial Intelligence Act or "AI Act"), 2025, https://digital-strategy.ec.europa.eu/en/policies/ai-board (02.06.2026)

BAKERHOSTETLER, *Case Tracker: Artificial Intelligence, Copyrights and Class Actions*, 2026, https://www.bakerlaw.com/services/artificial-intelligence-ai/case-tracker-artificial-intelligence-copyrights-and-class-actions/ (02.06.2026)

BARBERÀ, I., *AI Privacy Risks & Mitigations Large Language Models (LLMs)*, EDPB, março 2025, https://www.edpb.europa.eu/system/files/2025-04/ai-privacy-risks-and-mitigations-in-llms.pdf (02.06.2026)

BRYNJOLFSSON, ERIK *et al*., "Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence", 13.11.2025. *Stanford Digital Economy Lab*, https://digitaleconomy.stanford.edu/app/uploads/2025/12/CanariesintheCoalMine_Nov25.pdf (02.06.2026)

CARLINI, N. *et al*., *Extracting Training Data from Large Language Models*, arXiv:2012.07805v2 [cs.CR] 2021, https://arxiv.org/pdf/2012.07805 (02.06.2026)

CEDP, *Parecer 28/2024 sobre certos aspetos da proteção de dados relacionados com o tratamento de dados pessoais no contexto dos modelos de IA*, 17 de dezembro de 2024, https://www.edpb.europa.eu/system/files/2025-05/edpb_opinion_202428_ai-models_pt.pdf (02.06.2026)

CEN-CENELEC, *Joint Technical Committee 21 (JTC 21), About the Joint Technical Committee | CEN-CENELEC JTC 21*, 2026, https://jtc21.eu/about/ (02.06.2026)

COMISSÃO EUROPEIA, *Proposta de Regulamento do PE e do Conselho que altera os Regulamentos (UE) 2024/1689 e (UE) 2018/1139 no respeitante à simplificação da aplicação das regras harmonizadas em matéria de inteligência artificial (Omnibus Digital sobre a IA) [COM/2025/836/FINAL]*, https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A52025PC0837 (02.06.2026)

COMISSÃO EUROPEIA, *Proposta de Regulamento do Parlamento Europeu e do Conselho que altera os Regulamentos (UE) 2016/679, (UE) 2018/1724, (UE) 2018/1725 e (UE) 2023/2854 e as Diretivas 2002/58/CE, (UE) 2022/2555 e (UE) 2022/2557 no respeitante à simplificação do quadro legislativo digital e que revoga os Regulamentos (UE) 2018/1807, (UE) 2019/1150, (UE) -2022/868 e a Diretiva (UE) 2019/1024 (Omnibus Digital) [COM/2025/837 final]*, https://eur-lex.europa.eu/legal-content/pt/TXT/?uri=COM%3A2025%3A837%3AFIN (02.06.2026)

COMISSÃO EUROPEIA, *Comunicação da Comissão Orientações da Comissão sobre o âmbito das obrigações dos prestadores de modelos de IA de finalidade geral estabelecidas pelo Regulamento (UE) 2024/1689 (Regulamento IA)*, C(2025) 7719 final, Bruxelas, 19.11.2025, p. 8, https://digital-strategy.ec.europa.eu/en/library/guidelines-scope-obligations-providers-general-purpose-ai-models-under-ai-act (02.06.2026)

COMISSÃO EUROPEIA, *Comunicação da Comissão Nota explicativa e modelo para o resumo público dos conteúdos utilizados para o treino de modelos de IA de finalidade geral exigido pelo artigo 53.º, n.º 1, alínea d), do Regulamento (UE) 2024/1689 (Regulamento da Inteligência Artificial)*, C(2025) 8311 final, Bruxelas, 5.12.2025, https://digital-strategy.ec.europa.eu/en/library/explanatory-notice-and-template-public-summary-training-content-general-purpose-ai-models (02.06.2026)

COMISSÃO EUROPEIA, *Better regulation: guidelines and toolbox*, 3.11.2021 e 20.07.2023, https://commission.europa.eu/law/law-making-process/better-regulation/better-regulation-guidelines-and-toolbox_en (02.06.2026)

