European Journal of Risk Regulation (2024), **15**, 49–86
doi:10.1017/err.2023.81

ARTICLE

Regulating Gatekeeper Artificial Intelligence and Data: Transparency, Access and Fairness under the Digital Markets Act, the General Data Protection Regulation and Beyond

Philipp Hacker, Johann Cordes and Janina Rochon

Chair for Law and Ethics of the Digital Society, European New School of Digital Studies, European University Viadrina, Frankfurt, Germany
Corresponding author: Philipp Hacker; Email: hacker@europa-uni.de

Abstract

Artificial intelligence (AI) is not only increasingly being used in business and administration contexts, but a race for its regulation is also underway, with the European Union (EU) spearheading the efforts. Contrary to existing literature, this article suggests that the most far-reaching and effective EU rules for AI applications in the digital economy will not be contained in the proposed AI Act, but in the Digital Markets Act (DMA). We analyse the impact of the DMA and related EU acts on AI models and underlying data across four key areas: disclosure requirements; the regulation of AI training data; access rules; and the regime for fair rankings. We demonstrate that fairness, under the DMA, goes beyond traditionally protected categories of non-discrimination law on which scholarship at the intersection of AI and law has focused on. Rather, we draw on competition law and the FRAND criteria known from intellectual property law to interpret and refine the DMA provisions on fair rankings. Moreover, we show how, based on Court of Justice of the European Union jurisprudence, a coherent interpretation of the concept of non-discrimination in both traditional non-discrimination and competition law may be found. The final section sketches out proposals for a comprehensive framework of transparency, access and fairness under the DMA and beyond.

Keywords: AI; data; DMA

I. Introduction

Artificial intelligence (AI)[1] and, in particular, machine learning (ML)[2] are not only at the centre of countless economic applications[3] but have also triggered a broad regulatory

[1] In this paper, AI is understood based on the definition in Art 3(1) and Recitals 6a and 6b AI Act. This definition is far from perfect, see SJ Russell and P Norvig, *Artificial Intelligence: A Modern Approach* (3rd global edition, London, Pearson Education 2016) p 5; DL Poole and AK Mackworth, *Artificial Intelligence: Foundations of Computational Agents* (2nd edition, Cambridge, Cambridge University Press 2017) pp 3–7; M O'Shaughnessy, "One of the Biggest Problems in Regulating AI Is Agreeing on a Definition" (*Carnegie Endowment*, 2022) <https://carnegieendowment.org/2022/10/06/one-of-biggest-problems-in-regulating-ai-is-agreeing-on-definition-pub-88100> (last accessed 6 December 2022); P Hacker, "The European AI Liability Directives – Critique of a Half-Hearted Approach and Lessons for the Future" (2022) arXiv preprint arXiv:221113960, 11–12. However, it constitutes a workable definition for the purposes of this paper.

[2] For a definition of ML, see TM Mitchell, *Machine Learning* (1st edition, New York, McGraw Hill 1997) p 2: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

[3] I Goodfellow, Y Bengio and A Courville, *Deep Learning* (Cambridge, MA, MIT Press 2016) p 23 et seq.; M Javaid et al, "Artificial Intelligence Applications for Industry 4.0: A Literature-Based Study" (2022) 7 Journal of Industrial Integration and Management 83.

debate because of their partial lack of transparency[4] and tendencies to perpetuate discrimination[5] in some models and scenarios.[6] At the European Union (EU) level, the supposedly main pillar of AI regulation was proposed by the European Commission in April 2021[7]: the AI Act.[8] It is currently being hotly debated both in the European Council and Parliament[9] and will probably be adopted in late 2023 or 2024.[10] In addition, with the recently enacted Digital Services Act (DSA) and Digital Markets Act (DMA), the EU is pursuing the aim of regulating larger platforms to provide a more competitive environment for smaller providers and to support the creation of a safer digital space.[11] In September 2022, the final missing piece of European AI regulation was unveiled with the Commission proposal of two directives[12] concerning AI and software liability.[13]

With this package, the EU continues to build and refine the regulatory framework for the Digital Single Market. While much of the legal and technical debate concerning AI regulation has focused on the AI Act and the related AI liability provisions,[14] this paper

[4] ZC Lipton, "The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery" (2018) 16 Queue 31.

[5] X Ferrer et al, "Bias and Discrimination in AI: a cross-disciplinary perspective" (2021) 40 IEEE Technology and Society Magazine 72.

[6] See, eg, C Reed, "How should we regulate artificial intelligence?" (2018) 376(2128) Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 20170360; M Buiten, "Towards intelligent regulation of artificial intelligence" (2019) 10 European Journal of Risk Regulation 41; H Zech, *Entscheidungen digitaler autonomer Systeme: Empfehlen sich Regelungen zu Verantwortung und Haftung?* (Bonn, Deutscher Juristentag 2020); P Hacker, "Europäische und nationale Regulierung von Künstlicher Intelligenz" (2020) Neue Juristische Wochenzeitschrift 2142; N Smuha, "From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence" (2021) 13 Law, Innovation and Technology 57; M Ebers et al, "The European Commission's proposal for an artificial intelligence act – a critical assessment by members of the robotics and AI law society (RAILS)" (2021) 4 J 589; C Wendehorst and J Hirtenlehner, "Outlook on the Future Regulatory Requirements for AI in Europe" (2022) <https://ssrn.com/abstract=4093016> (last accessed 22 November 2023).

[7] For analysis and critique, see, eg, M Veale and FZ Borgesius, "Demystifying the Draft EU Artificial Intelligence Act – Analysing the good, the bad, and the unclear elements of the proposed approach" (2021) 22 Computer Law Review International 97, as well as some of the references in note 6, supra.

[8] European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence, COM(2021) 206 final; all references to the AI Act are to the following document: Council of the EU, Interinstitutional File: 2021/0106(COD), General Approach (= final version of the Council compromise text) of Nov. 25, 2022, Doc. No. 14954/22.

[9] L Bertuzzi, "Artificial Intelligence definition, governance on MEPs' menu" (*EURACTIV*, 10 November 2022) <https://www.euractiv.com/section/digital/news/artificial-intelligence-definition-governance-on-meps-menu> (last accessed 10 November 2022).

[10] See, eg, the remarks by MEP Axel Voss in L Bertuzzi, "The new liability rules for AI" (*EURACTIV*, 30 September 2022) <https://www.euractiv.com/section/digital/podcast/the-new-liability-rules-for-ai/> (last accessed 9 November 2022).

[11] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services, OJ L277/1 (DSA); Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector, OJ L265/1 (DMA).

[12] European Commission, Proposal for a Directive of the European Parliament and of the Council on Liability for Defective Products, COM(2022) 495 final (PLD Proposal); European Commission, Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence, COM(2022) 496 final (AILD Proposal).

[13] See, eg, Hacker, supra, note 1; G Spindler, "Die Vorschläge der EU-Kommission zu einer neuen Produkthaftung und zur Haftung von Herstellern und Betreibern Künstlicher Intelligenz" (2022) Computer und Recht 689.

[14] See, eg, Veale and Borgesius, supra, note 7; N Helberger and N Diakopoulos, "The European AI Act and how it matters for research into AI in media and journalism" (2023) 11 Digital Journalism 1751; F Sovrano et al, "Metrics, Explainability and the European AI Act Proposal" (2022) 5 J 126; N Smuha et al, "How the EU can achieve legally trustworthy AI: a response to the European Commission's proposal for an artificial intelligence act" (2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991> (last accessed 22 November 2023); P Hacker and J-H Passoth, "Varieties of AI Explanations Under the Law. From the GDPR to the AIA, and Beyond" in

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

seeks to show that the most far-reaching, most overlooked but potentially also most effective regulatory constraints for AI are ultimately erected by the DMA. The act contains numerous regulations for so-called gatekeepers: large online platforms such as Google or Amazon that exceed certain quantitative benchmarks and are thus essential for access to digital markets (Article 3 DMA). The possibility of fines – unmatched so far in any area of EU business law and exceeding those of the AI Act – underscore the importance of the DMA for legal and digital practice: authorities may punish a violation of the DMA with a fine up to 10% of the total annual global turnover (Article 30(1) DMA; 20% for repeat offences, Article 30(2) DMA). Hence, even if the DMA at times mirrors existing duties (eg transparency requirements largely repeating the provisions of the P2B Regulation), its unmatched enforcement apparatus opens a new chapter on AI regulation in the EU.

However, the DMA is currently being discussed primarily as an instrument to contain the market power of large online platforms.[15] In contrast, scholars have paid considerably less attention to the fact that the DMA contains rules that will probably have greater significance for the application of ML in the EU than the regulations provided for in the AI Act.[16] The central measures of the latter apply only to high-risk applications, which are, however, rare when ML deployments in the digital economy are considered.[17] Most ML applications of Google, Amazon and other large platforms will *not* be subject to the strict rules of the AI Act, but to those of the DMA. AI and ML power the core of the business and competitive edge of all of the large platforms: the ranking and scoring models used to present customers optimised, and often personalised, lists of items in return for a specific query. The quality of these rankings attracts customers, who in turn attract advertisers. Rankings are, therefore, key to platform success.

For example, when you search for a pair of shoes on Amazon, the resulting offers (the ranking) are created by a powerful ML learning-to-rank model that takes a variety of factors (so-called "features") into account.[18] Similarly, Google answers queries based on complex AI models.[19] The competitive edge of large platforms lies precisely in their capability to build such rankings in a fast, meaningful and personalised way.

A Holzinger et al (eds), *xxAI – Beyond Explainable AI, International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers* (Berlin, Springer 2022) p 343; ME Kaminski, "Regulating the Risks of AI" (2023) 103 Boston University Law Review (forthcoming); P Hacker, "A legal framework for AI training data – from first principles to the Artificial Intelligence Act" (2021) 13 Law, Innovation and Technology 257.

[15] See only M Eifert et al, "Taming the giants: The DMA/DSA package" (2021) 58 Common Market Law Review 987; J Laux, S Wachter and B Mittelstadt, "Taming the Few: Platform Regulation, Independent Audits, and the Risks of Capture Created by the DMA and DSA" (2021) 43 Computer Law & Security Review 105613; R Podszun, P Bongartz and S Langenstein, "The Digital Markets Act: Moving from Competition Law to Regulation for Large Gatekeepers" (2021) 11 EuCML: Journal of European Consumer and Market Law 60; A Davola and G Malgieri, "Data, Power and Competition Law: The (Im)Possible Mission of the DMA?" (2022) 2023 Research in Law and Economics, Forthcoming.

[16] See, eg, M Zehlike et al, "Beyond Incompatibility: Interpolation between Mutually Exclusive Fairness Criteria in Classification Problems" (2022) arXiv:2212.00469; P Hacker, "KI und DMA– Zugang, Transparenz und Fairness für KI-Modelle in der digitalen Wirtschaft " (2022) 75 Gewerblicher Rechtsschutz und Urheberrecht 1278.

[17] They are limited to medical AI, credit scoring, life and health insurance and employment; see Annexes II and III of the AI Act Proposal.

[18] For the "A9" algorithm, see V Sandeep and B Pohutezhini, "The e-commerce revolution of Amazon.com" (2019) 6 Splint Internationals Journal of Professionals 33, 37; N Maio and B Re, "How Amazon's e-commerce works" (2020) 2 International Journal of Technology for Business 8, 10; T Fries, "Amazon A9 – Amazon's ranking algorithm explained" (*Amalytix*, 10 December 2020) <https://www.amalytix.com/en/knowledge/seo/amazon-alogrithm-a9/> (last accessed 4 December 2022).

[19] A-J Su et al, "How to improve your Google ranking: Myths and reality" (2010) 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology 50; MG Southern, "Is RankBrain a Ranking Factor in Google Search?" (*SEJ*, 5 October 2022) <https://www.searchenginejournal.com/ranking-factors/rankbrain-ranking-factor> (last accessed 4 December 2022).

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

For the first time in history, the EU has now enacted, with the DMA, provisions explicitly and specifically regulating such rankings and other AI applications of gatekeepers. Due to the centrality of these applications to gatekeeper technology implementation and business, the provisions are likely to have a profound effect on the core systems of the digital economy. Although the term "artificial intelligence" is absent from the entire DMA, this paper identifies four regulatory complexes in the DMA that will significantly change the legal framework for AI in market transactions in the EU: prerequisites for the creation of fair rankings; information requirements; regulation of training data; and access rights.

In doing so, this paper proceeds in several steps. First, we focus on the provisions that have the potentially most far-reaching effect on the use of AI applications by gatekeepers: those governing fairness (ie the very nature and order of rankings; Section II). This section also includes an overview of the existing regulations and practices in the field, as well as an evaluation of the newly proposed rules. Next, the paper examines and critiques the new DMA rules regulating training data (Section III), access rights (Section IV) and information requirements (Section V). Concerning training data and transparency, the DMA rules are contrasted with the relevant counterparts in the AI Act. Regarding access rights, we compare the DMA provisions with the current proposal of an EU Data Act (DA). The final substantive section maps out a framework for fairness, transparency and access in gatekeeper AI going forward (Section VI). Section VII concludes.

II. Regulating AI-based rankings

Rankings have infiltrated virtually all areas of our lives. The prominence given to goods or services offered by online providers undeniably has the power to steer and potentially control our choices.[20] The vast amount of offerings fighting for our attention has created the need for an intermediary who supports us in our decision-making process by organising and prioritising these offerings. This directly results from the digital economy we live in. By controlling the demand, search and buying behaviour of individuals, rankings make up the boiler room of this new economy. Unsurprisingly, therefore, the business models of many tech giants use them as the basis for their competitive advantage.[21] This competitive advantage is even more prominent where the entity also directly sells or provides certain services or products, in addition to their role as an intermediary (vertical integration).[22]

The European Commission has identified the immense impact that the control of rankings has on market power as a potential problem and, therefore, included specific regulations on their creation and implementation in the DMA. After a short introduction to the differences between AI-based and traditional rankings (Section II.1), this section will provide an overview of the existing legal framework applicable to rankings and contrast it in detail with the new rules introduced by the DMA. In doing so, the section covers the most important regulatory dimensions of rankings: transparency (Section II.2), accuracy and rectification (Section II.3) and fairness/non-discrimination (Section II.4).

[20] See, eg, T-H Teh and J Wright, "Intermediation and steering: competition in prices and commissions" (2022) 14 AEJ: Microeconomics 281.

[21] See supra, notes 18 and 19.

[22] See, eg, J Padilla, J Perkins and S Piccolo, "Self-Preferencing in Markets with Vertically Integrated Gatekeeper Platforms" (2022) 70 The Journal of Industrial Economics 371; A Hagiu, T-H Teh and J Wright, "Should platforms be allowed to sell on their own marketplaces?" (2022) 53 The RAND Journal of Economics 297.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

1. Rankings: a cornerstone of the digital economy

A ranking is an ordered list[23] typically displayed in response to a search entry.[24] In the DMA, ranking is understood as the "relative prominence given to goods or services" or "the relevance given to search results by online search engines [. . .]" (cf. Article 2(22) DMA), irrespective of the technical means used for such presentation.[25] Such a ranking is created, for example, when a consumer searches for a refrigerator on a comparison or online shopping platform. In this sense, rankings establish a pre-selection of goods or services to facilitate consumers' purchasing decisions.[26]

a. Traditional versus AI-based rankings

In theory, a ranking can be manually created: a person could put specific entries on a list in a particular order based on specific predetermined criteria. One may conceive a shopkeeper who sorts certain goods (eg hard drives) according to their writing speed, then notes this on a piece of paper and displays it visibly in their shop to help their customers with their decisions. Indeed, according to Article 2(22) DMA, relative prominence or relevance qualifies a list as a ranking, regardless of the technical means by which one creates it. The legal definition of the DMA says nothing about the technical means themselves. To automate this process, one could also record the writing speed of the hard drives in a machine-readable way and rank it using a simple sorting algorithm.[27]

With vast amounts of data and accompanying large amounts of potentially relevant criteria for ordering, it has become possible and popular to resort to ML techniques to order the data and items efficiently.[28] Such rankings thus differ from the traditional rankings described above primarily in how the data are processed and rankings are personalised. With the release of ChatGPT[29] and implementing this model in Microsoft's search engine Bing,[30] there are now other ways to get a ranking: via dialogue with a chatbot.

b. The centrality of rankings for online platforms

Data processing creates specific legal challenges of data protection, transparency, accuracy and fairness, which the law must address. Regulating for transparent, accurate and fair

[23] R Ursu, "The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions" (2018) 37 Marketing Science 530; M Mohri, A Rostamizadeh and A Talwalkar, *Foundations of Machine Learning* (2nd edition, Cambridge, MA, MIT Press 2018) p 3.

