



Universiteit
Leiden
The Netherlands

Countering online hate speech: how to adequately protect fundamental rights?

Nave, E.V.R.

Citation

Nave, E. V. R. (2025, July 3). *Countering online hate speech: how to adequately protect fundamental rights?*. Meijers-reeks. Retrieved from <https://hdl.handle.net/1887/4252655>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4252655>

Note: To cite this publication please use the final published version (if applicable).

Countering online hate speech

Countering online hate speech

How to adequately protect fundamental rights?

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 3 juli 2025
klokke uur 16.00 uur

door

Eva Vaz Rodrigues da Nave

geboren te Guarda, Portugal

in 1991

Promotoren: prof. dr. S. van der Hof
prof. dr. T.E. McGonagle

Promotiecommissie: prof. dr. ir. B.H.M. Custers
prof. dr. L.A. van Noorloos
dr. H.C. Lahmann
prof. dr. drs. J.D. Temperman (Erasmus Universiteit)
prof. dr. N. Suzor (Queensland University of
Technology)

Opmaak binnenwerk: Anne-Marie Krens – Tekstbeeld – Oegstgeest
Omslagontwerp: Primo!Studio – Delft
Drukwerk: Ipskamp Printing – Amsterdam

© 2025 E. Nave

ISBN 978 94 6473 832 2

Behoudens de in of krachtens de Auteurswet gestelde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van reprografische verveelvoudigingen uit deze uitgave is toegestaan op grond van artikel 16h Auteurswet dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorecht (Postbus 3051, 2130 KB Hoofddorp, www.reprorecht.nl). Voor het overnemen van (een) gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (art. 16 Auteurswet) kan men zich wenden tot de Stichting PRO (Stichting Publicatie- en Reproductierechten Organisatie, Postbus 3060, 2130 KB Hoofddorp, www.stichting-pro.nl).

No part of this book may be reproduced in any form, by print, photoprint, microfilm or any other means without written permission from the publisher.

Para a Ana, o Fernando e o Poppy

Acknowledgements

I want to thank my supervisors, Simone van der Hof and Tarlach McGonagle, for sharing their valuable guidance and encouraging me to think critically, key to successfully complete this thesis. Within eLaw, I found inspiring and supporting colleagues. Nikki Vosters, thank you for being my Ph.D. sister, our thesis sessions and chats at Oba created authentic and warm memories. Dimitra Stefoudi, our online writing sessions in the first PhD year, during the pandemic, were life-saving. Kimia Heidary, thanks for contributing to a fun organization of the 2021 eLaw Colloquium and for your advice on the final steps leading up to the defence. Regina Noort and Francien Dechesne, thank you for your kindness and for your check-ins. Gianclaudio Malgieri, your commitment towards the research community at eLaw and your mentorship navigating publication and the research visit in Milan marked exciting times in my Ph.D. journey. Alan Sears, thanks for your patience to introduce me to the layered publishing American system as well as for your advice on conferences and prizes. Henning Lahmann, thanks for your overall support and for taking part of this research to your classes' discussions, that encouraged me to progress and to be more confident in my work at a challenging time. Amara García Adán and Zuzanna Gulczyńska, your visits at eLaw were refreshing and fun times as a more solid Ph.D. cohort.

I am forever grateful to Lottie Lane for accepting to collaborate with me in the second article; your availability, as well as your considerate and constructive feedback were key to decide to continue the Ph.D. journey and, now, I am also lucky to call you a friend. I also want to thank Stephan Raaijmakers and Thijs Veugen for the inspiring discussions during my secondment at the Netherlands Organization for Applied Scientific Research (TNO) and for the fruitful collaboration in this research. I would also like to thank Bart Custers, Marloes van Noorloos, Henning Lahman, Jeroen Temperman, and Nicolas Suzor for their time and thoughtful comments as members of the doctorate committee.