COMISSÃO EUROPEIA, *Code of Practice for General-Purpose AI Models*, 2025, https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai (02.06.2026)

COMISSÃO EUROPEIA, *Comission Opinion of 1.8.2025 on the assessment of the General-Purpose AI Code of Practice within the meaning of Article 56 of Regulation (EU) 2024/1689*, C(2025) 5361 final, Brussels, 1.08.2025, https://digital-strategy.ec.europa.eu/en/library/commission-opinion-assessment-general-purpose-ai-code-practice (02.06.2026)

COMISSÃO EUROPEIA, *Commission Staff Working Document Accompanying the documents Proposal for a Regulation of the European Parliament and of the Council Amending Regulations (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 and Directives 2002/58/EC, (EU) 2022/2555 and (EU) 2022/2557 as regards the simplification of the digital legislative framework, and repealing Regulations (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868, and Directive (EU) 2019/1024 (Digital Omnibus) Amending Regulations (EU) 2024/1689 and (EU) 2018/1139 as regards the simplification of the implementation of harmonised rules on artificial intelligence (Digital Omnibus on AI)*, SWD(2025) 836 final, Brussels, 19.11.2025, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52025SC0836 (02.06.2026)

COMISSÃO EUROPEIA, *Call for evidence and public consultation on the Apply AI Strategy*, 2025, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/14625-Apply-AI-Strategy-strengthening-the-AI-continent_en (02.06.2026)

COMISSÃO EUROPEIA, *General-Purpose AI Code of Practice*, July 10, 2025, https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai (02.06.2026)

COMISSÃO EUROPEIA, *Questions and answers on the code of practice for General-Purpose AI*, 2026, https://digital-strategy.ec.europa.eu/en/faqs/questions-and-answers-code-practice-general-purpose-ai (02.06.2026)

CONSEIL DE L'UNION EUROPÉENNE, *Dossier interinstitutionnel 2021/0106(COD)*, 13.05.2022, https://artificialintelligenceact.eu/wp-content/uploads/2022/05/AIA-FRA-Art-34-13-May.pdf (02.06.2026)

CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (CONVENTION 108), *Draft Guidelines on Privacy and Data Protection in the context of LLM-based systems*, T-PD(2025)3rev3, Council of Europe, Strasbourg, 12.05.2026, https://media.licdn.com/dms/document/media/v2/D4E1FAQHorwju8e1mrw/feedshare-document-sanitized-pdf/B4EZ4hHqZUKkBE-/0/1778672118494?e=1779530400&v=beta&t=N9oMFh0LPeNOmtNV-4hvWW6l3uuFmQdK9JCWNCrUFC4 (02.06.2026)

CORPORATE EUROPE OBSERVATORY, "Article by article, how Big Tech shaped the EU's roll-back of digital rights", *Corporate Europe Observatory* 14.01.2026, https://corporateeurope.org/en/2026/01/article-article-how-big-tech-shaped-eus-roll-back-digital-rights (02.06.2026)

COUNCIL OF THE EU, *Council agrees position to streamline rules on Artificial Intelligence* (Press Release), 13.03.2026, https://www.consilium.europa.eu/en/press/press-releases/2026/03/13/council-agrees-position-to-streamline-rules-on-artificial-intelligence/ (02.06.2026)

COWAN, J., 'How Much Is Your Data Worth?', *The New York Times*, 25.03.2019, https://www.nytimes.com/2019/03/25/us/newsom-hertzberg-data-dividend.html (02.06.2026)

CUSTERS, B. / VRABEC, H., "Tell me something new: data subject rights applied to inferred data and profiles", *Computer Law & Security Review*, Volume 52, Amsterdam, Elsevier, 2024, https://doi.org/10.1016/j.clsr.2024.105956 (02.06.2026)

DELINEA, *2026 Identity Security Report. Uncovering the Hidden Risks of the AI Race*, 2026, https://delinea.com/hubfs/Delinea/whitepapers/delinea-report-identity-security-report-2026.pdf (02.06.2026)

DEMIR, E., "The protection of human biodata: Is there any role for data ownership?", *Computer Law & Security Review*, Volume 51, Amsterdam, Elsevier, *2023*, pp. 1-13, https://doi.org/10.1016/j.clsr.2023.105905 (02.06.2026)