[24] Hacker, supra, note 16, 1281.

[25] The definition essentially corresponds to that of Art 2(8) of Regulation (EU) 2019/1150 on fairness and transparency for business users of online intermediary services (P2B Regulation); on the P2B Regulation, see Section II.2.a infra.

[26] Ursu, supra, note 23.

[27] See, eg, AD Mishra and D Garg, "Selection of best sorting algorithm" (2008) 2 International Journal of Intelligent Information Processing 363; however, the sorting algorithms that can be used differ, sometimes considerably, in their efficiency (eg in terms of their runtime and the memory they require). For an overview, see, eg, R Sedgewick and K Wayne, *Algorithms* (4th edition, Boston, MA, Addison-Wesley 2011) ch 2.

[28] Google, for example, uses ML techniques to best understand search queries and provide the most relevant answers; see B Schwartz, "How Google uses artificial intelligence in Google Search. From RankBrain, Neural Matching, BERT and MUM – here is how Google uses AI for understanding language for query, content and ranking purposes" (*Search Engine Land*, 3 February 2022) <https://searchengineland.com/how-google-uses-artificial-intelligence-in-google-search-379746> (last accessed 6 December 2022); see also supra, note 19. Amazon uses such techniques to, among other things, suggest ads based on the previous search behaviour of potential buyers, see supra, note 18.

[29] OpenAI, "Introducing ChatGPT" (*Official OpenAI Blog*, 30 November 2022) <https://openai.com/blog/chatgpt> (last accessed 5 May 2023).

[30] Y Mehdi, "Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web" (*Official Microsoft Blog*, 7 February 2023) <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/> (last accessed 27 March 2023).

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

rankings is crucial for three reasons. First, as mentioned, for many gatekeepers, the ability to produce high-quality rankings often constitutes a pivotal point of their business model, as their competitive advantage often lies precisely in producing fast and reliable rankings that are relevant.[31] This is also illustrated by the example of Microsoft's Bing search engine. By implementing ChatGPT,[32] the search engine has been seriously competing with its more popular competitor Alphabet and its search engine Google again for quite some time. It is not all that new for search engines to leverage the capabilities of language models; Google, for example, has been using language models to better understand user queries for a while now.[33] However, with implementing large generative AI models in search engines, it is now also possible to have a product or service displayed via a chat interface. Second, as empirical studies show, the ranking order has a considerable impact on decisions made by the consumer.[34] Finally, gatekeepers typically use ML techniques for the generation of such rankings.[35] Nevertheless, many applications of rankings in the digital economy, particularly in e-commerce or social media, do not qualify as high risk under the AI Act[36] and will therefore only be regulated by existing EU legislation, including the new Article 6(5) DMA. In this respect, rankings generated by large generative AI models are an exception.[37] However, large generative AI models do not create all rankings.

It is important, however, to point out that the DMA rules for rankings, such as Article 6(5) DMA, are based on and closely related to previously existing provisions. An increasing number of regulations are dedicated to the transparency (Section II.2), accuracy (Section II.3) and fairness (Section II.4) of online rankings.

2. *Transparency of rankings*

A stumbling block towards effective functioning and regulation of rankings is the information asymmetry between the platforms and all other stakeholders.[38] In recent years, the EU has therefore adopted several critical regulations on the transparency of rankings. These include the P2B Regulation,[39] the Consumer Rights Directive (CRD)[40] and the Unfair Commercial Practices Directive (UCPD),[41] as updated by the Omnibus

[31] B Martens, "An Economic Policy Perspective on Online Platforms" (2016) Institute for Prospective Technological Studies Digital Economy Working Paper 2016/05, 4, 20 et seqq.

[32] See supra, note 30.

[33] See supra, note 28, and also supra, note 19.

[34] Ursu, supra, note 23; M Derakhshan et al, "Product Ranking on Online Platforms" (2022) 68 Management Science 4024, 4028.

[35] T-Y Liu, "Learning to Rank for Information Retrieval" (2009) 3 Foundations and Trends in Information Retrieval 225.

[36] See Annexes II and III AI Act.

[37] P Hacker, A Engel and M Mauer, "Regulating ChatGPT and Other Large Generative AI Models", Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (Association for Computing Machinery 2023) <https://dl.acm.org/doi/10.1145/3593013.3594067> (last accessed 17 July 2023).

[38] Bongartz et al, supra, note 15, 61.

[39] Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services, OJ L 186, 11.7.2019, p 57.

[40] Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council, OJ L 304, 22.11.2011, p 64.

[41] Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council, OJ L 149, 11.6.2005, p 22.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

Directive,[42] respectively; the General Data Protection Regulation (GDPR)[43]; and the DMA, as well as the proposed AI Act. As we shall demonstrate, the DMA brings significant changes to this framework: not so much in terms of content, but enforcement.

a. P2B Regulation

The P2B Regulation has the aim of creating a fair, transparent and predictable environment for smaller businesses that are forced to rely on search engines and other online intermediation services in order to conduct their activities efficiently. The Regulation targets online brokerage services and search engines in general and is not limited to market-dominant undertakings. According to Article 5(1) and (2) of the P2B Regulation, both entities must disclose the main parameters that determine the ranking.[44] Mediation services must also disclose the reasons for the relative weighting of these parameters; search engines only have to disclose the relative weighting of the main parameters themselves. These requirements, while seemingly innocuous at first glance, are challenging concerning AI systems. For a long time, a dispute has raged about the extent to which current data protection law requires the disclosure of individual parameters (*features*) that the AI model analyses, and of their relative weighting.[45] This is relevant because general statements on the most important parameters determining the model's predictions are difficult or hardly possible to obtain for some advanced types of ML.[46] With support vector machines or artificial neural networks, in particular, techniques exist to explore, and only approximately, which features were decisive for an individual, concrete prediction (so-called *local* explanation).[47] It is technically much more difficult to determine those features that generally (ie concerning all decisions) are the most relevant ones (so-called *global* explanation).[48] However, Article 5 of the P2B Regulation arguably requires precisely this: the disclosure of general main parameters (ie global explanations of the AI model).[49] This requirement is not only a considerable legal innovation but also poses significant challenges for developers, especially when using artificial neural networks, which are often particularly potent.

[42] Directive 2019/2161/EU of 27 November 2019 on better enforcement and modernisation of Union consumer protection laws, OJ L 328. 18.12.2019, p 7.

[43] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC OJ L 119, 4.5.2016, p 1.

[44] The term "parameter" is poorly chosen, as in technical terms it is understood to mean the internal coefficients of the model and not the factors relevant to the decision, such as price, availability, etc. However, this seems to be meant in the case of the P2B Regulation and the CRD, see Recital 24 P2B Regulation and Recital 22 Omnibus Directive; see also C Alexander, "Neue Transparenzanforderungen im Internet – Ergänzungen der UGP-RL durch den 'New Deal for Consumers'" (2019) WRP 1235, marginal no. 30. Technically, the decision factors are rather called "features" (Goodfellow et al, supra, note 3, 3, 292 f).

[45] See, eg, LK Kumkar and D Roth-Isigkeit, "Erklärungspflichten bei automatisierten Datenverarbeitungen nach der DSGVO" (2020) JZ 277.

[46] Lipton, supra, note 4.

[47] SM Lundberg and S-I Lee, "A unified approach to interpreting model predictions" (2017) 30 Advances in Neural Information Processing Systems 4765; AB Arrieta et al, "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI" (2020) 58 Information Fusion 82, 92 et seqq.

[48] Alejandro Barredo Arrieta et al, supra, note 47, 90; for approaches, see, eg, S Lapuschkin et al, "Unmasking Clever Hans predictors and assessing what machines really learn" (2019) 10 Nature Communication 1; likewise, the calculation of an average of so-called Shapley values, which yield local feature relevance, is possible, cf. Lundberg and Lee, supra, note 47.

[49] Recital 24 P2B Regulation; M Grochowski et al, "Algorithmic Transparency and Explainability for EU Consumer Protection: Unwrapping the Regulatory Premises" (2021) 8 Critical Analysis of Law 43, 52; A Bibal et al, "Legal requirements on explainability in machine learning" (2021) 29 Artificial Intelligence & Law 149, 161; Hacker and Passoth, supra, note 14, 343, 364.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

b. Consumer Rights Directive

Similar considerations apply to the amendment of the CRD. As with the P2B Regulation, the CRD is applicable regardless of the market power of the addressees but sets obligations vis-à-vis consumers. According to the new Article 6a CRD, online marketplaces[50] must disclose the main parameters for rankings based on consumer search queries and their relative weighting. The provision's wording suggests an obligation to provide a local explanation, which would render explanations more feasible in case of advanced ML techniques, such as artificial neural networks, as seen. However, Recital 21 of the Omnibus Directive clearly states that the ranking-related transparency obligations of the CRD should mirror those of the P2B Regulation. Recital 23 of the Omnibus Directive also emphasises that traders owe no disclosure in individual cases.[51] Therefore, the obligation to provide global explanations remains, raising the aforementioned implementation challenges in handling artificial neural networks.

c. Unfair Commercial Practices Directive

The new Article 7(4a) of the UCPD accompanies the CRD disclosure obligation. It now qualifies the compulsory information on the main parameters laid down in Article 6a CRD as essential information in the sense of the prohibition of misleading information under unfair competition law. Hence, every violation of the CRD requirement automatically constitutes an unfair commercial practice, triggering, inter alia, legal action by competitors not foreseen under the CRD (Article 11(1) UCPD). In practice, these proceedings are among the most effective incentives for compliance by businesses with unfair competition law.

Simultaneously, Article 7(4a) UCPD applies to all entrepreneurs who enable a search for products of different suppliers.[52] Finally, according to point 11a of Annex I, it is considered unfair if payments or paid advertising used to achieve a higher ranking position are not clearly disclosed by the platform.[53] Such ranking categories must now be unambiguously identified as "sponsored" or marked similarly.

Overall, the UCPD provisions do not merely repeat the CRD obligations; rather, they significantly raise the likelihood that the CRD obligations do not remain a paper tiger but have a real effect in the digital economy.

d. The General Data Protection Regulation

Even though it predates the Omnibus Directive and the P2B Regulation, the GDPR has gone even further in establishing transparency as one of the main principles of data protection (Article 5(1)(a) GDPR). The principle is primarily operationalised through the rights to information (Articles 12–14 GDPR) and access (Article 15 GDPR). For any automated decision-making (including profiling), data controllers are required to disclose "meaningful information" about the logic involved as well as the significance and the envisaged consequences of such processing for the data subject (Articles 13(2)(f), 14(2)(g), 15(1)(h) GDPR). As has been previously outlined, most forms of truly efficient, high-quality AI-based rankings are currently created by analysing either the previous behaviour of the specific person to which the information is being presented or the average consumer

[50] According to the new Art 2(1)(n) of the UCP Directive, an online marketplace is a "service enabling consumers to conclude distance contracts with other traders or consumers through the use of software, including a website, part of a website or an application operated by or on behalf of the trader".

[51] See also Alexander, supra, note 44, para 34.

[52] ibid, para 29.

[53] In addition, the terms and conditions for sponsored ranking must be disclosed to business clients, Art 5(3) P2B Regulation.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

behaviour in general. Both of these determinations require the processing of personal data, thereby falling under the scope of the GDPR.[54]

*i. Additional requirements for automated decision-making***.** The existence of automated decision-making on its own is not necessarily sufficient to trigger the outlined transparency requirement, however. Articles 13–15 refer to Article 22(1) GDPR, which famously adds two additional requirements: the decision must be based *solely* on automated processing; and it needs to produce legal or *similarly significant* effects for the individual. In some cases of AI-based ranking, controllers may by now have included human involvement to avoid scrutiny under these provisions.[55] Typical website or product rankings will, however, be built purely automatically by gatekeeper AI, raising the question of significant effects.

Under the Article 29 Data Protection Working Party Guidelines on Automated Individual Decision-Making and Profiling (WP Guidelines), decisions qualify if they have the potential to significantly affect the circumstances, behaviour or choices of the individual or if they have a prolonged or permanent impact.[56] Since empirical studies have evidenced that rankings do, in fact, have considerable relevance to consumer decisions,[57] the possibility of such an impact cannot be excluded a priori. While searches conducted on trivial matters might not impact us relevantly, there are certainly instances in which the type and order of information presented can have serious ramifications because of the kind of product (eg insurance) or time sensitivity (eg information on poisons). Controllers might argue that such consequences are purely hypothetical. Still, if a misguided decision was, in fact, made based on a ranking, an individual could retroactively seek an explanation regarding the reasons (Article 15(1)(h) GDPR); and controllers reasonably expecting such effects need to proactively disclose the required information (Articles 13(2)(f), 14(2)(g) GDPR). In our view, some ranking models that are commonly used should be considered significant due to a cumulative effect. If individuals use Google every day, in almost all areas of their lives, an overall significant impact can hardly be denied. In that sense, gatekeeper AI would – barring human involvement – automatically fall under the definition of Article 22(1) GDPR.

*ii. Meaningful information***.** As a result, meaningful information must be provided on the logic involved. Like numerous terms in the GDPR, the formulation "meaningful information" is inherently vague, leaving vast room for (mis)interpretation. Since it is part of the transparency requirements, which are meant to provide to data subjects an understanding concerning the processing of their personal data, it should most likely be understood as in "meaningful to the data subject".[58] Hence, information about the "logic involved in the processing" should focus, at a minimum, on the general rationale behind the system phrased in a consumer-friendly way rather than making available a specific

[54] See, eg, FZ Borgesius, "Personal data processing for behavioural targeting: which legal basis?" (2015) 5 International Data Privacy Law 163.

[55] S Wachter, B Mittelstadt and L Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation" (2017) 7 International Data Privacy Law 76, 88.

[56] Article 29 Data Protection Working Party, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679" (2017) WP 251, 21.

[57] Ursu, supra, note 23; Derakhshan et al, supra, note 34.

[58] Article 29 Data Protection Working Party, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679" (2017) WP 251, 25; AD Selbst and J Powles, "Meaningful information and the right to explanation" (2017) 7 International Data Privacy Law 233, 236; B Custers and A-S Heijne, "The right of access in automated decision-making: The scope of Article 15(1)(h) GDPR in theory and practice" (2022) 46 Computer Law & Security Review 105727.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

algorithm or model.[59] Simultaneously, the controller would need to disclose enough information for the individual to form an understanding about the underlying logic and exercise their data subject rights under Articles 15–22 GDPR, where needed.[60]

While the wording in Articles 13–15 GDPR is identical, the rights should still be viewed separately.[61] In their privacy policies, controllers must opt for *general* descriptions, taking into consideration the average consumer (global explanations); this, however, arguably does not apply to access requests. Here, the data subject has the option to specify what kind of information would interest them (eg storage period, feature relevance) and hence, by selecting specific information, to shape the very meaning of "meaningful". Whether this, however, includes a subjective right to an explanation regarding a specific decision made by the AI is still subject to heated debate.[62] As is well known, the idea of a specific justification is supported by Sentence 4 of Recital 72, which lists a right of the data subject "to obtain an explanation of the decision reached after such assessment and to challenge the decision" as one of the safeguards to be implemented. Also, Article 29 WP Guidelines, while confirming that the disclosure of the full algorithm would not be required, still underlines that the provided information should enable the individual "to understand the reasons for the decision",[63] further supporting the idea of individual (= local) explanations. While none of these interpretations are binding, they do indicate a general tendency and are in line with the purpose of the access right: to provide data subjects with enough information to understand and contest data processing by exercising their data subject rights. Arguably, in numerous scenarios, an effective contestation presupposes a local explanation.[64] Hence, a flexible approach will have to be taken by controllers, adapting the information provided to the circumstances of each case, as long as the matter has not been decided by the Court of Justice of the European Union (CJEU).

iii. Emerging case law. Indeed, while the CJEU has not decided on algorithmic transparency yet, cases are starting to emerge in Member State courts.[65] Most notably, in March 2021, the District Court of Amsterdam ruled in a case against the ridesharing company Ola, which operates a platform similar to Uber in the Netherlands.[66] The company had

[59] G Zanfir-Fortuna, "Article 13: Information to be provided where personal data are collected from the data subject", in C Kuner et al (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (online edition, Oxford, Oxford Academic 2020) p 430.