This thesis is also the result of many thought-provoking interactions with networks of other scholars. A huge thank you to my Ph.D. colleagues and friends from the Network for Training on Hate (NETHATE) project; this was a difficult topic, especially for those personally impacted by it, and I will always thank the opportunity to have learnt from you. During these Ph.D. years, I also benefited from the events organized by the Netherlands Network for Human Rights Research (NNHRR), a dynamic community where I met

like-minded scholars across different universities in the Netherlands; this is no doubt an initiative to be praised and I hope it continues to serve future young researchers.

I am grateful for the opportunity to have completed two research visits during my Ph.D. First, I want to thank Marco Pedrazzi and Federica Favuzza for hosting me at the University of Milan La Statale, and Federica Falconi for her availability to discuss this work. Here, I was lucky to be invited by Angelica Bonfanti to attend her course on Business and Human Rights, which significantly influenced the last article of this thesis. During this research visit, I was also very fortunate to be welcomed by Oreste Pollicino and Giovanni De Gregorio, from the Bocconi University, to join local workshops around the topics of public and private values in the digital age where I witnessed senior scholars' moving commitment to guide and collaborate with junior researchers on the academic debate. A thanks to Pietro Dunn and to Francesca Cagossi for being engaging friendly peers in Milan.

Second, I want to thank Katharine Gelber for hosting me at the University of Queensland; this was a memorable experience that I will forever cherish, thank you for your availability and interest to engage with my thesis, for kindly sharing so many literature sources and for connecting me with inspiring scholars. I want to thank Nicholas Carah for the inviting atmosphere at the Center for Digital Cultures, at the University of Queensland, and to thank Axel Bruns and Daniel Angus for the prompt welcome to join events at the Digital Media Research Centre, at the Queensland University of Technology; here, I am also thankful for the stimulating exchanges with Ariadna Matamoros-Fernández. During this research visit, I was lucky to have met Nicolas Suzor; thank you for welcoming me in your group discussions at the ARC Centre of Excellence for Automated Decision-Making and Society and for the provocative literature references. In Brisbane, I was lucky to have Maria Brown and Sara Roetman as lively peer companions.

I want to express my sincere gratitude towards Ana Paiva and her team at the Cabinet of the Portuguese Secretary of State for Science; for almost a year, I was fortunate to count with your encouragement and with the flexibility to accommodate the needed time to complete this thesis. My working experiences prior to this Ph.D. also had a huge influence in successfully completing it. I want to thank extraordinary mentors such as Celine Francois, Camille Aubourg, Aida Ariño-Fernández, Flora Sutherland and Abigail Hartley, you are my best examples of caring and rigorous leadership and your imprint has permeated through this thesis.

I need to thank my therapist, who guided me out of difficult Ph.D. times and helped me choose effective coping strategies.

I would like to thank my friends. Over the past years living in Amsterdam, I am grateful for Nahal's friendship; our chats, work and yoga sessions reinstated a healthy sense of routine key in balancing the research lifestyle. I want to thank the Sabbatical-love crew, Elisa, Kim, Vítor, André and Clélia,

our food and dance gatherings were always heartwarming moments. For over 15 years, my university friends from CBT have remained a fundamental place of joy and playfulness; I hold Guidels, Marissol, Jujineide, Gostini, Tonia, Tomi, Vidalz, Sara, Marinel and Zezi deeply in my heart. I should note that the seed for this thesis was first planted after a conversation with Maria; thank you for keeping my thoughts in check and alive. Other essential friendships have grounded my personality and continue to be one of the love assets that I turn to, and were, naturally, key to finish this Ph.D. Having grown up together in Guarda, I want to thank Inês, Catarina, Rita, Sofia, Ana, Rui, and Yman, our shared teenage years and long lasting friendships into adulthood have been safe havens for laughter and easiness. Among others, I am also so lucky to have in my life Marilda, Inês and Nuno, Shatakshi, Diana and Mega, Tâmara, Leonor, Paradise by proxy, Tomás and Kathryn, Joep and Ararat, Giuliana and Angelo, Daan and Linda, Rose, Filipe, Galamba and Carla, Jane, Lara and Alex, Noffar and Chip, Nana, Chema, Rodrigo and Esteban.