DOMÍNGUEZ DE OLAZÁBAL, I., "Europe is dismantling its digital rights from within", *EDRI*, 2025, https://edri.org/our-work/europe-is-dismantling-its-digital-rights-from-within/ (02.06.2026)

EDPB, ARTICLE 29 DATA PROTECTION WORKING GROUP, *Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679*, 2017, http://ec.europa.eu/newsroom/document.cfm?doc_id=44137 (02.06.2026)EDPB, ARTICLE 29

DATA PROTECTION WORKING PARTY, *Guidelines on the right to data portability*, 16/EN WP 242 rev.01, 5.04.2017, https://ec.europa.eu/newsroom/article29/items/611233/en (02.06.2026)

EDPB, ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Personal data breach notification under Regulation 2016/679*, 2017/2018, https://ec.europa.eu/newsroom/article29/items/612052 (02.06.2026)

EDPB, ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, 17/EN WP251 rev.01, 2018, https://ec.europa.eu/newsroom/article29/items/612053/en (02.06.2026)

EDPB, *Report of the work undertaken by the ChatGPT Taskforce*, 23 May 2024, https://www.edpb.europa.eu/our-work-tools/our-documents/other/report-work-undertaken-chatgpt-taskforce_en (02.06.2026)

ENES, G., "A Estratégia Europeia para a Inteligência Artificial, *in* Maria Raquel Guimarães e Rute Teixeira Pedro (Coord.), *Direito e Inteligência Artificial*, Coimbra, Almedina, 2023, pp. 37-91, https://cij.up.pt//client/files/0000000001/9789894014348direito-e-inteligencia-artificial_2301.pdf (02.06.2026)

ENES, G., "Building a European Data Space", *in* Inês Quadros e Miguel Gorjão-Henriques (eds), *Revisiting the Liberal Constitutional Features of the European Union*, Cham, Springer, 247-285, pp. 251-252, https://doi.org/10.1007/978-3-031-95717-8_11

EU CHAMPIONS INITIATIVE, *Stop the Clock - Open Letter*, https://aichampions.eu/ (02.06.2026)

EUROPEAN ENVIRONMENT AGENCY, *Governance in complexity. Sustainability governance under highly uncertain and complex conditions*, *EEA Report 05/2024*, doi: 10.2800/597121, https://www.eea.europa.eu/en/analysis/publications/governance-in-complexity-sustainability-governance (02.06.2026)

EUROPEAN OMBUDSMAN, "How the European Commission ensures transparency, inclusiveness and accountability in the adoption of harmonised standards related to artificial intelligence. Case 1974/2025/MIK", 26.09.2025, https://www.ombudsman.europa.eu/en/opening-summary/en/212090#_ftn4 (02.06.2026)

EUROPEAN OMBUDSMAN, *Recommendation on the European Commission's compliance with 'Better Regulation' rules and other procedural requirements in preparing legislative proposals that it considered to be urgent (983/2025/MAS - the "Omnibus" case, 2031/2024/VB - the "migration" case, and 1379/2024/MIK - the "CAP" case)*, 25.11.2025, https://www.ombudsman.europa.eu/pt/recommendation/en/215920 (02.06.2026)

FERETZAKIS, G. et al., "GDPR and Large Language Models: Technical and Legal Obstacles", *Future Internet 2025*, 17(4), Basel, MDPI, p. 151, https://doi.org/10.3390/fi17040151 (02.06.2026)

FORTINET, *2025 Global Threat Landscape Report*, 2025, https://www.fortinet.com/resources/reports/threat-landscape-report#form-demo (02.06.2026)

GARANTE PER LA PROTEZIONE DEI DATI PERSONALE, "Intelligenza artificiale: il Garante blocca ChatGPT. Raccolta illecita di dati personali. Assenza di sistemi per la verifica dell'età dei minori", *Garante Per La Protezione Dei Dati Personali*, 31.03.2023, https://gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870847#english (02.06.2026)

GARANTE PER LA PROTEZIONE DEI DATI PERSONALE, *Comunicato Stampa - ChatGPT,* "il Garante privacy chiude l'istruttoria. OpenAI dovrà realizzare una campagna informativa di sei mesi e pagare una sanzione di 15 milioni di euro", *Garante Per La Protezione Dei Dati Personali*, 20.12.2024, https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10085432 (02.06.2026)

