

PART IX

THE DIGITAL REVOLUTION

CHAPTER 5

ALTERNATIVE DISPUTE RESOLUTION AND ARTIFICIAL INTELLIGENCE¹

Björn Laukemann²

<https://www.cplj.org/publications/9-5-alternative-dispute-resolution-and-artificial-intelligence>

TABLE OF CONTENTS

1 AI and Platform Justice: Algorithmic Rights and Standards Enforcement through Online Platforms	1
1.1 Fields of Application of Artificial Intelligence.....	1
1.1.1 Copyright enforcement (in the Context of Content ID).....	1
1.1.1.1 The Content-ID Procedure.....	1
1.1.1.2 The Functioning of AI with Content ID Claims.....	4
1.1.1.2.1 Conditions for Use	4
1.1.1.2.2 Functionality of the Algorithm (Upload Filter at Google Content ID).....	5
1.1.1.2.2.1 Excursus: Types and Functioning of Upload Filters	5
1.1.1.2.2.2 Google Content ID	6
1.1.1.2.2.3 Procedure in Case of Collision	9
1.1.1.2.3 Other Applications of AI in the Context of the Content ID-Procedure.....	9
1.1.1.3 Effectiveness and Efficiency Potentials in the Use of AI.....	10
1.1.1.3.1 Detection of Rights Violations	10
1.1.1.3.2 Costs.....	10

¹ The document is up to date as at 31 December 2023 with regard to sources and the legal situation. However, where relevant, changes resulting from the European Union's AI Act have been incorporated.

² Prof Dr Björn Laukemann (Maître en droit Aix-en-Provence) holds the Chair of Civil Law, German and International Law of Civil Procedure at the Eberhard Karls University of Tübingen (Germany).

1.1.1.3.3 Flexibilization and Individualization of the Sanction Regime	11
1.1.2 Algorithmic Content Moderation on Communication Platforms	11
1.1.2.1 Terminology	11
1.1.2.2 Incentives for Establishing Ex Ante Mechanisms of AI-Based Content Moderation	11
1.1.2.3 The Functioning of Algorithmic Content Moderation	14
1.1.2.3.1 Basic Features	14
1.1.2.3.2 The Technical Operation, <i>Inter Alia</i> , Machine Learning Systems	15
1.1.2.4 Excursus: The Content Moderation of the Meta Group as a Prototype of Individualized Process Design	16
1.1.2.4.1 The Emergence of Specific Types of Platform Procedures Using the Example of the Cross-Check System	17
1.1.2.4.1.1 Procedural Framework	17
1.1.2.4.1.2 Specifics of the Procedure	18
1.1.2.4.1.2.1 Context-Specific In-Depth Examination	18
1.1.2.4.1.2.2 Suspension Effect	18
1.1.2.4.1.3 Special Modalities of the Cross-Check Procedure	19
1.1.2.4.1.3.1 The Person-Related Strand of the Cross-Check Procedure: The 'Early Response Secondary Review (ERSR)'	19
1.1.2.4.1.3.2 The Content-Related Strand of the Cross-Check Procedure: The General Secondary Review (GSR)	20
1.1.2.4.1.4 Allocation of Resources in Favour of the Personal ERSR Procedure	21
1.1.2.4.1.5 Availability of Appeal Mechanisms in the Cross-Check Procedure	21
1.1.2.4.2 Other Specific Types of Procedures	22
1.1.2.4.2.1 The High Impact False Positive Override (HIPO)	22

1.1.2.4.2.2 Procedures for Responding to Political Crises and Viral Content	23
1.1.2.4.2.3 Procedure for Handling Government Requests	24
1.1.2.4.3 Conclusion.....	25
1.2 Dangers and Drawbacks of AI-Based Law and Standards Enforcement	26
1.2.1 The Economization of Intra-Platform Decisions: Manifestations and Technical Design of Platform Power	26
1.2.1.1 Structuring of Interactions on Platforms.....	26
1.2.1.2 Tailored Outcomes and Procedures	27
1.2.1.2.1 Individualized Process Design in Payment Transactions	28
1.2.1.2.2 The Content Moderation of the Meta Group as a Prototype of Individualized Process Design	28
1.2.1.3 Private Ordering and Monetization in Copyright Law.....	29
1.2.1.4 AI-Driven Structuring of Communication Processes and Separation Of User Groups.....	29
1.2.2 Merging Government and Private Regulatory Objectives: The Public- Private Divide	31
1.2.2.1 Privatization of Law Enforcement and Dispute Resolution.....	31
1.2.2.2 (Cross-Border) Influence of State Actors on Platforms	33
1.2.3 Technical Limitations and Restricted Traceability of Context-Dependent Content.....	35
1.2.4 Limited Transparency.....	38
1.2.5 Overblocking	39
1.2.6 Algorithmic Discrimination	41
1.2.7 Information Gaps and Insufficient Stakes.....	42
1.3 Procedural Answers.....	42
1.3.1 The Need for (Complementary) Procedural Resolution Schemes.....	42
1.3.1.1 Initial Situation: Proceduralization of Private Enforcement.....	42
1.3.1.2 Emergence and Justification of Platform-Based Law Enforcement through Procedure.....	44
1.3.1.2.1 Liability Law and Basic Procedural Structures.....	44
1.3.1.2.2 Reasons for Platform-Based Law Enforcement through Procedures.....	46

1.3.1.2.2.1 Structural Inferiority of the Addressee of Legal Enforcement: Empowerment of Effective Legal Protection through Procedure	46
1.3.1.2.2.2 Fast and Effective Protection against Irreversible Damage Caused by Procedures	47
1.3.1.2.2.3 Minimizing Erroneous Decisions: Procedural Law as an Instrument of Fact-Finding	47
1.3.1.2.2.4 Conclusion.....	47
1.3.2 Shaping Effective Legal Protection through Procedures: Potential and Limits of a Proceduralization of Platform-Based Legal Enforcement	48
1.3.2.1 Functions of Procedural Structures	48
1.3.2.2 Individual and Collective Law Dimensions of Effective Legal Protection	48
1.3.2.3 Intra-Company Legal Protection Proceedings: Basic Structures and Procedural Guarantees	50
1.3.2.3.1 Online Platforms and Attributions of their Procedural Function	51
1.3.2.3.2 Procedural Obligations under the Notification and Redress Procedure	53
1.3.2.3.3 Transparency Obligations	54
1.3.2.3.4 Information and Justification Obligations with Regard to Individual Moderation or Enforcement Decisions	56
1.3.2.3.5 Access to the Procedure	58
1.3.2.3.5.1 Access to the Notification and Redress Procedure	58
1.3.2.3.5.2 Trusted Whistleblowers (Trusted Flagging).	58
1.3.2.3.5.3 Excursus: The Independence of Platforms Vis- À-Vis The State.....	60
1.3.2.3.5.4 Uniform Access to Internal Complaints Procedures.....	62
1.3.2.3.5.5 Uniform Access to the Nature of Decision Making?	62
1.3.2.3.6 Proceduralization of Decision Making.....	64
1.3.2.3.6.1 Requirements for the Decision Maker	64
1.3.2.3.6.2 Decision Parameters and Decision Consistency	64

1.3.2.3.6.3 ‘Scaling’ of Decisions.....	66
1.3.2.3.7 Excursus: AI Act and Platform Enforcement.....	68
1.3.2.4 Necessary Dovetailing of Internal Complaint Procedures with Out-Of-Court Independent Dispute Resolution.....	72
1.3.2.4.1 Structural Bias of Online Platforms	72
1.3.2.4.2 Out-Of-Court Dispute Resolution	74
1.3.2.4.3 The Establishment of ‘Platform Courts’	76
2 Artificial Intelligence in Mediation and Conciliation.....	78
2.1 Fields of Application of Artificial Intelligence.....	78
2.1.1 Fields of Application According to Process Stages.....	78
2.1.1.1 Procurement of (Legal) Information by Artificial Intelligence.....	78
2.1.1.2 Automated Document Analysis, Fact-Finding and Case Management.....	80
2.1.1.3 AI-Tools for Preparing Settlement or Decision Proposals (Predictive Tools)	80
2.1.2 Sectoral Fields of Application of AI in Mediation and Arbitration.....	81
2.2 Use and Dangers of Using Artificial Intelligence	82
2.2.1 Benefit.....	82
2.2.2 Dangers	83
2.2.2.1 The Black Box Problem and Discrimination.....	83
2.2.2.2 ‘Petrification’ of Decision-Making through Strict Reference to the Past	84
2.2.2.3 ‘Legal Remoteness’ Decision-Making Standards and the Accumulation of Private Decision-Making Power	85
2.3 Procedural Answers.....	86
3 Arbitration and Artificial Intelligence.....	88
3.1 Fields of Application of Artificial Intelligence in Arbitration Proceedings	88
3.1.1 Document Analysis and Case Management	89
3.1.1.1 Automated Document Analysis	89
3.1.1.2 Automated Document Creation and Advanced Analysis Tools....	90
3.1.1.3 (Other) AI-Supported Case Management.....	90
3.1.1.4 Legal Admissibility	91
3.1.2 Selection of the Referee Using AI	91
3.1.3 Technologies for Risk Assessment and Prediction of Process Outcomes (Predictive Analytics).....	92

3.1.4 Fully Automated Process Management and Decision-Making ('Robo Arbitrator')	93
3.1.4.1 Computability of Law?	94
3.1.4.2 The problem of obtaining training data	95
3.1.4.3 Legal assessment of autonomous decision-making by AI	95
3.1.4.3.1 International arbitration law: UNCITRAL Model Law and New York Convention	96
3.1.4.3.2 National arbitration law	96
3.1.4.3.3 Formal and data protection hurdles	97
3.1.4.3.4 Outlook: Blockchain arbitration	98
3.2 Risks Associated with the Use of Artificial Intelligence in Arbitration Proceedings	98
3.2.1 Unconscious Bias and Risk of Discrimination	98
3.2.2 The Black Box Problem	99
3.2.3 Danger of 'Petrification'	100
3.2.4 Failure to Grant the Right to Be Heard	100
3.3 Procedural Answers	100
4 Summary of the Main Results	102
4.1 Platform-Based Enforcement of Rights and Standards	102
4.1.1 Economization of Platform Procedures	102
4.1.2 Merging of State and Private Enforcement Interests	103
4.1.3 Potential Effects of the Use of Artificial Intelligence	103
4.1.4 Proceduralization of Private Rights and Standards Enforcement	103
4.1.5 The Regulatory Dimension: Interaction Between (Individual) Rule Enforcement and (Collective) Setting of Rules	105
4.2 Use of Artificial Intelligence in Mediation, Dispute Resolution and Arbitration Proceedings	106
Abbreviations and Acronyms	107
Legislation	109
International/Supranational	109
National	110
Cases	112
International/Supranational	112
National	113
Bibliography	114

1 AI AND PLATFORM JUSTICE: ALGORITHMIC RIGHTS AND STANDARDS ENFORCEMENT THROUGH ONLINE PLATFORMS

1.1 Fields of Application of Artificial Intelligence

1.1.1 Copyright enforcement (in the Context of Content ID)

- 1 Playing a (technological) pioneering role in the field of algorithmic enforcement, Content ID is now, as a result of its success, serving as a blueprint for no less than resolving the problem of defamation, terrorist content, and other forms of harmful speech.³

1.1.1.1 The Content-ID Procedure

- 2 In order to enforce copyright infringements, the online platform YouTube⁴ uses the so-called 'Content ID' system.⁵ Content ID, forming part of YouTube's automated rights management system, is an upload filter that detects incriminated content when uploaded with the help of reference data stored by the respective rightsholder (filtering technology).⁶
- 3 In case of a detected infringement, the rightsholder has three options for action, apart from filing a deactivation request for copyright infringement:⁷ (i) blocking the content,⁸ (ii) monetizing the content by means of sharing in the advertising revenue,⁹ as well as

³ H Bloch-Wehba, 'Automation in Moderation' (2020) 53(1) Cornell International Law Journal 41, 74 f.

⁴ On the particular market dominance of the platform and its relevance to public discourse see M Perel and N Elkin-Koren, 'Accountability in Algorithmic Copyright Enforcement' (2016) 19(3) Stan. Tech. L. Rev. 473, 497.

⁵ This is why Google, YouTube's parent company, markets the Content ID System as not merely an upload filter, but even a 'copyright management system', see Google, 'How Google Fights Piracy' (2018) 24 <https://kstatic.googleusercontent.com/files/2bc15c350e6d8ba6363594195712a3c2528e56502c41c8a8a431746afce40adb9956ff837f9e54887c0277b413bceb8d79adc02ddae97c24969b55a30c70d836> accessed 31 December 2023; J Schillmöller and S Doseva, "'Chilling effects" durch YouTube's Content ID?' (2022) 25(3) MMR (Multimedia und Recht) 181, 182; J E Gray and N P Suzor, 'Playing with machines: using machine learning to understand automated copyright enforcement at scale' (2020) 7(1) Big Data & Society 1, 2 <https://doi.org/10.1177/2053951720919963> accessed 31 December 2023. – For general information on the Content ID procedure, see Perel and Elkin-Koren (n 4), (2016) 19(3) Stan. Tech. L. Rev. 473, 497; H Grosse Ruse-Kahn, 'Automated Copyright Enforcement Online: From Blocking to Monetization of User-Generated Content' (2020) PIJIP Research Paper Series 51 <https://digitalcommons.wcl.american.edu/research/51> accessed 31 December 2023.

⁶ For an overview of the technical functioning see Perel and Elkin-Koren (n 4), (2020) 19(3) Stan. Tech. L. Rev. 473 at fn 210-211 with further references; Gray and Suzor (n 5), (2020) 7(1) Big Data & Society 1, 2.

⁷ YouTube, 'What are policies?' https://support.google.com/youtube/answer/107383?hl=en&ref_topic=24332, https://support.google.com/youtube/answer/107383?hl=de&ref_topic=24332 accessed 31 December 2023.

⁸ Blocking can thus take place on the basis of so-called content ID claims as well as on the basis of an infringement of state copyright law.

⁹ Schillmöller and Doseva (n 5), (2022) 25(3) MMR (Multimedia und Recht) 181 f.

(iii) observing the audience figures for the purpose of a decision (so-called monitoring).¹⁰ From an economic point of view, monetizing the incriminated content is the biggest advantage of the Content ID system;¹¹ accordingly, this is the predominant choice helping to redress the so-called value gap.¹²

- 4 It is up to the respective user to initiate the subsequent complaint procedure: First, he has to dispute the content ID claim of the rightsholder.¹³ The rightsholder himself then decides whether he wishes to remedy the dispute: If he refrains from doing so, the user is given the ability to appeal.¹⁴ The specific procedural remedies of the uploader differ according to whether the rightsholder chooses to block or to monetize the incriminated

¹⁰ T Hess and H Waltermann, 'Upload-Filter für Content' (2019) 16(2) *MedienWirtschaft* 16, 19 f https://web.archive.org/web/20220225020059id_/https://www.beck-elibrary.de/10.15358/1613-0669-2019-2-16.pdf accessed 31 December 2023; L Solomon, 'Fair users or content abusers' (2015) 44(1) *Hofstra L. Rev.* 237, 256. Here, the rightsholder has the possibility to view data on the use of the video. There is no further information about the procedure. Google denies a legitimate reason for objection if the video is not monetized (YouTube, 'Dispute a Content ID claim' https://support.google.com/youtube/answer/2797454?hl=en&ref_topic=9282678#zippy=%2Coptionen-f%C3%BCr-den-anspruchsteller accessed 31 December 2023). However, the link to this notice makes it clear that the monetization of the uploaded video is at issue here, but not a possible restriction of the right to object in the case of monitoring: YouTube, 'What is copyright?' <https://support.google.com/youtube/answer/2797466#refrained&zippy=%2Cmissverst%C3%A4ndnis-nr-wenn-du-angibst-dass-deine-inhalte-nicht-kommerziellen-zwecken-dienen-kannst-du-jegliche-inhalte-verwenden> accessed 31 December 2023.

¹¹ This is specifically done by concluding a copyright licensing agreement, see Perel and Elkin-Koren (n 4), (2016) 19(3) *Stan. Tech. L. Rev.* 473, 512; Y Lev-Aretz, 'Second Level Agreements' (2012) 45(1) *AKRON L. REV.* 137, 152.

¹² The ability to monitor and monetize infringement redresses the 'value gap' between what YouTube pays for monetized content and what services such as Spotify or Pandora, which license content directly from rightsholders, pay, see Bloch-Wehba (n 3), (2020) 53(1) *Cornell International Law Journal* 41, 64. – Over the years, it has become clear that monetization is the preferred method: in the second half of 2021, monetization was selected in 90% of cases for a total volume of 759,540,199 Content ID claims. In less than one percent of the cases in which a Content ID claim was made, an objection (dispute) was raised at all, see YouTube, 'Copyright Transparency Report H2 2021' 3, 10 f https://storage.googleapis.com/transparencyreport/report-downloads/pdf-report-22_2021-7-1_2021-12-31_en_v1.pdf accessed 31 December 2023. In 2017 for example, monetization was chosen in 90% of all cases – in the music industry even in 95%. This has led to payments on the part of YouTube amounting to around 3 billion dollars, cf: Grosse Ruse-Kahn (n 5), (2020) *PIIIP Research Paper Series* 51, 1, 4.

¹³ Schillmöller and Doseva (n 5), (2022) 25(3) *MMR (Multimedia und Recht)* 181, 183 fn 25; YouTube, 'Monetization during Content ID disputes' https://support.google.com/youtube/answer/7000961?hl=en&ref_topic=9282678 accessed 31 December 2023. If the uploader remains inactive for five days, the revenue is paid to the rightsholder.

¹⁴ YouTube (n 13), 'Monetization during Content ID disputes'; YouTube, 'Appeal a Content ID claim' <https://support.google.com/youtube/answer/12104471> accessed 31 December 2023.

content. In the context of both complaint procedures, however, YouTube does not take a decision on the content;¹⁵ this is incumbent on the rightsholder alone.¹⁶

- 5 At the same time,¹⁷ the rightsholder is entitled to submit an official deactivation request based on copyright infringement (so-called takedown request),¹⁸ ie, outside the system of Content ID claims.¹⁹ The further procedure, in particular the user's right to file a counter notification,²⁰ is governed by US law, thus by the Digital Millennium Copyright Act (DMCA).²¹
- 6 The Content ID System is part of a larger ecosystem of private copyright enforcement. According to the 'YouTube Copyright Management Suite', the dispute system encompasses – apart from Content ID – a 'Copyright Match Tool' as well as a 'webform option'.²² These tools differ significantly in terms of accessibility and level of technical abilities: While only large rightsholders (eg, movie studios) are eligible for Content ID as the most sophisticated enforcement tool,²³ the 'Copyright Match Tool' is intended for the use of minor rightsholders²⁴ and users submitting a 'valid copyright removal

¹⁵ This is subject to strong criticism with regard to the fair design of the procedure (see below para 95-98). Instructive on the problem: Schillmöller and Doseva (n 5), (2022) 25(3) MMR (Multimedia und Recht) 181 f.

¹⁶ Accordingly, YouTube states: 'The initial objection and complaint will be reviewed by the claimant, as YouTube cannot make ownership decisions We do not know what content is properly licensed and therefore cannot determine when copyright exceptions such as fair use or fair dealing apply', YouTube, 'Dispute a Content ID claim' <https://support.google.com/youtube/answer/2797454> accessed 31 December 2023.

¹⁷ 'The whole process is dictated by the Digital Millennium Copyright Act. [...] YouTube also has Content ID, an automated copyright management system. It exists in parallel to the copyright takedown process and allows copyright owners to manage their content at scale on YouTube', YouTube, 'Frequently asked questions about copyright' <https://support.google.com/youtube/answer/2797449?hl=en> accessed 31 December 2023. The Notice and Takedown procedure must always be provided by YouTube due to other liability, K L Zawada 'The Emergence and Development of Content ID in Light of User-generated Law' (2017) How Deep is your Law?, 5th International Conference of PhD Students and Young Researchers Conference Papers 438.

¹⁸ Such a deactivation request leads to a copyright warning, Youtube 'copyright strike basics' <https://support.google.com/youtube/answer/2814000#zippy=%2Cfolgen-einer-urheberrechtsverwarnung> accessed 31 December 2023.

¹⁹ Grosse Ruse-Kahn (n 5), (2020) PIJIP Research Paper Series 51, 1, 4 at fn 7, according to which Content ID claims have outnumbered copyright takedowns by a ratio of 50 to 1 since 2014. In 2017, over 98% of copyright infringements were claimed via Content ID instead of notice-and-takedown: Google (n 5), 'How Google Fights Piracy' 23 f.

²⁰ See YouTube, 'Requirements for counter notifications' https://support.google.com/youtube/answer/6005919?hl=en&ref_topic=9282678 accessed 31 December 2023.

²¹ Schillmöller and Doseva (n 5), (2022) 25(3) MMR (Multimedia und Recht) 181, 183.

²² YouTube (n 12), 'Copyright Transparency Report H2 2021', 1.

²³ Regarding Content ID, YouTube describes the amount of automation as 'high'. According to the platform, 98% of all copyright actions on YouTube are handled through content ID, see YouTube (n 12), 'Copyright Transparency Report H2 2021', 1.

²⁴ This level, for instance, addresses creators in the YouTube's partner program and 'any channel that's filled out the copyright management tools application and shown a need for an advanced rights management tool': YouTube (n 12), 'Copyright Transparency Report H2 2021', 2.

request’²⁵ through the webform.²⁶ By contrast, ‘every [...] YouTube user has access to the webform’ whose standard of automation is broadly described as being ‘low’.²⁷

1.1.1.2 The Functioning of AI with Content ID Claims

1.1.1.2.1 Conditions for Use

- 7 The use of Content ID²⁸ requires the rightsholder to provide certain reference files (audio, visual or audiovisual)²⁹ and metadata to Google's database.³⁰ Segments from other rightsholders are to be excluded, as is copyright-free content.³¹ In addition, the platform operator provides ‘best practices’ notes.³² Moreover, the use requires a partnership agreement with Google.³³ There are also formal requirements for the content itself: For example, exclusive rights must exist for ‘a significant number of videos’ that are ‘frequently uploaded by the YouTube user community’.³⁴ In that

²⁵ This is the case since October 2021. Previously, the ‘Copyright match Tool’ was only available to the users of YouTube’s partner program, see YouTube (n 12), ‘Copyright Transparency Report H2 2021’, 5.