I would like to express my profound gratitude to my family. Tenho toda a sorte do mundo em ter dois incríveis sólidos núcleos familiares, sempre prontos para entrar em acção e para arranjar tempo para uma refeição juntos cheia de entretenimento e partilhas boas durante as visitas a Portugal; amo os Espigados e os Nave. Quero agradecer especialmente à minha irmã Sara e ao meu tio Júlio, que em tempos difíceis me deram especial consolo. Quero agradecer também aos meus avós Júlio, “Manela”, Dina, “Zé” António, João e bisavó Patrocínio, que pelo seu árduo trabalho conseguiram ir garantindo melhores condições de vida aos seus e que tanto me ensinaram a crescer. Um obrigada especial ao meu avô “Zé” António, um exemplo raro para a sua geração de patriarca congregador, carinhoso, trabalhador e curioso pela descoberta das novas possibilidades.

Dedico esta tese à minha mãe Ana. Palavras serão sempre insuficientes para te agradecer por me ensinares o amor, a resiliência, o compromisso, o movimento, as plantas, a música. A tese de mestrado não teria sido publicada se não fosse o teu encorajamento e, sem esse passo, dificilmente teria chegado a este doutoramento. És e serás sempre a minha referência. Que sorte ser tua filha!

Esta tese é dedicada também ao amor da minha vida e meu melhor amigo, o Fernando. Desde teres encontrado o anúncio desta posição, passando pela preparação para a entrevista, até todas as questões sobre o que é fazer investigação. *Garantidamente, esta tese – eu própria – não existiria sem ti.* Contigo aprendo a ponderação, o empenho, o humor e a coragem para experimentar o desconhecido. Que saibamos sempre preservar este lugar bonito que temos, no futuro com o nosso filho :)

Table of contents

LIST OF FIGURES AND TABLES	XV
ACRONYMS	XVII
1 INTRODUCTION	1
1.1 Context and social relevance	1
1.2 Problem statement and research questions	11
1.3 Methodology	13
1.3.1 Doctrinal legal research	13
1.3.2 Comparative legal research	16
1.3.3 Interdisciplinary legal research	17
1.4 Scope	18
1.5 Outline of the study	21
2 LEGAL CONCEPTUALIZATION OF HATE SPEECH	
Hate speech, historical or systematic oppressions, and European human rights	25
2.1 Introduction	26
2.2 Legal Theoretical Foundations of Hate Speech	30
2.2.1 First Legal Conceptualization	31
2.2.2 Impact and Harm	33
2.2.3 Regulation and Balancing Conflicting Rights	35
2.2.4 Current Legal Challenges	38
2.3 Approaches to Hate Speech at the Council of Europe	39
2.3.1 General Objectives	39
2.3.2 ECHR and ECtHR Jurisprudence	40
2.3.2.1 Hate Speech as a Clear Abuse of Rights	42
2.3.2.2 No Clear Abuse But Hate Speech is Prohibited	47
2.3.3 Other Treaties	50
2.3.4 Non-Treaty Initiatives	54
2.3.4.1 Observations Regarding the CM Recommendation on Combating Hate Speech	58
2.3.5 Overview Council of Europe, Hate Speech and Historical Oppression	64
2.4 Approaches to Hate Speech by the European Union	65
2.4.1 General Principles and Primary Sources	65
2.4.2 Secondary Sources	67
2.4.3 Soft Law	71