GARANTE PER LA PROTEZIONE DEI DATI PERSONALE, "Comunicato Stampa - AI: Il Garante sanziona la società che gestisce il chatbot "Replika"", *Garante Per La Protezione Dei Dati Personali*, 19.05.2025, https://gpdp.it/home/docweb/-/docweb-display/docweb/10132048 (02.06.2026)

GILLHAM, J., *Originality.ai*, December 7, 2025, https://originality.ai/blog/openai-chatgpt-lawsuit-list (02.06.2026)

GINART, A. *et al*., "Making AI forget you: data deletion in machine learning", in H. WALLACH et al., *Proc. 33rd Conference on Neural Information Processing Systems*, Vancouver, NeurIPS, 2019, https://arxiv.org/pdf/1907.05012 (02.06.2026)

GUREVICH, Y. *et al*., "Inverse Privacy", 2015, https://arxiv.org/abs/1510.03311 (02.06.2026)

GUTIERREZ, C. I. *et al*., "A Proposal for a Definition of General Purpose Artificial Intelligence Systems", Future of Life Institute – Working Paper, November 2022, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4238951 (02.06.2026)

HACKER, P., *Comentário no LinkedIn*, 8.05.2026, https://www.linkedin.com/feed/?highlightedUpdateUrn=urn%3Ali%3Aactivity%3A7458099365094674432&highlightedUpdateType=TOPIC_TRENDING_CONVERSATION_IN_YOUR_NETWORK&origin=inapp&showCommentBox=true (02.06.2026)

HACKER, P. / KILIAN, R. / COSTAS, J., *Simplifying European AI Regulation - An Evidence-based Study*, Gütersloh: Bertelsmann Stiftung, 2025, https://www.bertelsmann-stiftung.de/en/publications/publication/did/simplifying-european-ai-regulation-an-evidence-based-study (02.06.2026).IBM CORPORATION, *Cost of Data Breach Report 2025. The AI Oversight Gap*, 2025, p. 27, https://www.ibm.com/downloads/documents/us-en/131cf87b20b31c91 (02.06.2026)

ISO/IEC, *Information technology — Artificial intelligence — Artificial intelligence concepts and terminology*, 2022; ISO/IEC, *Information technology — Artificial intelligence — AI system life cycle processes*, 2023, https://www.iso.org/search.html (02.06.2026)

KRÖGER, J. L., "Recognizing Information Inferred about Individuals as Personal Data", Chapter adapted from: KRÖGER, J. L. , *Rogue Apps, Hidden Web Tracking and Ubiquitous Sensors*, Doctoral dissertation, Berlin, Technische Universität Berlin, 2022, pp. 222–225.

https://doi.org/10.14279/depositonce-16043, https://ssrn.com/abstract=4349200 or http://dx.doi.org/10.2139/ssrn.4349200 (02.06.2026)

KUHN, T. S., *A Estrutura das revoluções científicas*, (trad. portuguesa), Lisboa, Ed. Guerra e Paz, 2009

LANIER, J., *Jaron Lanier fixes the Internet*, New York, *New York* Times, The Privacy Project, 2019 (em especial, Episode 2), https://www.nytimes.com/interactive/2019/09/23/opinion/data-privacy-jaron-lanier.html (02.06.2026)

LIU, S. *et al*., "Rethinking machine unlearning for large language models*", Nature Machine Intelligence*, volume 7, Londres, Nature Portfolio, 2025, pp. 181–194, https://www.nature.com/articles/s42256-025-00985-0 (02.06.2026)

MALGIERI, G., Trade Secrets v Personal Data: a possible solution for balancing rights", *International Data Privacy Law*, Volume 6, Issue 2, Oxford, Oxford University Press, 2016, pp. 102–116, https://doi.org/10.1093/idpl/ipv030 (02.06.2026)

MULLER, C. *et al*., "AIA in-depth #1: Objective, Scope, Definition", *ALLAI*, 13.02.2022, https://allai.nl/wp-content/uploads/2022/03/AIA-in-depth-Objective-Scope-and-Definition.pdf (02.06.2026)