[60] Article 29 Data Protection Working Party, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679" (2017) WP 251, 25; Hacker and Passoth, supra, note 14, 343, 349; Selbst and Powles, supra, note 58, 236; Custers and Heijne, supra, note 58, 5.

[61] M Bäcker, "Article 15" in J Kühling and B Buchner (eds), *Datenschutzgrundverordnung BDSG* (3rd edition, Munich, CH Beck 2020) para 27; P Bräutigam and F Schmidt-Wudy, "Das geplante Auskunft- und Herausgaberecht des Betroffenen nach Article 15 der EU-Datenschutzgrundverordnung" (2015) CR 56, 62. The opposite is argued in LK Kumkar and D Roth-Isigkeit, "A Criterion-Based Approach to GDPR's Explanation Requirements for Automated Individual Decision-Making" (2021) 12 JIPITEC, the Journal of Intellectual Property, Information Technology and Electronic Commerce Law 289, 296.

[62] G Malgieri and G Comandé, "Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation" (2017) 7 International Data Privacy Law 243; Wachter et al, supra, note 55, 76–99; Selbst and Powles, supra, note 58, 233–42. M Brkan, "Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond" (2019) 27(2) International Journal of Law and Information Technology 91, 110–19.

[63] Article 29 Data Protection Working Party, supra, note 56, 25.

[64] See also sources cited supra, note 60.

[65] See, eg, (French) Conseil Constitutionnel, Décision n° 2020-834 QPC du 3 avril 2020, Parcoursup; (Dutch) Rechtbank Den Haag, Case C-09-550982-HA ZA 18-388, SyrRI, ECLI:NL:RBDHA:2020:1878; (Italian) Corte Suprema di Cassazione, Judgment of 25 May 2021, Case 14381/2021,

[66] District Court of Amsterdam, Case C/13/689705/HA RK 20-258, Ola, ECLI:NL:RBAMS:2021:1019 (Ola Judgment); see also R Gellert, M van Bekkum and FZ Borgesius, "The Ola & Uber judgments: for the first time a court recognises a GDPR right to an explanation for algorithmic decision-making" (*EU Law Analysis*, 28 April 2021)

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

established an algorithmic model for automatically penalising drivers if a ride was cancelled or invalid. The drivers sued to enforce their right of access under Article 15(1)(h) GDPR. In its judgment, the court interpreted the clause in favour of the applicants and required the platform to explain the logic involved in the decisions. More specifically, it recurred on Article 29 WP Guidelines, according to which meaningful information about the logic involved implies the disclosure of "criteria relied on in reaching the decision".[67] The court concluded that "Ola must communicate the main assessment criteria and their role in the automated decision to [the drivers], so that they can understand the criteria on the basis of which the decisions were taken and they are able to check the correctness and lawfulness of the data processing".[68]

In the view of the court, the GDPR access right implies an explanation of algorithmic decisions by feature relevance (main assessment criteria and role). The wording ("automated decision") and purpose of the judgment (correctness check) suggest that these explanations must be individually personalised. This means that, as suggested by our above analysis, local explanations of the individual decisions are required. As mentioned, such explanations can be furnished by post-hoc explanation models, such as the Shapley Additive Explanation (SHAP), even in the case of highly complex, "black box" artificial neural networks.[69] Furthermore, the judgment rightly highlights that explanations must be adapted to the cognitive and educational background of data subjects.

The judgment was until recently on appeal with a positive outcome for the plaintiffs – the Court of Appeals predominantly upheld the District Court's decision, supporting the notion that platforms were required to explain the logic involved in their decisions.[70] Regarding the requirements to comply with the right to be informed under Article 15(1)(h) GDPR, the appeals court clarified in paragraph 3.48, referring to the WP Guidelines, that the information must be provided in such a way that the data subject can decide with sufficient knowledge whether to exercise their rights guaranteed by the GDPR. This information includes, in particular, that about the factors that were taken into account in the decision-making process and their respective "weighting" at an aggregate level. The information must also be complete enough for the data subject to understand the reasons for the decision. The court also says that it follows from the guidelines that this does not have to be a complex explanation of the algorithms used or a presentation of the entire algorithm. However, it says that what is required is at least information about what factors and the weighting of those factors Ola used to arrive at the ratings at issue, as well as a statement of information necessary to understand why.

This underlines the importance of purpose-driven interpretation of Article 15(1)(h) GDPR, whose obligations require information meaningful for the respective data subjects to enable them to exercise their respective rights. More specifically, in our view, explanations must conform to reasonable expectations of the data controller concerning the data subjects' understanding. In sum, Article 15(1)(h) GDPR demands explanations of

<https://eulawanalysis.blogspot.com/2021/04/the-ola-uber-judgments-for-first-time.html/> (last accessed 12 December 2022).

[67] Article 29 Data Protection Working Party, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679" (2017) WP 251, 25.

[68] Ola Judgment, para 4.52; translation according to Anton Ekker, "Dutch court rules on data transparency for Uber and Ola drivers" (*Ekker Blog*) <https://ekker.legal/en/2021/03/13/dutch-Court-rules-on-data-transparency-for-uber-and-ola-drivers/> (last accessed 12 December 2022).

[69] See supra, note 45.

[70] The judgment can be found here: <https://uitspraken.rechtspraak.nl/#!/details?id=ECLI:NL:GHAMS:2023:804>; an unofficial translation can be found here: <https://5b88ae42-7f11-4060-85ff-4724bbfed648.usrfiles.com/ugd/5b88ae_de414334d89844bea61deaaebedfbbfe.pdf>; see also J Turner, "Amsterdam Court Upholds Appeal in Algorithmic Decision-Making Test Case: Drivers v Uber and Ola" (*Fountain Court Blog*, 4 June 2023) <https://www.fountaincourt.co.uk/2023/04/amsterdam-court-upholds-appeal-in-algorithmic-decision-making-test-case-drivers-v-uber-and-ola/> (last accessed 22 November 2023).

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

automated decision-making systems, including AI, that are concrete, local and adapted to the respective audience.

*iv. Large generative AI models***.** Also noteworthy in the context of reviewing the emerging case law is the temporary ban put on ChatGPT by the Italian Data Protection Authority. In March 2023 the Garante per la protezione dei dati personali imposed an immediate temporary limitation on the processing of Italian users' data by OpenAI.[71] While the service was quickly reinstated for Italian users, based on a list of measures deployed by the provider,[72] concerns with regards to the compliance of such services with the GDPR remain.[73]

Large generative AI models will not always create a ranking in the context of providing the user with an answer. As has been outlined above, however, the interaction with the chatbot will still frequently require the generation of rankings for the purpose of submitting a reasonable response.[74] The question of transparency in the context of large generative AI models such as ChatGPT is also not limited to the above considerations regarding the provisions of meaningful information about the logic involved in the automated decision-making process under Article 22(1), but also affects the general question of transparency in accordance with Articles 13 and 14 of the GDPR. On the one hand, the service provider is collecting personal data directly from the user, in the form of processing their queries and requests (Article 13), while on the other hand personal data have likely been used for the training of the model itself without any prior contact to the data subject (Article 14). Whereas the provision of relevant information concerning the processing of user and account-related data should be relatively easy to achieve, defining parameters such as the affected data subjects, the categories of data concerned or the length of the processing operation in the context of AI training data will be virtually impossible. Based on that, controllers might be inclined to use the exemption of "disproportionate effort" established in Article 14(5)(b).[75] Should that be the case, the same Article would, however, still require them to introduce "appropriate measures to protect the data subject's rights and freedoms and legitimate interests" and arguably perform a balancing test between the effort and the interest of the individual.[76] How this could be achieved by the provider and to what extent they would even be able to use the mentioned exemption, considering that their data processing is not done for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, is highly questionable. The general lack of transparency of those models is particularly problematic when taking into account the identified risk of inversion attacks leading to the reproduction of the used training data.[77] In that sense, further steps from the relevant data protection authorities would be welcome.

[71] See: Garante per la protezione dei dati personali, "Intelligenza artificiale: il Garante blocca ChatGPT. Raccolta illecita di dati personali. Assenza di sistemi per la verifica dell'età dei minori" <https://www.garanteprivacy.it:443/home/docweb/-/docweb-display/docweb/9870847> (last accessed 17 July 2023).

[72] <https://www.gpdp.it/home/docweb/-/docweb-display/docweb/9881490#english> (last accessed 10 May 2023).

[73] Hacker et al, supra, note 37; J Möller-Klapperich, "ChatGPT und Co. – aus der Perspektive der Rechtswissenschaft" (2023) 4 Neue Justitz 144; I Goodfellow et al, "Generative adversarial networks" (2020) 63 Communications of the ACM 11, 139–44; "Press Release: LfDI informiert sich bei OpenAI, wie ChatGPT datenschutzrechtlich funktioniert" <https://www.baden-wuerttemberg.datenschutz.de/lfdi-informiert-sich-bei-openai-wie-chatgpt-datenschutzrechtlich-funktioniert/> (last accessed 11 May 2023).

[74] See, eg, Hacker et al, supra, note 37.

[75] ibid.

[76] Art 29 WP, Guidelines on transparency under Regulation 2016/679, 17/EN WP260 rev.01, para 64.

[77] Goodfellow et al, supra, note 73; N Carlini et al, "Extracting Training Data from Diffusion Models" (arXiv, 30 January 2023) <http://arxiv.org/abs/2301.13188> (last accessed 17 July 2023); R Plant et al, "You Are What You Write: Preserving Privacy in the Era of Large Language Models" (April 2022) <https://www.researchgate.net/publication/360079388_You_Are_What_You_Write_Preserving_Privacy_in_the_Era_of_Large_Language_Models> (last accessed 17 July 2023).

e. The Digital Markets Act

The recently enacted DMA adds to the plethora of transparency requirements under EU law. Like the P2B Regulation and the CRD, the DMA contains specific transparency requirements for rankings in its Article 6(5)(2): the gatekeeper must carry out the ranking based on transparent conditions. However, the concept of transparency is not further defined or elaborated upon. This raises the question of how the DMA's transparency obligation should be interpreted: does it go beyond those of the P2B Regulation, does it fall short of them or are they rather congruent? In our view, the latter interpretation is the most convincing.[78]

To start with, nothing in the Recitals of the DMA suggests that specific types of transparency or explanation (eg local or global explanations, sensitivity lists or contrasting explanations[79]) are required under Article 6(5) DMA. Rather, Recital 52 DMA does not refer to the GDPR but to the Commission's Guidelines on transparency of rankings under the P2B Regulation, which is intended to facilitate the implementation and enforcement of Article 6(5) DMA. If, however, the P2B Regulation guidelines are supposed to provide guidance for the implementation of the DMA transparency provisions as well, both requirements need to be identical.

The controversies surrounding the transparency requirements for AI-based models in Article 13 AI Act[80] (see the next subsection) and Articles 13–15 GDPR[81] show that various approaches are conceivable in factual and legal terms. A completely independent interpretation of the DMA transparency requirements, detached from the surveyed existing obligations, would entail significant legal uncertainty and practical implementation difficulties. Hence, overall, the DMA transparency requirement, in our view, mirrors the P2B Regulation requirement and is not influenced by the – differently worded – GDPR rules just analysed.

This is not the end of the story, though. The P2B Regulation leaves enforcement to the Member States (Article 15 P2B Regulation). It only has to be, as always, "effective, proportionate and dissuasive". The DMA, by contrast, spells out the most significant enforcement apparatus in the entire digital economy, including the AI Act. First, the Commission is tasked with enforcement, being a highly experienced and well-funded body regarding the prosecution of competition rule violations. Second, fines range up to 10% of global annual turnover, and even 20% in cases of repeat infringements (Article 30 DMA) – amounts unmatched by any other instrument, including the AI Act (which will be limited to fines in the order of 6–7% of global annual turnover; see Article 71(3) AI Act European Parliament Version). Third, the Commission is endowed with sweeping powers, including for conducting market investigations (eg Article 16 DMA). Representative actions may also be brought, unlike under the AI Act (Article 42 DMA). Overall, transparency provisions are endowed with unmatched bite by the DMA, even though the substance refers back to the P2B Regulation.

[78] See also D Brouwer, "Towards a ban of discriminatory rankings by digital gatekeepers? Reflections on the proposal for a Digital Markets Act" (*Internet Policy Review*, 11 January 2021) <https://policyreview.info/articles/news/towards-ban-discriminatory-rankings-digital-gatekeepers-reflections-proposal-digital> (last accessed 7 December 2022), according to which fairness was already to be understood as transparency in the sense of the P2B Regulation under the Commission draft.

[79] See, eg, the overview in A Holzinger et al, "xxAI – Beyond Explainable Artificial Intelligence" in A Holzinger et al (eds), *xxAI – Beyond Explainable AI, International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers* (Berlin, Springer 2022) p 13; Arrieta et al, supra, note 47; for an account of their accuracy, see A Alonso and JM Carbó, "Accuracy of Explanations of Machine Learning Models for Credit Decision" (2022) Banco de España Working Paper 2222.

[80] Hacker and Passoth, supra, note 14, 343, 358 et seqq.

[81] See only Wachter et al, supra, note 55; Selbst and Powles, supra, note 58; Kumkar and Roth-Isigkeit, supra, note 45.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

f. The AI Act

The AI Act was designed to foster trustworthy AI,[82] of which transparency is a key element.[83] Hence, one would have expected the AI Act to provide precise and meaningful requirements for transparency and explanations in AI systems that go beyond the status quo in general EU law and the generic reference in Article 6(5) DMA. Unfortunately, however, the opposite is the case.[84]

The key transparency requirement for high-risk AI systems is contained in Article 13 AI Act. According to this provision, some specific information needs to be disclosed, such as "the level of accuracy, including its metrics, robustness and cybersecurity" as well a description of the input and training data and expected output.[85]

These provisions, however, are generic, even though they do go beyond the current disclosure framework in many respects. Regarding rankings, however, Article 11, in conjunction with Annex IV AI Act, provides for more detailed regulation. For example, a description of the system architecture needs to be offered, as well as an assessment of potentially discriminatory impacts and information on the "relevance of the different parameters" (Annex IV(2)(b) AI Act). Taken together, these provisions therefore establish a disclosure regime for fairness metrics (discriminatory impact) and global explanations of the model in terms of the relevance of its features (ie the factors that mainly contribute to its output, averaged over all cases).[86] While the explanation requirement matches the provisions of the P2B Regulation and the updated CRD and UCPD (see above) but also replicates their implementation problems, the novel fairness disclosure hints at the increased relevance of fairness in AI systems and rankings (see Section II.4 infra).

All of these disclosure obligations improve the ability of business users and end users, as well as associations, non-governmental organisations (NGOs) and public authorities, to gain insight into the ranking criteria of online intermediary services, search engines and other ranking providers. However, they do not formulate any further substantial fairness requirements for the arrangement of the elements of the ranking itself. Such substantive obligations may, however, be found in the GDPR and other parts of the DMA.

3. Accuracy: rectification of rankings

Another data subject right found in the GDPR, which is linked to the principle of accuracy (Article 5(d) GDPR), rather than transparency, but which also potentially affects AI-based ranking, is the right to rectification enshrined in Article 16 GDPR. It provides that data subjects may obtain rectification of their personal data if they are inaccurate. This, of course, would apply across the board to any datasets processed by gatekeepers, or other data controllers, including those that rankings are based on. What is unclear, however, is the extent to which the right of rectification could apply to the ranking itself.