²⁶ As far as the ‘Copyright Match Tool’ is concerned, YouTube describes the level of automation as ‘medium’: YouTube (n 12), ‘Copyright Transparency Report H2 2021’, 1. However, the meaning of this classification as well as the precise difference between the technical capabilities of the Copyright Match Tool, on the one hand, and Content ID, on the other, remains opaque.

²⁷ YouTube (n 12), ‘Copyright Transparency Report H2 2021’, 5.

²⁸ As of July 2015, more than 8,000 ‘partners’ were using the Content ID tool, Zawada (n 17), (2017) *How Deep is your Law?*, 5th International Conference of PhD Students and Young Researchers Conference Papers 438, 445; Google (n 5), ‘How Google Fights Piracy’ 18. As of November 2018, more than 9.000 ‘partners’ were using Content ID, Google (n 5), ‘How Google Fights Piracy’ 13.

²⁹ As of November 2018, there were already more than 80 million reference files on Google's servers: Google (n 5), ‘How Google Fights Piracy’ 25.

³⁰ This is information for managing copyrights. The information consists of the reference file, metadata, ownership notices, and established policies: YouTube, ‘Create an asset’ https://support.google.com/youtube/answer/3011552?hl=en&ref_topic=3011550 accessed 31 December 2023. In addition, the information has to indicate the (local) scope of the exclusive rights: YouTube, ‘Using Content ID’ <https://support.google.com/youtube/answer/3244015?hl=en> accessed 31 December 2023; cf Schillmöller and Doseva (n 5), (2022) 25(3) MMR (Multimedia und Recht) 181, 182.

³¹ To increase the number of hits, YouTube recommends ‘full length’ references: YouTube, ‘What is a reference?’ <https://support.google.com/youtube/answer/107004?hl=en> accessed 31 December 2023; YouTube, ‘Best practices for references’ <https://support.google.com/youtube/answer/107008> accessed 31 December 2023. – For more information on using a CSV template: see YouTube, ‘Deliver content using spreadsheet templates’ <https://support.google.com/youtube/answer/6066171> accessed 31 December 2023; for the DDEX feed, see YouTube, ‘Using the YouTube DDEX feed’ <https://support.google.com/youtube/topic/3505247> accessed 31 December 2023.

³² YouTube, ‘Best practices for references’ <https://support.google.com/youtube/answer/107008> accessed 31 December 2023.

³³ For this, a user must have at least 1,000 subscribers, the channel (among other requirements) must have a playback time of more than 4,000 hours in the last 12 months: YouTube, ‘YouTube Partner Program overview & eligibility’ <https://support.google.com/youtube/answer/72851> accessed 31 December 2023; YouTube, ‘Qualify for Content ID’ <https://support.google.com/youtube/answer/1311402> accessed 31 December 2023. Otherwise, content creators must rely on other programs such as the ‘Copyright Match Tool’, the ‘Content Verification Tool’, or even the ‘Copyright Complaint Web Form’. – Cf also Grosse Ruse-Kahn (n 5), (2020) PIJIP Research Paper Series 51, 1, 8 with further references.

³⁴ YouTube (n 30), ‘Using Content ID’; referring to YouTube (n 33), ‘Qualify for Content ID’.

context, Google restricts the existence of individual rights for certain content such as collages, trailers, best-ofs or gameplay videos.³⁵ In addition, the platform operator does not check whether these rights actually exist, but rather refers to the responsibility of the respective user.³⁶

1.1.1.2.2 Functionality of the Algorithm (Upload Filter at Google Content ID)³⁷

- 8 Generally, upload filters match content by using a database built on predefined rules, checking whether the new content is a copy or whether the content violates other rules or legal provisions.³⁸

1.1.1.2.2.1 Excursus: Types and Functioning of Upload Filters

- 9 (aa) *Hash-based upload filters.* In hash-based upload filters (hashing algorithms), the algorithm converts content into so-called hashes. Hashes are numerical representations of a file that are significantly smaller than the original file.³⁹ Hashes are also created for uploads. Subsequently, the hashes of the uploads are matched with the hashes of the database. If a match is found, the upload is blocked.⁴⁰
- 10 The goal of hash algorithms is to generate unique 'keys'. However, this is not always successful: it is possible that the same or similar hashes are created for two different contents ('collisions' or 'clashes'). With all hash algorithms, there is a theoretical probability of this happening. To reduce this probability, hashes are created using 'robust features'.⁴¹
- 11 Even though only content that has already been defined as undesirable in the database can be discerned, machine learning (ML) algorithms promise a remedy: Content that is potentially undesirable compared to existing rules is blocked as a precaution.

³⁵ YouTube (n 33), 'Qualify for Content ID'.

³⁶ YouTube, 'Content eligible for Content ID' <https://support.google.com/youtube/answer/2605065?hl=en> accessed 31 December 2023. However, YouTube notes that appeals are monitored 'continuously', see YouTube (n 30), 'Using Content ID'.

³⁷ Details about the Content ID matching process have been kept secret by Google so far; more precise statements about how the algorithm works are therefore difficult to make, Perel and Elkin-Koren (n 4), (2016) 19(3) Stan. Tech. L. Rev. 473, 514.

³⁸ Hess and Waltermann (n 10), (2019) 16(2) MedienWirtschaft 16, 18.

³⁹ E Engstrom and N Feamster, 'The Limits of Filtering: A Look at the Functionality & Shortcomings of Content Detection Tools' (2017) Engine, 12 <https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/58d058712994ca536bbfa47a/1490049138881/FilteringPaperWebsite.pdf> accessed 31 December 2023.

⁴⁰ Hess and Waltermann (n 10), (2019) 16(2) MedienWirtschaft 16, 18 f.

⁴¹ T Lester and D Pachamano, 'The Dilemma of False Positives: Making Content ID Algorithms More Conducive to Fostering Innovative Fair Use in Music Creation' (2017) 24(1) UCLA Entertainment Law Review 51, 62 f <https://escholarship.org/content/qt1x38s0hj/qt1x38s0hj.pdf?t=ovwl6c> accessed 31 December 2023.

Nonetheless, this also blocks permissible content, as it is difficult to make a clear distinction – for example, between insults/defamation and satire.⁴²

- 12 *(bb) Search Algorithms.* Unlike hashing algorithms, which convert the original content into hashes, search algorithms decompose content: for instance, a piece of music into a sequence of audio events. This creates ‘music phonemes’ (elementary units of music) and a sequence of these ‘phonemes’ that best represent the piece. The total features (‘dimension’) of a piece are thus reduced to a smaller set of representative features (‘alphabet’). These representative features, when put together in the right combination, can generate (‘transcribe’) the piece. As new pieces are added, the set of ‘phonemes’ is revised so that the pieces are better represented in the database. Thus, the search algorithms ‘learn’ continuously. The algorithm calculates the probability that the pieces in the database represented by the music units are the ‘producers’ of the new uploads. If the probability is higher than a certain predefined threshold, the upload is considered a copy. The threshold can be changed to lower the percentage of false positives, with a higher threshold resulting in fewer false positives.⁴³

1.1.1.2.2.2 Google Content ID

- 13 *(aa) Digital fingerprinting.* Google Content ID is a so-called digital fingerprinting system.⁴⁴ Fingerprints (comparable to hashes) which are created using an algorithm are stored in a database for content.⁴⁵ For this purpose, the algorithm examines characteristics of the file (for example, notes and their volume in a piece of music).⁴⁶ The

⁴² Hess and Waltermann (n 10), (2019) 16(2) *MedienWirtschaft* 16, 19 f. – Examples of hashing algorithms are: (i) terrorist content screening databases. One of them was developed by the EU Internet Forum (including Microsoft, Google, Facebook and Twitter) together with Europol. The program identifies terrorist propaganda content and reports it. The content is then reviewed by staff. This is a staged process based on the ‘human-in-the-loop principle’, see M Monroy, ‘EU-Kommission droht mit „gesetzgeberischen Maßnahmen“ zur Entfernung von Internetinhalten’ (2018) *Netzpolitik.Org* <https://netzpolitik.org/2018/eu-kommission-droht-mit-gesetzgeberischen-massnahmen-zur-entfernung-von-internetinhalten/#netzpolitik-pw>. However, the filter can only identify uploads that have already been uploaded once and subsequently deleted, M Monroy, ‘“EU-Internetforum”: Viele Inhalte zu “Extremismus” werden mit Künstlicher Intelligenz aufgespürt’ (2017) *Netzpolitik.Org* <https://netzpolitik.org/2017/eu-internetforum-viele-inhalte-zu-extremismus-werden-mit-kuenstlicher-intelligenz-aufgespuert/> accessed 31 December 2023 – (ii) PhotoDNA (from Microsoft): used by Google, Facebook, and Twitter, among others, to identify child pornography content.

⁴³ Lester and Pachamano (n 41), (2017) 24(1) *UCLA Entertainment Law Review* 51, 63 f.

⁴⁴ Solomon (n 10), (2015) 44(1) *Hofstra L. Rev.* 237, 238 (2015). – On the distinction of fingerprinting from so-called watermarking, see D Milano, *Content control: Digital Watermarking and Fingerprinting* (Rhozet 2013) 2 ff; Jacques S, Garstka K, Hviid M and Street J, ‘An empirical study of the use of automated anti-piracy systems and their consequences for cultural diversity’ (2018) 15(2) *Script-Ed* 277, 287 at fn 28.

⁴⁵ Hess and Waltermann (n 10), (2019) 16(2) *MedienWirtschaft* 16, 19; Solomon (n 10), 44(1) *Hofstra L. Rev.* 237, 256 (2015).

⁴⁶ E Engstrom and N Feamster (n 39) 13.

fingerprints of new uploads are compared with the content stored in the database.⁴⁷ If a threshold of similarities is exceeded, the upload is considered to infringe copyright.⁴⁸

- 14 An algorithm that works with fingerprints can also identify files that are only partial matches or have been slightly modified or manipulated; an algorithm based on hashes, on the other hand, cannot.⁴⁹ However, fingerprinting technologies cannot be used for all file formats. Reason being that the algorithm of eg an audio fingerprinting tool examines, among other things, the frequency values in a music file and thus cannot be used to identify copyrighted photos. If a technology is to examine all files on a web page, there would need to be a different fingerprinting tool for each type of media. Given that the range of copyrightable content is very broad, there is no fingerprinting tool for many types of content (eg, for architectural designs).⁵⁰
- 15 Content ID can be used for classic uploads as well as for live streams and thus in real time.⁵¹
- 16 *(bb) Possible results and accuracy of the algorithm.* The following results are conceivable:⁵²
 - i. The content is identified as a copy, the identification is correct ('true positive').
 - ii. The content is identified as a copy, the identification is false ('false positive').
 - iii. The content is not identified as a copy, the identification is correct ('true negative').
 - iv. The content is not identified as a copy, the identification is false ('false negative').

⁴⁷ Cf Google (n 5), 'How Google Fights Piracy' 25; H Grosse Ruse-Kahn, 'Global Content Protection through Automation' (2018) 49(9) IIC (International Review of Intellectual Property and Competition Law) 1017, 1018 f; G Nolte, 'Three Theses on the Current Debate on Liability and Distributive Justice in Hosting Services with User-Generated Content (the so-called "Value Gap" Debate)' (2017) 61(4) ZUM (Zeitschrift für Urheber- und Medienrecht) 304, 309; Schillmöller and Doseva (n 5), (2022) 25(3) MMR (Multimedia und Recht) 181, 182.

⁴⁸ Hess and Waltermann (n 10), (2019) 16(2) MedienWirtschaft 16, 19; Solomon (n 10), (2015) 44(1) Hofstra L. Rev. 237, 256.

⁴⁹ Engstrom and Feamster (n 39) 14.

⁵⁰ Ibid.

⁵¹ The scope of Content ID technology is nevertheless limited in the context of live streaming: For example, the content must be 'time-sensitive live content' (such as a sporting event) that guarantees a 'high probability that users will live stream copies of your content'. Also, the rights ownership must be global and exclusive. In the case of live streaming, the sanctioning is naturally different from the so-called standard matching of classic uploads: A warning message is first displayed to the streamer. The live stream is then replaced by a standard image without sound and finally interrupted, see YouTube, 'Use Content ID matching on live streams' <https://support.google.com/youtube/answer/9896248?hl=en> accessed 31 December 2023.

⁵² Cf Lester and Pachamano (n 41), (2017) 24(1) UCLA Entertainment Law Review 51, 65.

- 17 The accuracy of an algorithm is calculated with the sum of ‘true positives’ and ‘true negatives’ divided by the total number of evaluated contents.⁵³ Thus, an accuracy of 99% means that 99% of all content is correctly identified. However, the value says nothing about how well the algorithm performs in terms of ‘false positives’ and ‘false negatives’, ie whether the remaining 1% of cases are more ‘false negatives’ (disadvantage for the rightsholder) or more ‘false positives’ (disadvantage for the uploader). The ability of an algorithm to discern false positives and false negatives can be found out by using other metrics.⁵⁴
- 18 According to Google, the algorithm works with 99,7% accuracy in the music section.⁵⁵ Uploaded content can be blocked or monetized by Content ID within minutes up to hours of publication.⁵⁶ Efforts to circumvent the filter, such as by rotating the image, slowing or speeding up the audio track, or changing speed, pitch, or sound quality, are also detected.⁵⁷ No automated system, however, can evaluate content contextually, for example, to determine whether the ‘fair use’ principle is relevant.⁵⁸

⁵³ Lester and Pachamano (n 41), (2017) 24(1) UCLA Entertainment Law Review 51, 65.

⁵⁴ ‘The True Positive Rate’ (TPR) (‘sensitivity’) describes the ratio of ‘true positives’ to ‘true positives’ plus ‘false negatives’ and thus the probability that the algorithm will find infringing content. Accordingly, the ‘sensitivity’ expresses the percentage of all infringing cases that the algorithm identifies as infringing.

‘The True Negative Rate’ (TNR) (‘specificity’) indicates the ratio of ‘true negatives’ to ‘true negatives’ plus ‘false positives’ and, therefore, the probability that the algorithm identifies non-infringing content as non-infringing. The ‘specificity’ stands for the percentage of all non-infringing cases that the algorithm identifies as non-infringing.

‘The Positive Predictive Value’ (PPV) (‘precision’) describes the ratio of ‘true positives’ to ‘true positives’ plus ‘false positives’ and, therefore, the probability that content classified as infringing is actually infringing. ‘Precision’ indicates the percentage of all content identified as infringing that is actually infringing.

‘The Negative Predictive Value’ (NPV) expresses the ratio of true negatives to true negatives plus false negatives, and, consequently, the probability that content classified as non-infringing is actually non-infringing. This paraphrases a percentage of all content identified as non-infringing that is actually non-infringing, Lester and Pachamano (n 41), (2017) 24(1) UCLA Entertainment Law Review 51, 65 f.

⁵⁵ This applies to Content ID claims, see C Muller, ‘Setting the Record Straight’ <https://blog.youtube/news-and-events/setting-record-straight/> accessed 31 December 2023; referencing Nolte (n 47), (2017) 61(4) ZUM (Zeitschrift für Urheber- und Medienrecht) 304, 309.

⁵⁶ Cf the study by Gray and Suzor (n 5), (2020) 7(1) Big Data & Society 1, 4.

⁵⁷ Google (n 5), ‘How Google Fights Piracy’ 27. In addition, melodies and compositions can now be recognized, J P Titlow, ‘Youtube is using AI to police copyright to the tune of \$2 billion in payouts’ <https://www.fastcompany.com/4013603/youtube-is-using-ai-to-police-copyright-to-the-tune-of-2-billion-in-payouts> accessed 31 December 2023; Nolte (n 47), (2017) 61(4) ZUM (Zeitschrift für Urheber- und Medienrecht) 304, 309.

⁵⁸ R Andrea, ‘No Safe Harbor: YouTube’s Content ID and Fair Use’ (2020) Boston College Intellectual Property & Technology Forum 1, 5.

1.1.1.2.2.3 Procedure in Case of Collision

- 19 In YouTube's rights management system, the so-called assets can be administered.⁵⁹ In individual cases, reference files may overlap or several rightsholders may claim content ID rights to a video.
- 20 In the former case, in the event of a reference file collision, the rightsholder who uploaded the 'latest reference file' is notified of the collision.⁶⁰ Within 30 days, the latter can either 'claim ownership rights' or 'exclude a reference overlap'.⁶¹ If both rightsholders insist on exclusive rights to their reference file, the claims remain with the partner who provided the reference file first.⁶²
- 21 If multiple Content ID claims from competing partners relate to a video (eg, one claim relates to the audio track, the other to the image), the most 'restrictive' legal consequence (blocking instead of mere monetization) is applied.⁶³ By contrast, the legal consequence of observation intervenes when insufficient information is available for certain assets.⁶⁴ If two different rightsholders have content ID claims to the same asset, the policy of the rightsholder who owns the rights in the country of upload is applied; if both partners are supposed to own an asset in the same country, the strongest legal consequence is applied again.⁶⁵

1.1.1.2.3 Other Applications of AI in the Context of the Content ID-Procedure

- 22 In the context of the Content ID process, artificial intelligence is also used to search for incorrect reference material.⁶⁶ This is accompanied by human control, also with regard

⁵⁹ Assets are composed of the reference file, metadata, ownership information, and established policies, Google, 'Asset erstellen' https://support.google.com/youtube/answer/3011552?hl=de&ref_topic=3011550 accessed 31 December 2023.

⁶⁰ YouTube, 'Fix reference overlaps' https://support.google.com/youtube/answer/3022604?hl=en&ref_topic=3013248 accessed 31 December 2023. This is the case 'when two reference files have segments that collect audio, video, or audiovisual content'.

⁶¹ Ibid.

⁶² Ibid. – For the rest, YouTube refers to a mutually agreeable solution ('If an asset has two or more rightsholders, you must resolve the conflict together with other holders'): YouTube (n 60), 'Fix reference overlaps'.

⁶³ YouTube has determined that the least restrictive is when none of the policies (monetize, watch, block, or disable) are in place, followed by monetize. More restrictive is watching, then blocking, and finally disabling: YouTube, 'How policies are applied' <https://support.google.com/youtube/answer/3369929> accessed 31 December 2023.

⁶⁴ This is the case, for example, when no rights information has been provided in a particular country, YouTube, 'How policies are applied' <https://support.google.com/youtube/answer/3369929> accessed 31 December 2023.

⁶⁵ In the latter case, if both rightsholders have chosen the same legal consequence, revenue is withheld for music assets until an appeal has been decided; if it is not a music asset, the legal consequence of monitoring applies to both rightsholders, albeit only after the conflict has been resolved: YouTube (n 64), 'How policies are applied'.

⁶⁶ Google (n 5), 'How Google Fights Piracy' 28.

to the correct application of the Content ID system.⁶⁷ Google does not disclose exactly how this control is carried out.

1.1.1.3 Effectiveness and Efficiency Potentials in the Use of AI

1.1.1.3.1 Detection of Rights Violations

- 23 Effectiveness and efficiency gains are the most significant benefits of using algorithms to detect copyright infringement. For example, according to a study by *Gray/Suzor*, content is blocked within ‘minutes to hours’ after it is first published.⁶⁸ Without automatic matching against YouTube’s database, rightsholders would not be able to detect anywhere near as many copyright infringements in a correspondingly short period of time. With ongoing investments in the further development of the Content ID system,⁶⁹ the platform operator promises to make the algorithmic law enforcement system technically more reliable and more accurate on an ongoing basis,⁷⁰ but also to improve the complaint procedure (in the interest of users).⁷¹

1.1.1.3.2 Costs

- 24 From the perspective of rightsholders, the Content ID procedure is a cost-effective way of (provisionally) enforcing their own rights, especially in comparison to state court proceedings.⁷² However, shifting the burden of complaint⁷³ and cost to users makes them refrain from filing appeals against enforcement measures of the platforms before state courts – out of fear of high legal costs.⁷⁴

⁶⁷ YouTube, ‘Update: Improving Content ID for creators’ <https://blog.youtube/news-and-events/update-improving-content-id-for-creators/> accessed 31 December 2023.

⁶⁸ Gray and Suzor (n 5), (2020) 7(1) *Big Data & Society* 1, 4. Post-publication means that the video is ‘uploaded’ to the YouTube site but not retrievable. In contrast, blocking can also occur before ‘upload’ (publication). According to Gray and Suzor, however, there is no data on this from YouTube.

⁶⁹ By its own account, YouTube/Google invested more than \$100 million in 2018 in the development of Google (n 5), ‘How Google Fights Piracy’ 27. In addition, there are maintenance costs: D L Burk, ‘Algorithmic Fair Use’ (2019) 86(2) *University of Chicago Law Review* 283, 289. Conversely, Google retains 45% of advertising revenue, see Zawada (n 17), (2017) *How Deep is your Law?*, 5th International Conference of PhD Students and Young Researchers Conference Papers 438, 442.

⁷⁰ See already above para 22 fn 67.

⁷¹ For example, with respect to reducing the response time to a complaint in the Content ID process, or with respect to the possibility of a ‘direct complaint’: YouTube, ‘Accelerated Content ID & Complaint Process’ <https://support.google.com/youtube/thread/171619847> accessed 31 December 2023.

⁷² This aspect applies whether or not the accusation of uploading copyrighted material is true: Burk (n 69), (2019) 86(2) *University of Chicago Law Review* 283, 289.

⁷³ See B Laukemann, ‘Private Rechtsdurchsetzung zwischen (digitaler Selbsthilfe) und gerichtlichem Rechtsschutz’ (2022) 8(3) *ZfPW (Zeitschrift für die gesamte Privatrechtswissenschaft)* 357, 380.

⁷⁴ Cf Burk (n 69), (2019) 86(2) *University of Chicago Law Review* 283, 289. Burk also believes that, in principle, costs do not disappear, but are only ever redistributed; *ibid* 293.