2.4.4	Overview European Union, Hate Speech and Historical Oppressions	72
2.5	Main Elements of Hate Speech in the European Context	72
2.5.1	What Is Hate Speech and What Does It Do?	72
2.5.2	How to Substantively Regulate Hate Speech	76
2.5.2.1	Hate Speech Is Always Criminal or Illegal Speech	76
2.5.2.2	People Targeted By Hate Speech Have Been Historically or Systematically Oppressed	76
2.5.2.3	Hate Speech Should Only Be Criminalized In Its Most Serious Forms	77
2.5.2.4	If Not Criminalized, Then Need to Balance Human Rights	78
2.5.2.5	Challenges With the Regulation of Online Hate Speech	78
2.6	Conclusion	79
3	HUMAN RIGHTS RESPONSIBILITIES OF ONLINE PLATFORMS TO PREVENT CRIMINAL HATE SPEECH How do European corporate preventive human rights responsibilities impact terms of service?	81
3.1	Introduction	82
3.2	Online hate speech is always illegal, sometimes criminalised	86
3.2.1	Original conceptualisation	87
3.2.2	Key conceptual elements in European regulation	88
3.2.2.1	Hate speech is always illegal	91
3.2.2.2	The most serious forms of hate speech are criminally actionable	92
3.3	Broader framework: AI and the corporate responsibility to respect human rights	93
3.3.1	United Nations Guiding Principles on Business and Human Rights	93
3.3.2	Initiatives by the Organization for Economic Cooperation and Development	95
3.3.3	EU Directive on Corporate Sustainability Due Diligence	97
3.3.4	EU Artificial Intelligence Act	99
3.4	Specific framework: preventive corporate responsibilities to counter online hate speech	101
3.4.1	EU Regulation on a Single Market for Digital Services	102
3.4.2	EU Audiovisual Media Services Directive	105
3.4.3	EU Code of conduct on countering illegal hate speech online	106
3.4.4	Council of Europe Committee of Ministers' Recommendation CM/Rec(2022)16	107
3.4.5	Proposal of a legal standard	108
3.5	Case studies: Compliance of 'Terms of Service' with the conceptualization of criminal hate speech	110
3.5.1	Facebook	111
3.5.2	Twitter (now X)	114
3.5.3	YouTube	116
3.6	Conclusion	121

4	HUMAN RIGHTS RESPONSIBILITIES OF ONLINE PLATFORMS TO MITIGATE CRIMINAL HATE SPEECH Disrupting incitement to violence in large groups on end-to-end encrypted services in Europe	123
4.1	Introduction	124
4.2	Criminal hate speech as cybercrime	128
4.2.1	Incitement to violence as criminal hate speech	128
4.2.2	Implications on end-to-end encrypted services	131
4.2.3	Key human rights safeguards in countering criminal hate speech in E2EE	134
4.2.3.1	Freedom of expression, assembly and association	134
4.2.3.2	Privacy and data protection	135
4.3	Corporate human rights due diligence (HRDD) to counter hate speech in E2EE	137
4.3.1	Internet intermediaries' responsibility to protect human rights	137
4.3.2	Corporate HRDD to counter criminal hate speech online	139
4.3.3	Corporate HRDD to counter illegal content in E2EE services	143
4.3.3.1	EU Regulation on Terrorist Content Online	143
4.3.3.2	EU Proposal Regulation on Child Sexual Abuse Material	144
4.4	Digital technologies: content moderation in E2EE	146
4.4.1	Metadata	147
4.4.2	Hashing	148
4.4.3	Homomorphic encryption	150
4.5	Standard proposal: expanding HRDD to counter incitement to violence in E2EE services	151
4.5.1	Substantive regulation: Incitement to violence	151
4.5.2	Procedural regulation	153
4.5.2.1	HRDD responsibilities of E2EE to counter incitement to violence	153
4.5.2.2	Technical implementation: disruption as the minimum legal standard	154
4.5.2.3	Legal implementation	158
4.5.3	Critical analysis: human rights safeguards	158
4.6	Conclusion	162
5	HUMAN RIGHTS RESPONSIBILITIES OF ONLINE PLATFORMS TO REMEDIATE CRIMINAL HATE SPEECH A call for a thorough corporate remedial responsibilities framework in Europe for criminal hate speech attributable to online platforms	165
5.1	Introduction	165
5.2	Criminal hate speech on online platforms	169
5.2.1	European standards on criminal hate speech	169
5.2.2	The role of online platforms	171
5.3	Right to remedy for criminal hate speech online	175
5.3.1	Harm caused by hate speech	175