For example, the underlying information may have been correct but the individual in question might still consider the *outcome* to be wrong. Imagine that a person frequently travels to London for work and spends significant time online researching opportunities to buy tickets for a Chelsea FC football match. She is simply intrigued by the new setup of the team after a change of ownership. As a result, whenever she searches for pullovers online, the first items on the ranking invariably include some merchandise from Chelsea. This, however, is perceived as deeply offensive by the London traveller as she is, in reality, an

[82] Recitals 5 and 62 AI Act.

[83] W Pieters, "Explanation and trust: what to tell the user in security and AI?" (2011) 13 Ethics and Information Technology 53.

[84] See, eg, Hacker and Passoth, supra, note 14, 343, 357 et seqq.

[85] See Arts 13(3)(b)(ii) and (iv)–(vi) AI Act.

[86] See, eg, Lapuschkin et al, supra, note 48.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

inveterate fan of Bayern Munich but never wears any football merchandise. She is not opposed to personalised rankings but would like them to be accurate. Does she then have a right for these personalised rankings to be rectified to better match her preferences?

a. Rankings as personal data

Since the right to rectification only applies to personal data, the first question is whether the ranking would even fall under that definition. As per Article 4(1) GDPR, the term includes *any information relating to an identified or identifiable natural person.* Non-personalised rankings in the sense of prominence given to goods or services offered online or the relevance of search results (Article 2(22) DMA) do not generally fall under that definition. While rankings may *contain* personal data (eg if one googles a person), the ranking model itself solely constitutes a grouping of the relevant elements, based on the preferences of the average user.[87]

To achieve more accurate results, however, providers mostly rely on personalised recommendations, based on criteria such as preferences disclosed by the user, previous activities or social connections.[88] The presented outcome is therefore created *specifically* for that individual. To be personal data, though, it also has to *relate* to that person, whereas in this case it is rather *presumed to relate*, according to the model. While the simple fact of potential inaccuracy is not a reason for exclusion (after all, most personal data can be incorrect as well),[89] the legal status of inferences is questionable.[90]

A wide concept of personal data, including personalised rankings, seems to be supported by the general trend in interpreting that concept. Article 29 WP Guidelines, in its Opinion on the Concept of Personal Data, suggests that information can relate to a natural person based on content, purpose or result.[91] That same broad understanding of the term was supported by the CJEU in *Nowak*.[92] Based on this categorisation, the element of "content" is present where information is "about" a person; "purpose" where the data are used to evaluate that individual; and "result" where it is likely to impact their rights and interests.[93]

Arguably, a ranking is not primarily used for evaluation of the data subject; if anything, it *constitutes* an evaluation. As already outlined in discussing its significant effect, however, the

[87] Tie-Yan, supra, note 35; T Joachims et al, "Accurately interpreting clickthrough data as implicit feedback" in *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (New York, Association for Computing Machinery 2005) p 154.

[88] G Adomavicius and A Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions" (2005) 17 IEEE Transactions on Knowledge and Data Engineering 734; G Takács and D Tikk, "Alternating least squares for personalized ranking" in *RecSys '12: Proceedings of the Sixth ACM Conference on Recommender Systems* (New York, Association for Computing Machinery 2012) p 83; T Zhao, J McAuley and I King, "Leveraging Social Connections to Improve Personalized Ranking for Collaborative Filtering" in *CIKM'14: Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management* (New York, Association for Computing Machinery 2014) p 261.

[89] The fact that personal data can be incorrect is explicitly recognised by the GDPR through the introduction of the right to rectification (see: Art 29 Working Party, "Opinion 4/2007 on the concept of personal data", WP 136 (01248/07/EN) 6 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf> (last accessed 17 November 2022)). Also, arguably opinions do not need to be held to the standards of accuracy at all, see: D Hallinan and FZ Borgesius, "Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle" (2020) 10 International Data Privacy Law 1.

[90] S Wachter and B Mittelstadt, "A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI" (2019) 2 Columbia Business Law Review 494.

[91] Art 29 Working Party, "Opinion 4/2007 on the concept of personal data", WP 136 (01248/07/EN) 10 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf> (last accessed 17 November 2022).

[92] Case C-434/16 *Nowak v Data Protection Commissioner*, ECLI:EU:C:2017:994, para 35.

[93] ibid, 10–11.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

potential impact on the individual cannot be denied. Also, the entire purpose of the ranking is, in fact, the provision of information most relevant, or otherwise optimised, for a specific person. Additionally, it might be considered whether, in fact, the ranking does not constitute new information *about* that person. This at least is the opinion of the Article 29 WP Opinion regarding profiling, which is described as "creating derived or inferred data about individuals – 'new' personal data that has not been provided directly by the data subjects themselves".[94] In our view, personalised rankings, therefore, do generally constitute personal data.

b. Rectification

This raises the question of the ranking's rectifiability. While the Article 29 WP Opinion considered the right to apply to profiles and scores as well,[95] the practical implementation poses significant challenges. After all, the ranking does not exist independently and objectively but stems from the individual's personal data, as analysed by the algorithmic model. If the accuracy of the underlying information is not in question, one element that could be challenged is the ranking method itself. The applied method, however, is generally the same for each data subject and thus cannot constitute personal data.

This differentiation between the underlying data and the analytical method used is also supported by the CJEU's ruling on the joint cases of *YS and M. and S. vs. Minister door Immigratie, Integratie en Asiel.*[96] Here, the Court decided that while the applicant's data on a resident permit, which constitutes the basis for the decision, do constitute personal data, the legal analysis itself does not.[97] The same logic could be applied to rankings: whereas the user's data constitute personal data, as well as the outcome, the used ranking algorithm (the analysis) does not and is, therefore, not subject to rectification.

If an individual seeks to obtain a rectification of such a ranking, they would be left with the option of requesting rectification for the other two elements (analysed data, outcome). On the one hand, the direct revision of the outcome could be demanded. In that case, the data subject could submit a manually created, "correct" ranking to the controller or indicate which items are incorrectly highlighted in the ranking. The controller would then need to revise the ranking algorithm for the data subject to exclude such items. The user would constitute, so to speak, an additional supervision instance for training the model.

Another option would be to revise the underlying data. If their accuracy is not in question, however, that would in essence leave the option of completion, as Article 17 also provides individuals with the right to have "incomplete personal data completed, including by means of providing a supplementary statement". Both solutions hold the potential to correct personalised rankings, but practical problems in challenging incorrect and implementing novel rankings are likely to persist.

c. Article 15 AI Act and the AI Liability Directive

Recently, the legislator has added another instrument for contesting inaccurate AI decisions of all sorts. Article 15(1) AI Act holds that high-risk systems need to achieve an appropriate level of accuracy.[98] While the AI Act itself does not contain a private enforcement mechanism, private rights of action complementing the AI Act are provided

[94] Article 29 Data Protection Working Party, supra, note 56, 9.

[95] ibid, 18.

[96] Case C-141/12 *YS v Minister voor Immigratie*, ECLI:EU:C:2014:2081.

[97] ibid, paras 45–46.

[98] Technically speaking, the correct term would be "performance", with accuracy being only one of several relevant performance measures, see A Lindholm et al, *Machine Learning – A First Course for Engineers and Scientists* (Cambridge, Cambridge University Press 2022) p 88.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

by the new AI liability package unveiled by the Commission in September 2022.[99] Hence, if the AI developer violates Article 15(1) AI Act, the causal link between this fault and the AI output is presumed, for tort law claims noted in Member State law, according to Article 4(2) of the AI Liability Directive (AILD) Proposal. While this provision is obsolete in the case of a breach of the performance requirements – as the incorrect AI output *is* the fault[100] – the data subject can nevertheless claim damages based on national tort law. They may also obtain relevant evidence from the AI provider to the extent that this is necessary to prove the claim (Article 3 AILD Proposal). Moreover, the updated Product Liability Directive (PLD) can be invoked, under which a violation of the AI Act will probably indicate the defectiveness of the AI product.[101]

The damage would consist, inter alia, in the provision of an incorrect ranking; remedying the damage would include providing a correct ranking, or at least one with an appropriate level of accuracy in the sense of Article 15(1) AI Act. This includes not only cases of personalised rankings but also non-personalised ones, such as candidate lists outputted by hiring algorithms. However, unfortunately, such types of damage are not covered by the PLD framework (Article 4(6) PLD Proposal). Hence, such rectification can only be sought under national tort law regimes,[102] unless the PLD Proposal is changed to incorporate certain types of pure economic or immaterial damage.[103]

4. Fair Rankings

The final and perhaps most comprehensive pillar of ranking regulation beyond transparency and accuracy concerns fairness (ie non-discrimination in various guises).

a. Lessons from competition law

On 10 November 2021, the CJEU dismissed an action by Google challenging a €2.4 billion fine imposed by the Commission and confirmed the allegation that Google had intended to weaken the competitors' market position by favouring results from its own shopping comparison service over those of competitors.[104] Similarly, the Commission has been conducting several proceedings against Amazon since November 2020, among other things to examine whether the company gives preferential treatments to its own offers and offers from sellers using Amazon's logistics and shipping services.[105]

These cases exemplify the business models of many online platforms. On the one hand, they attempt to establish a business relationship between their business users (ie merchants and end users; *matching*). On the other hand, the companies behind the platforms sometimes also place their own production on these platforms, which they sell to the end users and thus enter direct competition with the business users (*dual mode*).[106] As a result, the platforms can shape the rankings to their advantage, disadvantaging their competitors and thus distorting competition.[107] Under Article 102 TFEU, such

[99] See, eg, Hacker, supra, note 1, 5 et seqq.

[100] Cf. Hacker, supra, note 1, 34, 37.

[101] See Art 6 PLD Proposal; Hacker, supra, note 1, 22, 53.

[102] European Commission, "Questions and answers on the revision of the Product Liability Directive" (QANDA/22/5791, 2022), under 9.

[103] For an argument in this vein, see Hacker, supra, note 1, 44.

[104] CJEU, Case T-612/17, *Google Shopping*, ECLI:EU:T:2021:763; on the procedure, see, eg, A Lohse, "Marktmissbrauch durch Internetplattformen?" (2018) 182 ZHR 321, 348–53.

[105] European Commission, Procedure AT.40703 (Amazon Buy Box).

[106] Hagiu et al, supra, note 22; Bongartz et al, supra, note 15; J Padilla, J Perkins and S Piccolo, "Self-Preferencing in Markets with Vertically Integrated Gatekeeper Platforms" (2022) 70 The Journal of Industrial Economics 371.

[107] I Graef, "Differentiated Treatment in Platform-to-Business Relations: EU Competition Law and Economic Dependence" (2019) 38 Yearbook of European Law 453.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

self-preferential treatment is (correctly) classified as exclusionary conduct,[108] as it leads to palpable distortions of competition and welfare losses.[109]

Article 102 TFEU is the central provision in proceedings such as those described above, which prohibits the abuse just described and represents the oldest and perhaps best-known rule on fairness of rankings in e-commerce. Thus, current competition law already prevents companies from favouring their products over competitors' ones on their own platforms. This prohibition of self-preference is a core component of substantial fairness regulations in rankings.[110]

However, the first prerequisite for the applicability of Article 102 TFEU is proof of a dominant position.[111] These proceedings necessitate an encompassing, economically orientated ex-post review, are often time-consuming and thus generally come too late, particularly for smaller competitors.[112] For example, the proceedings described above did not prevent Google from significantly improving its economic position[113] – a result that does not effectively counteract the abuse of a dominant position. The DMA addresses, among other things, precisely this time lag of competition law (Recital 5 DMA).

b. The Digital Markets Act

With Article 6(5), the DMA has now introduced a central provision for regulating AI-based rankings – a provision that is probably the most far-reaching one for regulating AI applications by gatekeepers. The provision comprises three components: first, the gatekeeper may not favour its products or services over those of third parties in a ranking (Article 6(5)(1) DMA); second, the ranking must be transparent in general (see above); and third, it must be fair and non-discriminatory (Article 6(5)(2) DMA).

*i. Prohibition of self-preference***.** Like general competition law, the DMA addresses the prohibition of self-preference. However, Article 6(5)(1) DMA draws consequences from the weaknesses of enforcement of Article 102 TFEU outlined above[114]: unlike there, sanctions for self-preferential treatment under Article 6(5) DMA are not tied to an (elaborate) procedure to prove the dominant position, the anti-competitive effects and the market definition; and justification by positive welfare effects is also impossible.[115] This sets the stage for more effective enforcement and, as a result, compliance by deterrence. Further, compared to the Commission's proposal, the final version of the DMA has extended this prohibition of self-preference to indexing and crawling.

*ii. Non-discrimination***.** Next to the prohibition of self-preference and the transparency requirement, Article 6(5)(2) DMA stipulates that rankings must be based on "fair and non-discriminatory conditions".[116] General non-discrimination provisions for rankings such as

[108] ibid.

[109] Hagiu et al, supra, note 22, 300.

[110] Bongartz et al, supra, note 15, 62.

[111] This can lead to a problematic market definition, particularly in the case of digital companies; see, eg, Bundeskartellamt, decision of 6 February 2019, B6-22/16, para 166 et seq.

[112] P Marsden, "Google Shopping for the Empress's New Clothes – When a Remedy Isn't a Remedy (and How to Fix it)" (2020) 11 Journal of European Competition Law & Practice 553; Digital Competition Expert Panel, "Unlocking Digital Competition" (Report, 2019) para 2.46.

[113] Cf. TJ Gerpott, "Das Gesetz über digitale Märkte nach den Trilog-Verhandlungen" (2022) CR 409, according to which Google's share value has increased by a factor of five during the almost eleven-year duration of the proceedings (citing Macrotrends, "Alphabet Market Cap 2010–2021" <https://www.macrotrends.net/stocks/charts/GOOGL/alphabet/market-cap> (last accessed 12 September 2022), for stock market value data).

[114] See Section II.4.a.

[115] Eifert et al, supra, note 15, 1003 f.

[116] Brouwer, supra, note 78.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

these have already been discussed in the anti-trust literature[117] and offer a point of reference for the DMA clause as well.

(1) Relationship to FRAND conditions

Fairness and non-discrimination cannot be reduced to transparency in Article 6(5) DMA.[118] Rather, the wording of Article 6(5)(2) DMA explicitly distinguishes between transparency, fairness and non-discrimination. However, the fairness and non-discrimination obligation can, from a systematic perspective, be linked to criteria known from general anti-trust law, according to which dominant undertakings must treat their corporate customers in fair, reasonable and non-discriminatory ways (in short: FRAND) in certain areas.[119] This obligation typically follows from Article 102 TFEU.[120] It is particularly relevant in intellectual property law, where owners of, for example, standard-essential patents (ie intellectual property rights that are essential for market access) must observe FRAND criteria when licensing.[121]

In the DMA itself, the FRAND formulation appears explicitly in Article 6(11), according to which gatekeepers must grant search engine operators access to ranking, search, click and display data of search results on FRAND terms.[122] Article 6(12) DMA requires FRAND conditions for access to app stores, among others. It is striking, however, that the wording of Article 6(5) DMA incorporates the FRAND language but, in contrast to Articles 6(11) and (12) DMA, removes the attribute of "reasonableness". In the FRAND context, this criterion regularly refers to the access or licence conditions, particularly the amount of the license fee.[123] Rankings created by gatekeepers may, of course, include items free of charge, so that no fee is paid by business users. However, the reasonableness control refers not only to the amount of the fee but also to other access conditions.[124] Indeed, special sections of the ranking are often reserved for paid content ("sponsored content").[125] In this respect, there would undoubtedly have been room for a reasonableness test. Ultimately, the decision to exclude a reasonableness review can only be explained by the reluctance of the EU legislator to trigger a complex control of conditions and, above all, prices in the context of rankings.[126] Incidentally, the transparency obligations described above do apply to sponsored rankings, as seen.

Hence, all rankings, regardless of whether they are sponsored, must be fair and non-discriminatory (FAND) under the DMA. However, the fairness element typically has no independent meaning within FRAND law.[127] Admittedly, Article 12(5) DMA lists practices

[117] Graef, supra, note 107, 463 f.