1.1.1.3.3 Flexibilization and Individualization of the Sanction Regime

- 25 The Content ID system is also characterized by a high degree of individualization. It enables rightsholders to use ‘guidelines’ to decide in advance exactly which legal consequence a Content ID claim should be directed at.⁷⁵ Rightsholders can, for example, specify that Content ID claims addressing users from the USA are to be monetized, while videos are to be blocked if the uploader is located in Germany.⁷⁶ It is also possible to allow short uploads by fans for promotional purposes.⁷⁷

1.1.2 Algorithmic Content Moderation on Communication Platforms

1.1.2.1 Terminology

- 26 The term ‘content moderation’ is not defined by law. In general, content moderation describes the process online platforms typically use (i) to set own standards of public discourse, information flow, and individual freedom of expression; (ii) to enforce those standards and (statutory) rights, particularly by means of screening, ranking, filtering, and blocking user-generated (unlawful) content, including proactive (and reactive) algorithmic techniques; and (iii) to establish internal proceedings with the primary purpose of protecting user rights more effectively.⁷⁸ According to a broader definition, content moderation is meant to be ‘the governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse.’⁷⁹

1.1.2.2 Incentives for Establishing Ex Ante Mechanisms of AI-Based Content Moderation

- 27 From the perspective of online platforms, there are multiple reasons for establishing *ex ante* automated technology allowing to find, remove, and prevent illicit or standard-infringing user generated content. First and foremost, internet service providers are driven to install such means in order to avoid or, at any rate, minimize potential risks of liability.⁸⁰ By the same token, platforms might be tempted to pre-empt potential

⁷⁵ YouTube, ‘What are policies?’ https://support.google.com/youtube/answer/107383?hl=en&ref_topic=24332 accessed 31 December 2023.

⁷⁶ Ibid.

⁷⁷ Grosse Ruse-Kahn (n 5), (2020) PIJIP Research Paper Series 51, 1, 10 at fn 44.

⁷⁸ Cf Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 42, 51; see also Y Nahmias and M Perel, ‘The oversight of content moderation by AI: Impact assessment and their limitations’ (2021) 58(1) Harvard Journal on Legislation 145, 171: ‘[...] the organized practice of screening online content based on the characteristics of the website, its targeted audience, and jurisdictions of user-generated content to determine whether such content is appropriate’.

⁷⁹ J Grimmelmann, ‘The Virtues of Moderation’ (2015) 17 YALE J.L. & TECH. 42, 47, differentiating between hard and soft moderation, defining ‘moderation’ as ‘the governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse’.

⁸⁰ Vividly expressed by Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 78: ‘[...] the new wave of Internet regulation and the emergence of “voluntary” filtering illustrates the risk that governments will *informally* pressure platforms to adopt limitations on speech’.

regulatory constraints tightening (intermediate) liability, or even to bow to the pressure of powerful rightsholders to stem the flood of infringing content more effectively compared to a reactive notice and takedown system.⁸¹

- 28 Even though there is no explicit statutory obligation for online platforms to actively seek facts indicating infringing user content – neither under the EU DSM Directive,⁸² the E-Commerce Directive,⁸³ and the Digital Services Act⁸⁴ nor under the US DMCA⁸⁵ –, (large)

⁸¹ With respect to copyright enforcement: J E Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (Oxford University Press 2019) 123 f. As a reaction to unsatisfying copyright enforcement based on notice and takedown, commercial copyright owners and providers of user-generated content (UGC) services entered, in 1997, into the ‘UGC Principles’, a non-binding set of principles calling, inter alia, for to use of ‘effective content identification technology’ – as, for example, was the case, in 2007, with YouTube’s fingerprinting technology (Content ID). Correspondingly, large rightsholders developed automated mechanisms to detect, track, and report online infringement and generate takedown requests, see Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 63 f, referring also to N P Suzor, *Lawless: The Secret Rules that Govern our Digital Lives* (Cambridge University Press 2019) 76-78.

⁸² See Art 17(8) of the EU Directive 2019/790 of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, Official Journal L 130/92: ‘The application of this Article shall not lead to any general monitoring obligation’. In its recent judgment regarding the annulment of Art 17(4) lit b) and c) of the DSM Directive, the CJEU stated that ‘a filtering system which might not distinguish adequately between unlawful content and lawful content, with the result that its introduction could lead to the blocking of lawful communications, would be incompatible with the right to freedom of expression and information [...] and would not respect the fair balance between that right and the right to intellectual property’. Regarding the prohibition of a general monitoring obligation under Art 17 DSM Directive, the service providers ‘cannot be required to prevent the uploading and making available to the public of content which, in order to be found unlawful, would require an independent assessment of the content by them in the light of the information provided by the rightholders and of any exceptions and limitations to copyright’: CJEU, 26 April 2022, C-401/19 – *Poland v Parliament and Council*, ECLI:EU:C:2022:297, para 86, 90.

⁸³ See Art 15(1) of the Directive 2000/31/EC of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, Official Journal L 178/1.

⁸⁴ Regulation on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC, COM (2020) 825 final. While Art 14(1) and 15 of the E-Commerce Directive prohibit general monitoring, Member States may oblige host providers to ‘detect and prevent certain types of illegal activities’ (recital 48) and impose ‘monitoring obligations in a specific case[s]’ (recital 47). To this, see K Kaesling, ‘Privatising Law Enforcement in Social Networks: A Comparative Model Analysis’ (2018) (3) Erasmus Law Review 151, 154 f. In sharp contrast to 47 U.S.C. § 230, there is no Good Samaritan Privilege under the E-Commerce Directive. With regard to the CJEU ruling in *L’Oréal*, denying the liability privilege to be applied to platforms when an ‘active role’ (CJEU, 12 July 2011, C-324/09 – *L’Oréal v eBay*, ECLI:EU:C:2011:474, para 113 and CJEU, 23 March 2010, C-236/08 und C-238/08 – *Google France v Luis Vuitton Malletier*, ECLI:EU:C:2010:159, para 120) and Art 14(1) E-Commerce Directive referring to the threshold of actual and constructive knowledge, platforms run the serious risk of liability for user generated content. – Art 6 DSA Regulation now introduces a Good Samaritan Clause. This provision applies to measures taken in accordance with EU law as well as to a platform’s own terms and conditions. Therefore, Art 6 DSA Regulation might ‘not oblige platforms to monitor but rather invite them to do so’, N Gielen and S Uphues, ‘Digital Markets Act und Digital Services Act’ (2021) 32(14) EuZW (Europäische Zeitschrift für Wirtschaftsrecht) 627, 632, thus incentivizing platforms to perform more removals on their own initiative, see A Kuczerawy, ‘The Good Samaritan that wasn’t: voluntary monitoring under the (draft) Digital Services Act’ <https://verfassungsblog.de/good-samaritan-dsa/> accessed 31 December 2023.

⁸⁵ See 17 U.S.C. § 512(m) DMCA.

Internet platforms actually developed corresponding monitoring and filtering techniques.⁸⁶ Rather, the EU legislator⁸⁷ and the CJEU⁸⁸ tolerate the use of algorithmic filter technologies.⁸⁹ On the contrary, intermediary immunities and safe harbors⁹⁰ stimulated the proliferation of innovative moderation technologies.⁹¹

- 29 At least, developing such programs helps service providers to promote their own business and reputation while, at the same time, protecting their users.⁹² In this lens, platforms have, in the course of the (recent) past, undertaken substantial investments in techniques enabling them to proactively monitor or filter illegal and standard-violating online content.⁹³

⁸⁶ Cf also Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 63 with further references.

⁸⁷ Neither the German Netzwerkdurchsetzungsgesetz (NetzDG) nor the EU Code of Conduct on Countering Illegal Hate Speech (this Code of Conduct is a self-regulatory, joint act of internet service companies initiated by the European Union: European Commission, 'Tackling online disinformation: Commission proposes an EU-wide Code of Practice' https://ec.europa.eu/commission/presscorner/detail/en/IP_18_3370 accessed 31 December 2023) explicitly demand proactive automated measures. In doing so, the short time frames (less than 24 hours) and the mere scale of affected content prompt platforms to take such measures. Similar see Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 70-72.

⁸⁸ See, eg, CJEU, 3 October 2019, C-18/18 – *Glawischnig-Piesczek*, ECLI:EU:C:2019:821, para 46, imposing the obligation to prevent defamatory content of equivalent nature and explicitly stating that this is not '[...] an excessive obligation being imposed on the host provider, in so far as the monitoring of and search for information [...] does not require the host provider to carry out an independent assessment, since the latter has recourse to automated search tools and technologies'.

⁸⁹ See Commission Recommendation (EU) 2018/334 on measures to effectively tackle illegal content online, 1 March 2018, declaring that 'in addition to notice-and-action mechanisms, proportionate and specific proactive measures taken voluntarily by hosting service providers, including by using automated means in certain cases, can also be an important element in tackling illegal content online, without prejudice to Article 15(1) of Directive 2000/31/EC', C [2018] 1177 final <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018H0334&from=EN>.

⁹⁰ Cf Art 14(1) of the E-Commerce Directive 2000/31/EC; Art 3-5 DSA Regulation, and the 47 U.S.C. § 230. – Art 14(1) of the E-Commerce Directive is described as the 'European equivalent' to 47 U.S.C. § 230, see Gielen and Uphues (n 84), (2021) 32(14) EuZW (Europäische Zeitschrift für Wirtschaftsrecht) 627, 632.

⁹¹ K Klonick, 'Why the History of Content Moderation Matters' <https://www.techdirt.com/articles/20180129/21074939116/whyhistory-content-moderation-matters.shtml> accessed 31 December 2023; Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 51.

⁹² See S.D. Cal. (USA), 25 April 2012, 11cr0938 JM – *United States v Green*, 857 F. Supp. 2d 1015, quoting testimony of Don Colcolough, AOL's Director of Investigations and Cyber Security.

⁹³ See Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 58, referring to the PhotoDNA technique developed by Microsoft. This tool, which is licensed for free to technology companies and law enforcement, can match the hash values of photos or videos uploaded by individual users against a database of hash values of other photos or videos containing illegal images of child sexual abuse. Cf also J Kosseff, 'Private Computer Searches and the Fourth Amendment' (2018) 14(2) I/S: A Journal of Law and Policy 187, 209.

1.1.2.3 The Functioning of Algorithmic Content Moderation

1.1.2.3.1 Basic Features

- 30 Content moderation can take place both automatically and by human moderators.⁹⁴ In communication platforms, both variants are usually used (hybrid moderation system), albeit the proportion of automated content moderation varies considerably depending on the platform.
- 31 In hybrid moderation systems, automated content moderation has three main functions: (i) detecting violations of law or platform standards;⁹⁵ (ii) supporting human content moderators, for example by prioritizing submitted content⁹⁶ or by means of a preliminary assessment of the content in question;⁹⁷ and finally (iii) independent decision-making by the moderation system.⁹⁸ In addition, automated content moderation is meant to prevent the re-upload of content that violates the law or the terms and conditions of the platform (so-called ‘stay down’).⁹⁹
- 32 Automated content moderation can take place both proactively – through the use of filtering technology at the time between upload and publication of the content¹⁰⁰ – and

⁹⁴ K Klonick, ‘The New Governors: The People, Rules and Process Governing Online Speech’ (2018) 131(6) *Harvard Law Review* 1598, 1635; R Gorwa, R Binns and C Katzenbach, ‘Algorithmic content moderation’, (2020) 7(1) *Big Data & Society* 1, 7, tabulating the use of human moderators in the automated moderation systems of various platforms.

⁹⁵ N Elkin-Koren and M Perel, ‘Separation of Functions for AI: Restraining Speech Regulation by Online Platforms’ (2020) 24(3) *Lewis & CLARK L. REV.* 857, 878.

⁹⁶ Gorwa, Binns and Katzenbach (n 94), (2020) 7(1) *Big Data & Society* 1, 6; Meta, ‘How we review content’ <https://about.fb.com/news/2020/08/how-we-review-content/> accessed 31 December 2023. To Wikipedia’s ‘Huggle’ bot, human Wikipedia moderators by prioritizing suspicious content for human content review: Content with the highest likelihood of abusive editing is thus reviewed first, see E Katsh and O Rabinovich-Einy, *Digital Justice: Technology and the Internet of Disputes* (Oxford University Press 2017) 125.

⁹⁷ Google’s Perspective technology calculates a score about the ‘impact a comment might have on a conversation’. This score could be used by content moderators or provide real-time ‘feedback’ to the posting user about their content; Gorwa, Binns and Katzenbach (n 94), (2020) 7(1) *Big Data & Society* 1, 9.

⁹⁸ Meta (n 96), ‘How we review Content’: ‘Our AI systems automate decisions for certain areas where content is highly likely to be violating’.

⁹⁹ Elkin-Koren and Perel (n 95), (2020) 24(3) *Lewis & CLARK L. REV.* 857, 879.

¹⁰⁰ Differentiation according to Klonick (n 94), (2018) 131(6) *Harvard Law Review* 1598, 1635 f; similarly: Katsh and Rabinovich-Einy (n 96) 52; on *Facebook*: see Meta, ‘How Facebook uses super-efficient AI models to detect hate speech’ <https://ai.facebook.com/blog/how-facebook-uses-super-efficient-ai-models-to-detect-hate-speech/> accessed 31 December 2023; on *Airbnb*: AIRBNB, ‘Scoring the user to prevent “suspicious” activity before it occurs: What Does It Mean When Someone’s ID Has Been Checked?’ <https://www.airbnb.com/help/article/2356/what-does-it-mean-when-someones-id-has-been-checked> accessed 31 December 2023; cf R Van Loo, ‘Federal Rules of Platform Procedure’ (2021) 88(4) *University of Chicago Law Review* 829, 845.

reactively after publication of the incriminated content on the platform (eg, in response to flagging of the content by users).¹⁰¹

1.1.2.3.2 The Technical Operation, *Inter Alia*, Machine Learning Systems

- 33 Automated content moderation can be static – as in copyright law – in order to classify known content by means of fingerprinting (hash matching),¹⁰² as well as dynamic on the basis of machine learning tools: The goal here is to detect and sanction new, unknown content. In addition, there are hybrid systems that have elements of both static hashing/matching and dynamic classification methods.¹⁰³
- 34 Concrete technical tools for dynamic detection of incriminated content are primarily so-called classification tools, which use statistical methods to predict the probability of an infringement.¹⁰⁴
- 35 The decision-making process in machine-learning based on content moderation is highly dynamic: The artificial intelligence identifies certain patterns and correlations and adjusts its sanctioning reaction (allowing the upload, blocking or removal) accordingly.¹⁰⁵ This can be done by labeling ('offensive/not offensive') as part of a supervised learning.¹⁰⁶ In that respect, analysis techniques are applied to large amounts of data. Based on valid training data, the artificial intelligence is capable of learning, recognizing patterns and classifying on the basis of probabilities with regard to the installed labels. Subsequently, a sanction can be enforced automatically.¹⁰⁷ Finally, the

¹⁰¹ It is by no means mandatory to focus on the time of publication, nor is it primarily descriptive in nature. It is also conceivable, for example, to differentiate between action by the platform before the infringed party becomes aware of the content (proactive) or only in response to a report (such as flagging) of the incriminated content by users or other actors (reactive). Nevertheless, in the case of content relevant to the law of expression, the time of publication (as the initial possibility of third parties taking notice) is sensitive to fundamental rights and thus particularly relevant, see Klonick (n 94), (2018) 131(6) Harvard Law Review 1598, 1635 f.

¹⁰² As regards the technical functioning of fingerprinting, see already above para 13-15.

¹⁰³ Gorwa, Binns and Katzenbach (n 94), (2020) 7(1) Big Data & Society 1, 5: 'There are also systems which blur the lines between the two. For instance, a series of photos taken milliseconds apart might be something that a matching system ought to classify as similar, even though the underlying images are different and therefore technically not matches Facial recognition technologies may serve the dual purpose of inducing patterns from many faces and matching particular faces belonging to the same person. In these cases, the distinction between identity-matching and classification is a matter of degree'.

¹⁰⁴ Gorwa, Binns and Katzenbach (n 94), (2020) 7(1) Big Data & Society 1, 4 f; Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 885 f.

¹⁰⁵ N Elkin-Koren, 'Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence', (2020) 7(2) Big Data & Society 1, 5.

¹⁰⁶ Gorwa, Binns and Katzenbach (n 94), (2020) 7(1) Big Data & Society 1, 4 f.

¹⁰⁷ Elkin-Koren (n 105), (2020) 7(2) Big Data & Society 1, 5; Gorwa, Binns and Katzenbach (n 94), (2020) 7(1) Big Data & Society 1, 4 f.

use of deep learning technologies with the aid of artificial neural networks is also conceivable.¹⁰⁸

- 36 Information about content identified as illegal is fed back into the machine learning system. Through such a feedback mechanism, the system ‘learns’; a so-called feedback loop is created.¹⁰⁹
- 37 Machine learning requires considerable amounts of data.¹¹⁰ On one thing, these are fed by the data on the platform, for example about user activities.¹¹¹ Certainly, data available outside the specific platform can also be included, eg, by connecting third-party applications via application programming interfaces (APIs).¹¹²
- 38 In the light of different objectives amalgamating in an AI system, machine-learning based content moderation finds application both to detect and sanction incriminated content and with the aim of commercializing content, eg, by personalized matching of this content.¹¹³

1.1.2.4 Excursus: The Content Moderation of the Meta Group as a Prototype of Individualized Process Design¹¹⁴

- 39 While originally, Meta’s content moderation was only rudimentarily regulated and based exclusively on unpublished guidelines,¹¹⁵ it henceforth draws on an increasingly complex

¹⁰⁸ N Elkin-Koren (n 105), (2020) 7(2) Big Data & Society 1, 55 f.

¹⁰⁹ In this regard: Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 885 f.

¹¹⁰ On AI and data protection, see M Valkanova, ‘Trainieren von KI-Modellen’ in M Kaulartz and T Braegelmann (ed), *Rechtshandbuch Artificial Intelligence und Machine Learning* (2020) 336 ff.

¹¹¹ Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 862, 876.

¹¹² For example, Facebook incorporates data from its users who identify themselves to other services using Facebook accounts into the Facebook Social Graph, Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 862, 876; J-C Plantin, C Lagoze, P N Edwards and C Sandvig, ‘Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook’ (2018) 20(1) New Media & Society 293, 304; A Helmond, ‘The Platformization of the Web: Making Web Data Platform Ready’ (2015) 1(2) Social Media and Society 1 ff. Airbnb has patented an AI technology that screens users’ online activities outside the platform to rank individual users in a ‘trustworthiness score’ or a ‘compatibility score’. This is done based on user behavior and ‘personality trait metrics’ on the basis of a scoring system. Deploying these tools serves to prevent ‘suspicious’ activities before they actually occur, see Van Loo (n 100), (2021) 88(4) University of Chicago Law Review 829, 844.

¹¹³ Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 862, 886.

¹¹⁴ This part (1.1.2.4) Excursus: The Content Moderation of the Meta Group as a Prototype of Individualized Process Design) was written independently by Helena Müller, former research assistant at the chair of Prof. Dr. Björn Laukemann.

¹¹⁵ According to Dave Willners, author of the first draft of the Facebook Community Standards, when he joined the company in 2009, all content moderation was based on one page of internal rules These were applied globally. There was also very limited guidance on content moderation, see Klonick (n 94), (2018) 131(6) Harvard Law Review 1598, 1630 f; see also A Heldt, *Intensivere Drittwirkung* (Mohr Siebeck 2023) 198 f.

construct of hierarchically structured private rules. These are enforced as part of a multi-track – now largely automated – internal dispute resolution process.¹¹⁶

1.1.2.4.1 The Emergence of Specific Types of Platform Procedures Using the Example of the Cross-Check System

1.1.2.4.1.1 Procedural Framework

- 40 Meta even has its own internal procedures for certain groups of users (‘entitled entities’). Content posted by these persons, which is subjected to a review as part of the moderation process, is treated separately as part of the so-called cross-check system.¹¹⁷ In doing so, users are accorded a more prominent position in terms of the procedural and substantive legal framework.¹¹⁸ The specific procedural position of the user is in some cases largely determined by its economic relevance for the Group.¹¹⁹ This does not only apply in comparison to the regular platform procedure. Such a differentiation is also found within the cross-check system.
- 41 The procedure as such is tailored to the correction of so-called ‘false positives’ – ie content decisions that were wrongly classified as a violation of the in-house community standards during the regular moderation procedure, but are in fact to be judged as compliant.¹²⁰ As a result, a cross-check procedure is only opened if the content in question is found to violate the community standards during the regular moderation procedure (‘at scale’). However, the opposite constellation of so-called ‘false negatives’ is not covered. This is content that violates the company’s own standards but is incorrectly qualified as compliant during the regular moderation process.¹²¹ This means that the procedure is unilaterally geared towards preventing overenforcement.¹²² This

¹¹⁶ Meta Transparency Center, ‘How technology detects violations’ <https://transparency.meta.com/de/de/enforcement/detecting-violations/technology-detects-violations/> accessed 31 December 2023.

¹¹⁷ The cross-check system received significant public attention in the wake of the Facebook paper scandal. The cross-check system was recently the subject of a Policy Advisory Opinion by Meta’s Oversight Board, see Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*. The process also goes by the name ‘X-Check’, see J Horwitz, ‘Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That’s Exempt’ <https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353> accessed 31 December 2023.

¹¹⁸ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, Part 1, 4.

¹¹⁹ This was also harshly criticized by the Oversight Board (decision of 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, Part 1, 3): ‘While Meta told the Board that cross-check aims to advance Meta’s human rights commitments, we found that the program appears more directly structured to satisfy business concerns. The Board understands that Meta is a business, but by providing extra protection to certain users selected largely according to business interests, cross-check allows content which would otherwise be removed quickly to remain up for a longer period, potentially causing harm’.

¹²⁰ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 12, 18; Meta Transparency Center, ‘Reviewing high-impact content accurately via our cross-check system, 12 May 2023’ <https://transparency.fb.com/enforcement/detecting-violations/reviewing-high-visibility-content-accurately/> accessed 31 December 2023.

¹²¹ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 12.