NEUMAN, M., "Deal over EU's AI omnibus package represents political victory for business", *in MLex*, Newsletter May 7, 2026, https://www.mlex.com/mlex/artificial-intelligence/articles/2474987?nl_pk=94e4dd0b-c9bf-4432-85b4-95a382f04fa2&utm_source=newsletter&utm_medium=email&utm_campaign=mlex/artificial-intelligence&utm_content=2026-05-08&read_main=1&nlsidx=0&nlaidx=0 (02.06.2026)

NOLTE, N. / WERKMEISTER, CHRISTOPH, "DS-GVO Art. 17 Recht auf Löschung ("Recht auf Vergessenwerden")", Peter Gola (ed.), *Datenschutz-Grundverordnung VO (EU) 2016/679*, 1.ª Ed., Munique, CHBeck, 2017

NOVELLI, C. *et al*., "AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act", in *Digital Society*, Volume 3, n.º 13, Londres, Springer Nature, 2024, https://link.springer.com/article/10.1007/s44206-024-00095-1 (02.06.2026)

NOYB, "Digital Omnibus: EU Commission wants to wreck core GDPR principles", *NOYB,* 19.11.2025, https://noyb.eu/en/digital-omnibus-eu-commission-wants-wreck-core-gdpr-principles (02.06.2026)

OECD, *Emerging Systemic Risks in the 21st Century: An Agenda for Action*, 2003, https://www.oecd.org/content/dam/oecd/en/publications/reports/2003/04/emerging-risks-in-the-21st-century_g1gh2feb/9789264101227-en.pdf (02.06.2026)

OECD, "OECD Framework for the Classification of AI systems", OECD Digital Economy Papers, 22.02.2022, https://www.oecd.org/en/publications/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en.html (02.06.2026)

OECD, "Explanatory Memorandum on the updated OECD definition of an AI System", OECD Artificial Intelligence Papers, OECD Publishing, 2024, No. 8, https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/03/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_3c815e51/623da898-en.pdf (02.06.2026)

OPENAI, *How your data is used to improve model performance*, 2026, https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance (02.06.2026)

OPENAI, *How ChatGPT and our foundation models are developed*, 2025, https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed#h_6b40d4df3c (02.06.2026)

OPENAI, *Política de Privacidade*, 2026, https://openai.com/pt-PT/policies/row-privacy-policy/ (02.06.2026)

OPENAI, *GPT-4 technical report*, Dec. 2023, Chapter 6, https://arxiv.org/abs/2303.08774 (02.06.2026)

OUYANG, L., *Training language models to follow instructions with human feedback*, 2022, https://arxiv.org/pdf/2203.02155 (02.06.2026)

PAGALLO, U. / SCIOLLA, J. C., "Anatomy of web data scraping: ethics, standards, and the troubles of the law", *European Journal of Privacy Law & Technologies*, n.º 2, Nápoles, Suor Orsola Benincasa Università Editrice, 2023 pp. 1- 9, https://doi.org/10.57230/EJPLT232PS (02.06.2026)

PE, *Alterações aprovadas pelo Parlamento Europeu, em 26 de março de 2026, sobre a proposta de regulamento do Parlamento Europeu e do Conselho que altera os Regulamentos (UE) 2024/1689 e (UE) 2018/1139 no respeitante à simplificação da aplicação das regras harmonizadas em matéria de inteligência artificial (pacote Omnibus Digital sobre a IA)*, https://www.europarl.europa.eu/doceo/document/TA-10-2026-0098_PT.html (02.06.2026)

PE, *AI Act: deal on simplification measures, ban on "nudifier" apps*, 7.05.2026, https://www.europarl.europa.eu/news/en/press-room/20260427IPR42011/ai-act-deal-on-simplification-measures-ban-on-nudifier-apps (02.06.2026)

SCHOKRI, R. *et al.*, *Membership Inference Attacks Against Machine Learning Models*, S/D, https://www.cs.cornell.edu/~shmat/shmat_oak17.pdf (02.06.2026)

SMITH, M., "AI Infringement Case Updates: July 28, 2025", *McKool Smith*, 28.07.2025, https://www.mckoolsmith.com/newsroom-ailitigation-33 (02.06.2026)