[118] But see Brouwer, supra, note 78.

[119] See only J Block and B Rätz, "The FRAND offer – attempt at an international definition" (2019) GRUR 797; J-S Borghetti, I Nikolic and N Petit, "FRAND licensing levels under EU law" (2021) 17 European Competition Journal 205; JG Sidak, "The Meaning of FRAND, Part I: Royalties" (2013) 9 Journal of Competition Law and Economics 931; M Heim and I Nikolic, "A FRAND Regime for Dominant Digital Platforms" (2019) 10 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 38.

[120] ibid.

[121] See, eg, C Ann, *Patentrecht* (8th edition, Munich, CH Beck 2022) § 43 marginal no. 35 et seqq.

[122] PG Picht and H Richter, "EU Digital Regulation 2022: Data Desiderata" (2022) GRUR International 395.

[123] Sidak, supra, note 119.

[124] Ann, supra, note 121, § 43 marginal no. 36.

[125] C Busch, "Mehr Fairness und Transparenz in der Plattformökonomie? Die neue P2B-Verordnung im Überblick" (2019) GRUR 788.

[126] See, eg, Ann, supra, note 121, § 43 marginal no. 39 ff on the various calculation approaches.

[127] Sidak, supra, note 119, 968; Brouwer (supra, note 78) would like to load fairness in the ranking context, albeit still on the basis of the original Commission draft, with the transparency rules of the P2B Regulation. This is now prohibited, as transparency is an independent criterion.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

that must each be considered as limiting the contestability of platform services or as unfair. The second and third practices mentioned there[128] lack any reference to rankings – unlike the third practice, which addresses barriers to market entry.[129] However, the wording of this variant clearly refers to the contestability, and thus not to the fairness, of platform services. Hence, the FRAND jurisprudence and the DMA do not offer any support for a separate fairness criterion for rankings. Therefore, rankings under the DMA need only be non-discriminatory – which may be difficult enough, as we shall presently see.

(2) Non-discriminatory rankings: between non-discrimination and competition law

Platforms regularly create rankings based on specific product attributes. Importantly, the non-discrimination requirement arguably refers to these same product attributes.[130] The FAND condition in Article 6(5) DMA is not limited to mere non-discriminatory *access* to the ranking results. This follows from a comparison with Article 6(11)(1) DMA: here, FRAND conditions are restricted to access to certain data. *E contrario*, the FAND clause in Article 6(5) DMA must encompass more than access (ie the order of the ranking itself).

The link to the FRAND literature provides a starting point for delimiting the meaning of non-discrimination with respect to DMA rankings. As in the case of FRAND conditions, what exactly constitutes a protected attribute under Article 6(5) DMA will probably depend to a certain degree on the individual case. However, certain guidelines can be established based on (non-exhaustive) groups of cases: in the first group, discrimination refers to the anthropocentric attributes protected by classical non-discrimination law,[131] such as gender, religion, racial or ethnic origin and nationality. This group of attributes also corresponds to the ongoing research effort in computer science on fair (ie non-discriminatory) rankings.[132] In the second group, discrimination refers to the understanding of the term under competition law, where, for example, making the ranking of a product dependent on the relationship between the companies rather than objective criteria would fall within that term.

(a) Coherence between non-discrimination and competition law

Within this first group of cases, one may differentiate again: the list may directly rank persons or groups of persons (such as the providers of services or works or users listed in social media feeds). In this case, significantly, non-discrimination doctrine should guide the interpretation of Article 6(5) DMA.[133] This corresponds to a coherent interpretation of

[128] This concerns access to essential input data and equivalence violations in the contractual structure.

[129] Art 12, para 5, lit b(i) DMA.

[130] L Cabral et al, "The EU Digital Markets Act" (Luxembourg, Joint Research Centre, JRC122910, Publications Office of the European Union 2021) p 13.

[131] See, eg, Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (Race Equality Directive); Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation (Framework Directive); Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation; Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services.

[132] See, eg, the overview in A Asudeh et al, "Designing Fair Ranking Schemes" in *SIGMOD '19: Proceedings of the 2019 International Conference on Management of Data* (New York, Association for Computing Machinery 2019) p 1259; A Singh and T Joachims, "Fairness of Exposure in Rankings" in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (New York, Association for Computing Machinery 2018) p 2219; M Zehlike, K Yang and J Stoyanovich, "Fairness in Ranking, Part II: Learning-to-Rank and Recommender Systems" (2022) 55 ACM Computing Surveys 117.

[133] See, eg, S Wachter, "The Theory of Artificial Immutability: Protecting Algorithmic Groups Under Anti-Discrimination Law" (2022) arXiv preprint arXiv:220501166; S Wachter, B Mittelstadt and C Russell, "Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI" (2021) 41 Computer Law & Security Review 105567; FJZ Borgesius, "Strengthening legal protection against discrimination by algorithms and

EU law that the CJEU has repeatedly demanded even across legal fields in cases such as *Pereničová and Perenič*,[134] *Bankia*[135] and *Pillar Securitisation*.[136] The key takeaway from these cases is that a concept such as non-discrimination, transplanted from one area of EU law to another, should be interpreted coherently, while acknowledging the idiosyncrasies of the respective fields. In the words of AG Trstenjak, "what is needed is a coherent interpretation of the relevant rules of law so as to avoid conflicting assessments".[137] Therefore, in our view, a finding of discrimination under non-discrimination law does not automatically imply a violation of Article 6(5) DMA but offers a strong indication of the use of discriminatory conditions according to the DMA unless DMA-specific justifications can be found.[138]

Hence, if persons are ranked, direct or indirect discrimination may occur in the ranking.[139] For example, if ranking parameters include attributes of these persons or groups of persons protected by classical non-discrimination law, such as gender, this will generally constitute direct discrimination, unless a justification applies. Such a justification may be found in traditional non-discrimination law or in an implicit, DMA-specific justification modelled on the FRAND literature and jurisprudence. For example, restricting the ranking to female persons may be justified if the female gender constitutes a genuine and determining occupational requirement.[140] This implies that the protected attribute is essential for the task at hand.[141] Search for justification fails, for example, if a platform intermediates jobs for merely historically female jobs (eg cleaning) that are now performed by persons of any gender.

In our view, another subcategory of this group of cases concerns the ranking of goods or companies that is influenced by traditionally protected attributes of persons related to these items. For example, rankings of companies and their products may generally not be influenced by the religion or gender of their CEOs (or other company members). Importantly, even customer preferences do not provide a justification, according to the CJEU, in such cases.[142]

(b) Beyond non-discrimination law

The second group of cases concerns discrimination as it is understood in competition law. For example, the ranking position of a product may depend on whether the offering

artificial intelligence" (2020) 24 The International Journal of Human Rights 1572; M Zehlike, P Hacker and E Wiedemann, "Matching code and law: achieving algorithmic fairness with optimal transport" (2020) 34 Data Mining and Knowledge Discovery 163; FZ Borgesius, *Discrimination, Artificial Intelligence, and Algorithmic Decision-Making* (Strasbourg, Council of Europe, Directorate General of Democracy 2018); P Hacker, "Teaching fairness to artificial intelligence: existing and novel strategies against algorithmic discrimination under EU law" (2018) 55 Common Market Law Review 1143.

[134] CJEU, Case C-453/10, *Pereničová and Perenič*, ECLI:EU:C:2012:144.

[135] CJEU, Case C-109/17, *Bankia*, ECLI:EU:C:2018:201.

[136] CJEU, Case C-694/17, *Pillar Securitisation*, ECLI:EU:C:2019:44, para 35.

[137] AG Trstenjak, Case C-453/10, *Pereničová und Perenič*, ECLI:EU:C:2011:788, para 90 (discussing legal acts relating to EU consumer law).

[138] See also P Hacker, *Datenprivatrecht* (Tubingen, Mohr Siebeck 2020) p 335 et seqq.

[139] See, eg, J Adams-Prassl, R Binns and A Kelly-Lyth, "Directly Discriminatory Algorithms" (2023) 86(1) The Modern Law Review 144.

[140] See, eg, Art 4(1) of the Race Equality Directive; Art 4(1) of the Framework Directive; Art 14(2) of the recast Gender Equality Directive 2006/54/EC; CJEU, Case C-229/08, *Wolf*, ECLI EU:C:2010:3, para 35.

[141] Hacker, supra, note 133, 1166.

[142] CJEU, Case C-188/15, *Bougnaoui*, EU:C:2017:204, para 40; Joined Cases C-804/18 and C-341/19, *Wabe and Müller*, ECLI EU:C:2021:594, para 65; see also E Howard, "Headscarves and the CJEU: Protecting fundamental rights or pandering to prejudice" (2021) 28 Maastricht Journal of European and Comparative Law 648, 255 et seqq.; E Cloots, "Safe Harbour or Open Sea for Corporate Headscarf bans? Achbita and Bougnaoui" (2018) 55 Common Market Law Review 589, 613.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

company has concluded an exclusivity agreement with the gatekeeper. Such a ranking also constitutes discrimination within the meaning of Article 6(5) DMA.[143]

This example indicates that the concept of discrimination in Article 6(5) DMA does not merely correspond to that of classical non-discrimination law but must go beyond it. Products do not exhibit any of the traditionally protected anthropocentric attributes, such as gender. Clearly, it cannot be considered discrimination if the query itself, in a legitimate way, restricts the search to goods particularly valuable to certain protected groups, such as female or male shoes. Limiting the scope of the prohibition of discrimination in Article 6(5) DMA to traditionally protected attributes, however, would not meet the aim of the DMA, and in particular the ranking provisions, to ensure the contestability of markets.[144] Hence, a more comprehensive understanding of the prohibition of discrimination is needed. Indeed, it may link back to the requirements for entities covered by FRAND obligations: these may not treat similarly situated business users differently without justification.[145]

At the outset, this implies that the covered entities must offer the same conditions for comparable circumstances.[146] However, the purpose of a ranking is precisely to *distinguish* comparable products or services from each other to facilitate the customer's selection decision. To award the same ranking position twice would contradict this purpose. Hence, freedom from discrimination in the sense of Article 6(5) DMA means that distinctions must be objectively justifiable.[147] In our view, this entails that the ranking must be based on criteria that are relevant for the comparison of ranked products to ultimately facilitate the economic decision of consumers. If, for example, the ranking is based on whether business users use the logistics channels of the gatekeeper ("Fulfilment by Amazon"), the ranking constitutes prohibited discrimination that can only be justified in exceptional cases, according to the criteria just mentioned.

iii. Technical possibilities and difficulties. The legal requirements just described also need to be technically implemented in the ranking models. The prohibition of self-preference, the transparency and the non-discrimination requirements must be operationalised so that the algorithm creating the ranking takes these into account and outputs a ranking that meets the requirements.

While implementing the prohibition of self-preference appears rather unproblematic,[148] the first major difficulties arise in the operationalisation of the transparency obligations. This is due to the lack of the congruence and coordination of the transparency requirements under the DMA/P2B Regulation on the one hand and the GDPR as well as the AI Act on the other. Finally, technical challenges arise in implementing non-discrimination. To create a non-discriminatory AI-based ranking, the model must be audited for its impact on protected groups (eg gender and religion). A comprehensive exploration of algorithmic fairness techniques already exists in computer science.[149] If an unjustified unequal treatment is detected, adequately chosen constraints from the algorithmic fairness literature ensure that

[143] Cabral et al, supra, note 130, 13.

[144] In particular the contestability of ranked products and services, cf. Recitals 7, 11 and 51 DMA.

[145] See, eg, District Court of Düsseldorf, [2018] *GRUR-RS 37930*, ECLI:DE:LGD:2018:1212.4B.O4.17.00, para 202; CJEU, Case C-313/04, *Egenberger*, ECLI:EU:C:2008:728, para 33; *Unwired Planet v Huawei* [2018] EWCA Civ 2344, para 162; European Commission, "Setting Out the EU Approach to Standard Essential Patents" (Communication) COM(2017) 712 final, 7; see also Brouwer, supra, note 78; Art 102(c) TFEU.

[146] *Unwired Planet v Huawei* [2018] EWCA Civ 2344, para 162 et seq.

[147] See also Brouwer, supra, note 78; *Unwired Planet v Huawei* [2018] EWCA Civ 2344, paras 169–70.

[148] The fact that a company has entered into an exclusivity agreement with Amazon is not relevant to the accuracy of prediction, which is orientated towards the needs of consumers.

[149] D Pessach and E Shmueli, "Algorithmic Fairness" (2020) arXiv preprint arXiv:2001.09784.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

at least a similar number of persons or products of each of the protected classes is also represented in the upper part of the ranking.[150]

However, classical non-discrimination law defines only a finite, low number of protected attributes. In the case of Article 6(5) DMA, however, an infinite number of possibly prohibited differentiating attributes exists, which redoubles the complexity of the auditing procedure. On the one hand, the large number of possible discriminatory attributes makes it more difficult to find discrimination within the meaning of Article 6(5) DMA in the first place.[151] On the other hand, a larger number of protected attributes may also mean greater protection for the discriminated groups. Once a discriminatory attribute is identified, one could continuously expand the list of protected attributes by adding similar instances and intersections with other protected attributes. Thus, a higher and more differentiated level of protection could be achieved in the future. This, however, makes finding and remedying discrimination in the sense of the DMA computationally complex and costly.

iv. Compliance requirements. Against this background, clear compliance guidelines must be offered. Article 8(1) DMA stipulates that gatekeepers do not only have the obligation to ensure compliance with the Regulation but should also be able to demonstrate this fact. Such compliance and documentation obligations are also prevalent in other highly regulated areas, such as banking law, data protection law (Article 5(2) GDPR) and the proposed AI Act (eg Articles 9 and 16 AI Act). However, particular challenges arise for gatekeepers in the context of compliance with the fair ranking provision under Article 6(5) DMA.[152]

As seen, at least theoretically, an infinite amount of potentially discriminatory attributes exists, which would need to be considered. Fines, however, can only be imposed in cases of intent or negligence (Articles 30(1) and (3) DMA). Hence, gatekeepers must have violated a duty of care to be liable. In our view, such a duty implies three distinctive compliance obligations.

The first one links Article 6(5) DMA even more closely to the transparency requirements in the P2B Regulation[153]: gatekeepers need to constantly monitor the main parameters of the ranking, which also have to be disclosed. These features may not relate to any attributes protected under classical anti-discrimination law or to other illegal differentiating attributes (affiliation with the gatekeeper, etc.).

Second, gatekeepers must investigate evidence of possible discrimination in rankings brought to their attention and remove the incriminated practice if warranted. This procedure essentially corresponds to the notice-and-takedown procedure applied in the context of potential copyright infringement in intellectual property law.[154] Its purpose lies in the establishment of a clearly structured and rapid procedure for checking and removing an unpredictable amount and type of potential legal violations.[155]

Intellectual property law, however, contains clear limitations of liability for hosting providers,[156] which the DMA lacks. This suggests, third, that gatekeepers are held to a

[150] M Zehlike et al, "FA*IR: A Fair Top-k Ranking Algorithm" in *Proceedings of the 2017 ACM on Conference on Information and Knowledge, Management* (New York, Association for Computing Machinery 2017) p 1569; Zehlike et al, supra, note 133.

[151] Cf. Hacker, supra, note 16, 1284.

[152] See also Cabral et al, supra, note 130, 13.

[153] See also Brouwer, supra, note 78.

[154] See, eg, S Bar-Ziv and N Elkin-Koren, "Behind the scenes of online copyright enforcement: empirical evidence on notice & takedown" (2018) 50 Connecticut Law Review 339; J Cobia, "The digital millennium copyright act takedown notice procedure: Misuses, abuses, and shortcomings of the process" (2008) 10 Minnesota Journal of Law Science & Technology 387.

[155] JM Urban and L Quilter, "Efficient process or chilling effects – takedown notices under Section 512 of the Digital Millennium Copyright Act" (2005) 22 Santa Clara Computer & High Tech Law Journal 621, 622.