¹²² Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 11-13.

is a procedural design that the Meta Group explicitly supports with regard to the content covered by the system, especially that of important advertising customers.¹²³

- 42 The cross-check procedure can only be opened by the Meta Group (quasi ‘ex officio’). Therefore, it is designed as an internal review procedure. Users have no possibility to initiate such a procedure (eg, by filing an in-house appeal).

1.1.2.4.1.2 Specifics of the Procedure

1.1.2.4.1.2.1 Context-Specific In-Depth Examination

- 43 Content is prioritized as part of the cross-check system and reviewed by particularly expert moderators located within the company (‘Early Response Secondary Review [ERSR]’). This review is always performed by human moderators. The moderators carry out an in-depth review of the content, considering the specific context of the statement, and have special decision-making powers. For example, at an advanced stage of the cross-check process, they can grant exceptions for content that is contrary to Facebook community standards. This is done by applying specific policies¹²⁴ and exceptions¹²⁵. Both are not publicly accessible¹²⁶ and can only be applied within the scope of such special procedures – ie, not in the regular procedural course (‘at scale’)¹²⁷ – by selected bodies within the company.¹²⁸ As a consequence, they are denied to the vast majority of Facebook users. The application of such specific internal regulations and exceptions conflicts with Meta’s statement that identical standards are used as the basis for content moderation on all platforms.

1.1.2.4.1.2.2 Suspension Effect

- 44 In terms of timing, the cross-check procedure is downstream of the regular moderation procedure (‘at-scale’). It takes place immediately after the original moderation decision (‘initial decision’), but before the intended private sanction is enforced. In this process,

¹²³ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 16.

¹²⁴ The ‘Spirit of Policy Allowance’ is a well-known example. Structurally, this equates to a teleological interpretation of the Community Standards Based on these teleological considerations, a dispensation from the regular moderation decision can be granted. In the regular moderation process (‘at scale’), this possibility does not exist. Regular moderators must decide strictly on the basis of the wording of the policies (‘letter of the policy’); Oversight Board, 17 June 2022, 2022-001-FB-UA – *Knin Cartoon*, Part 8.1, 17.

¹²⁵ In this context, the so-called ‘Newsworthiness Allowance’ has been the subject of much discussion. According to press reports, the soccer player Neymar is considered to be a beneficiary of this exemption, see Horwitz (n 117), ‘Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That’s Exempt’.

¹²⁶ In light of this, Meta’s Oversight Board criticizes the divergence in content between publicly available policies and non-publicly available internal moderation standards and policies, see Oversight Board, 28 January 2021, 2020-005-FB-UA – *Nazi Quote*, Key Findings and Part 8.1.

¹²⁷ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 18.

¹²⁸ For example, the ‘Early Response Team’, see Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, charts at 14, 21.

the content to be reviewed remains on the platform until the cross-check procedure is completed¹²⁹ – at least within the scope of the person-related Early Response Secondary Review.¹³⁰ From a structural perspective, initiating the cross-check procedure has a kind of suspension effect. This represents a significant difference from the regular moderation procedure and other special procedural steps of the moderation process. In this case, the sanction imposed in the initial decision remains enforced until the final decision, even in the event of a renewed review of the content. This applies both on the basis of an in-house appeal filed by the user and on the basis of an autonomous review initiated by the Meta Group.

1.1.2.4.1.3 Special Modalities of the Cross-Check Procedure

- 45 Within the cross-check procedure, a differentiation is made between a person-related (so-called Early Response Secondary Review) and a content-related procedure (General Secondary Review).

1.1.2.4.1.3.1 The Person-Related Strand of the Cross-Check Procedure: The ‘Early Response Secondary Review (ERSR)’

- 46 Content from users (‘entitled entities’) that violates community standards, which is on an internal list (‘cross-check list’) specifically intended for this purpose, is always subject to a follow-up check as part of the cross-check procedure (Early Response Secondary Review, in the following: ERSR).¹³¹ This particularly includes business partners of the Meta Group, advertising customers and users who represent a particular legal or regulatory risk for the Group – for example, in the case of ongoing legal disputes – as well as state actors.¹³² Also addressed are prominent groups of persons such as journalists.¹³³ The qualification of a user as an ‘entitled entity’ is at the discretion of the

¹²⁹ Under the content-based General Secondary Response (GSR) process, this is only the case under certain conditions; see below para 40-44 for both procedural modalities specific to Meta’s General Secondary Review System: Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*, Part 6, 9.

¹³⁰ Oversight Board, 9 January 2023, 2022-013-FB-UA – *Iran Protest Slogan*, Part 8.1 II., 14: ‘[...] Cross-check [...] enable[s] users’ content to be reviewed on escalation prior to removal’. See Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, Part 1, 4. This applies without limitation in terms of timing only to the ERSR process: Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 47; see on this point below para 50-52.

¹³¹ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 19, 24-39. In technical terms, this is done by tagging the data subjects, see para 24.

¹³² See below para 73 for more information.

¹³³ The cross-check procedure only provides information about the internal categorization of the Meta Group. The exact identity of the beneficiary users is unclear. Exemplary named are: Users associated with ‘significant world events’, members of groups of people disproportionately affected by overenforcement, and media organizations, see: Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 24 f. For example, the Al Jazeera news network and Donald Trump are part of the cross-check system: Oversight Board, 14 September 2021, 2021-009-FB-UA – *Shared AI*

Meta Group. According to the company's own Oversight Board, economic parameters in particular are decisive for this.¹³⁴

1.1.2.4.1.3.2 The Content-Related Strand of the Cross-Check Procedure: The General Secondary Review (GSR)

- 47 In addition, a cross-check procedure can also be initiated for content-related reasons, irrespective of the identity of the posting party.¹³⁵ This procedural modality – the so-called General Secondary Review, in the following: GSR – was only created retrospectively in 2021,¹³⁶ presumably as a reaction of the group to the Facebook Files scandal.¹³⁷
- 48 The detection of such GSR content is carried out by an algorithm ('cross-check-ranker'), resembling an automated triage process. In principle, all content on the platform is suitable, regardless of the identity of the posting user.¹³⁸ The content in question must already have been classified by the regular moderation system as a violation of the community standards and be intended for enforcement of a corresponding platform sanction.¹³⁹ Trivial is whether this is done by a human moderator or automatically.¹⁴⁰
- 49 The algorithm's key decision-making criteria include the sensitivity of the post and the user concerned, the potential enforcement severity of a platform sanction, the probability of a false positive decision, and the potential spread of the content in

Jazeera Post, Part 6; also Oversight Board, 5 May 2021, 2021-001-FB-FBR – *Former President Trump's suspension*, Part 2. In addition, The Oversight Board failed to gain access to the company's internal cross-check lists in the course of the cross-check process, despite multiple requests to the Meta Group, Horwitz (n 117), 'Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt'.

¹³⁴ Explicitly Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 24. The categorization of the groups of persons already suggests this.

¹³⁵ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 20, 40-55.

¹³⁶ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, Part I, 3.

¹³⁷ The scandal originated from internal documents of the Meta Group ('Facebook Files'), which were leaked to the Wall Street Journal and the US Senate by whistleblower Frances Haugen. The documents reveal, among other things, that the Meta Group was aware as early as 2019, following an internal investigation, that the platform's recommendation system was significantly encouraging the spread of hate and disinformation. The documents also made the existence of the cross-check program public knowledge, see the Wall Street Journal's article series 'The Facebook Files' <https://www.wsj.com/articles/the-facebook-files-11631713039> accessed 31 December 2023; esp D Seetharaman, J Horwitz and J Scheck, 'Facebook Says AI Can Enforce Its Rules, but the Company's Own Engineers Are Doubtful' <https://www.wsj.com/articles/facebook-ai-enforce-rules-engineers-doubtful-artificial-intelligence-11634338184> accessed 31 December 2023. Specific to the cross-check program: M Reuter, 'Facebook Knew What All Was Going Wrong' <https://netzpolitik.org/2021/facebook-files-facebook-wusste-was-alles-schieflaeuft/> accessed 31 December 2023.

¹³⁸ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 40.

¹³⁹ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 43.

¹⁴⁰ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 43.

question.¹⁴¹ The cross-check ranker continuously identifies new content suitable for GSR. As a result, previously unprocessed, older and lower-priority content is downgraded in the internal processing order of the GSR process and, due to the time limit of the process, eventually drops out of the GSR process altogether.¹⁴²

1.1.2.4.1.4 Allocation of Resources in Favour of the Personal ERSR Procedure

- 50 A gradation also takes place within the cross-check procedure. Content in the personal ERSR process is prioritized and given priority for review. In that regard, it is guaranteed that the content of the 'entitled entities' is checked.¹⁴³ In contrast, in the content-based GSR procedure, content is only checked once there is any residual capacity:¹⁴⁴ If there is no such capacity within a certain period,¹⁴⁵ this GSR-related content falls outside the scope of the cross-check procedure. This ends the privilege of this content remaining on the platform during the ongoing procedure and the sanction of the regular ('at scale') procedure takes effect. This weighs heavily in particular due to the very high overturn rate in the GSR procedure: For example, in February 2022, 80% of the content that was classified as a violation of the community standards by the regular moderation system was classified as compliant with those standards in the GSR procedure.¹⁴⁶
- 51 Contrarily, there is no such time limit in the ERSR procedure: Here, the content remains on the platform until the final decision is made.¹⁴⁷ As a result, content from users who are part of the ERSR system is always reviewed by a human moderator.
- 52 The allocation of resources in favor of the ERSR procedure also means that content from the GSR procedure reaches the higher instances of the cross-check procedure less frequently. However, since the specific policies and exceptions described above can only be granted at an advanced stage of the cross-check procedure, this means that the vast majority of GSR procedures do not benefit from them.¹⁴⁸

1.1.2.4.1.5 Availability of Appeal Mechanisms in the Cross-Check Procedure

- 53 The divergent availability of appeal mechanisms within Meta's content moderation continues within the cross-check procedure. In that regard, there is only partial internal

¹⁴¹ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 42 f.

¹⁴² Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 46.

¹⁴³ Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*, Part 8.1, II, 12.

¹⁴⁴ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 44 f.

¹⁴⁵ The Oversight Board cites a two to four-day time frame here: Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 46.

¹⁴⁶ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 48.

¹⁴⁷ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 46.

¹⁴⁸ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 47.

‘legal protection’ against decisions made in the cross-check procedure.¹⁴⁹ Firstly, this applies to legal remedies for the affected users themselves (the posting party). Consequently, third parties who wish to appeal against the content of a user who benefits from the cross-check procedure (such as an advertising customer) are denied internal legal protection.¹⁵⁰

- 54 Under both procedural modalities, there is no notification that the content (or user) in question is part of the cross-check procedure.¹⁵¹ Users who are part of the cross-check system also have no opportunity to defend themselves against inclusion in the special procedure.¹⁵²

1.1.2.4.2 Other Specific Types of Procedures

- 55 There are also other specific types of proceedings. However, considerably less public information exists about how they work. As far as can be seen, these procedures also pursue specific purposes, such as preventing overenforcement or enabling the Meta Corporation to react to acute political crises and content that is likely to reach a viral level on the platform.

1.1.2.4.2.1 The High Impact False Positive Override (HIPO)

- 56 The High Impact False Positive Override procedure (in the following: HIPO) is also used to correct false positives.¹⁵³ Therefore, it pursues the same objective as the cross-check procedure. HIPO is also a verification procedure initiated by Facebook itself. Users do not have the option of initiating it. Contrary to the cross-check procedure, however, the opening of a HIPO procedure does not result in a suspension effect: The sanction imposed in the initial decision remains in place during the subsequent HIPO procedure; a review is therefore only carried out ex post.¹⁵⁴
- 57 The detection of HIPO-suitable content can be automated.¹⁵⁵ A check is then performed as part of the separate HIPO procedure track. The order of the review is based on an automatically assigned score that depends on the priority of the content (‘HIPO ranker’). Decisive criteria for this prioritization are also the sensitivity of the content and the user,

¹⁴⁹ For example, Meta states that approximately 35% of the content that went through the cross-check system had no ‘legal recourse’ to Meta’s Oversight Board, see decision from 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 174.

¹⁵⁰ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 174.

¹⁵¹ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 114.

¹⁵² Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta’s Cross-Check Program*, para 114.

¹⁵³ Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*, Part 2, 5.

¹⁵⁴ Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*, Part 6, 9.

¹⁵⁵ Cf also recently: Oversight Board, 18 December 2023, 2023-054-FB-UA, 2023-055-FB-UA, 2023-056-FB-UA, 2023-057-FB-UA – *Goebbels Quote*.

as well as the anticipated degree of dissemination of platform content.¹⁵⁶ Meta's Oversight Board criticized the functioning of this HIPO ranker as 'highly inaccurate'.¹⁵⁷

- 58 The review in the HIPO procedure is carried out by external (out-sourced) moderators.¹⁵⁸ A review in the HIPO procedure only takes place if sufficient personnel capacities are available. If this is not the case, there is no additional human review at all; the moderation decision is made automatically.¹⁵⁹
- 59 This structure is similar to the General Secondary Review (GSR) under the cross-check procedure.¹⁶⁰ Nonetheless, it remains completely unclear according to which criteria content is proposed for review in the GSR procedure and when merely a HIPO procedure takes place. This classification has significant consequences: For example, the content in question remains on the platform during the GSR procedure, while a HIPO review only enables a downstream review.¹⁶¹ It can probably be assumed that content deemed less relevant by the Meta Group will only be dealt with subsidiarity in the context of the HIPO review.

1.1.2.4.2.2 Procedures for Responding to Political Crises and Viral Content

- 60 The Meta Group provides an option for subjecting potentially viral content to additional human review (so-called High-Risk Early Review Operations: HERO).¹⁶² The same internal

¹⁵⁶ Ibid, 9.

¹⁵⁷ According to the board, users who regularly report on the activities of dangerous organizations or individuals are exposed to an increased risk of sanctions. This results from the policy of deleting content related to such dangerous organizations or persons in case of doubt, if it is not quite clear that the content in question is merely a factual report about the events. This leads to such content being removed disproportionately often. In the facts underlying the 'Board's decision, the HIPO ranker did not recognize the weightiness of the content. The subject of the decision was a post by an Indian magazine that reported on the Taliban's school closures in Afghanistan. The post was blocked by the Meta Group on the basis of its 'Dangerous Individuals and Organizations policy'. This policy forbids 'praising' such 'dangerous entities' as terrorist organizations. A human moderator also classified the content as a violation. Against this, the Indian newspaper filed a user appeal, whereupon the post in question was added to the queue of the HIPO proceedings. The content was classified as non-priority by the HIPO ranker. In addition, there was a lack of capacity of Urdu-speaking HIPO moderators. For these reasons, the content fell out of the HIPO system and was adjudicated 'at scale'. In its decision, the Board criticized (in addition to insufficient staff capacity in the HIPO process) the HIPO ranker's lack of sensitivity to press coverage under the Dangerous Individuals and Organizations policy. Due to the high relevance of press coverage for freedom of expression, such content was particularly weighty. See: Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*, Part 8.3, III, 16.

¹⁵⁸ Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*, Part 6, III, 10.

¹⁵⁹ Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*, Part 6, III, 9.

¹⁶⁰ Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*, Part 6, III, 9.

¹⁶¹ Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*, Part 6, III, 9.

¹⁶² Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 181 f.

team that also carries out checks as part of the cross-check procedure is responsible for this.¹⁶³ However, it remains unclear for which procedure the personnel resources are allocated. In any case, it seems likely that the person-related ERSR procedure will be prioritized.

1.1.2.4.2.3 Procedure for Handling Government Requests

- 61 The so-called escalation process provides for a separate procedure for reports by prosecution agencies.¹⁶⁴ This applies both to reports based on potential illegality or relevance to criminal law, but also in particular to reports based on alleged violations of Meta's own community standards.¹⁶⁵ The sanctioning of content reported on the basis of a potential law infringement is regularly only carried out locally,¹⁶⁶ for example through geo-blocking.¹⁶⁷ In contrast, content that violates the community standards is blocked globally.
- 62 In terms of its objective, this procedure is to a certain extent the 'mirror image' of the cross-check procedure: In the government request procedure, the identity of the reporter is decisive, whereas in the context of the cross-check procedure, the identity of the posting party is the relevant criterion. In the escalation procedure, an in-depth review of the content is also carried out by internal company moderators.¹⁶⁸ Likewise, there is no possibility of redress against these decisions:¹⁶⁹ Users whose content is sanctioned by government agencies on the basis of a report are thus denied legal protection within the company solely on the basis of the special status of the complainant. The situation is further aggravated by the fact that the user concerned is not informed of this fact either.¹⁷⁰ The escalation procedure also opens up the possibility

¹⁶³ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 181 f.

¹⁶⁴ Oversight Board, 22 November 2022, 2022-007-IG-MR – *UK Drill Music*, Part 6, 16.

¹⁶⁵ Oversight Board, 22 November 2022, 2022-007-IG-MR – *UK Drill Music*, Part 6, 16.

¹⁶⁶ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 56; Meta Transparency Center, 'How we assess reports of content violating local law' <https://transparency.fb.com/data/content-restrictions/> accessed 31 December 2023. If ordered by the state, (potentially) infringing content is also blocked worldwide. According to Meta, this was the case in 14 cases in the period from January to June 2022. Meta generally refers to such orders as 'extraterritorial jurisdiction', Meta Transparency Center, 'Global restrictions' <https://transparency.fb.com/data/content-restrictions/> accessed 31 December 2023.

¹⁶⁷ For instructions on how this works, see J Hörnle, *Internet Jurisdiction. Law and Practice* (Oxford University Press 2021) 448-450.

¹⁶⁸ Oversight Board, 22 November 2022, 2022-007-IG-MR – *UK Drill Music*, Part 1, 6. – The Oversight Board had already requested the Meta Group on several occasions to formalize the procedure for handling government requests and to list their number in transparency reports: Oversight Board, 14 September 2021, 2021-009-FB-UA – *Shared AI Jazeera Post*, Part 10.

¹⁶⁹ Meta justifies this by saying that decisions made in the Escalation process are not made by content moderators.

¹⁷⁰ Oversight Board, 22 November 2022, 2022-007-IG-MR – *UK Drill Music*, Part 6, 16. This had already been criticized by the Board in an earlier decision, cf Oversight Board, 8 July 2021, 2021-006-IG-UA – *Ocalan's Isolation*, Part 10.

of applying specific policies and exceptions.¹⁷¹ Furthermore, government agencies are not required to substantiate their reports – for example, by naming the community standard presumed to be affected, by justifying the enforcement decision or by providing a sufficient basis of evidence.¹⁷²

- 63 The procedure for reporting by government agencies always takes precedence over the cross-check procedure. In this case, the company's internal decision-makers are in a position to make sanction decisions directly – irrespective of the specific content or the identity of the posting party. This applies to both modalities of the cross-check procedure, in particular also to 'entitled entities' under the ERSR procedure.¹⁷³
- 64 Overall, sovereign actors are thus privileged in two ways. In their favour, separate procedures apply both to content posted by them and to content reported by them.

1.1.2.4.3 Conclusion

- 65 Meta's internal procedures show a progressive automation of content moderation. In that regard, rapid technical progress offers the possibility of ever greater flexibility and individualization of platform procedures.¹⁷⁴ Their design is apparently based on economic interests: namely a preference for advertising and business partners, the minimization of liability risks and the avoidance of negative public perception. This is particularly clear from the privileged treatment of user groups in the ERSR procedure and state actors in the government request procedure. Moreover, the design of the internal platform procedures focuses on preventing overenforcement. In contrast, Meta's efforts to prevent underenforcement¹⁷⁵ are only weakly developed – probably to avoid the impression of censorship.¹⁷⁶
- 66 The possibility of a company-internal user appeal depends on the discretion of the platform. Its predominant role becomes particularly apparent in cases of purely automated decisions (and technical deficits in the system), not least if a user has no other effective legal protection. In view of the fact that platform-based dispute resolution substitutes traditional legal protection, the design of internal company legal remedies should not be subject solely to the private design power of platforms, but – as

¹⁷¹ In the facts underlying the Oversight Board's decision in *Drill Music*, the 'veiled threats analysis' was applied. In the context of this analysis, it is again considered whether the report is made by the state. Thus, in situations such as this one, where government entities report content that falls within the scope of this exception, there is a de facto double consideration of the government identity of the reporter: Oversight Board, 22 November 2022, 2022-007-IG-MR – *UK Drill Music*, Part 6, 15.

¹⁷² Oversight Board, 22 November 2022, 2022-007-IG-MR – *UK Drill Music*, Part 8.1, II b, 26.

¹⁷³ Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*, para 56 f.

¹⁷⁴ See in detail below para 71-73.

¹⁷⁵ The so-called *reporter appeals* serve this purpose, see: Oversight Board, 6 December 2022, 2021-002-F-PAO – *Meta's Cross-Check Program*, para 181 f.

¹⁷⁶ Oversight Board, 6 December 2022, 2021-002-F-PAO – *Meta's Cross-Check Program*, para 16.

recently in the DSA (esp in Art 20) – should be accompanied by an effective guarantee of basic procedural safeguards in platform proceedings and corrected where necessary.

1.2 Dangers and Drawbacks of AI-Based Law and Standards Enforcement

- 67 An AI-based enforcement of private rights is often viewed critically. Using Content ID as an example, the following section describes the main dangers and disadvantages that can arise from a computer-based (copy-)rights check or from the private enforcement of (copy-)rights. Legal issues arising solely from the ‘regulatory framework’ of the Content ID process will be excluded.

1.2.1 *The Economization of Intra-Platform Decisions: Manifestations and Technical Design of Platform Power*

1.2.1.1 Structuring of Interactions on Platforms

- 68 As an integral and dominant part of the market, digital platforms enable the exchange of goods and services, the processing of payments, and global communication by means of their technical and economic infrastructure.¹⁷⁷ The platform economy is essentially based on the use of continuously refined algorithms.¹⁷⁸ In addition to providing and structuring the infrastructure, platforms as economic actors primarily pursue their own economic interests – whether through participation in generated profits,¹⁷⁹ by placing advertisements¹⁸⁰ or also for the purpose of avoiding liability. Both functions of the platform merge in a technical system.¹⁸¹
- 69 As a consequence of the law-enforcing and dispute-settling role that platforms are increasingly assuming, classic procedural guarantees – such as the right to be heard – are being dosed on the basis of economic parameters, ie, limited or even denied comprehensively.¹⁸² This is driven in particular by tendencies toward privatization of tasks traditionally incumbent on state actors.¹⁸³

¹⁷⁷ Introducing the legal economic foundations of digital platforms: A Engert, ‘Digitale Plattformen’ (2018) 218(2-4) AcP (Archiv für die civilistische Praxis) 218, 304; J K Mendelsohn, ‘Die “normative Macht” der Plattformen – Gegenstand der zukünftigen Digitalregulierung’ (2021) 24(11) MMR (Multimedia und Recht) 857, 858.

