

De procedurele wending van de DSA: het recht op menselijke tussenkomst in content moderation

Column

Ljubiša Metikoš*

Met de hedendaagse vloedgolf aan *mediacontent* worstelen grote socialemediaplatformen met de uitvoering van *content moderation*. Onder druk van de wetgever, aandeelhouders en platformgebruikers, moeten platformen illegale of anderszins problematische content snel en op grote schaal aan kunnen pakken. Het zou echter financieel en praktisch onmogelijk zijn om mensen alle besluiten hierover te laten nemen. Platformen zetten daarom vaak zelflerende algoritmen in die, semiautonom, content verwijderen, lager rangschikken, of uitsluiten van advertentie-inkomsten.¹ Deze algoritmen zijn doorgaans ondoorzichtig, en bevatten vaak ook discriminatoire vooroordelen die moeilijk vooraf uit te bannen zijn.² Regelmatig vindt er dan ook onterechte en soms onbedoelde *content moderation* plaats, die in strijd kan zijn met de vrijheid van meningsuiting, de vrijheid van ondernemerschap, of het discriminatieverbod.

In zulke gevallen kan een onterecht benadeelde *content creator* klagen bij een mens. Artikel 13 van het Europees Verdrag van de Rechten van de Mens en artikel 47 van het Europees Handvest geven eenieder het recht op een effectief rechtsmiddel wanneer diens fundamentele rechten worden geschonden. Een toegankelijke en effectieve rechtsprocedure is echter vaak niet weggelegd voor kleine *content creators*.³ Het ligt daarom voor de hand platformen te verplichten om platformgebruikers te voorzien van een toegankelijk intern klachtafhandelingssysteem, waar men een oordeel krijgt van een menselijke geschilbeslechter.

Een dergelijk *recht op menselijke tussenkomst* is oorspronkelijk terug te vinden in het Europees privacy- en gegevens-

beschermingsrecht. De *travaux préparatoires* van de Data-beschermingsrichtlijn laten zien dat de Europese wetgever zich al in 1994 zorgen maakte omdat besluitnemers hun eigen verantwoordelijkheid zouden verloochenen wanneer procedures té geautomatiseerd worden.⁴ Thans staat in artikel 22 lid 3 AVG de verplichting voor de verwerkingsverantwoordelijke om te waarborgen dat de betrokkene het recht heeft op betekenisvolle menselijke tussenkomst, een recht op een verklaring en een recht van verzet.⁵

Voor grote socialemediaplatformen bevat artikel 20 DSA inmiddels ook specifieke verplichtingen voor het instellen van klachtafhandelingssystemen, die niet *volledig* geautomatiseerd mogen zijn.⁶ Platformen hebben nu de verplichting om te voorzien in procedures die tot “snelle, niet-discriminerende, niet-willekeurige en eerlijke resultaten leiden, en die aan menselijke controle worden onderworpen”.⁷

Maar wat maakt menselijke tussenkomst *betekenisvol*?⁸ Hoe moet een mens geautomatiseerde processen controleren? Wat als een mens juist méér discrimineert,⁹ of anderszins besluiten minder accuraat maakt?¹⁰ Hoe ervaren platformgebruikers klachtafhandelingsprocedures wanneer een mens (niet) betrokken is?¹¹ Deze onderwerpen werden vroeger vaak slechts besproken binnen de technologie- en rechtsfilosofie, en het privacy- en procesrecht. De *procedurele wending*¹² van de DSA brengt echter met zich mee dat voortaan óók mediarechtjuristen moeten reflecteren op de relatie tussen mens en machine, de waarborgen van een eerlijk proces en de verplichtingen van menselijke geschilbeslechters.

* Mr. Lj. Metikoš is promovendus ‘Digital Justice’ aan het Instituut voor Informatierecht, Paul Scholten Center for Jurisprudence, RPA Human(e)AI, Universiteit van Amsterdam.

1 Tarleton Gillespie, Hoofdstuk 4, ‘Three Imperfect Solutions to the Problem of Scale’, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, Yale University Press 2018, <https://www.degruyter.com/document/doi/10.12987/9780300235029-004/html>, geraadpleegd op 14 oktober 2021.
2 Reuben Binns and others, ‘Like Trainer, Like Bot? Inheritance of Bias in Algorithmic Content Moderation’, Springer International Publishing 2017.
3 Naomi Appelman e.a., *Access to Digital Justice: In Search of an Effective Remedy for Removing Unlawful Online Content*, Edward Elgar Publishing 2022, <https://www.elgaronline.com/display/book/9781802203820/book-part-9781802203820-18.xml>, geraadpleegd op 20 maart 2023.
4 Art. 16 Databeschermingsrichtlijn, Explanatory Memorandum for a Council Directive, on the protection of individuals with regard to the processing of personal data and on the free movement of such data, COM(92) 422 final – SYN 287, Brussel, 15 Oktober 1992, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:51992PC0422&from=DE>, geraadpleegd op 20 maart 2023.
5 EDPB, ‘Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679’ (2018), https://edpb.europa.eu/our-work-tools/our-documents/guideline/automated-decision-making-and-profiling_en, geraadpleegd op 20 maart 2023. NB: de betekenis, inhoud en reikwijdte van het ‘recht op een verklaring’ in de AVG, is zeer omstreven in het hedendaagse academische discours. Zie bijv.: Bryan Casey, Ashkon Farhangi and Roland

Vogl, ‘Rethinking Explainable Machines: The GDPR’s “Right to Explanation” Debate and the Rise of Algorithmic Audits in Enterprise’ (2019) 34 *Berkeley Technology Law Journal*, p. 143.

6 Art. 20 lid 6 DSA.

7 Overweging 58 DSA.

8 Riikka Koulu, ‘Proceduralizing Control and Discretion: Human Oversight in Artificial Intelligence Policy’ (2020) 27 *Maastricht Journal of European and Comparative Law*, p. 720.

9 Marco Almada, ‘Human Intervention in Automated Decision-Making: Toward the Construction of Contestable Systems’, *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*, Association for Computing Machinery 2019, <https://doi.org/10.1145/3322640.3326699>, geraadpleegd op 20 maart 2023.

10 Ben Green, ‘The Flaws of Policies Requiring Human Oversight of Government Algorithms’, Social Science Research Network 2021) SSRN Scholarly Paper ID 3921216, <https://papers.ssrn.com/abstract=3921216>, geraadpleegd 10 februari 2022.

11 Sarah Myers West, ‘Censored, Suspended, Shadowbanned: User Interpretations of Content Moderation on Social Media Platforms’ (2018) 20 *New media & society* 4366.

12 Christoph Busch & Vanessa Mak, ‘Putting the Digital Services Act into Context: Bridging the Gap between EU Consumer Law and Platform Regulation’ (1 May 2021), <https://papers.ssrn.com/abstract=3933675>, geraadpleegd op 20 maart 2023; Pietro Ortolani, ‘If You Build It, They Will Come: The DSA’s “Procedure Before Substance” Approach’ [2022] *Verfassungsblog*, <https://verfassungsblog.de/dsa-build-it/>, geraadpleegd op 20 maart 2023.