[156] See, eg, Arts 5 and 14 DSA and Art 17 C-DSM Directive.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

higher level of care. Thus, in our view, the compliance requirements under the DMA also oblige gatekeepers to regularly examine their ML models for possible discrimination-relevant distinctions (test and audit), regardless of any specific allegations or clues that would suggest their existence.[157] This could, for example, be achieved by regularly creating two rankings for test purposes: one with and without the identification of the business users.[158] Should the two differ from each other, this could be construed as an indication of possible unjustified differentiations, as the ranking would therefore be more dependent on the mere identity of the customer rather than on their attributes or those of their products.

v. Interim summary. Overall, the new ranking regulation in the DMA, while representing a step in the right direction, still leaves room for improvement concerning both regulatory agencies and addressees (see Section VI.3 infra). Nonetheless, with its wide-ranging provisions, its reference to the FRAND landscape and the accompanying fines, it opens a new chapter –in terms of both substance and enforcement (see Section II.2 supra) – in the regulation of gatekeepers' core business models and AI functionalities.

III. Training data regulation

Whereas the previous section focused on the potential impacts Article 6(5) DMA and other pertaining regulations might have on AI-based rankings, a number of other obligations in the DMA will also significantly affect the development and deployment of AI by gatekeepers and competitors. The first category in that context relates to restrictions regarding the use of certain datasets for gatekeepers' own purposes. This is relevant for AI training, validation and test data, which are crucial components of AI modelling.[159]

1. Article 5(2) DMA

Article 5(2) DMA contains a number of provisions relating to the cross-service use of personal end user data (PED). For example, gatekeepers may not use such data for personal advertising if they are obtained via third-party services that make use of gatekeepers' central platform services (lit. a). Furthermore, PED cannot be combined across services (lit. b), which specifically precludes the practice of combining datasets for the enhanced training of AI systems. Even the registration of end users with other gatekeeper services by the gatekeepers themselves for the purpose of combining PED is prohibited (lit. d). Finally, even without combining the datasets, PED may not be used by gatekeepers across services (lit. c). This would also include so-called federated learning strategies, in which the data remain formally separate but the information obtained from them is combined in one single AI model.[160] Such strategies are, however, particularly privacy-preserving and sustainable in terms of energy consumption,[161] which is why their encumbrance should be reconsidered.

While these new rules have the potential to significantly affect gatekeepers' data management practices, the practical impact is considerably diminished by the exemption introduced in the same article, which allows these processing activities to occur if GDPR-

[157] See also, for a practical proposal in this vein, M Veale and R Binns, "Fairer machine learning in the real world: mitigating discrimination without collecting sensitive data" (2017) 4 Big Data & Society 2053951717743530.

[158] Cabral et al, supra, note 130, 13.

[159] Lindholm et al, supra, note 98, p 67 ff, 299 et seqq.; Hacker, supra, note 14, 259.

[160] Q Yang et al, "Federated Machine Learning: Concept and Applications" (2019) 10 ACM Transactions on Intelligent Systems and Technology 12.

[161] B Güler and A Yener, "Sustainable federated learning" (2021) arXiv preprint arXiv:210211274.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

compliant consent has been collected from the end user.[162] This invites legal uncertainty since numerous legal[163] and behavioural[164] problems related to the collection of valid and meaningful consent under the GDPR have been outlined in the literature. These mainly concern practices such as nudging,[165] bundling of purposes,[166] rational ignorance[167] and "consent fatigue", which effectively causes individuals to simply agree to any given form of processing without even seeking to understand its implications.[168] The general inconsistency of users' statements regarding their privacy preferences and their actual online behaviour is also referred to as the "privacy paradox".[169]

Considering how easily consent is often obtained in practice, the only significant restriction for gatekeepers therefore results from the modification of Article 6(1) GDPR, contained in Article 5(2)(3) DMA. It provides that, where consent is refused, cross-service processing of personal data will only be possible where it can be based on compliance with a legal obligation, the protection of vital interests of a natural person or performance of a task in the public interest. This implies that the legal bases of the performance of contract and legitimate interest (Articles 6(1)(b) and (f) GDPR) can no longer be applied. Such a restriction is significant as these two legal bases are typically used as "fallbacks" when obtaining valid consent is impossible or challenging. Nevertheless, obtaining consent from rationally ignorant data subjects will often be quite feasible for gatekeepers.

For the moment, the overall impact of the rules seeking to prevent the accumulation of data by gatekeepers is, therefore, rather limited, especially since the gatekeepers will probably use the loopholes just discussed to their advantage.[170] While coupling the requirement for consent with the use of a service is not possible (Article 5(2)(2) DMA in conjunction with Article 7(4) GDPR[171]), businesses are still likely to present consent requests in a form that will ultimately lead a substantial amount of users to agree to cross-service processing.[172] The regulation does, however, look more promising when considering current developments concerning the prevention of "dark patterns" seeking

[162] See already R Podszun, "Should Gatekeepers Be Allowed to Combine Data? Ideas for Article 5(a) of the Draft Digital Markets Act" (2022) 71 GRUR International 197, 199.

[163] See, eg, N Richards and W Hartzog, "The pathologies of digital consent" (2018) 96 Wash University Law Review 1461; P Blume, "The inherent contradictions in data protection law" (2012) 2 International Data Privacy Law 26, 29 et seqq.

[164] SY Soh, "Privacy Nudges: An Alternative Regulatory Mechanism to 'Informed Consent' for Online Data Protection Behaviour" (2019) 5 European Data Protection Law Review 65; Y Hermstrüwer, "Contracting around privacy: the (behavioral) law and economics of consent and big data" (2017) 8 JIPITEC, the Journal of Intellectual Property, Information Technology and Electronic Commerce Law 9.

[165] C Utz et al, "(Un)informed Consent: Studying GDPR Consent Notices in the Field" in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19)* (New York, Association for Computing Machinery) p 973.

[166] D Machuletz and R Böhme, "Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR" in *Proceedings on Privacy Enhancing Technologies* (2020) pp 481–98; B Kostic and EV Penagos, "The freely given consent and the 'bundling' provision under the GDPR" (2017) 153 Computerrecht 217.

[167] C Van Slyke et al, "Rational ignorance: a privacy pre-calculus" (2021) WISP 2021 Proceedings 12.

[168] M Nouwens et al, "Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence" in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, Association for Computing Machinery 2020) p 1.

[169] S Barth and MDT de Jong, "The privacy paradox – investigating discrepancies between expressed privacy concerns and actual online behavior – a systematic literature review" (2017) 34 Telematics and Informatics 1038–58.

[170] D Geradin, K Bania and T Karanikioti, "The interplay between the Digital Markets Act and the General Data Protection Regulation" (29 August 2022) <http://dx.doi.org/10.2139/ssrn.4203907> (last accessed 6 November 2022).

[171] Kostic and Penagos, supra, note 166; FZ Borgesius et al, "Tracking walls, take-it-or-leave-it choices, the GDPR, and the ePrivacy regulation" (2017) 3 European Data Protection Law Review 353, 361.

[172] Cf. Utz et al, supra, note 165, 973.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

to steer users towards excessive consent.[173] The specification in Recital 67 DSA on dark patterns is aimed at a more neutral presentation of request, trying to assure that consent is in fact freely given and specific; ultimately, however, it defers to the UCPD and the GDPR.[174] Should this succeed, gatekeepers might be confronted with notable drops in consent rates, potentially adding relevance to the DMA in the future. Importantly, the recent decision of the European Data Protection Board (EDPB) against Meta points exactly in this direction.[175] This adds to a general trend by data protection authorities and NGOs to monitor and enforce the requirements for valid consent more aggressively.[176] To facilitate informed decisions by users, it would be worthwhile to combine this more robust enforcement with a traffic light system for different types of data sharing by gatekeepers.[177]

2. Article 6(2) DMA

A corresponding prohibition, aimed at business user data, can be found in Article 6(2) DMA: gatekeepers are prohibited from using data in competition with business users that these or their end users have generated or provided in the context of relevant core platform services. This rule, however, does not apply to data that are already publicly accessible. To delineate that concept, Article 6(2)(2) DMA specifies that non-public data also include information that can be "inferred from, or collected through, the commercial activities of business users or their customers". Consequently, it will be possible neither to use end user or business user data for AI-based inferences nor to further harness them in competition with business users.

The introduction of the new restriction makes sense from the perspective of workable competition as it precludes a further entrenchment of the gatekeeping position via data-based inferences. The EU, with this rule, moves onto largely uncharted territory, since lawmakers have, until this point, largely refrained from regulating AI-based inferences, in spite of their economic and informational importance[178] and criticism from the literature.[179]

The overall impact of Article 6(2) DMA will probably be stronger than that of Article 5(2) DMA as the former applies to all types of data, not just personal data. Given the practical difficulties in distinguishing personal from non-personal data, especially in the context of AI training data,[180] this simplification is a step in the right direction. Additionally, it should

[173] See, eg, J Luguri and L Strahilevitz, "Shining a light on dark patterns" (2021) 13 Journal of Legal Analysis 43; M Martini et al, "Dark patterns" (2021) 1 Zeitschrift für Digitalisierung und Recht 47.

[174] See the analysis in P Hacker, "Manipulation by algorithms. Exploring the triangle of unfair commercial practice, data protection, and privacy law" (2022) European Law Journal.

[175] noyb, "noyb win: Personalized Ads on Facebook, Instagram and WhatsApp declared illegal" (*noyb*, 6 December 2022) <https://noyb.eu/en/noyb-win-personalized-ads-facebook-instagram-and-whatsapp-declared-illegal> (last accessed 8 December 2022).

[176] J Bryant, "Belgian DPA fines IAB Europe 250K euros over consent framework GDPR violations" (*iapp*, 2 February 2022) <https://iapp.org/news/a/belgian-dpa-fines-iab-europe-250k-euros-over-consent-framework-gdpr-violations/> (last accessed 2 September 2022); noyb, "226 Complaints Lodged Against Deceptive Cookie Banners" (*noyb*, 9 August 2022) <https://noyb.eu/en/226-complaints-lodged-against-deceptive-cookie-banners> (last accessed 21 September 2022).

[177] See also Podszun, supra, note 162, 201 et seq.; Hacker, supra, note 138, 627 et seqq. on privacy scores.

[178] See, eg, A Goldfarb, S Greenstein and C Tucker, "Introduction to Economic Analysis of the Digital Economy" in A Goldfarb, S Greenstein and C Tucker (eds), *Economic Analysis of the Digital Economy* (Chicago, IL, University of Chicago Press 2015) p 1.

[179] See only Wachter and Mittelstadt, supra, note 90.

[180] See, eg, Hacker, supra, note 14, 265–68; see also, more generally, M Finck and F Pallas, "They who must not be identified – distinguishing personal from non-personal data under the GDPR" (2020) 10 International Data Privacy Law 11.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

be noted that Article 6(2) DMA, in contrast to Article 5(2) DMA, cannot be waived based on consent or any other grounds, expanding its impact even further.

3. Comparison with Article 10 AI Act

Contrasting the DMA data governance regime with Article 10 AI Act, the provision detailing specific requirements for data used in training high-risk AI applications,[181] reveals strikingly different objectives. Article 10 AI Act compels developers of high-risk AI systems to only use high-quality training data to facilitate the creation of accurate predictions and to mitigate bias in AI systems. For that purpose, several quality criteria are outlined in Articles 10(2)–(5) AI Act, inter alia specifying the representativeness of training data for the target group and statistical appropriateness. In contrast, the DMA restrictions seek to prevent gatekeepers from gathering high-quality datasets by tapping into the data trove accumulating on the platform; if anything, this will reduce the predictive accuracy of gatekeeper models.

The DMA is nonetheless right in explicitly blocking gatekeepers from leveraging their specific position to build better models. This would most likely have the effect of even further cementing their position on the market, thereby continuing to hinder the possibility of workable competition. This points to an inherent conflict of objectives in the area of AI and platform regulation: technological tools and AI systems are supposed to be high-performing, but it is precisely this capability that may lead to further market concentration and the weakening of competitive processes. The AI Act and the DMA, therefore, rightly seek to accommodate this tension by allocating specific and, prima facie, strikingly divergent duties to gatekeepers on the one hand and developers of high-risk AI systems on the other.

IV. Access rights

Another set of AI-relevant rules targeting the competitive position of gatekeepers is represented by the access rights contained in Articles 6(10) and (11) DMA.[182] The idea of harnessing access rules instead of a data producers' right to strengthen innovation and competition in data-driven markets has already received much scholarly attention.[183] The DMA has now, for the first time, introduced general access rules for gatekeepers independently of the business sector in which they are active.

In a sense, access rights are the flipside of the restrictions concerning the use of training data. While these limitations are supposed to prevent gatekeepers' AI systems from becoming too powerful and thereby creating an unfair advantage, access rights are intended to provide business users with the necessary tools to develop high-performing algorithms themselves, including ML models. Since access rights are not considered in the current draft of the AI Act, the importance of such within the DMA is even more pronounced.

[181] For a critical analysis of Art 10 AI Act, see Hacker, supra, note 14, 296–300; M van Bekkum and FZ Borgesius, "Using sensitive data to prevent discrimination by artificial intelligence: does the GDPR need a new exception?" (2023) 48 Computer Law & Security Review 105770, 11–12.

[182] The right of access in Art 6, para 12 DMA, on the other hand, has no specific reference to AI.

[183] DL Rubinfeld and MS Gal, "Access Barriers to Big Data" (2017) 59 Arizona Law Review 339; W Kerber, "Governance of Data: Exclusive Property vs. Access" (2016) 47 IIC – International Review of Intellectual Property and Competition Law 759; see also H Schweitzer, "Datenzugang in der Datenökonomie: Eckpfeiler einer neuen Informationsordnung" (2019) 121 Gewerblicher Rechtsschutz und Urheberrecht 569.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

1. Article 6(10) DMA

According to Article 6(10) DMA, business users or authorised third parties may have access to data provided for or generated in the context of the use of the respective core platform services by the business users themselves or their end users. Access must be granted free of charge and in a way that is effective, high-quality, continuous and real-time. In that context, AI could be particularly helpful in forecasting demand and optimising product design.

The provision does, however, also include important restrictions regarding access to personal data. Given the broad interpretation of the concept of personal data by the CJEU and legal scholarship,[184] this concerns a large share of the data eligible to be accessed. First of all, they may only be provided where the information is directly related to the use of the business user's services and products by the end user through the relevant core platform service. This, in turn, means that businesses will still only receive information concerning individuals who are already part of their client base. In that sense, the competitive effect will be significantly limited: information about consumers who the business user was not yet able to reach would arguably be more valuable in that respect. The limited scope of Article 6(10) DMA cuts against the level-playing-field rationale of Article 6 DMA as a whole and paragraph 2 in particular.

Second, gatekeepers are only allowed to share such personal data where the user has provided their consent. As has already been mentioned, businesses have by now found a number of ways to assure the obtainment of consent from data subjects through the use of means such as nudging, specific framings or bundling of purposes.[185] If gatekeepers design and steer user consent, they should be able to guide them into excluding their data use by business users. Such behaviour would arguably be legal but would render Article 6(10) DMA effectively futile, as all data generated by a customer will fall under the definition of personal data and, hence, the outlined restriction. This issue will be taken up again in at the end of the paper (Section VI.2).

2. Article 6(11) DMA

The provisions introduced by Article 6(11) DMA, by contrast, intend to remedy a central weakness of the market for search engine operators. It is common knowledge that the quality of a search engine is predominantly determined based on the delivered results. These are, however, largely being optimised by means of analysing the historical search and click behaviour of end customers, with the support of ML systems.[186] Therefore, whichever provider is able to generate more end users up front will naturally be able to increase their competitive advantage even further, potentially creating a positive feedback loop of extended competitive advantage on the side of the gatekeeper.

To mitigate the impact of this phenomenon, Article 6(11) DMA conveys upon search engine operators the right to access the dataset of gatekeepers who themselves operate search engines. The access to these ranking, query, click and view data must be provided on FRAND terms, which have already been outlined above, the purpose again being the enablement of business users to optimise their own AI models (Recital 61 DMA).