¹⁷⁸ Engert (n 177), (2018) 218(2-4) AcP (Archiv für die civilistische Praxis) 218, 304, 307.

¹⁷⁹ Engert (n 177), (2018) 218(2-4) AcP (Archiv für die civilistische Praxis) 218, 304, 307 f.

¹⁸⁰ Klonick (n 94), (2018) 131(6) Harvard Law Review 1598, 1627.

¹⁸¹ Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 875-879.

¹⁸² Instructive on this point: R Van Loo, (2016) ‘The Corporation as Courthouse’ 33(2) Yale Journal on Regulation 547.

¹⁸³ H Askani, *Private Rechtsdurchsetzung bei Urheberrechtsverletzungen im Internet* (Nomos 2021) 162. For more details, see below para 78-82.

- 70 The agglomeration of users and the tendency to form monopolies is inherent in the economy of platforms.¹⁸⁴ Platforms structure their markets (explicitly) and act as private rule-setters.¹⁸⁵ For example, e-commerce platforms create digital market orders and structure the exchange of goods between their customers;¹⁸⁶ communication platforms act on the communication of their users. In the context of content moderation, social media platforms set their own (behavioural) standards and implement them on the basis of their own sanctioning system. In doing so, they act functionally comparable to a court.¹⁸⁷ The structuring of interactions on platforms also – if not primarily – takes place implicitly through the architecture, the code of the platform.¹⁸⁸ In technical terms, this behaviour control is highly individualized and – through the use of algorithms and machine learning tools¹⁸⁹ – adaptive (so-called hyper nudging).¹⁹⁰ By means of behavioral microtargeting,¹⁹¹ it is possible to address users in a personalized manner on the basis of behavior and personality-based user profiles. These profiles are generated by algorithmic analysis of large volumes of data. In particular, the user's behaviour on the platform is considered.

1.2.1.2 Tailored Outcomes and Procedures

- 71 The alignment of the platform infrastructure based on economic parameters takes place at the level of access to the procedure and throughout all stages of the procedure. This

¹⁸⁴ For more details, see: Engert (n 177), (2018) 218(2-4) AcP (Archiv für die civilistische Praxis) 218, 304, 307.

¹⁸⁵ Mendelsohn (n 177), (2021) 24(11) MMR (Multimedia und Recht) 857, 858.

¹⁸⁶ See in detail M Glogowski, *Plattformbedingungen* (Mohr Siebeck 2022) 3-5.

¹⁸⁷ Furthermore: D Wielsch, 'Die Ordnungen der Netzwerke. AGB – Code – Community Standards' in M Eifert and T Gostomzyk (ed), *Netzwerkrecht* (Nomos 2018) 61 f. Cf also Askani (n 183) 173; Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 871.

¹⁸⁸ L Lessig, *Code and Other Laws of Cyberspace* (Basic Books 1999), 382. Yeung criticizes the fact that the relevance of the design or code of technical environments generally receives too little attention in legal research, and criticizes the fact that too much thought is still given to the classical structure of legal prohibitions and their enforcement ('command and control'), K Yeung, "'Hypernudge': Big Data as a mode of regulation by design', (2017) 20(1) Information, Communication & Society 118, 120; see also B Wagner, *Global Free Expression – Governing the Boundaries of Internet Content* (Springer 2016), 130.

¹⁸⁹ Mendelsohn (n 177), (2021) 24(11) MMR (Multimedia und Recht) 857, 859.

¹⁹⁰ S Katyal, 'Private Accountability in the Age of Artificial Intelligence' (2019) 66(1) UCLA Law Review 55, 93 f.

¹⁹¹ On the technical functioning of microtargeting and the implications of this technology for private autonomy, see M Ebers, '§ 3 Regulierung von KI und Robotik' in M Ebers, C Heinze and B Steinrötter (ed), *Künstliche Intelligenz und Robotik* (Beck 2020) 75 para 101 ff. For a list of data that Facebook uses to generate personalized ads, see: C Dewey, '98 personal data points that Facebook uses to target ads to you' <https://www.washingtonpost.com/news/the-intersect/wp/2016/08/19/98-personal-data-points-that-facebook-uses-to-target-ads-to-you/> accessed 31 December 2023.

makes it possible to tailor both the concrete design of the procedure and the substantive (enforcement) decision to the individual user.¹⁹²

1.2.1.2.1 Individualized Process Design in Payment Transactions

- 72 In the case of payment service providers, it is known that an individually created score for the user, which is based on a user's economic profitability, for example, determines the level of expertise of the decision-maker that the platform uses to moderate or settle a dispute in each case.¹⁹³ The material content of decisions – eg, when granting goodwill – is also based individually on the 'value', ie, the economic relevance of a customer. Admittedly, this is not a new development. However, the possibilities for analyzing large volumes of data, which have been considerably increased by means of artificial intelligence, are making it easier to examine and evaluate customer behaviour and to react to it (in a timely manner).¹⁹⁴

1.2.1.2.2 The Content Moderation of the Meta Group as a Prototype of Individualized Process Design

- 73 The development outlined becomes particularly vivid around communication platforms. The Meta Group serves as the inspiration for this. It has been known for some time that in the moderation process of Facebook's own platform,¹⁹⁵ the factors of virality, the potential danger of the content and the probability of a violation of the company's proper community standards determine whether decisions are made automatically by an AI system or by a human content moderator.¹⁹⁶ In particular, this development can be observed on the basis of the Group's special platform procedural traits. The impression suggests itself that the establishment and design of these procedures primarily serves the pursuit of economic interests – for example through the preferential treatment of advertising and business partners –, furthermore the minimization of liability risks and the avoidance of negative public perception. This becomes evident in

¹⁹² Van Loo (n 182), (2016) 33(2) Yale Journal on Regulation 547, 602. On the individualization of the customer relationship: J Taeger and S Kremer, *Recht im E-Commerce und Internet* (Fachmedien Recht und Wirtschaft 2021) 10.

¹⁹³ As a consequence, this score determines whether the customer is served by an internal 'executive customer relations' department of the bank or by an external call center: Van Loo (n 182), (2016) 33(2) Yale Journal on Regulation 547, 564 f.

¹⁹⁴ On Amazon, see: H Eidenmüller and G Wagner, *Law by Algorithm* (Mohr Siebeck 2021) 239 ff; Van Loo (n 182), (2016) 33(2) Yale Journal on Regulation 547, 564-566.

¹⁹⁵ For more details on the procedure, see above para 39-66.

¹⁹⁶ In 2020, the Meta Group stated that this prioritization process was used to automatically decide on content that obviously violates standards. The goal is to create capacities to use human moderators primarily for deciding complex or context-dependent constellations of facts, see: Meta, 'How we review Content – Prioritization' <https://about.fb.com/news/2020/08/how-we-review-content/> accessed 31 December 2023.

the privileged treatment of certain user groups under the cross-check procedure and the provision of a specific procedure for notifications by government agencies.¹⁹⁷

1.2.1.3 Private Ordering and Monetization in Copyright Law

- 74 The alignment of the platform infrastructure with economic self-interest is not less vivid in the area of copyright, as is exemplified by YouTube's content ID module.¹⁹⁸ With the U.S. DMCA, the DSM Directive of the European Union, and most recently the German Copyright Service Provider Act (UrhDaG), there is (by now) a narrow regulatory framework for dispute resolution – compared to the law of expression. Nevertheless, there is a clear shift from the classic notice-and-takedown system to the monetization of content.¹⁹⁹ Primary beneficiaries, as shown, are platforms and rightsholders.²⁰⁰ The design of Content-ID as well as technically continuously optimized (upload) filter technologies have a structural effect to the detriment of the uploader (encroachment on freedom of expression,²⁰¹ danger of overblocking or chilling effects).²⁰² Last but not least, the obligation to use filter technology opens up new markets and thus investment incentives in the economic interest of platforms and software manufacturers to develop increasingly precise filter technologies.²⁰³ The shift toward monetization shows that platforms know how to use the regulatory leeway provided by the state to their own economic advantage. Corresponding technological developments are based on state regulations, insofar as they have a mandatory character. Beyond this, they still often deviate from the legal model in crucial areas and become independent. YouTube's content ID is paradigmatic for the further development of private law enforcement and dispute resolution systems from state guidelines and principles of copyright law (such as uniform law enforcement and the application of exemptions) to complex systems of private orders.²⁰⁴

1.2.1.4 AI-Driven Structuring of Communication Processes and Separation Of User Groups

- 75 The social, political and private dimensions of platform use cannot always be clearly separated from the underlying commercial transaction, especially on social media

¹⁹⁷ In detail below para 83-89.

¹⁹⁸ For the technical procedure, see above para 1-22.

¹⁹⁹ For the concept and operation of monetization, see above para 2-6.

²⁰⁰ Also in that regard: Schillmöller and Doseva (n 5), (2022) 25(3) MMR (Multimedia und Recht) 181 f.

²⁰¹ G Frosio, 'Algorithmic Enforcement Online' in P Torremans (ed), *Intellectual Property and Human Rights* (4th edn, Kluwer Law International 2020) 24; J Lennartz and V Kraetzig, 'Filtering fundamental Rights' <https://verfassungsblog.de/filtering-fundamental-rights/> accessed 31 December 2023. On the DSM Directive's incentives for the use of proactive filtering technologies, see: Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 883.

²⁰² For this, see below para 99-102.

²⁰³ This can be seen not least in Audible Magic's lobbying for the enshrinement of an obligation to use filtering technologies within the framework of Art 17 DSM Directive, see: Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 85 f.

²⁰⁴ Grosse Ruse-Kahn (n 5), (2020) PIJIP Research Paper Series 51, 1, 7.

platforms.²⁰⁵ Here, the moderation of harmful or even (criminally) illegal content merges with the commercial goal of encouraging users to stay on the platform as long as possible in order to collect data and expose personalized advertisements.²⁰⁶

- 76 The described mixed situation is reinforced by the technical functioning of algorithms: In particular, the concept of ‘filter bubbles’ assumes that algorithms, due to their technical mode of operation, tend to display to the user only information that matches his or her previous views and to hide information that does not fit.²⁰⁷ The phenomenon of algorithmic segmentation of groups of people into so-called echo chambers – which is quite controversial in its scope – is also discussed.²⁰⁸ In this context, Internet users interact only with like-minded people. Such separation harbours the risk of intensifying discriminatory ideas and radicalizing communication processes (‘algorithmic radicalization’).²⁰⁹ Adaptive platform environments and the use of feedback loops in turn encourage user behavior influenced in this way to be ‘fed back’ into the system. This shows: The configuration of the algorithm determines which content and information is displayed.²¹⁰ In doing so, the algorithmically controlled architecture of platforms has a very significant influence on the freedom to form and express opinions, as well as on the guarantee of media pluralism as a whole.²¹¹

²⁰⁵ Mendelsohn (n 177), (2021) 24(11) MMR (Multimedia und Recht) 857, 858.

²⁰⁶ Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 875 f.

²⁰⁷ Ebers (n 191), 75 para 112.

²⁰⁸ On echo chambers, see C R Sunstein, *#Republic – Divided Democracy in the Age of Social Media* (Princeton University Press 2017) 122-124. – The studies on the existence of both phenomena are not clear, on this Ebers (n 191), 75 para 112-114. Critically also J Lüdemann, ‘Warum und wie reguliert man digitale Informationsintermediäre?’ in J Lüdemann and Y Hermstrüwer (ed), *Schutz der Meinungsbildung im digitalen Zeitalter* (Mohr Siebeck 2021) 15-19. However, studies suggest that the influence of personalized media offerings on opinion formation may be less than previously assumed, see: E Dubois and G Blank, ‘The echo chamber is overstated: the moderating effect of political interest and diverse media’ (2018) 21(5) Information, Communication & Society 729.

²⁰⁹ Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 889. Affirmatively on the influence of personalization algorithms on radicalization processes: S Musa and S Bendett, ‘Islamic Radicalization in the United States – New Trends and a Proposed Methodology for Disruption’ (2010), National Defense University, Washington DC Center for Technology and National Security Policy 17 ff <https://appsdtic.mil/sti/pdfs/ADA532696.pdf> accessed 31 December 2023. In particular, recent studies on the YouTube algorithm assume that it is even capable of counteracting radicalization, see M Ledwich and A Zautsev, ‘Algorithmic Extremism: Examining YouTube’s Rabbit Hole of Radicalization’, <https://arxiv.org/abs/1912.11211> accessed 31 December 2023. M Wolfowicz, D Weisburd and B Hasisi, ‘Examining the interactive effects of the filter bubble and the echo chamber on radicalization’ (2023) 19(1) Journal of Experimental Criminology, 119.

²¹⁰ T Gillespie vividly refers to this algorithmic separation and structuring process as ‘calculated publics’: ‘The Relevance of Algorithms’ in T Gillespie, P Boczkowski and K Foot (ed), *Media Technologies* (The MIT Press 2014) 188.

²¹¹ Ebers (n 191) 75 para 111.

- 77 Algorithmic separation of certain groups of people also takes place on platforms with a direct link to the market,²¹² especially in e-commerce²¹³ and on sharing platforms²¹⁴. However, the encroachment on areas sensitive to fundamental rights is less pronounced here.

1.2.2 Merging Government and Private Regulatory Objectives: The Public-Private Divide

1.2.2.1 Privatization of Law Enforcement and Dispute Resolution

- 78 In the area of law enforcement and dispute resolution in particular, it is becoming apparent that platforms are increasingly performing tasks that have traditionally been the responsibility of state actors,²¹⁵ including the judiciary.²¹⁶ Complex privatization processes are emerging across all regulatory and legal areas – whether in copyright law, the law of expression, data protection law²¹⁷ or the settlement of contractual disputes in e-commerce or payment transactions.
- 79 The causes of this development are intricate. Some of them stem from complex governance problems, especially in the case of cross-border conflicts.²¹⁸ At the same time, state regulation is shifting dispute resolution to private actors:
- 80 In copyright law, such a shift of responsibility²¹⁹ manifests itself, for example, in the imposition of procedural obligations,²²⁰ which in turn are implemented and (excessively) concretized by private forms of dispute resolution such as YouTube's Content ID.²²¹ The DSM Directive sets incentives, if not a de facto obligation, to use filtering systems.²²²

²¹² Engert (n 177), (2018) 218(2-4) AcP (Archiv für die civilistische Praxis) 218, 304, 307 f.

²¹³ On the individualization of matching on platforms: M Berberich and A Conrad, '§ 30 Plattformen und KI' in M Ebers, C Heinze and B Steinrötter (ed), *Künstliche Intelligenz und Robotik* (Beck 2020) para 28.

²¹⁴ Specifically on Airbnb, see Katsh and Rabinovich-Einy (n 96) 68 ff.

²¹⁵ Hörnle (n 167) 450.

²¹⁶ Detailed on the phenomenon of privatization of the judiciary, including in the context of the right to be forgotten: E Haber, 'Privatization of the Judiciary', (2016) 40(1) *Seattle University Law Review* 115, 120 ff; Hörnle (n 167) 450. Specifically, in the context of expression and copyright law, see Elkin-Koren and Perel (n 95), (2020) 24(3) *Lewis & CLARK L. REV.* 857, 871; Askani (n 183) 177 ff, 251 f, 269.

²¹⁷ E Haber (n 216), (2016) 40(1) *Seattle University Law Review* 115, 118 ff.

²¹⁸ Hörnle (n 167) 450. On the causes of the increase in private enforcement in copyright law, see Askani (n 183) 174-176.

²¹⁹ Elkin-Koren and Perel (n 95), (2020) 24(3) *Lewis & CLARK L. REV.* 857, 871.

²²⁰ F Hofmann, 'Prozeduralisierung der Haftungsvoraussetzungen im Medienrecht – Vorbild für die Intermediärhaftung' (2017) 61(2) *ZUM (Zeitschrift für Urheber- und Medienrecht)* 102, 104 f.

²²¹ According to Askani, the adaptation of intermediaries can be understood as a reaction to the existing legal system that the development took place, so to speak, 'in the shadow of the law', Askani (n 183) 170; see also Frosio (n 201) 6; Grosse Ruse-Kahn (n 47), (2018) 49(9) *IIC (International Review of Intellectual Property and Competition Law)* 1017, 1018.

²²² Lennartz and Kraetzig (n 201), *Filtering fundamental Rights*.

- 81 In the area of the law of expression, platforms regulate the boundaries of what can be said as part of their content moderation by means of private standard-setting and technical infrastructure²²³ – even beyond the area of illegal content and content relevant to criminal law.²²⁴ This development is also driven by economic motives, such as maintaining contractual relationships with advertising customers. For example, in response to criticism from major advertising customers, Google announced that it would refine the content classifiers it uses to detect extremist and terrorist content.²²⁵ This reaction was preceded by the placement of ads on websites with antisemitic and homophobic content.²²⁶
- 82 As the German Network Enforcement Act (Netzwerkdurchsetzungsgesetz, in the following: NetzDG) exemplifies, state regulation shifts the enforcement of state law to private parties.²²⁷ Where legal leeway remains, the German legislator explicitly relies on self-regulation²²⁸ and accepts that platforms transfer their (primarily automated) dispute resolution systems, which they have developed for complaints against their general terms and conditions, to the enforcement of state law – ie, matters within the scope of the NetzDG.²²⁹ The concrete modes of control are manifold: In addition to procedural rules, the NetzDG also relies on indirect behavioral control by means of liability obligations²³⁰ and the threat of fines.²³¹

²²³ Wielsch (n 187), 61, 65; Klonick (n 94), (2018) 131(6) Harvard Law Review 1598, 1603 f, 1669 f.

²²⁴ See only BGH (Germany), 29 July 2021, III ZR 179/20, BGHZ 230, 347 = (2021) 24(11) MMR (Multimedia und Recht) 903 para 78 (deletion of posts and account blocking by Facebook in the case of hate speech).

²²⁵ As a result of the incident, major advertisers, including the British government and L'Oréal Group, withdrew their ads, also on YouTube, see: M Murgia, H Warell and D Bond, 'YouTube revenues under threat over ads alongside extremist videos' <https://www.ft.com/content/04f8bf56-0b12-11e7-97d1-5e720a26771b> accessed 31 December 2023; K Walker, 'Four ways Google will help to tackle extremism' <https://www.ft.com/content/ac7ef18c-52bb-11e7-a1f2-db19572361bb> accessed 31 December 2023; Kent Walker is Google's senior vice-president and general counsel.

²²⁶ Murgia, Warell and Bond (n 225), 'YouTube revenues under threat over ads alongside extremist videos'.

²²⁷ Critically: J Lüdemann, 'Privatisierung der Rechtsdurchsetzung in sozialen Netzwerken?' in M Eifert and T Gostomzyk (ed), *Netzwerkrecht* (2018) 165.

²²⁸ See Report of the German Federal Government on the Evaluation of the Network Enforcement Act, Bundestag-Drucksache 19/22610, 8: 'In this context, the specifications contain implementation leeway for the providers of the social networks with regard to the implementation of the specifications'.

²²⁹ Report of the German Federal Government on the Evaluation of the Network Enforcement Act, Bundestag-Drucksache 19/22610, 10, 29, 86. However, the provision of § 2(2) no 2 NetzDG imposes transparency obligations on network operators to provide information on procedures for automated content recognition, see J-C Kalbhenn, 'Design Specifications for Chatbots, Deepfakes, and Emotion Recognition Systems' (2021) 65(8/9) ZUM (Zeitschrift für Urheber- und Medienrecht) 663, 672.

²³⁰ See also Elkin-Koren and Perel (n 95), (2020) 24(3) Lewis & CLARK L. REV. 857, 871. – On the copyright liability model: Askani (n 183) 192-194.

²³¹ § 4 NetzDG (Network Enforcement Act).

1.2.2.2 (Cross-Border) Influence of State Actors on Platforms

- 83 A reflection of this is a growing state influence on globally effective private regulatory systems of large online platforms. By exerting regulatory impact on the governance structures of such platforms, states are able – whether intentionally or unintentionally – to achieve a global effect of their regulatory acts (which in themselves only have a national impact).²³² The main reason for this is the position of digital platforms as gatekeepers for key social and economic needs. By adapting their technical infrastructure to the regulatory requirements (of the European Union, for example), platforms are in a position to help these requirements achieve transnational, even global validity.²³³
- 84 The CJEU's decision in *Glawischnig-Piesczek v Facebook* serves as a model for the de facto extension of national law. The ruling approved both the worldwide blocking of the incriminated content and the obligation of the platform to prevent future infringements with content similar in wording and meaning by the Austrian court of instance. The CJEU explicitly stated that hosting operators are allowed to use 'automated techniques and means of investigation' – such as keyword filtering²³⁴ – and are therefore not obliged to make 'autonomous assessments'.²³⁵
- 85 Such 'regulatory reinforcement' can take place explicitly, for example in the form of legislative acts. Increasingly, however, it is also possible to observe the exertion of informal pressure.²³⁶ This 'jawboning'²³⁷ is expressed, for example, in the form of threats by political actors or the activity of so-called Internet Referral Units (IRUs).²³⁸ These entities are state agencies that report content to the service providers on the basis of violations of the platform's own standards – and thus not on the basis of violations of state law – and encourage removal, usually on a global basis.²³⁹ Platforms are not

²³² Also: Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 86.

²³³ Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 86.

²³⁴ Hörnle (n 167) 38.

²³⁵ CJEU, 3 October 2019, C-18/18 – *Glawischnig-Piesczek*, ECLI:EU:C:2019:821, para 46.

²³⁶ D Kaye, *Speech Police* (Columbia Global Reports 2019) 79.