STAAB, R. *et al.*, *Beyond Memorization: Violating Privacy Via Inference with Large Language Models*, 2024, https://arxiv.org/abs/2310.07298 (02.06.2026)

STATISTA, *Statista Consumer Insights*, 28.3.2025, https://cdn.statcdn.com/Infographic/images/normal/31706.jpeg (02.06.2026)

116TH U.S. CONGRESS (2019-2020), Sen. Kennedy, J., *Own Your Data Act, S. 806,* 14th of March 2019, https://www.congress.gov/bill/116th-congress/senate-bill/806/text?format=txt (02.06.2026)

TASKAN, B. *et al*., "Clarifying the conceptual map of VUCA: a systematic review", *International Journal of Organizational Analysis*, Vol. 30, Nº. 7, Leeds, Emerald Publishing, 2022, pp. 196-217, https://doi.org/10.1108/IJOA-02-2022-3136 (02.06.2026)

TEGMARK, M., *Life 3.0: Being Human in the Age of Artificial Intelligence*, Nova Iorque, Alfred A. Knopf, 2017

THE NEW YORK TIMES, Motion to vacate the Order, 12.11.2025, https://fingfx.thomsonreuters.com/gfx/legaldocs/dwpkqdwwwpm/OPENAI%20NEW%20YORK%20TIMES%20COPYRIGHT%20LAWSUIT%20reconsider.pdf (02.06.2026)

TÕNURIST, P., HANSONE, A., *Anticipatory Innovation Governance. Shaping the future through proactive policy making*, OECD Working Papers on Public Governance No. 44, 24 December 2020, https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/12/anticipatory-innovation-governance_d1aded4e/cce14d80-en.pdf (02.06.2026)

VASSILEV, A. *et al*., *Adversarial Machine Learning A Taxonomy and Terminology of Attacks and Mitigations*, Gaithersburg, National Institute of Standards and Technology, 2024, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2023.pdf (02.06.2026)

VEALE, M. / BINNS, R. / EDWARDS, L., "Algorithms that remember: model inversion attacks and data protection law", *Philosophical Transactions of the Royal Society A*, Vol 376, Ed. 2133, Londres, Royal Society Publishing, 2018, http://dx.doi.org/10.1098/rsta.2018.0083 (02.06.2026)

VELLI, F., "Processing Personal Data in the Context of AI Models: EDPB's Opinion 28/2024", *European Papers*, Vol. 9, 2024, No 3, pp. 1497-1500. doi: 10.15166/2499-8249/822, https://www.europeanpapers.eu/europeanforum/protecting-personal-data-in-context-of-ai-models (02.06.2026)

WACHTER, S. / MITTELSTADT, B., "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI", *Columbia Business. Law Review*, n.º 2, Nova Iorque, Columbia School of Law, 2019, pp. 1-135, https://ssrn.com/abstract=3248829 (02.06.2026)

WASIL, A. R. *et al*., "Affirmative safety: An approach to risk management for high-risk AI", 24.04.2024, https://arxiv.org/pdf/2406.15371 (02.06.2026)

WEIDINGER, L. *et al*., Sociotechnical Safety Evaluation of Generative AI Systems, 2023, https://arxiv.org/pdf/2310.11986 (02.06.2026)

WEIZENBAUM, J., "ELIZA—a computer program for the study of natural language communication between man and machine", *Communications of the ACM*, Volume 9, Issue 1, Nova Iorque, ACM, 1966, https://dl.acm.org/doi/epdf/10.1145/365153.365168 (02.06.2026)

WORLD ECONOMIC FORUM, *Global Risks Report 2026*, 21st Report, 2026, https://reports.weforum.org/docs/WEF_Global_Risks_Report_2026.pdf (02.06.2026)

ZHANG, R. *et al.*, "The Dark Side of AI Companionship: A Taxonomy of Harmful Algorithmic Behaviors in Human-AI Relationships", *CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems,* pp. 1–17, https://doi.org/10.1145/3706598.3713429 (02.06.2026)

ZUBOFF, S., *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York, PublicAffairs, 2019

Jurisprudência

Ac. do TJ de 20 de dezembro de 2017, *Nowak*, Proc. C-434/16, ECLI:EU:C:2017:994, https://infocuria.curia.europa.eu/tabs/document/C/2016/C-0434-16-00000000RP-01-P-01/ARRET/198059-PT-1-html (02.06.2026)