Significantly, Article 6(11) DMA also provides that any personal data that are part of the ranking, query, click and view data must be offered in an anonymised form, meaning that

[184] See, eg, Finck and Pallas, supra, note 180; N Purtova, "The law of everything. Broad concept of personal data and future of EU data protection law" (2018) 10 Law, Innovation and Technology 40.

[185] D Machuletz and R Böhme, "Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR" in (2020) Proceedings on Privacy Enhancing Technologies 481–98; Kostic and Penagos, supra, note 166; Utz et al, supra, note 165, 973–90; Nouwens et al, supra, note 168, 1–13.

[186] Rubinfeld and Gal, supra, note 183, 353; Martens, supra, note 31, 4, 24 et seq.; see also supra, note 19.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

they cannot be related to an identified or identifiable natural person.[187] While the introduction of such an obligation is definitely reasonable from the perspective of protecting the individuals' rights and freedoms, it does create significant implementation issues on the side of the gatekeeper. The anonymisation of personal data is, after all, not an uncomplicated task, as research has repeatedly shown that supposedly anonymised data can be de-anonymised through a range of strategies.[188]

Consequently, it will be necessary for gatekeepers to use strong, state-of-the-art anonymisation strategies to comply with the requirements of the DMA and the GDPR. Since, however, strong anonymisation also requires a fair amount of data to be either removed from or modified in the existing set,[189] such measures can also lead to a reduction in value of the dataset for the purpose of AI training.[190] In that sense, so-called privacy-preserving ML (PPML) strategies,[191] which attempt to strike a balance between data protection and performance, are likely to gain more practical and regulatory relevance as a result of Article 6(11) DMA. On the other hand, the extent to which gatekeepers would be compelled to go above and beyond in the search for privacy-preserving techniques, assuring a high-quality dataset for their competitors, is questionable. At least in the initial phase, businesses are likely to simply provide extensively altered and anonymised datasets, with limited practical use. As PPML progresses, so will the requirement for gatekeepers to offer datasets that are not only anonymised but also performance-enabling due to state-of-the-art PPML.

3. Comparison with the Data Act proposal

Another legal basis introducing extensive access rights at the EU level is the DA,[192] which serves the purpose of removing barriers in data-sharing.[193] The access rights under the DA are, however, significantly more consumer-driven than those under the DMA.[194] Articles 4 and 5 DA regulate the right of users to access and use data generated by the use of products or related services and to share them with third parties, respectively. Other businesses will therefore potentially be given the option of using such information, but only "upon request by a user, or by a party acting on behalf of a user".[195] Hence, the consumer would need to act first, or at least clearly communicate their preferences. The DA seeks to balance the acquisition of data by competing businesses with an individual's right to

[187] Finck and Pallas, supra, note 180, 15.

[188] See, eg, L Rocher, JM Hendrickx and Y-A de Montjoye, "Estimating the success of re-identifications in incomplete datasets using generative models" (2019) 10 Nature Communications 3069.

[189] See, eg, Irish Data Protection Commission, "Guidance on Anonymisation and Pseudonymisation" (June 2019) <https://www.dataprotection.ie/sites/default/files/uploads/2019-06/190614%20Anonymisation%20and%20Pseudonymisation.pdf> (last accessed 7 December 2022); Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, WP 216, 2014.

[190] EF Villaronga, P Kieseberg and T Li, "Humans forget, machines remember: artificial intelligence and the right to be forgotten" (2018) 34 Computer Law & Security Review 304, 310.

[191] See, eg, P Mohassel and Y Zhang, "SecureML: A System for Scalable Privacy-Preserving Machine Learning" (2017) IEEE Symposium on Security and Privacy (SP) 1.

[192] European Commission, Proposal for a Regulation on the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act) COM(2022) 68 final.

[193] ibid, 2 et seq.

[194] Regarding the relationship of the DA to other legal acts, including the DMA, cf. also L Specht-Riemenschneider, "Der Entwurf des Data Act – Eine Analyse der vorgesehenen Datenzugangsansprüche im Verhältnis B2B, B2C und B2G" (2022) 25 Zeitschrift für IT – Recht und Recht der Digitalisierung 809, 810 et seq; critical with regard to the achievement of consumer empowerment is W Kerber, "Governance of IoT Data: Why the EU Data Act Will not Fulfill Its Objectives" (2022) GRUR International 1.

[195] See Art 5(1) DA.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

self-determination.[196] The only instances in which a right to access is directly recognised for a legal entity are regulated in Chapter V and benefit public-sector bodies or union institutions, agencies or bodies (eg in cases of emergencies).[197]

This dependency on the consumer's initiative, while conducive to the individual's rights and freedoms, also makes the access rights under the DA significantly less useful when it comes to providing business users with the necessary tools to develop powerful analytical methods themselves, including ML models.[198] After all, the value of the datasets kept by gatekeepers lies largely in their comprehensiveness. If business users are dependent on first agreeing on the conditions of the processing with the end user (Article 6(1) DA), the collection of a dataset, with a wide enough scope to be relevant, will be considerably more difficult due to simple transaction costs.

Significantly, though, the aims pursued by the DMA are also recognised by the DA.[199] First, they are recognised directly through Article 5(2) DA, which excludes businesses designated as gatekeepers from the possibility of making use of the outlined access rights.[200] Second, they are recognised indirectly through Article 7(1) DA, which provides that micro and small enterprises do not need to accommodate such access. This clearly underlines the overarching purpose of both acts to even out current imbalances in the market. The extent to which they will achieve their common goal, however, remains doubtful, as Section V.2 explores further.

V. Information requirements regarding advertising

Last but not least, Articles 5 and 6 DMA also contain several information requirements for gatekeepers regarding their advertising practices. Their primary purpose is to counteract the information asymmetry between platforms and business users concerning the conditions and functioning of advertisements.[201] While the requirements themselves are arguably the least intrusive of those concerning data and AI, they are still crucial for many business users. Under the current market realities, advertising constitutes the central source of revenue for most platforms, including gatekeepers, as well as their competitors.[202] The Commission is currently considering tightening the EU acquis even further with a view to fair advertising.[203] Given their economic significance, the potential impact of the DMA rules on advertising will be examined in the following subsections and compared to the information requirements in the AI Act.

[196] See, for a critique regarding the effectiveness of the DA's guarantee of self-determination, Specht-Riemenschneider, supra, note 194, 816 et seqq.

[197] On the B2G ("business to government") relationship, see ibid, 824 et seqq.

[198] For further critique on the dependency on the consumer's initiative, see, eg, R Podszun and P Offergeld, "The EU Data Act and the Access to Secondary Markets" (24 October 2022) <https://ssrn.com/abstract=4256882> (last accessed 22 November 2023) 45 et seq.

[199] Data Act, supra, note 192, 5.

[200] See also I Graef and M Husovec, "Seven Things to Improve in the Data Act" (2022) <https://ssrn.com/abstract=4051793> (last accessed 22 November 2023) 2 et seq., who mention the possibility of bypassing the exclusion by relying on Art 20 GDPR instead.

[201] See, eg, D Geradin and D Katsifis, "An EU competition law analysis of online display advertising in the programmatic age" (2019)15 European Competition Journal 55, 62; Recital 45 DMA.

[202] Geradin and Katsifis, supra, note 201, 55 f.

[203] L Bertuzzi, "Dark patterns, online ads will be potential targets for the next Commission, Reynders says" (*EURACTIV*, 9 December 2022), <https://www.euractiv.com/section/digital/interview/dark-patterns-online-ads-will-be-potential-targets-for-the-next-commission-reynders-says> (last accessed 9 December 2022).

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

1. Articles 5(9) and (10) DMA

Articles 5(9) and (10) DMA oblige gatekeepers to disclose to both advertisers and publisher, respectively, upon request, the metrics used to calculate prices, fees and remunerations for each advertisement placed or displayed. Metrics used in the field of advertisement, and e-commerce in particular, are both numerous and varied.[204] Also, based on their crucial influence on the business model, enterprises are constantly working on their optimisation.[205] The impact of a disclosure requirement on gatekeepers should not be underestimated, especially when taking into consideration the fact that prices for advertising are typically determined automatically within fractions of a second by real-time auctions, which in turn are backed by complex algorithms or ML techniques.[206] Concerning AI, this new obligation therefore indirectly compels gatekeepers, and ad tech networks,[207] to use explainable AI systems.[208]

The rule also has another quite significant impact. Since the requirement refers to each individual advertisement, this effectively creates an obligation to deliver local explanations, which disclose the relevant features for each individual decision.[209] While this can be quite burdensome, it is increasingly possible even with complex, "black box" systems such as artificial neural networks, as seen.[210] A similar obligation might have been considered potentially too burdensome for regular companies.[211] Its fulfilment by gatekeepers should, however, be considered a proportionate measure.

2. Article 6(8) DMA

Similarly, Article 6(8) DMA requires gatekeepers to disclose, to advertisers and publishers, the tools they use to measure the performance of advertising. In instances in which AI systems are being deployed, these will often be performance metrics, such as predictive accuracy.[212] Additionally, the data necessary for business users to perform their own verifications must be made available. In the context of advertising, these will include data such as the conversion rate (*click-through rate*).[213]

3. Alignment with the AI Act (Articles 11, 13, 15)

In comparison with the transparency requirements in Articles 13 and 11 AI Act in conjunction with Annex IV AI Act, it is noteworthy that the information owed according to

[204] X Liu et al, "Neural Auction: End-to-End Learning of Auction Mechanisms for E-Commerce Advertising" (2021) Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining 3354-64 <https://doi.org/10.1145/3447548.3467103> (last accessed 9 December 2022); Z Zhang et al, "Optimizing Multiple Performance Metrics with Deep GSP Auctions for E-Commerce Advertising" (2021) Proceedings of the Fourteenth ACM International Conference on Web Search and Data Mining (WSDM '21) 993–1001 <https://doi.org/10.1145/3437963.3441771> (last accessed 9 December 2022); SC Geyik et al, "Joint Optimization of Multiple Performance Metrics in Online Video Advertising" (2016) Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16) 471–80. <https://doi.org/10.1145/2939672.2939724> (last accessed 9 December 2022).

[205] Cf. ibid.

[206] Geradin and Katsifis, supra, note 201, 61.

[207] See, on their significance, see D Geradin, T Karanikioti and D Katsifis, "GDPR myopia: how a well-intended regulation ended up favouring large online platforms – the case of ad tech" (2021) 17 European Competition Journal 47, 50 et seqq.

[208] See the overview in note 79, supra.

[209] See supra, note 46, and accompanying text.

[210] ibid.

[211] Cf. Geradin supra, note 207, 48–49.

[212] Liu et al, supra, note 204; Zhang et al, supra, note 204.

[213] For details on the click-through rate and other procedures, see Zehlike et al, supra, note 16.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

the DMA has a significantly narrower scope of application, as it only concerns advertising. Conversely, the type of information that needs to be provided is more detailed, as businesses will be entitled to receive local explanations concerning each individual case.[214] In the DMA, only global explanations, concerning the entire model, are foreseen (see Section II.2.f supra).

However, a right to a more "concrete" explanation will not necessarily benefit gatekeepers' competitors, who might have been more interested in an overall explanation of the functioning of the underlying AI system. On the other hand, the DMA provides businesses with a completely new claim, as a right to a local explanation is either not foreseen or heavily debated under other EU acts (see Section II.2 supra). While the option to receive information on individual advertisements provides advertisers and publishers with more choice, it is, from a systematic perspective, surprising to see this included in a piece of legislation aimed at remediating distortions in competition, as local approaches are generally preferred by users, whereas business users and developers should have a greater interest in global explanations.[215] Furthermore, it should be acknowledged that the verification efforts enabled by the DMA partly presuppose considerable technical prowess on the part of advertisers and publishers, which will not always be present to a sufficient extent.

Nevertheless, the overall aim of the DMA to ensure greater transparency in the advertising market and in the AI systems used in it should be welcomed from both a competition and a due process perspective.[216] In its current form, it has potential to facilitate the comparison of advertising conditions and the verification of a platform's promises, fostering the contestability of gatekeeper positions in the advertising and publishing market.

VI. Regulating gatekeeper data and AI going forward

The preceding sections have revealed that the DMA complements a growing and increasingly elaborate regime regulating the data, algorithms and models used by gatekeepers in performing their core business functions. However, the analysis has also shown that significant shortcomings remain. The final section of the paper therefore sketches amendments and policy proposals regarding the three areas covered here: transparency, access and fairness.

1. Transparency

With respect to transparency, the DMA adds to an already copious but incoherent regime demanding various types of transparency activities from platforms and AI providers or users. The AI Act is bound to complicate this regime with further disclosure and explainability duties.

What is lacking, so far, is a unifying framework for transparency and explainability with respect to complex software systems, including AI. This gap is detrimental for both end users and AI developers as it entails legal uncertainty, raises the cost of compliance and litigation and fails to meet the purpose of the transparency regime: balancing the

[214] Cf., eg, Annex IV, para 2, lit b AI Act: meaning of the various parameters; see also Hacker and Passoth, supra, note 14, 343, 357 et seqq.

[215] Cf. The Royal Society, "Explainable AI: The Basics – Policy Briefing" (November 2019) <https://royalsociety.org/topics-policy/projects/explainable-ai/> (last accessed 2 October 2022) 14.

[216] Cf. also D Citron and F Pasquale, "The scored society: due process for automated predictions" (2014) 89 Washington University Law Review 1; D Citron, "Technological due process" (2007) 85 Washington University Law Review 1249.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

fundamental right of data protection and access to information, both of consumers and business users, with countervailing rights and interests of gatekeepers.

a. Relationship to explainable AI

With the DMA now enacted, the AI Act would have the unique opportunity to consolidate the EU algorithmic transparency regime, specifically for AI and complex software falling under the broad AI definition of Article 3(1) AI Act, read in conjunction with Recitals 6a and 6b AI Act.[217] To foster innovation and legal certainty – whose absence is quite detrimental to AI development and deployment – the requirements for explanations should be further specified and adapted to varying recipients.

*i. Opening the black box***.** As mentioned, many techniques have been developed over recent years in computer science research for opening so-called "black box" AI systems long thought to be particularly opaque.[218] However, not all of these techniques fit the needs of all audiences; rather, they must be actionable.[219] End users, for example, will primarily be interested in ascertaining that the rankings created by gatekeepers reflect their interests and preferences and not those of the gatekeeper. To this end, feature salience can indeed be helpful. If the main features list "affiliation with gatekeeper" or other categories unrelated to consumer preference, end users may switch to other providers.

Even though many customers will probably ignore such disclosures,[220] they will nevertheless be analysed by consumer associations, journalists or even regulatory agencies, who may then act as information intermediaries or enforcement entities for any suboptimal or illegal features found.[221] Increasingly, such analysis is automated using ML as well, so that the amount of data analysed and the policies flagged as problematic increase substantially.[222] This, in turn, increases deterrence and compliance pressure.

*ii. Actionable explanations for business users***.** For business users, in turn, feature salience is important as well: they may deduce whether their products were ranked in a meaningful and fair way. Business users also have a much greater incentive than consumers to monitor ranking conditions, as these are essential for commercial success on the platform. Opaque rankings or main features unrelated to product performance will raise incentives to switch platforms, fostering the contestability of rankings markets.

Moreover, however, business users will also be interested in how they may improve the ranking of their products. Technically, this may be achieved by delivering so-called

[217] See supra, note 1.

[218] For an overview, see references in note 79, supra.

[219] Hacker and Passoth, supra, note 14, 343, 362 et seqq.

[220] Cf. O Ben-Shahar and A Chilton, "Simplification of privacy disclosures: an experimental test" (2016) 45 The Journal of Legal Studies S41; JA Obar and A Oeldorf-Hirsch, "The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services" (2020) 23 Information, Communication & Society 128.

[221] See, eg, P Bischhoff, "Comparing the privacy policy of Internet giants side-by-side (*comparitech*, March 2017) <https://www.comparitech.com/blog/vpn-privacy/we-compared-the-privacy-policies-of-internet-giants-side-by-side/> (last accessed 6 September 2022); Forbrukerrådet, Deceived by Design, Report, 2018, <https://fil.forbrukerradet.no/wp-content/up/loads/2018/06/2018-06-27-deceived-by-design-final.pdf> (last accessed 6 September 2022).