²³⁷ H Bloch-Wehba, 'Global Platform Governance: Private Power in the Shadow of the State' (2019) 72(1) SMU Law Review 27, 63; Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 78; D E Bambauer, 'Against Jawboning' (2015) 100(1) MINN. L. REV. 51, 57-58.

²³⁸ Subdivisions of (also national) police forces are often affected: for example, the British Counter-Terrorism Internet Referral Unit (CTIRU) was established by Scotland Yard, see Kaye (n 236) 79. The European equivalent is located at Europol, see Europol, EU Internet Referral Unit – EU IRU <https://www.europol.europa.eu/about-europol/european-counter-terrorism-centre-ectc/eu-internet-referral-unit-eu-iru> accessed 31 December 2023. On the genesis of the EU Internet Referral Unit, see R Eghbariah and A Metwally, 'Informal Governance: Internet Referral Units and the Rise of State Interpretation of Terms of Service' (2021) 23 Yale J.L. & Tech. 545, 574 f.

²³⁹ According to the Meta Group, sanctioning of content reported for potential violation of state law regularly occurs only locally, for example through geo-blocking. In contrast, blocking of content violating community standards has a global effect: Oversight Board, 6 December 2022, 2021-002-FB-

obligated to remedy the deletion requests and remove the content; rather, they act voluntarily.²⁴⁰ Nevertheless, these requests represent a vehicle of state influence: For example, mass and disproportionate reporting of specific content risks imposing a state-preformed understanding of platform standards on platforms.²⁴¹ The often broad and open wording of these standards benefits state actors in this regard.²⁴²

- 86 Moreover, the service providers are quite cooperative: The Meta Group, for instance, prioritizes the processing of inquiries from such internet referral units within the framework of a separate, so-called government request procedure, with the involvement of expert decision-makers equipped with special competencies.²⁴³ In addition, there is no internal legal protection against moderation decisions made in response to a report from government agencies.²⁴⁴ Although such reports are also reviewed exclusively due to a company's own community standards, it is possible to observe a shortening of the company's internal legal protection simply because the complainant is a sovereign actor. YouTube also offered to classify the Internet Referral Unit of the EU as a Trusted Flagger.²⁴⁵
- 87 By participating in industry-specific hash databases of service providers, state internet referral units also explicitly exercise 'definitional power'.²⁴⁶ In that regard, the regulation

PAO – *Meta's Cross-Check Program*, para 56; Meta Transparency Center, 'How we assess reports of content violating local law' <https://transparency.meta.com/reports/content-restrictions/content-violating-local-law/> accessed 31 December 2023. – Instructive on how geo-blocking works: Hörnle (n 167) 448-450. Critical of this form of extraterritorial governance: Eghbariah and Metwally (n 238), (2021) 23 Yale J.L. & Tech. 545, 599.

²⁴⁰ For example, Europol emphasizes: 'The decision to remove the referred content is taken by the concerned service provider in accordance with their policies and terms of service', in EU Internet Referral Unit Transparency Report (2021) 3 https://www.europol.europa.eu/cms/sites/default/files/documents/EU_IRU_Transparency_Report_2021.pdf; see also B Chang, 'From Internet Referral Units to International Agreements: Censorship of the Internet by the UK and EU' (2018) 49(2) Columbia Human Rights Law Review 114, 135; see also L Helfer and M K Land, 'The Meta Oversight Board's Human Rights Future' (2023) 44(6) Cardozo Law Review 2233, 2275 ff.

²⁴¹ Eghbariah and Metwally (n 238), (2021) 23 Yale J.L. & Tech. 545, 592, 601-606; Kaye (n 236) 81.

²⁴² Kaye (n 236) 82.

²⁴³ Moderators in the 'escalation process', for instance, may apply special policies and exceptions not available to the public. Such moderators have special expertise, and undertake an in-depth, contextual review of the content in question, see: Oversight Board, 22 November 2022, 2022-007-IG-MR – *UK Drill Music*, Part 6, 15 f.

²⁴⁴ Oversight Board, 22 November 2022, 2022-007-IG-MR – *UK Drill Music*, Key Findings, 3. For more details on the Government Request process, see above para 61-64.

²⁴⁵ Chang (n 240), (2018) 49(2) Columbia Human Rights Law Review 114, 122, 135; Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 61; Bloch-Wehba (n 237), (2019) 72(1) SMU Law Review 27, 45 f, 62 f.

²⁴⁶ The EU Internet Referral Unit, for instance, states: 'The EU IRU participated in the EU Internet Forum Senior Officials meetings [...] and provided relevant contents to feed the database of hashes', Europol, 2018 Consolidated Annual Activity 44 https://www.europol.europa.eu/cms/sites/default/files/documents/consolidated_annual_activity_report_2018.pdf accessed 31 December 2023; see also Eghbariah and Metwally (n 238), (2021) 23 Yale J.L. & Tech. 545, 604.

of platforms varies from country to country: The United States, Great Britain and Germany are considered particularly influential.²⁴⁷

- 88 Given that the entire content moderation of a platform takes place within the same technical system, the adaptation of state regulation has resulted in an increasing fusion of two – functionally fundamentally different – forms of enforcement: the enforcement of state regulations in the guise of private, platform-specific standards. The technical peculiarities of machine learning used in content moderation further encourage this development.²⁴⁸
- 89 In copyright law, too, the obligations of the U.S. DMCA are being implemented on a de facto global scale by incorporating the relevant regulations into the technical infrastructure of the platform. Accordingly, blocking because of DMCA notices often takes place globally and not in the form of geo-blocking limited to local effects.²⁴⁹

1.2.3 Technical Limitations and Restricted Traceability of Context-Dependent Content

- 90 AI systems are subject to technical limitations: While AI-based upload filters already work reliably when it comes to identical sound and image files, the result for the recognition of static image files in ‘dynamic audiovisual files’ is (still) significantly less favorable.²⁵⁰
- 91 A central point of criticism is furthermore the difficult tangibility of context specific, time-sensitive and locally dependent content by AI.²⁵¹ Systems based on the

²⁴⁷ Wagner (n 188) 6.

²⁴⁸ See above para 33-38; Elkin-Koren and Perel (n 95), (2020) 24(3) *Lewis & CLARK L. REV.* 857, 887.

²⁴⁹ A Bridy, ‘Intellectual Property’ in D Keller (ed), *Law, Borders, and Speech: Proceedings and Materials* (2017) 13; Bloch-Wehba (n 3), (2020) 53(1) *Cornell International Law Journal* 41, 86; D Holznagel, *Notice and Take-Down-Verfahren als Teil der Providerhaftung* (Mohr Siebeck 2013) 125 (stating that the notice-and-takedown process in Germany is purely self-regulatory); Askani (n 183) 169.

²⁵⁰ A Conrad and G Nolte, ‘Schrankenbestimmungen im Anwendungsbereich des UrhDaG’ (2021) 65(2) *ZUM* (Zeitschrift für Urheber- und Medienrecht) 111, 118, referring to the statement of Google/YouTube from 8 November 2020 on the draft bill of the German Federal Ministry of Justice and Consumer Protection (BMJV) for a law to adapt copyright law to the requirements of the digital single market 15 https://www.bmj.de/SharedDocs/Downloads/DE/Gesetzgebung/Stellungnahmen/2020/110820_Stellungnahme_Google_RefE_Urheberrecht-ges.pdf?__blob=publicationFile&v=3 accessed 31 December 2023.

²⁵¹ Nahmias and Perel (n 78), (2021) 58(1) *Harvard Journal on Legislation* 145, 176 ff, with respect to ‘offensive language’; further Engstrom and Feamster (n 39) 18: ‘[...] such technologies are not sufficient to consistently identify infringements with accuracy, as they can only indicate whether a file’s contents match protected content, not whether a particular use of an identified file is an infringement in light of the context within which the media was being used’. Suggesting design improvement for the Content ID process: L D Shinn, ‘YouTube’s Content ID as a Case Study of Private Copyright Enforcement Systems’ (2015) 43(2/3) *AIPLA Quarterly Journal* 359, 386 ff.

fingerprinting method have – at least so far²⁵² – difficulties in distinguishing copyright-legitimate from certain unlawful reproductions, primarily regarding the scope of application of the fair use exception under US copyright law.²⁵³ The same applies to typically context-dependent content such as criticism, caricatures, parodies, pastiche, or quotation.²⁵⁴

- 92 Considering this, parts of the literature assume that algorithms are generally not capable of legally adequate assessments and scrutiny.²⁵⁵ The deficient recognizability of permitted content is seen as problematic, especially in view of a ‘remix culture’ whose creative approach consists precisely in alienating already existing content and protected works (see, for example, memes, supercuts or mash-ups).²⁵⁶
- 93 Particularly in view of the limited algorithmic recognizability of fair use exceptions, the fundamental question arises as to whether a system that learns on the basis of comparative materials will be able to replicate abstract legal and statutory situations or, on the basis of case law, selectively decided legal questions of the past. Also questionable is whether a correspondingly static ‘law recognition and application program’ will be able to either subsume unknown facts (not yet fed into the system) under applicable, context-based legal provisions in accordance with the principles of legal methodology, further to develop the law accordingly, or ‘simply’ to represent

²⁵² Conrad and Nolte (n 250), (2021) 65(2) ZUM (Zeitschrift für Urheber- und Medienrecht) 111, 118 (In the foreseeable future, it is not to be expected that, for example, quotations or parodies will be correctly included in a legal assessment of AI).

²⁵³ Engstrom and Feamster (n 39), (2017) The Limits of Filtering: A Look at the Functionality & Shortcomings of Content Detection Tools 18: ‘It is often permissible to excerpt or otherwise refer to copyrighted content in contexts that are permitted by fair use [...]. Although an automated algorithm could determine whether the content (or excerpt) matched known copyrighted content, such an algorithm would not be able to determine whether the particular use of a given file is infringing or not’. See also Shinn (n 251), (2015) 43(2/3) AIPLA Quarterly Journal 359, 364; Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 65. – YouTube itself recognizes that Content ID cannot decide ‘fair use’: YouTube, ‘Frequently asked questions about fair use’ <https://support.google.com/youtube/answer/6396261#zippy=%2Ci-posted-a-disclaimer-on-my-video%2Ci-gave-credit-to-the-copyright-owner%2Cim-using-the-content-for-entertainment-or-non-profit-uses%2Cwhen-does-fair-use-apply%2Cwhat-constitutes-fair-use%2Cchow-does-fair-use-work%2Cchow-does-content-id-work-with-fair-use> accessed 31 December 2023.

²⁵⁴ Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 65. See further S Jacques, K Garstka, M Hviid and J Street, ‘An empirical study of the use of automated anti-piracy systems and their consequences for cultural diversity’, (2018) 15(2) Script-Ed 277, 298, which found that in a sample of 1.839 parodies, videos were five times more likely to be blocked by Content ID than by a DMCA proceeding; Gray and Suzor (n 5), (2020) 7(1) Big Data & Society 1, 6; A Metzger and M Senftleben, ‘Selected Aspects of Implementing Article 17 of the Directive on Copyright in the Digital Single Market into National Law – Comment of the European Copyright Society’ (20 April 2020) 1, 16 <https://ssrn.com/abstract=3589323> or <http://dx.doi.org/10.2139/ssrn.3589323> accessed 31 December 2023.

²⁵⁵ Nolte (n 47), (2017) 61(4) ZUM (Zeitschrift für Urheber- und Medienrecht) 304, 310.

²⁵⁶ Grosse Ruse-Kahn (n 5), (2020) PIJIP Research Paper Series 51, 1, 2; M Becker, ‘Von der Freiheit, rechtswidrig handeln zu können’ (2019) 63(8/9) ZUM (Zeitschrift für Urheber- und Medienrecht) 636, 644.

dynamically developing legal situations in an accurate manner.²⁵⁷ This explains why fingerprinting and hashing technologies are quite effective modes of public-private cooperation on policing content that is predetermined to be unlawful, while they are unlikely to helpfully address highly context-dependent questions.²⁵⁸

- 94 Finally, as far as can be seen, there is no adaptation to national (copyright) law, including the fair use exception, within the algorithm either.²⁵⁹ This aspect is reinforced by an ‘Americanization’ of the Content ID procedure given that YouTube²⁶⁰ seems to apply the US principle of fair use de facto worldwide.²⁶¹

²⁵⁷ Cf Burk (n 69), (2019) 86(2) *University of Chicago Law Review* 283, 297 f: ‘The common law evolves, whether from purely judicial reasoning or from judicial riffing off of legislative enactments’ (at 298).

²⁵⁸ Bloch-Wehba (n 3), (2020) 53(1) *Cornell International Law Journal* 41, 65; D Keller, ‘Internet Platforms: Observations on Speech, Danger, and Money’ (2018) *Hoover Inst. Aegis Paper Series* no 1807 6, 7: ‘an ISIS video looks the same, whether used in recruiting or in news reporting’.

²⁵⁹ Grosse Ruse-Kahn (n 5), (2020) *PIJIP Research Paper Series* 51, 1, 9.

²⁶⁰ YouTube, ‘What Does Fair Use Mean’ <https://support.google.com/youtube/answer/9783148?hl=de> accessed 31 December 2023. With respect to deactivations for copyright infringement, YouTube makes reference to the US DMCA (cf the reference to the formal counter notification requirements of 17 USC § 512(g)(3), which requires submission to US jurisdiction in 17 USC § 512(g)(3)(D) as a condition of a counter notification, see https://support.google.com/youtube/answer/6005919?hl=de&ref_topic=9282678 accessed 31 December 2023. With respect to other jurisdictions, the YouTube website contains only a reference to where ‘useful information on copyright outside the U.S.’ can be found and refers to the websites of the European Commission and WIPO in this regard. Here, it is expressly clarified that these references serve only ‘informational purposes’ and do not constitute a ‘binding recommendation’ by YouTube. No reference contains specific information. The link to the WIPO’s website refers to a list of National IP Offices, see YouTube, ‘Where can I get more information about copyright outside the U.S.?’ https://support.google.com/youtube/answer/2797449?hl=de&ref_topic=2778546#zippy=%2Cwo-erhalte-ich-weitere-informationen-zum-urheberrecht-außerhalb-der-usa accessed 31 December 2023.

²⁶¹ To this aspect see Nahmias and Perel (n 78), (2021) 58(1) *Harvard Journal on Legislation* 145, 178, referring to R Radu, *Negotiating Internet Governance* (1st edn, Oxford University Press 2019) 179.

1.2.4 Limited Transparency

- 95 Another point of criticism concerns the lack of transparency of AI-based law enforcement.²⁶² In general, algorithms that are not disclosed or cannot be traced hamper creative user activity²⁶³ and generate incentives to circumvent the algorithm.²⁶⁴
- 96 Moreover, there is a serious risk of bias with respect to the ‘preselected’ data, but also with respect to the algorithm itself.²⁶⁵ It is well known that even state judges are sometimes guided by arbitrary motives and dependencies. However, the decisive factor in state civil proceedings is whether a judicial decision, on account of its justification, is comprehensible in terms of its content: If this is the case, the actual motives of a judge are irrelevant as long as they do not give rise to any concern of bias or partiality. However, it is precisely this comprehensibility that is in question in the case of AI-based law enforcement:²⁶⁶ According to the so-called black box problem, the developers of an AI regularly do not themselves know or understand extensively how the algorithm works in detail,²⁶⁷ nor can they predict or subsequently provide information about what an AI

²⁶² I S Nathenson, ‘The Procedural Foundations of Information Regulation’ (2020) 24(1) *Lewis & Clark Law Review* 109, 129; Gray and Suzor (n 5), (2020) 7(1) *Big Data & Society* 1, 2; D K Citron, ‘Technological due process’ (2008) 85(6) *Washington University Law Review* 1249, 1250, 1254; N Elkin-Koren, ‘After twenty years: revisiting copyright liability of online intermediaries’ in S Frankel and D Gervais (ed), *The Evolution and Equilibrium of Copyright in the Digital Age* (1st edn, Cambridge University Press 2014) 29, 47; Burk (n 69), (2019) 86(2) *University of Chicago Law Review* 283, 301; C Castets-Renard, ‘Algorithmic Content Moderation on Social Media in EU Law: Illusion of Perfect Enforcement’ (2020) (2)2 *University of Illinois Journal of Law, Technology & Policy* 283, 308.

²⁶³ Bloch-Wehba (n 3), (2020) 53(1) *Cornell International Law Journal* 41, 81 ff: ‘But overconfidence in technical solutions can have damaging effects. Far from serving as a neutral arbiter, the algorithms that Internet intermediaries use to rank and prioritize content often reflect and encode social bias’.

²⁶⁴ Cf the example of T Zhou, ‘Postmortem: Every Frame a Painting’ (2 December 2017) <https://perma.cc/U5WU-M6ZZ> accessed 31 December 2023. – On so-called reverse engineering, see Burk (n 69), (2019) 86(2) *University of Chicago Law Review* 283, 303; Bloch-Wehba (n 3), (2020) 53(1) *Cornell International Law Journal* 41, 83.

²⁶⁵ See only Burk (n 69), (2019) 86(2) *University of Chicago Law Review* 283, 296: ‘Algorithms do not make judgments; they are rather the products of human judgment’.

²⁶⁶ Art 13 of the EU AI Act provides for transparency obligations for so-called high-risk systems in the form of users being able to ‘appropriately interpret and use the results of the system’. The problem of lacking comprehensibility of AI decisions is the starting point of the research field of so-called ‘Explainable AI’, see D Bomhard and M Merkle, ‘Regulation of Artificial Intelligence’ (2021) 10(6) *EuCML (Journal of European Consumer and Market Law)* 257, 260; S Heiss, ‘Artificial Intelligence Meets European Union Law’ (2021) 10(6) *EuCML (Journal of European Consumer and Market Law)* 252, 258; D Gunning et al, ‘XAI-Explainable artificial intelligence’ (2019) 4(37) *Science Robotics* DOI: 10.1126/scirobotics.aay7120 accessed 31 December 2023.

²⁶⁷ Nathenson (n 262), (2020) 24(1) *Lewis & Clark Law Review* 109, 129 f. Consequently, even experts in the field might not understand the formulas if the algorithm were fully disclosed. For a detailed discussion of the black box problem, see F Pasquale, *The Black Box Society* (1st edn, Harvard University Press 2015) 3 ff.

has actually learned²⁶⁸ – this circumstance makes it considerably more difficult to deal with the AI and to disclose the internal processes.²⁶⁹

- 97 This finding is reinforced not only by the fact that YouTube leaves the decision on the dispute largely to the parties or, primarily, to the rightsholder.²⁷⁰ In addition, since the actions of the online platforms are structurally guided by interests and liability,²⁷¹ it is hardly possible to reliably act as a dispute moderator in a, in principle, independent manner.
- 98 Finally, non-transparent AI-based legal enforcement makes it difficult for the uploader to find out whether notifications were sent willfully or without sufficient legal research, thus weakening his effective defense options.²⁷²

1.2.5 Overblocking

- 99 Technical limitations of AI, such as in particular a limited ability to map context-dependent content, can, in principle, lead to both underenforcement²⁷³ and overenforcement in the context of private law enforcement. According to the predominant assessment of the Content ID procedure, the limited recognizability of context-dependent permissions leads to structural overblocking, ie, to the sanctioning of user content that is wrongly recognized as illegal.²⁷⁴

²⁶⁸ Even full disclosure of an algorithm would be insufficient to the extent that AI-based results may also depend on the algorithm's technical infrastructure (hardware and other software): Burk (n 69), (2019) 86(2) University of Chicago Law Review 283, 302.

²⁶⁹ J Burrell, 'How the machine thinks: Understanding opacity in machine learning algorithms' (1/2016) 3(1) Big Data & Society 3 ff DOI: 10.1177/2053951715622512 accessed 31 December 2023.

²⁷⁰ Schillmöller and Doseva (n 5), (2022) 25(3) MMR (Multimedia und Recht) 181, 185; Perel and Elkin-Koren (n 4), (2016) 19(3) Stanford Technology Law Review 473, 483; S Bar-Ziv and N Elkin-Koren, 'Behind the Scenes of Online Copyright Enforcement: Empirical Evidence on Notice & Takedown' (2018) 50(2) Connecticut Law Review 339, 382; D Leenheer Zimmerman, 'A Tale of Legislative Abdication' (2014) 35(1) Pace Law Review 260, 273 f.

²⁷¹ See already B Laukemann, 'Private law enforcement and intellectual property: Regulatory challenges in a digital era' in B Hess, E Jayme and H-P Mansel (ed), *Europa als Rechts- und Lebensraum: Liber Amicorum für Christian Kohler zum 75. Geburtstag* (Giesecking 2018) 269, 276 f; cf also: Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 46, with respect to content moderation ('Rather, content moderation rules – and the technologies that apply them – reflect corporate, social, and legal values'); Pasquale (n 267) 61.

²⁷² Cf Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 65.

²⁷³ Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 83.

²⁷⁴ A study of US law by J Urban, J Karagani and B Schofield found that about 30% of notifications were probably unfounded. According to another study, which looked at Google Images, this affected 70% of notices: 'Notice and Takedown in Everyday Practice' (2016) UC Berkeley Public Law Research Paper No 2755628 <https://ssrn.com/abstract=2755628> or <http://dx.doi.org/10.2139/ssrn.2755628>, 11 f; D C Nunziato, 'The Beginning of the End of Internet Freedom' (2014) 45 Georgetown Journal of International Law 383, 383: '[...] such Internet filtering regimes [...] inevitably lead to overblocking of harmless Internet content'; M Senftleben, 'Institutionalized Algorithmic Enforcement – The Pros and

- 100 If an algorithm detects a copyright infringement even if only quantitatively minor parts are affected,²⁷⁵ and if the platform operator (as YouTube probably does in the context of the Content ID procedure²⁷⁶) sanctions this infringement by the (alleged) infringer losing the possibility of monetization *comprehensively* to the (alleged) rightsholder, such an incentive reinforces the phenomenon of overblocking.²⁷⁷ It is obvious that in these cases the legal sphere of platform users is encroached upon, first and foremost their right to freedom of expression, thereby silencing legitimate or marginalized speech.²⁷⁸
- 101 Not all categories are equally affected by content blocking: For instance, rightsholders decide significantly more often to block the incriminated content than to monetize it, especially in the case of illegal interference with exploitation rights to films or to content in the field of sports. In contrast, content ID claims relating to video games are enforced less frequently.²⁷⁹
- 102 On an overall basis, YouTube is extremely cautious in taking action against abusive Content ID claims. In that regard, the company merely reserves the right to take

Cons of the EU Approach to UGC Platform Liability' (2020) 14(2) FIU Law Review 299, 312: 'Filtering more than necessary is less risky than filtering only clear-cut cases of infringement'; Gray and Suzor (n 5), (2020) 7(1) Big Data & Society 1, 2; Bar-Ziv and Elkin-Koren (n 270), (2018) 50(2) Connecticut Law Review 339: 'Analysis of the data reveals that the N&TD procedure has been extensively used to remove non-infringing materials'; H Maier, *Remixes on Hosting Platforms* (Mohr Siebeck 2018) 152; C Katzenbach, 'The "Algorithmic turn" in platform governance' (2020) 74(1 supp) Cologne Journal of Sociology and Social Psychology 283, 297: 'For the copyright field, the few existing studies point to clear overblocking'; Gorwa, Binns and Katzenbach (n 94), (2020) 7(1) Big Data & Society 1, 5.