5.3.2	State's duty to ensure access to remedy	177
5.3.2.1	European standards on remedies	177
5.3.2.2	Remedies for gross human rights violations	179
5.4	General framework: corporate remedial responsibilities for online platforms	180
5.4.1	Modes of corporate responsibility	181
5.4.2	Remedial processes	182
5.4.3	Remedial outcomes	183
5.5	European framework: online platforms remedial responsibilities for criminal hate speech	184
5.5.1	Challenges with current framework	184
5.5.1.1	Corporate remedial responsibilities in the EU	185
5.5.1.2	Remedial responsibilities in the Digital Services Act	186
5.5.1.3	Complementary corporate remedial frameworks for hate speech online	188
5.5.2	Proposed standards for a comprehensive framework	189
5.5.2.1	Restitution and satisfaction as amplification of survivors' speech	191
5.5.2.2	Compensation and rehabilitation beyond area of services	192
5.5.2.3	Guarantees of non-repetition as business models' change	193
5.6	Conclusion	194
6	CONCLUSION	197
6.1	Findings related to the problem statement and research questions	197
6.1.1	Legal conceptualization of hate speech	198
6.1.2	Human rights responsibilities of online platforms to prevent criminal hate speech	203
6.1.3	Human rights responsibilities of online platforms to mitigate criminal hate speech on end-to-end encrypted services	205
6.1.4	Human rights responsibilities of online platforms to remediate criminal hate speech	208
6.2	Recommendations	211
6.3	Areas for future research	214
	SUMMARY	219
	SAMENVATTING (DUTCH SUMMARY)	223
	BIBLIOGRAPHY	229
	LIST OF PUBLICATIONS AND TALKS	263
	CURRICULUM VITAE	267

List of figures and tables

Figure 1	Overview of regulatory and policy human rights frameworks impacting the regulation of online hate speech	19
Figure 2	OECD Due diligence process and supporting measures	95
Figure 3	OECD Due diligence process, including precautionary measures to counter criminal hate speech	109
Figure 4	Homomorphic approach to secure message analysis	157
Figure 5	Corporate remedial responsibilities for adverse human rights impacts	182
Table 1	Case studies' compliance with the conceptualisation of criminal hate speech	119
Table 2	Summary of proposed HRDD standard	161

Acronyms

ADI/MSI-DIS	Committee of Experts on combatting hate speech, Council of Europe
AI	Artificial Intelligence
AI Act	Artificial Intelligence Act
APCC	Additional Protocol to the Convention on Cybercrime
AVMSD	Audiovisual Media Services Directive
CFREU	Charter of Fundamental Rights of the EU
CM	Committee of Ministers, Council of Europe
Code or CoC	EU Code of Conduct on countering illegal hate speech online
CoE	Council of Europe
CRT	Critical Race Theory
CSAM	Child Sexual Abuse Material
CSAR	EU Regulation on Child Sexual Abuse Material
CSDDD	Directive on corporate sustainability due diligence
DSA	Digital Services Act
E2EE	End-to-end encryption
EC	European Commission
ECHR	European Convention for the Protection of Human Rights and Fundamental Freedoms
ECJ	European Court of Justice
ECRI	European Commission against Racism and Intolerance
ECtHR	European Court of Human Rights
EDRi	European Digital Rights
EP	European Parliament
EU	European Union
FHE	Fully Homomorphic Encryption
GDPR	General Data Protection Regulation
GIFs	Graphic Interchange Format
GPR	General Policy Recommendation
HRDD	Human Rights Due Diligence
ICC	International Criminal Court
ICCPR	International Covenant on Civil and Political Rights
ICERD	International Convention on the Elimination of All Forms of Racial Discrimination
IMCO	Committee on Internal Market and Consumer Protection
LGBTIQ+	Lesbian, Gay, Bisexual, Trans, Intersex and Queer+
NETHATE	Network of Excellence for Training on Hate
NetzDG	Network Enforcement Law
OECD	Organization for Economic Cooperation and Development

TCOR	EU Regulation on Terrorist Content Online
TEU	Treaty of the European Union
TFEU	Treaty on the Functioning of the European Union
ToS	Terms of Service
UDHR	Universal Declaration of Human Rights
UN	United Nations
UNGPs	United Nations Guiding Principles on Businesses and Human Rights
VLOPs	Very Large Online Platforms
VLOSEs	Very Large Online Search Engines