[222] See, eg, S Zimmeck and SM Bellovin, "Privee: an architecture for automatically analyzing web privacy policies" (2014) 23rd USENIX Security Symposium 1; L Austin et al, "Towards Dynamic Transparency: The AppTrans (Transparency for Android Applications) Project" (Working Paper, 2018) <https://ssrn.com/abstract=3203601> (last accessed 22 November 2023).

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

counterfactual explanations.[223] In this situation, however, they raise the question of possible manipulations of the ranking, to which we now turn.

b. Trade secrets and manipulation

From a legal point of view, the mentioned disclosure requirements also raise the problem of trade secrets and the manipulability of the ranking. These issues need to be considered in transparency rules concerning AI and software more generally.

i. Platform problems. First, far-reaching transparency rules may undermine incentives to invest in innovation in the first place and counteract the protection afforded by intellectual property rights. Competitors may reverse engineer algorithms, models or datasets and free-ride on discoveries made by the holders of trade secrets.[224] However, the empirical evidence of the fact of trade secrets on innovation is mixed. While some results suggest that trade secrets spur investment in research and development, particularly in the information technology sector,[225] another recent empirical study suggests that strengthened trade secret regimes may in fact hamper innovation.[226] In markets dominated by informal networks of learning and collaboration, strong trade secrets protection often does more harm than good by preventing the exchange of ideas and knowledge.[227] AI research and development are, arguably, to a considerable extent based on such networks. Hence, claims to protect trade secrets for the sake of AI development should be taken *cum grano salis*. In the specific situation of business users seeking information about the data and models used by gatekeepers, the argument of protecting trade secrets is weakened even further, as the market would generally benefit from greater contestability and competition, challenging the entrenched position of gatekeepers.

Second, providing information about the main factors relevant for ranking provides opportunities for their manipulation.[228] This problem is particularly virulent if the predictive features are merely correlated with and not causal for the target variable. In this case, they can be artificially changed by business users and cause an improvement in the ranking position without any concurrent improvement in the characteristics that the target variable seeks to capture. For example, in one study, the purchase of felt tips (for preventing damage to floors caused by the moving of furniture) was found to be highly predictive for creditworthiness. The relationship is obviously a mere correlation, not a causal one. If it was disclosed that a credit-scoring model takes this into account, candidates could order felt tips on purpose to improve their creditworthiness.

Ultimately, however, this does not speak against transparency per se but in favour of using causal inference instead of models based primarily on correlations. Nevertheless,

[223] S Wachter, B Mittelstadt and C Russell, “Counterfactual explanations without opening the black box: Automated decisions and the GDPR” (2017) 31 Harvard Journal of Law & Technology 841; RK Mothilal, A Sharma and C Tan, “Explaining machine learning classifiers through diverse counterfactual explanations” (2020) Proceedings of the 2020 conference on Fairness, Accountability, and Transparency 607.

[224] DD Friedman, WM Landes and RA Posner, “Some economics of trade secret law” (1991) 5 Journal of Economic Perspectives 61.

[225] I Png, “Law and innovation: evidence from state trade secrets laws” (2017) 99 Review of Economics and Statistics 167.

[226] A Contigiani, DH Hsu and I Barankay, “Trade secrets and innovation: evidence from the ‘inevitable disclosure’ doctrine” (2018) 39 Strategic Management Journal 2921.

[227] L Pedraza-Fariña, “Spill your (trade) secrets: knowledge networks as innovation drivers” (2016) 92 Notre Dame Law Review 1561.

[228] J Bambauer and T Zarsky, “The Algorithm Game” (2018) 94 Notre Dame Law Review 1.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

while causal ML is making steady progress,[229] it cannot be deployed across the board yet.[230] As long as correlational models persist as the state-of-the-art technology in many areas, manipulability needs to be taken into account when designing transparency rules.

ii. Legal solutions. Article 5(6) of the P2B Regulation explicitly considers trade secrets and manipulability. It exempts the disclosure of such information from the general transparency obligation that makes it possible to manipulate the ranking with sufficient probability and mentions the Trade Secrets Directive.[231] Still, it does not provide guidelines for balancing the need for transparency with these countervailing interests. What needs to be disclosed are the features and their relative importance. The described tension can be resolved, in our view, in such a way that no precisely quantified weights are divulged but only intervals or even only ordinally ordered lists of the relevance of the individual parameters.[232] In this way, it will be distinctly more difficult to reverse engineer the model, and the risk of manipulability is also lowered.

Importantly, one will have to apply the manipulation and trade secret protection of Article 6(5) P2B Regulation to Article 7 of the P2B Regulation by analogy, since manipulation and free-riding can occur just as much based on the information provided in the terms and conditions.

Turning to the GDPR, Article 15 GDPR, as with all fundamental and GDPR rights, does not apply without restrictions. Next to the general exemptions of manifestly unfounded and excessive requests contained in Article 12(5) GDPR, Article 15(4) GDPR also postulates that the rights and freedoms of others should not be adversely affected. While the clause technically only refers to the right to obtain a copy of the personal data undergoing processing (Article 15(3) GDPR), it should by analogy and where necessary apply to the details provided under Article 15(1) GDPR as well. This understanding is also supported by Sentence 5 of Recital 63 GDPR and Article 29 WP Guidelines.[233]

This analogy is of particular relevance to the scope of access to the underlying model. While rights of others in the sense of different data subjects are unlikely to be affected, the trade secrets or intellectual property rights of the gatekeeper might in fact be engaged and warrant a restriction of the individual's rights.[234] Significantly, this does not imply that, in such instances, no information should be provided. Rather, the applicable rights need to be balanced. For example, less intrusive means of access, such as partial access or less granular information, should be considered, like in the case of information according to the P2B Regulation.

2. Access

For consumers, transparency (eg an explanation of a decision) is an important prerequisite to contesting data practices. Business users, however, additionally need access to data, and potentially models, to use them for products that may eventually challenge the

[229] See, eg, J Pearl, "The seven tools of causal inference, with reflections on machine learning" (2019) 62 Communications of the ACM 54; J Richens et al, "Improving the accuracy of medical diagnosis with causal machine learning" (2020) 11 Nature Communications 1.

[230] J Kaddour et al, "Causal machine learning: a survey and open problems" (2022) arXiv preprint arXiv: 220615475.

[231] See also C Busch, "Mehr Fairness und Transparenz in der Plattformökonomie? Die neue P2B-Verordnung im Überblick" (2019) 121 Gewerblicher Rechtsschutz und Urheberecht 788, 793.

[232] See also Bäcker, supra, note 61, Art 13, para 54; Hacker and Passoth, supra, note 14, 343, 350.

[233] Article 29 Data Protection Working Party, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679" (2017) WP 251, 17.

[234] Cf. also ibid.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

gatekeepers' competitive position. The access rights contained in the DMA and the DA are steps in the right direction. However, they still do not go far enough.

First, as seen, access to end user data under Article 6(10) DMA hinges on end user consent, which is generally collected by gatekeepers. They have, however, an incentive to discourage consent in this respect in order to block access requests by potential competitors. Hence, business users should be allowed to review and veto the design of consent requests by gatekeepers insofar as consent to the reuse of personal data by their own end users according to Article 6(10) DMA is concerned. In such a setting, business users would be able to reject consent designs under which gatekeepers, via framing or other behavioural effects,[235] seek to obtain consent for their own data-sharing practices but to nudge users to withhold consent regarding data-sharing with business users.

Furthermore, an amendment to Article 6(11) DMA should specify that personal data cannot be anonymised arbitrarily by gatekeepers but only in a way that preserves the utility of the dataset to a sufficient degree, using state-of-the-art PPML techniques.[236] Gatekeepers should have to document their choice of anonymisation technique and the reasons for choosing it.[237] This would ensure that gatekeepers have to strike an explicit, documented and auditable balance between safeguarding the data protection rights of affected persons and the interests of competitors in receiving useful datasets.

3. Fairness

Our third proposal relates to rankings. They are now at the heart of the digital economy. Rankings are the logical answer to the digitally mediated, excessive supply of information and products.[238] By their ordering and prioritising effect, they control demand, search and buying behaviour.[239] They represent the core of the business models of many gatekeepers and are therefore rightly regulated even more comprehensively in the final version of the DMA than in the Commission's draft. In particular, the expansion to crawling and indexing seems sensible, because these activities can decisively influence the visibility of products and thus, for example, the rankings created by search engines (cf. Recital 51 DMA). The explicit inclusion of virtual assistants as possible creators of rankings (Articles 6(5) and 2(22) DMA) is also fully justified.

Already now, classical anti-discrimination law is in principle applicable to rankings.[240] However, it often falls short, since the protected characteristics listed there, such as gender or ethnic origin, generally only refer to persons and not to objects. Therefore, in principle, it is necessary to operate, in the area of e-commerce, with a broadened equal treatment rule. This provides an important building block for fairness in e-commerce.[241] Article 6(5) DMA further develops the case constellations known from competition law. However, the rule must remain operationalisable for gatekeepers, especially in view of the significant threat of sanctions. Furthermore, the economic core function of rankings – to enable the realisation of preferences through selective ordering – must not

[235] See, eg, J Sakshaug et al, "The effect of framing and placement on linkage consent" (2019) 83 Public Opinion Quarterly 289.

[236] See supra, note 184 and accompanying text.

[237] Cf. Art 11 AI Act.

[238] Ursu, supra, note 23, 530.

[239] ibid, 549.

[240] Hacker, supra, note 133, 1159.

[241] On the concept of fairness in the DMA, see H Schweitzer, "The Art to Make Gatekeeper Positions Contestable and the Challenge to Know What Is Fair: A Discussion of the Digital Markets Act Proposal" (2021) 29 Zeitschrift für Europäisches Privatrecht 2021, 503; Bongartz et al, supra, note 15, 62; see also W Fikentscher, P Hacker and R Podszun, *FairEconomy* (Berlin, Springer 2013).

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

be undermined. At the same time, however, the competitive effects of rankings must be considered precisely because of their selection and steering effects.

This is epitomised by currently hotly debated popularity-based rankings, according to which the order is defined by the presumed attractiveness of the items to users.[242] On the one hand, a differentiation of ranked products according to the expected purchase and click rates does provide a feasible shortcut for approximating rankings to user preferences. On the other hand, the technique may have anti-competitive effects insofar as products that have already been on the market for a longer time tend to be favoured over new ones, since only the former can demonstrate successful historical purchase and click rates.[243] This may entrench the position of historically successful brands and companies to the detriment of newcomers. Therefore, it seems reasonable to compel gatekeepers, by way of a teleological interpretation of Article 6(5) DMA, to shuffle their popularity rankings – for example, by reserving some attractive ranking positions for new products.[244] In this way, in our view, the interest of consumers in receiving product recommendations that are as predictive of preferences as possible could be combined with the interest in dynamic competition.

The same can be said for voice commerce, which is essential especially in the area of virtual assistants. Here, usually only one product is selected, which is ordered via voice control.[245] Behind this selection, however, is a ranking,[246] which is typically only visible in the app.[247] This ranking must also conform to Article 6(5) DMA for gatekeepers. Recital 48 DMA now makes this unequivocally clear. However, since a single product is highlighted, the gatekeeper ought to ensure a shuffling of this highlighted position to comply with Article 6(5) DMA. This could be achieved, for example, by permuting (across all comparable queries) the items occupying the top three ranking positions: each of the three highest-ranked elements' criteria are randomly placed in the top position for one third of the queries. The underlying ranking must, of course, be transparent, fair and non-discriminatory as per Article 6(5) DMA. In this way, catering to consumer preferences is combined with a technologically mediated process for fostering workable competition and preventing winner-takes-all markets.[248]

VII. Summary

The DMA will not only provide more competitive opportunities on, alongside and between large online platforms but will also decisively shape the way gatekeepers and their competitors deal with AI. Especially in the digital economy, the DMA's impact on AI systems is likely to be much more noticeable than that of the future AI Act, unless its list of high-risk applications is significantly expanded in the trilogue.

The provisions of the DMA that are relevant for AI can be divided into four areas. First, new rules for fair rankings are introduced. With this, the DMA ventures into the core of the

[242] See, eg, Y Shmargad and S Klar, "Sorting the news: how ranking by popularity polarizes our politics" (2020) 37 Political Communication 423; F Germano et al, "The few-get-richer: a surprising consequence of popularity-based rankings?" (2019) The World Wide Web Conference 2764.

[243] S de Brouwer, "Privacy self-management and the issue of privacy externalities: of thwarted expectations, and harmful exploitation" (2020) 9 Internet Policy Review 1, 17.

[244] S Pandey et al, "Shuffling a Stacked Deck: The Case for Partially Randomized Ranking of Search Engine Results" (2005) Proceedings of the 31st VLDB Conference, DOI:10.48550/arXiv.cs/0503011.

[245] A Ezrachi and ME Stucke, "Is your digital assistant devious?" Oxford Legal Studies Research Paper 52/2016.

[246] A Mari, "Voice Commerce: Understanding Shopping-Related Voice Assistants and their Effect on Brands" (IMMAA Annual Conference, 2019) 4.

[247] Busch, supra, note 231, 792.

[248] Cf. B Kuchinke and M Vidal, "Exclusionary strategies and the rise of winner-takes-it-all markets on the Internet" (2016) 40 Telecommunications Policy 582.

https://doi.org/10.1017/err.2023.81 Published online by Cambridge University Press

AI-based business model of gatekeepers. Article 6(5) DMA consolidates the prohibition of self-preference known from competition law and transparency rules for rankings already existing in other EU law instruments. However, the inclusion of F(R)AND criteria for rankings is new and potentially groundbreaking. They point significantly beyond existing anti-discrimination law and, in our view, introduce a need for justifying differentiations between comparable products in the ranking. From a technical point of view, techniques developed in the computer science research on algorithmic fairness can be used. However, adapting this framework to the DMA is complex due to the potentially unlimited number of protected attribute combinations – unlike in classical anti-discrimination law. The compliance requirements must take this into account.

Second, the use of data and thus also, in particular, their collection and use for AI training by gatekeepers are significantly restricted. The thrust here is diametrically opposed to that of Article 10 AI Act. To put it bluntly: the DMA does not, in contrast to the AI Act, seek to foster high-performing AI but to prevent additional improvements by gatekeepers' models based on the specific competitive setting in which gatekeepers operate. Third, access rights are created for business users to enable them to develop high-performance AI models themselves. Fourth, the DMA harnesses information obligations to reduce the information asymmetry between gatekeepers and their business users, especially in the area of advertising.

We complement these findings with policy suggestions in three main areas. First, the AI Act should spell out a coherent and precise transparency regime. It must clarify the relationship to various technical strategies to implement explainable AI and take trade secrets and the manipulability of rankings into account via a smart design of disclosures. Second, access rights for users need to be expanded and data protection safeguards be balanced with the interests of gatekeepers' competitors, and of society at large, in the provision of meaningful datasets that do allow for the development of products contesting gatekeeper positions. Third, fair rankings also need to balance the rankings' original economic function – selecting items and thus facilitating the fulfilment of consumer preferences – with the broader competitive interest in preventing winner-takes-all markets in which newcomers fight an uphill battle to climb in popularity-based rankings.

All in all, this paper shows that the DMA seeks to bridge a variety of economic and non-economic discourses and combines crucial societal interests that necessitate delicate balancing exercises at many points. The recently enacted regulation furthermore points to currently underexplored questions, at the intersection of law and computer science, surrounding the optimal degree of transparency and fairness of e-commerce rankings – one of the key competition ingredients of the digital economy.

Acknowledgments. We are grateful for research assistance from Sarah Großheim and Marco Mauer.

Competing interests. The authors declare none.

Cite this article: P Hacker, J Cordes, and J Rochon (2024). "Regulating Gatekeeper Artificial Intelligence and Data: Transparency, Access and Fairness under the Digital Markets Act, the General Data Protection Regulation and Beyond". *European Journal of Risk Regulation* **15**, 49–86. https://doi.org/10.1017/err.2023.81