²⁷⁵ For example, only one minute of a total 15-minute video, as exemplified by Grosse Ruse-Kahn (n 5), (2020) PIJIP Research Paper Series 51, 1, 5.

²⁷⁶ This appears to be YouTube's current approach, Grosse Ruse-Kahn (n 5), (2020) PIJIP Research Paper Series 51, 1, 5 at fn 19. YouTube, for its part, cites case law on so-called fair use, according to which there is no minimum time (of exploitation of others' works) that would be allowed under copyright law (concerning a few seconds of a sample): YouTube, 'Answers to common questions about Copyright claims on YouTube' <https://support.google.com/youtube/thread/1281991> accessed 31 December 2023.

²⁷⁷ Schillmöller and Doseva (n 5), (2022) 25(3) MMR (Multimedia und Recht) 181, 186 f.

²⁷⁸ Nahmias and Perel (n 78), (2021) 58(1) Harvard Journal on Legislation 145, 173; R Tushnet, 'All of this has happened before and all of this will happen again: Innovation in copyright licensing' (2014) 29(3) Berkeley Technology Law Journal 1447, 1460; T Spoerri, 'On Upload Filters and other Competitive Advantages for Big Tech Companies under Article 17 of the Directive on Copyright in the Digital Single Market' (2019) 10(2) Journal of Intellectual Property, Information Technology and Electronic Commerce Law 173, 176; F Mostert, 'Free Speech and Internet Regulation' (2019) 14(8) Journal of Intellectual Property Law & Practice 607, 612: 'Over-blocking and excessive filtering could too easily lead to censorship'.

²⁷⁹ Gray and Suzor (n 5), (2020) 7(1) Big Data & Society 1, 6 f: When content related to video games is removed, it is usually because of a music rightsholder's demand.

‘appropriate measures’²⁸⁰ and refers to its ‘Best Practices for Claims’.²⁸¹ Whether YouTube checks references from rightsholders in advance is not clear.²⁸²

1.2.6 Algorithmic Discrimination

- 103 Platforms are non-neutral decision-makers due to the pursuit of their own economic interests.²⁸³ This problem is reinforced and perpetuated by the use of AI-based decision-making systems. In the field of communication platforms, algorithms by means of which content is curated and prioritized are said to contain stereotypical and discriminatory assumptions.²⁸⁴ In any case, this encourages discrimination against marginalized groups and minorities: For example, expressions frequently used by English speaking African-American are flagged significantly more often; posts by Muslim users have disproportionately often been blocked in the past on the grounds that they are terrorist content.²⁸⁵ In contrast, the major platforms Facebook and YouTube have a significantly more lenient moderation practice with regard to racist and right-wing terrorist content, ignoring such content to a greater extent.²⁸⁶
- 104 Algorithmic discrimination is usually not a consequence of an (unknowing or even conscious) implementation on the part of the programmers, but often occurs through the interaction of the AI system with the technical environment and the available data.²⁸⁷ Such biases can occur both at the level of the training process – for example, through the use of incomplete or unrepresentative data – and at the classification level, for example, by linking to forbidden distinguishing characteristics such as race or gender.²⁸⁸ Especially in AI-driven content moderation, the complicating factor is that a

²⁸⁰ YouTube states: ‘For example, we may disable certain reference files or segments and remove associated claims entirely. Manual review is also required for certain reference categories. In cases of serious infringement, we may revoke access to Content ID or terminate the partnership between YouTube and the copyright owner’: YouTube, ‘Content eligible for Content ID’ <https://support.google.com/youtube/answer/2605065#zippy> accessed 31 December 2023.

²⁸¹ YouTube, ‘Best practices for claims’ <https://support.google.com/youtube/answer/4352063> accessed 31 December 2023.

²⁸² In this respect the assumption of Grosse Ruse-Kahn (n 5), (2020) PIJIP Research Paper Series 51, 1, 11, referring to YouTube, ‘Review potentially invalid references’ <https://support.google.com/youtube/answer/6013183> accessed 31 December 2023.

²⁸³ See in more detail above para 68-77 and below para 167-170. Further: Askani (n 183) 178 f.

²⁸⁴ Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 82. For example, evaluating a user's online activities enables accurate predictions about ethnicity, partisan political views, religion, substance use, sexual orientation, extraversion, intelligence, or emotional stability.

²⁸⁵ Ibid 41, 76.

²⁸⁶ Ibid 41, 77.

²⁸⁷ In the Google Ads algorithm, this learning process occurs by assigning weights or statistical probabilities based on the call history of ads, see L Sweeney, ‘Discrimination in Online Ad Delivery’ (2013) 56(5) Comm. ACM 44, <http://cacm.acm.org/magazines/2013/5/163753-discrimination-in-online-ad-delivery/> accessed 31 December 2023; A Chander, ‘The Racist Algorithm?’ (2017) 115(6) Michigan Law Review 1023, 1037.

²⁸⁸ Ebers (n 191) 75 para 162-168. In-depth on the technical causes of algorithmic discrimination also S Barocas and A D Self, ‘Big Data’s Disparate Impact’ (2016) 104(3) California Law Review 671, 680 f.

platform environment tailored to users feeds interactions of a user with this environment back into the system in the form of so-called feedback loops.²⁸⁹

- 105 Due to the possibilities of individualizing procedures,²⁹⁰ discrimination takes place in all described procedural stages and thus also at the level of the decision itself. The problem of algorithmic discrimination is intensified by the lack of transparency of internal company decisions. In the case of AI-based moderation, this is further perpetuated by the opacity of the AI system itself ('black box problem').²⁹¹

1.2.7 Information Gaps and Insufficient Stakes

- 106 Given the opaque nature of AI systems and a lack of mandatory publication requirements pertaining to impact-assessment, users often lack insight into the way an AI-based content moderation system operates, but also in terms of whether a specific removal or other enforcement measure has affected them.²⁹²
- 107 Further, in cases of legitimate speech being erroneously removed or blocked, the perceived harm might seem too small for a speaker to act upon and challenge the enforcement decision. In doing so, a platform operator does not solely cause harm to the individual user concerned but equally to the public which might be deprived of parts of the public discourse and of access to information.²⁹³

1.3 Procedural Answers

1.3.1 The Need for (Complementary) Procedural Resolution Schemes

1.3.1.1 Initial Situation: Proceduralization of Private Enforcement

- 108 Internet platforms appear to be natural enforcers of rights: Their technical dominance over access to and use of networks ('modern public square'²⁹⁴) predestines them for the role of enforcing private rights – their own and those of others. This happens every day in a myriad of cases: In the period between April and June 2023 alone, Facebook sanctioned 8.9 million cases of alleged violations of its community standards on physical

²⁸⁹ More generally on the technical workings of feedback loops, see Yeung (n 188), (2017) 20(1) *Communication & Society* 118, 121 f.

²⁹⁰ See already above para 68-70.

²⁹¹ See already above para 95-98. Bloch-Wehba (n 3), (2020) 53(1) *Cornell International Law Journal* 41, 83; Chander (n 287), (2017) 15(6) *Michigan Law Review* 1023, 1037.

²⁹² Nahmias and Perel (n 78), (2021) 58(1) *Harvard Journal on Legislation* 145, 181 ff.

²⁹³ See M E Kaminski, 'Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability' (2019) 92(6) *S. CAL. L. REV.* 1529, 1580-1582; Nahmias and Perel (n 78), (2021) 58(1) *Harvard Journal on Legislation* 145, 182 with further references at fn 231.

²⁹⁴ See Nahmias and Perel (n 78), (2021) 58(1) *Harvard Journal on Legislation* 145, 178 with further references at para 235; Van Loo (n 100), (2021) 88(4) *University of Chicago Law Review* 829, 830, 863.

abuse and sexual exploitation of children worldwide;²⁹⁵ add enforcement actions for violations of legally conferred subjective rights or even specific norms of criminal law.²⁹⁶

- 109 The intermediaries are involved in a process of clarification and evaluation of unclear factual and legal position. The German Federal Court of Justice, which recently had to rule in two high-profile decisions²⁹⁷ on whether Facebook's deletion of suspected hate comments was lawful, even speaks of a binding *procedural law* that platforms with market power would have to establish in their terms and conditions as they are bound by fundamental rights due to their structural superiority.²⁹⁸ Under the topos of effective 'protection of fundamental rights through procedure',²⁹⁹ the Federal Court of Justice developed individual procedural requirements in order to collaterally give effect to the fundamental rights of platform users, in particular their fundamental right to freedom of expression, Art 5(1) Grundgesetz (Fundamental Law, hereinafter: GG), even in the context of private legal relationships.³⁰⁰
- 110 The EU legislator is also vigorously driving forward the proceduralization of private enforcement. He is doing so by establishing procedures for the enforcement of rights and platform standards (general terms and conditions) as well as internal complaints, thereby assigning the role of 'objective' dispute resolvers to platforms: for example, in 2019 in the DSM Directive on the enforcement of copyrights in the digital single market,³⁰¹ in the P2B Regulation³⁰² or more recently, across legal areas, in the Digital

²⁹⁵ Facebook (Meta), 'Community Standards Enforcement Report: Child Endangerment: Nudity and Physical Abuse and Child Sexual Exploitation' <https://transparency.fb.com/data/community-standards-enforcement/child-nudity-and-sexual-exploitation/facebook/> accessed 31 December 2023.

²⁹⁶ See, for example, in German law the complaint and counter-proposal procedure in §§ 3 ff NetzDG (Network Enforcement Act).

²⁹⁷ BGH (Germany), 29 July 2021, III ZR 179/20, BGHZ 230, 347 = (2021) 65(11) ZUM (Zeitschrift für Urheber- und Medienrecht) 953; BGH (Germany), 29 July 2021, III ZR 192/20, (2021) 25 (11) ZUM-RD (Zeitschrift für Urheber- und Medienrecht – Rechtsprechungsdienst) 612.

²⁹⁸ The BGH (Germany) argued for a competence of network operators to prohibit forms of 'hate speech' on the basis of their terms and conditions also on this side of punishable or right-infringing expressions of opinion, see only BGH (Germany), 29 July 2021, III ZR 192/20, (2021) 25(11) ZUM-RD (Zeitschrift für Urheber- und Medienrecht – Rechtsprechungsdienst) 612 para 91.

²⁹⁹ BGH (Germany), 29 July 2021, III ZR 192/20, (2021) 25(11) ZUM-RD (Zeitschrift für Urheber- und Medienrecht – Rechtsprechungsdienst) 612 para 96.

³⁰⁰ BGH (Germany), 29 July 2021, III ZR 192/20, (2021) 25(11) ZUM-RD (Zeitschrift für Urheber- und Medienrecht – Rechtsprechungsdienst) 612 para 66, 77, 80. In this light, the BGH does not develop the procedural rights on the basis of 'contractual types', ie in relation to the – in any case difficult to identify – legal model of usage agreements between platform and user, cf also D Holznagel, 'Nutzerrechte bei Facebook: Klärung durch den BGH und bevorstehende Irrwege des EU-Gesetzgebers' (2021) 37(11) CR (Computer und Recht) 733, 735.

³⁰¹ Directive (EU) 2019/790 of 17 April 2019 on copyright and related rights in the digital single market and amending Directives 96/9/EC and 2001/29/EC.

³⁰² On the internal complaints procedure pursuant to Art 11 Regulation (EU) 2019/1150 of 20 June 2019 on promoting fairness and transparency for business users of online intermediary services (hereinafter: P2B Regulation).

Services Act of the European Union.³⁰³ In the scope of the E-Commerce Directive,³⁰⁴ the CJEU even recognized in the *Glawischnig-Piesczek* case the possibility for Member State courts to order hosting providers such as Facebook to delete or block illegal content worldwide.³⁰⁵

- 111 This trend can most recently be observed in the regulatory strategy of U.S. law: The draft Platform Accountability and Consumer Transparency Act ('PACT Act'), for instance, also requires the establishment of a complaint system for users, processing deadlines for incoming notices, information and justification obligations after a platform sanction has been imposed, and the publication of transparency reports.³⁰⁶ Texas and Florida have already passed similarly structured laws.³⁰⁷

1.3.1.2 Emergence and Justification of Platform-Based Law Enforcement through Procedure

1.3.1.2.1 Liability Law and Basic Procedural Structures

- 112 The trigger for the emergence of procedural structures is the liability – indirect or perpetrator – of Internet platforms, for example, for copyright, trademark and personal rights infringements and for violations of the law of fair dealing by third parties.³⁰⁸ In this

³⁰³ Cf Art 16(6), 14(4), 23(3) and recitals 24 s 3, 26 s 2 of the Regulation (EU) 2022/2065 of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act; hereinafter: DSA Regulation), OJ L 277, 27 October 2022, 1-102.

³⁰⁴ Directive 2000/31/EC of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce').

³⁰⁵ CJEU, 3 October 2019, C-18/18 – *Glawischnig-Piesczek*, ECLI:EU:C:2019:821, para 48 f. The case was submitted by the Austrian Supreme Court. In casu, the issue was the assessment of defamatory statements in accordance with § 78 Austrian Copyright Act (öUrhG), § 1330 Austrian General Civil Code (öABGB). – Critically: T Hoeren, 'Sperrpflichten eines Hosting-Anbieters bei rechtswidrigen Informationen sowie wort- und sinngleichen Inhalten' (2020) LMK (Leitsätze mit Kommentierung) 425949. – In contrast, the CJEU had ruled for search engines that they are obliged to remove links from result lists only on a Europe-wide basis when assessed from a data protection perspective, see CJEU (Grand Chamber), 24 September 2019, C-507/17 – *Google (Spatial scope of delisting)*, ECLI:EU:C:2019:772, para 44 f.

³⁰⁶ Platform Accountability and Consumer Transparency Act, 116th Cong. § 5(2) (2020) ('PACT Act'); On 3 June 2023, the revised Constitution was introduced in the U.S. Senate, see Platform Accountability and Consumer Transparency Act, 118th Congress.

³⁰⁷ H.B. 20 (Tx. 2021); S.B. 7072 (Fl. 2021); see also E Douek, 'Content Moderation as Systems Thinking' (2022) 136(2) Harvard Law Review 528, 566 f.

³⁰⁸ Examples include: (i) EU law, Art 11 s 2 of the Enforcement Directive (2004/48/EC of April 29, 2004 on the enforcement of intellectual property rights) or Art 8(3) of the InfoSoc Directive (2001/29/EC of 22 May 2001 on the harmonization of certain aspects of copyright and related rights in the information society). – (ii) In German law, so-called 'Störerhaftung' applies in part; in copyright law, platforms have recently even been held liable as perpetrators: Art 17 of the DSM Directive, for example, imposes perpetrator liability ('täterschaftliche Haftung') on the platform operators concerned if certain traffic duties are violated: F Hofmann, 'Fünfzehn Thesen zur Plattformhaftung nach Art 17 DSM-RL' (2019) 121(12) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 1219. An overview of German law is provided, for example, by F Hofmann, 'Mittelbare Verantwortlichkeit im Internet' (2017) 57(8) JuS (Juristische Schulung) 713 ff.

context, verification obligations, at least outside copyright law (DSM Directive), do not arise unsolicited – § 7(2) of the German Telemedia Act,³⁰⁹ Art 14(1), 15(1) E-Commerce Directive³¹⁰ –, but in principle only as soon as the platform operator is made aware of a – regularly clear³¹¹ – infringement (so-called notice-and-takedown procedure).³¹² As a consequence, there is no general monitoring obligation:³¹³ The right holder is burdened with a kind of obligation to present and substantiate³¹⁴ the legal facts of the infringement and its factual basis in order to identify the legal position to be enforced³¹⁵

³⁰⁹ § 7(2) of the German Telemedia Act (Telemediengesetz: TMG) states: ‘Service providers within the meaning of Sections 8 to 10 are *not* obliged to monitor the information they transmit or store or to investigate circumstances that indicate illegal activity’.

³¹⁰ On permitted ‘specific monitoring obligations’ see also recital 47 Directive 2000/31/EC as well as CJEU, 3 October 2019, C-18/18 – *Glawischnig-Piesczek*, ECLI:EU:C:2019:821, para 31 ff., 34. In Austria, the provision was implemented in § 18 E-Commerce Act.

³¹¹ The situation is different with regard to liability for setting hyperlinks: In this regard, the Federal Court of Justice (Bundesgerichtshof) allows a simple reference to the infringement to suffice: BGH (Germany), 18 June 2015, I ZR 74/1 – *Liability for Hyperlink*, BGHZ 206, 103 = (2016) 69(11) NJW (Neue Juristische Wochenschrift) 804 para 27.

³¹² See, for example, Art 16(1), (3), Art 6(1) lit b) DSA Regulation. Furthermore, BGH (Germany), 17 August 2011, I ZR 57/09 – *Stiftparfüm*, BGHZ 191, 19 = (2011) 113(11) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 1038 para 21 ff, 26 (‘Störerhaftung’ for Internet auction houses); BGH (Germany), 12 July 2012, I ZR 18/11 – *Alone in the dark*, BGHZ 194, 339 = (2013) 66(11) NJW (Neue Juristische Wochenschrift) 784 para 28 (‘Störerhaftung’ of file hosting services). – An infringement of rights that gives rise to liability (and thus indicates a risk of repetition) only exists if the intermediary does not comply with a justified request for deletion. Only with this breach of duty do the costs of a notice of infringement become recoverable, see § 97a(3) s 1 German Copyright Act (UrhG): BGH (Germany), 17 August 2011, I ZR 57/09 – *Stiftparfüm*, BGHZ 191, 19 = (2011) 113(11) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 1038 para 39; F Hofmann (n 308), (2017) 57(8) JuS (Juristische Schulung) 713, 715.

³¹³ The EU legislator also exempts service providers from a general obligation to monitor or actively investigate under Art 8 of the DSA Regulation. However, Art 4(3), 5(2) and 6(4) DSA Regulation allow service providers to be required by Member State law to ‘cease or prevent infringements’ in response to judicial or administrative orders (*stay down*). In this respect, questions of demarcation also arise in the context of the DSA Regulation with regard to the prohibition of general monitoring and the duty to prevent future infringements, cf on this (with regard to the DSA draft): R Janal, ‘Haftung und Verantwortung im Entwurf des Digital Services Acts’ (2021) 29(2) ZEuP (Zeitschrift für Europäisches Privatrecht) 227, 248-252.

³¹⁴ See recently under the Digital Services Act: Art 16(2) s 1 (‘sufficiently precise and adequately substantiated notification’), s 2 lit a) (explanation of the basis for classifying information as unlawful) and s 2 lit d) (confirmation of the accuracy and completeness of the notification). – The obligation to substantiate also varies depending on the law concerned: For example, the registry for ‘.de-domains’ (DENIC) is only liable for the breach of obligations of conduct in the case of domain registrations that infringe the law if the infringement is readily apparent. For this to be the case, the German Federal Court of Justice requires that either DENIC has a legally enforceable title or that the infringement is so clear that it must impose itself on [DENIC]: BGH (Germany), 27 October 2011, I ZR 131/10 – *regierung-oberfranken.de* (2012) 65(31) NJW (Neue Juristische Wochenschrift) 2279 para 26, on § 12 BGB (German Civil Code); clear abuse.

³¹⁵ On the identification function of the notice, see F Hofmann (n 220), (2017) 61(2) ZUM (Zeitschrift für Urheber- und Medienrecht) 102, 104 f.

and to trigger an obligation on the part of the intermediary to respond.³¹⁶ The platform operators are, in turn, required to cooperate seriously in clarifying the facts.

1.3.1.2.2 Reasons for Platform-Based Law Enforcement through Procedures

113 The basic task of private enforcement of rights through proceedings is to (provisionally) reconcile legal positions that conflict with each other in an uncertain legal and factual situation – in the interest of the alleged infringer and the presumptive rightsholder alike. This individual-protective function of proceduralization is collectivized by scaling processes.³¹⁷

1.3.1.2.2.1 Structural Inferiority of the Addressee of Legal Enforcement: Empowerment of Effective Legal Protection through Procedure

114 Functionalizing Internet platforms for the enforcement of private rights decisively improves the enforcement chances of affected rightsholders, especially in procedural terms. The following aspects are crucial in this respect: Under the conditions of the Internet, state courts are often unable to enforce private rights within a reasonable period of time – even in case of provisional legal protection – or to cope with the systemic task of handling the uncountable number of cases.³¹⁸ In addition, rigorous *ex ante* due diligence and verification obligations of platforms create additional incentives for excessive enforcement of private rights (overenforcement).³¹⁹

115 Such effects are amplified by an innate structural information advantage of the online platform over its users. Information asymmetries can already have an effect in the run-up to a sanction or legal enforcement measure and thus create strategic incentives for platforms to prevent the emergence of a legal dispute or the use of legal protection from the outset. Thereby making it more difficult, if not de facto impossible, for affected users to access internal complaints procedures or judicial legal protection.³²⁰

³¹⁶ This is also accompanied by a (considerable) reduction of due diligence costs, see G Wagner, 'Haftung von Plattformen für Rechtsverletzungen (Teil 2)' (2020) 122(5) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 447, 448.