Ac. do TJ de 24 de setembro de 2019, *Google (Territorial scope of de-referencing)*, Proc. C-507/17, ECLI:EU:C:2019:772, https://infocuria.curia.europa.eu/tabs/document/C/2017/C-0507-17-00000000RP-01-P-01/ARRET/218105-PT-1-html (02.06.2026)

Ac. do TJ de 1 de agosto de 2022, *Vyriausioji tarnybinės etikos komisija*, Proc. C-184/20, ECLI:EU:C:2022:601, https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:62020CJ0184 (02.06.2026)

Ac. do TJ de 4 de maio de 2023, *Österreichische Datenschutzbehörde e CRIF*, Proc. C-487/21, EU:C:2023:369, https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:62021CJ0487 (02.06.2026)

Ac. do TJ de 4 de julho de 2023, *Meta v. Bundeskartellamt*, Proc. C-252/21, ECLI:EU:C:2023:537, https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:62021CJ0252 (02.06.2026)

Ac. do TJ de 14 de dezembro de 2023, *Natsionalna agentsia za prihodite*, Proc. C-340/21, ECLI:EU:C:2023:986, https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:62021CJ0340 (02.06.2026)

Ac. do TJ de 21 de dezembro de 2023, *Krankenversicherung Nordrhein*, C-667/21, EU:C:2023:1022, https://infocuria.curia.europa.eu/tabs/document/C/2021/C-0667-21-00000000RP-01-P-01/ARRET/280768-PT-1-html (02.06.2026)

Ac. do TJ de 4 de outubro de 2024, *Schrems v. Meta*, Proc. C-446/21, ECLI:EU:C:2024:834, https://infocuria.curia.europa.eu/tabs/document/C/2021/C-0446-21-00000000RP-01-P-01/ARRET/290674-PT-1-html (02.06.2026)

Ac. do TJ de 9 de janeiro de 2025, Österreichische Datenschutzbehörde (Pedidos excessivos), C-416/23, ECLI:EU:C:2025:3, https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=ecli:ECLI%3AEU%3AC%3A2025%3A3 (02.06.2026)

Ac. do TJ de 4 de setembro de 2025, *AEPD/CUR (Notion de données à caractère personnel)*, Proc. C-413/23 P, ECLI:EU:C:2025:645, https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:62023CJ0413 (02.06.2026)

Ac. do TJ de 19 de março de 2026, Brillen Rottler, Proc. C-526/24, ECLI:EU:C:2026:216, https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:62024CJ0526 (02.06.2026)

Ac. do TG de 26 de abril de 2023, *CUR/AEPD*, Proc. T-557/20, EU:T:2023:219, https://infocuria.curia.europa.eu/tabs/document/T/2020/T-0557-20-00000000RD-01-P-01/ARRET/272910-PT-1-html (02.06.2026)

United States Northern District California, P.M. v. OpenAI LP (3:23-cv-03199) District Court, N.D. California, https://www.courtlistener.com/docket/67535351/pm-v-openai-lp/ (02.06.2026)

United States Northern District California, Order Granting Motions to Dismiss at 1, Cousart v. OpenAI LP, No. 3:23-cv-4557 (N.D. Cal. May 24, 2024), ECF No. 78, https://storage.courtlistener.com/recap/gov.uscourts.cand.417880/gov.uscourts.cand.417880.78.0.pdf (02.06.2026)

United States District Court Norhern California, Order Graning Motion to Dismiss with Leave to Amend and Denying Administrative Motion to relate without Prejudice, JL et al. v. ALPHABET INC., et al., No. Case 3:23-cv-03440-AMO, https://storage.courtlistener.com/recap/gov.uscourts.cand.415223/gov.uscourts.cand.415223.46.0.pdf (02.06.2026)

United States Southern District of New York, The New York Times Company v. Microsoft Corporation, 1:23-cv-11195, (S.D.N.Y.), https://www.courtlistener.com/docket/68117049/the-new-york-times-company-v-microsoft-corporation/ (02.06.2026)

(texto submetido a 5.06.2026 e aceite para publicação a 15.09.2026)