³¹⁷ For more details, see below para 155-160.

³¹⁸ This idea is also echoed by Wagner (n 316), (2020) 122(5) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 474, 455, when he formulates that effective law enforcement by state courts is 'illusory' under the conditions of the Internet. Cf from a U.S. perspective: Van Loo (n 100), (2021) 88(4) University of Chicago Law Review 829, 830, 889.

³¹⁹ In addition, if – reinforced by increasingly strict *ex ante* due diligence and auditing obligations of platforms – there is at least a de facto liability differential, because the preventive control function of liability law is less pronounced in the relationship of the intermediary to the enforcement addressee than vis-à-vis the affected holder of the right to be enforced, this in turn creates additional incentives for excessive enforcement of private rights.

³²⁰ On corresponding practices of the sharing platform Airbnb: Van Loo (n 100), (2021) 88(4) University of Chicago Law Review 829, 844 f, 860, 879.

116 In view of these initial conditions, there is a serious risk of a loss of effective legal protection. Without further procedural safeguards, platform users would be severely limited in their legal power to defend themselves effectively in and out of state courts and *structurally* inferior to the enforcement power of the rightsholder concerned. This is particularly true in view of the gatekeeper role large online intermediaries take over in participation in social life or in economic transactions – often in particularly sensitive areas such as access to housing, credit or jobs.³²¹

1.3.1.2.2.2 Fast and Effective Protection against Irreversible Damage Caused by Procedures

117 However, there are also mirror-inverted legal enforcement deficits: It is well known that the realization of rights on the Internet is particularly at risk. The reason for this is their disproportionately greater exposure as well as a faster and wider spread of infringements.³²² This is accompanied by the difficulty of preventing infringements or reversing infringements that have already occurred.³²³

1.3.1.2.2.3 Minimizing Erroneous Decisions: Procedural Law as an Instrument of Fact-Finding

118 Finally, effective legal protection is realized upstream in minimizing the risk of a factual or legal misjudgment of the enforcement decision, despite its *normative* provisional nature. In this context, it will be necessary to ask how procedural elements of factual clarification can be made usable for the platform operator to improve the decision-making basis of private legal enforcement.

1.3.1.2.2.4 Conclusion

119 Given the initial situation outlined above, especially the strong de facto absorption effect of platform-based legal enforcement vis-à-vis state legal protection, and in view of the inadequate protective and interest-balancing effects of a purely liability-based

³²¹ Cf BGH (Germany), 29 July 2021, III ZR 179/20, BGHZ 230, 347 = (2021) 65(11) ZUM (Zeitschrift für Urheber- und Medienrecht) 953 para 66; also Van Loo (n 100), (2021) 88(4) University of Chicago Law Review 829, 863 (2021).

³²² See also CJEU, 3 October 2019, C-18/18 – *Glawischnig-Piesczek*, ECLI:EU:C:2019:821, para 28, 36; also EU Commission, ‘Staff Working Document Impact Assessment, Accompanying the document Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online’ SWD (2018), 408 final, at 2.4.3 (‘Generally speaking, the longer the content is able to survive online, the more views it may receive, and the more harm it may cause’).

³²³ As a result, especially in the case of personality rights, there is a risk of deepening or even irreversible damage; Wagner (n 316), (2020) 122(5) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 447, 455, speaks of ‘Schadensvertiefung durch Zeitablauf’.

integration of platforms, it seems necessary to integrate the latter into effective legal protection through procedural structures and proceedings.

- 120 Considering the dangers of algorithmic platform behavior already described, this finding also applies to the enforcement of the platform's own community standards. In the following, it will be shown whether and to what extent procedural structures go beyond guaranteeing legal protection for the individual and for a large number of individual affected parties.

1.3.2 *Shaping Effective Legal Protection through Procedures: Potential and Limits of a Proceduralization of Platform-Based Legal Enforcement*

1.3.2.1 Functions of Procedural Structures

- 121 The core task of corresponding procedural structures is to react quickly to particular risk situations, to create trustworthy (transparent) and error-minimizing decision-making conditions for this purpose, and to ensure the provisional nature (and correctability) of enforcement measures. In addition, there is a need for procedural structures capable of handling the mass of platform users – as potential violators of rights and terms and conditions as well as those affected by such violations.
- 122 In this orientation, proceduralization can strengthen the legitimacy of private law enforcement overall.³²⁴ Corresponding gains in trust, in the form of strengthened user loyalty, often also have an economic impact in favor of platforms.³²⁵

1.3.2.2 Individual and Collective Law Dimensions of Effective Legal Protection

- 123 In view of the dangers (and potential) of algorithmic enforcement of rights and terms and conditions, the question arises as to which procedural instruments and guarantees can be used to adequately implement the basic tasks assigned to a platform procedure. Normative reference points from state civil proceedings are the requirements of effective legal protection and transparent procedures, which in principle can be applied to online platforms.
- 124 In its individual-protection dimension, the right to effective legal protection should apply regardless of whether platforms implement state-imposed rights of third parties or their own community standards. Thus, the Digital Services Act also recently requires platforms

³²⁴ Cf generally on the contribution of procedural structures to the increased legitimacy of the results produced by be: T R Tyler, *Why people obey the law* (Princeton University Press 2006) 5, 9; further N Luhmann, *Legitimation durch Verfahren* (11th edn, Suhrkamp 2019) 55 ff.

³²⁵ Van Loo (n 100), (2021) 88(4) University of Chicago Law Review 829, 865; see also C Rule, 'Quantifying the Economic Benefits of Effective Redress: Large ECommerce Data Sets and the Cost-Benefit Case for Investing in Dispute Resolution' (2012) 34(4) U ARK LITTLE ROCK L REV 767, 776, finding that buyers who reached amicable dispute resolutions were more likely to return than buyers who simply achieved a full refund in their dispute.

to take into account the fundamental rights guaranteed by the EU Charter in the context of both forms of enforcement, including the right to an effective legal remedy.³²⁶ As will be explained in more detail below, this requirement for legal protection should be shaped by various guarantees related to the platform. Particularly by effective access to complaints procedures (both internal and external and open to decision-making) and to a human decision-maker,³²⁷ furthermore by a guarantee of a fair hearing and procedural equality of arms and opportunities³²⁸ as well as by protection of the rights of the defense.³²⁹ Moreover, enforcement addressees must be protected against unjustified reports and the improper use of platform-based enforcement actions.³³⁰ In this respect, there are functional and structural parallels to state interim legal protection.³³¹

- 125 The multitude of potentially parallel, similar or even identical situations of impairment in relation to legitimate interests and rights of users and other affected persons, which are regularly caused by platform structures, also requires effective legal protection to be granted in a supra-individual, ie ‘collective’ dimension. In this context, procedural instruments and standards must be developed which, on the one hand, enable rapid and effective legal protection for a large number of affected persons and, on the other, provide procedural safeguards against the regulatory actions of platforms – for example, in the form of the scaling of interpretation standards or decisions.³³²
- 126 Under EU law, elements of private self-monitoring and third-party control accompany the procedural monitoring mechanisms – in the form of company compliance departments to be set up³³³ or an annual independent audit.³³⁴ In addition, certain

³²⁶ Thus, for the notification and redress procedure vis-à-vis ‘all data subjects’: recital 52 s 1, 2 DSA Regulation; for general terms and conditions of providers of intermediary services vis-à-vis users: Art 14(4) DSA Regulation.

³²⁷ For more details, see below para 150-151.

³²⁸ On this point, H-J Blanke in C Callies and M Ruffert (ed), *EUV/AEUV* (6th edn, Beck 2022) Art 47 EU-GRCh para 15 f.

³²⁹ Cf on this in extrajudicial dispute resolution: B Hess, ‘Prozessuale Mindestgarantien in der Verbraucherschlichtung’ (2015) 70(11) JZ (Juristenzeitung) 548 ff.

³³⁰ For more details, see below para 129-132.

³³¹ Cf, for example, the instruments of provisional legal protection under the German Code of Civil Procedure (§§ 916 ff ZPO), such as measures to protect against unjustified claims and an abusive use of state legal protection, for example by means of prima facie evidence, §§ 920(2), 294 ZPO. Furthermore, the provision of security by the claimant (§ 921 s 2 ZPO); effective downstream legal protection (order to bring an action, § 926 ZPO; legal remedies, §§ 924, 927 ZPO) as well as protection in the event of unjustified recourse to interim legal protection (for example, through compensatory damages under § 945 ZPO).

³³² For more details, see below para 155-160.

³³³ Art 41(3) lit e) DSA Regulation with regard to providers of very large online platforms – On the classification of Amazon Services Europe Sàrl as a ‘very large online platform’ within the meaning of the Digital Services Act, see most recently the Order of the European General Court, 27 September 2023, T-367/23 – *Amazon Services Europe v Commission*, ECLI:EU:T:2023:589.

³³⁴ Art 37(1) lit a) DSA Regulation in relation to providers of very large online platforms

violations of the provisions of the Digital Services Act can be sanctioned with fines.³³⁵ This results in an overall multilayered regulatory arrangement.³³⁶

1.3.2.3 Intra-Company Legal Protection Proceedings: Basic Structures and Procedural Guarantees

127 In light of this, it is firstly important to provide the enforcement addressee with an effective, ie easily and for all affected parties accessible (ie transparent and user-friendly³³⁷), incidentally free of charge and fast³³⁸ legal protection procedure.³³⁹ This procedure, to be set up by the platform itself and designed for a large number of complaints, is intended to enable the enforcement addressee and affected right holders to effectively assert their own rights that are presumably affected by the enforcement of rights (enabling function of procedures). Governmental (or supranational) regulation has prescribed the establishment of corresponding procedures in various forms over the course of the past few years: for example, in the European Union through the Digital Services Act – particularly in its Art 20³⁴⁰ – or the P2B Regulation,³⁴¹ but also in the USA³⁴² and in Germany³⁴³.

³³⁵ See on the sanctions regime of the Digital Services Act: Art 52(1), Art 74(1) lit a).

³³⁶ On the whole, T Mast, 'AGB-Recht als Regulierungsrecht' (2023) 78(7) JZ (Juristenzeitung) 287, 291.

³³⁷ These are the programmatic functional descriptions of Art 16(1) and 20(4) of the DSA Regulation. – For the principle of transparency in out-of-court dispute resolution, see Art 7 Directive 2013/11/EU of the European Parliament and of the Council of 21 May 2013 on alternative dispute resolution for consumer disputes and amending Regulation (EC) No 2006/2004 and Directive 2009/22/EC (Directive on consumer ADR), OJ L 165, 18 June 2013, 63.

³³⁸ For example, the deadline for a decision under the German UrhDaG is one week after the complaint has been filed, § 14(3) no 3 UrhDaG. – The German NetzDG establishes a graduated time limit model that is essentially based on the complexity of the decision: In the case of complaints involving 'obviously illegal content', blocking must take place within 24 hours; in the case of merely 'illegal content', it must generally occur within seven days, § 3(2) NetzDG. The subsequently introduced counternotification procedure within the meaning of § 3b NetzDG does not stipulate a time limit requirement, nor does the Digital Services Act, see Art 20(4): 'zeitnah'/'timely'; cf also G Spindler, 'Der Vorschlag für ein neues Haftungsregime für Internetprovider – der EU-Digital Services Act (Teil 1)' (2021) 123(4) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 545, 553. – Pursuant to § 89b(5) no 5 Austrian UrhG (Copyright Act), complaints must generally be concluded within two weeks.

³³⁹ Cf the German § 14(1) UrhDaG (internal complaints procedure): 'The service provider must provide users and rightsholders with an effective, free of charge and expeditious complaints procedure about blocking and about the communication to the public of protected works' Similarly, the internal complaints procedure for commercial users under Art 11 P2B Regulation; also §§ 3-3b NetzDG. – Pursuant to § 3(1) of the Austrian Communications Platforms Act (KoPiG), service providers must establish an 'effective and transparent procedure for dealing with and settling reports of allegedly illegal content available on the communications platform'. Apart from this, § 3(4) of the KoPiG stipulates an 'effective and transparent' counternotification procedure.

³⁴⁰ See also Art 14(1), Art 16(1) DSA Regulation.

³⁴¹ See Art 11 P2B Regulation.

³⁴² Texas: H.B. 20 (Tx. 2021); Florida: S.B. 7072, Subchapter C (Fl. 2021).

³⁴³ §§ 14, 15 UrhDaG; §§ 3-3b NetzDG.

128 In the following, individual procedural guarantees for the implementation of effective in-house legal protection are discussed in more detail below – detached from any previous affirmative action.

1.3.2.3.1 Online Platforms and Attributions of their Procedural Function

129 Various procedural models are available to reduce erroneous enforcement decisions and ensure the realization of rights in endangered situations. They vary, depending on the role or corresponding obligations that can be attributed to online platforms (and correspondingly to users and rightsholders or whistleblowers) in clarifying the facts relevant to enforcement.

130 According to the concept of ‘clarification responsibility’³⁴⁴, platforms act as de facto dispute mediators (*Streitmittler*). Hereafter, a host provider should be regularly obligated to forward the complaint of the affected rightsholder to the responsible party for comment before deleting an allegedly infringing blog entry. If a substantiated counterstatement raises justified doubts about an infringement, the affected party must in turn be given the opportunity to respond and, if necessary, be ordered to submit further required evidence.³⁴⁵ A corresponding ‘shuttle procedure’ makes the parties in the primary infringement relationship more responsible for clarifying the facts³⁴⁶ – such as the counternotification procedure under § 3b NetzDG (Network Enforcement Act).³⁴⁷

³⁴⁴ German: *Aufklärungsverantwortung*.

³⁴⁵ In that regard, the so-called ‘shuttle procedure’ of the VI Civil Senate of the German Federal Court of Justice: BGH (Germany), 25 October 2011, VI ZR 93/10 – *Blog Eintrag*, BGHZ 191, 219 = (2012) 65(3) NJW (Neue Juristische Wochenschrift) 148 para 27; also: BGH (Germany), 1 March 2016, VI ZR 34/15 – *Ärztewertungsportal III (jameda.de)*, BGHZ 209, 139 = (2016) 69(29) NJW (Neue Juristische Wochenschrift) 2106 para 21 ff, 37 ff, 41 ff.

³⁴⁶ As a consequence, failure to respond on the part of the parties to the dispute leads to disadvantages in enforcing the law: If, for example, a blogger remains silent in response to a request for comments, this upholds the prior deletion of his (presumably incriminated) blog entry; conversely, a corresponding failure to act on the part of the rightsholder precludes a (final) deletion of this content: BGH (Germany), 25 October 2011, VI ZR 93/10 – *Blog-Eintrag*, BGHZ 191, 219 = (2012) 65(3) NJW (Neue Juristische Wochenschrift) 148, para 27; also: BGH (Germany), 1 March 2016, VI ZR 34/15 – *Ärztewertungsportal III (jameda.de)*, BGHZ 209, 139 = (2016) 69(29) NJW (Neue Juristische Wochenschrift) 2106 para 21 ff, 37 ff, 41 ff. – Cf in German law furthermore the counternotification procedure under § 3b NetzDG.

³⁴⁷ After a decision has been made pursuant to § 3 NetzDG, the complainant or, in a mirror image, the user must file an application for review of the decision pursuant to § 3b of the NetzDG. This application must be substantiated (without making ‘too high demands’ on this, see Bundestags-Drucksache 19/18972, 47). In the event of a planned remedy on the part of the platform, the other party is, in turn, granted an opportunity for a countermotion, on this also F Hofmann and L Specht-Riemenschneider, ‘Verantwortung von Online-Plattformen (Responsibility of Online Platforms)’ (2021) 13(1) ZGE (Zeitschrift für geistiges Eigentum) 48, 99, who refer to these mechanisms as ‘notice-and-negotiation procedures’ and assume that §§ 3, 3b NetzDG map the procedure of blog entry jurisdiction. – For the most part, the affected third party is not involved in the decision under § 3 NetzDG. The possibility of involving the user does exist in the context of the review of ‘illegal’ content pursuant to § 3(2) no 3 lit a) NetzDG. However, according to the wording, this review is merely optional and is not actually carried out by the platforms, see M Eifert et al, *Netzwerkdurchsetzungsgesetz in der Bewährung* (Nomos 2020) 80 f. – In contrast to the procedure under § 3(4) Austrian KoPiG, the counternotification procedure pursuant to § 3b KoPiG is much more comprehensive.

At the same time, the moderate model of the Federal Court of Justice borrows from a premise of the principle of submission in civil proceedings (*Beibringungsgrundsatz*), namely that the opposing interests of the parties grant the correctness of congruent statements of fact to a greater extent.

- 131 Nevertheless, this model reveals striking weaknesses: On the one hand, it leads to an expansion of the factual material relevant to enforcement through regularly disputed party submissions. This makes the clarification of the facts more complex and costlier, without online platforms being provided with the clarification tools of state civil proceedings or judges,³⁴⁸ in order to reliably clear up any ambiguities that often remain even after the parties to the dispute have been heard. Furthermore, the shuttle model – which is time-consuming in individual cases – is not sufficiently tailored to the urgent enforcement needs of affected, known and, above all, unknown rightsholders, nor to a large number of similar enforcement situations.³⁴⁹ Above all, however, it forces – especially in the case of the highly context-dependent facts of an infringement of personality rights, copyright or trademark rights³⁵⁰ – a weighing of conflicting legal positions in individual cases that is both time-consuming and prone to error, and thus leaves the forecasting risk for the actual existence of ‘clear’ infringements with the platform operator.³⁵¹
- 132 These disadvantages could be partially countered by further reducing or standardizing the inspection obligations of platforms – (also) in the interest of their increased predictability and fulfillment (‘model of procedural responsibility for action’³⁵²): On the one hand, with regard to the substantive standard of review, in the form of simplified decision parameters, for example concretized by legal presumptions, or a typification of illegal user behavior.³⁵³ On the other hand, platforms should generally be allowed to rely

³⁴⁸ In German law, for example, §§ 138(3), 142, 144 ZPO (Code of civil procedure).

³⁴⁹ Similarly as with the state regime on provisional legal protection the objective of optimized clarification of facts recedes with special platform-typical endangerment situations behind the request to protect a realization of the endangered right effectively, by enforcing platforms this immediately, at least on basis of a smaller measure of conviction, provisionally.

³⁵⁰ Keller (n 258), (2018) Hoover Inst. Aegis Paper Series no 1807 6, 7, speaks of ‘context-blindness’; cf further Bloch-Wehba (n 3), (2020) 53(1) Cornell International Law Journal 41, 65 with reference to the difficulty of detecting ‘hate speech’ offenses using fingerprinting technology.

³⁵¹ This is evident, for example, in the German NetzDG proceedings: the figures for ‘obviously unlawful’ and ‘unlawful’ content differ greatly depending on the platform. While the rate of ‘obviously illegal’ content was 95% for one provider, it was 8,69% (!) for another, see Eifert et al (n 347), (2020) 69 f. To make matters worse, the examination in the NetzDG counternotification proceedings is subsequently not limited to the reasons stated in the counternotification, but rather takes place ‘under all legal aspects that come into consideration’, see Bundestags-Drucksache 19/18972, 47.

³⁵² In German: *Modell der prozeduralen Handlungsverantwortung*.

³⁵³ Cf under German law: § 9(2) UrhDaG, which for clearly defined facts of user-generated content – eg, content that contains less than half of a work of a third party or several works of third parties (no 1) or that is marked as legally permitted pursuant to § 11 UrhDaG (no 3) – rebuttably presumes that its use is legally permitted according to § 5 UrhDaG (so-called presumed permitted uses). – In such a case, the

on the substantiated factual submissions of the parties to the dispute when examining whether there has been a clear violation of the law, and this, in parallel to the state's regime on provisional legal protection,³⁵⁴ on the basis of a lower degree of conviction (predominant).³⁵⁵ In urgent cases, it is also recommended that platform operators establish clear procedural obligations to act and react, which the affected rightsholder can trigger by means of a – non-automated – notification (by independent actors, so-called trusted flaggers³⁵⁶) of special risk situations.³⁵⁷

1.3.2.3.2 Procedural Obligations under the Notification and Redress Procedure

- 133 In the context of the notification and redress procedure, the Digital Services Act assumes a fundamentally increased responsibility for clarification on the part of the reporter or rightsholder:³⁵⁸ According to this, the illegality of the content must be 'sufficiently precise and appropriately substantiated.'³⁵⁹ In addition to unambiguous information on the electronic location (URL address) where a presumably illegal content is stored,³⁶⁰ the reporter basically owes the disclosure of his identity (name and e-mail address).³⁶¹ Furthermore, he must confirm his good faith with regard to the accuracy and completeness of his information.³⁶²

service provider shall immediately inform the rightsholder of the public communication and of the right to file a complaint pursuant to § 14 UrhDaG in order to have the presumption reviewed pursuant to § 9(2) UrhDaG.

³⁵⁴ Cf in German law §§ 920(2), 936, 294 ZPO (Code of civil procedure).

³⁵⁵ The Digital Services Act has recently moved in this direction in Art 16(3), as a liability standard for content-related inspection obligations: 'Notices referred to in this Article shall be considered to give rise to actual knowledge or awareness for the purposes of Article 6 in respect of the specific item of information concerned where they allow a diligent provider of hosting services to identify the illegality of the relevant activity or information without a detailed legal examination'.

³⁵⁶ For more details, see below para 141-143.

³⁵⁷ Recently, the German legislator has formulated a comparable obligation to respond to internal complaint procedures of the platforms (service providers) in § 14(4) UrhDaG, thereby implementing the DSM Directive on copyrights in the digital single market, although only for so-called trustworthy rightsholders – The trustworthiness of the rightsholder is to be assessed by the service provider and can result, for example, from the scope of the valuable repertoire deposited with the service provider, the associated deployment of particularly qualified personnel or also from the successful completion of quite a few complaint procedures in the past. In the event of disputes about the trustworthiness of a rightsholder (and thus at the same time about access to the 'red button'), this question can also be clarified in court, see Bundestag-Drucksache 19/27426, 144. – A repeated, clearly incorrect blocking request under § 14(4) UrhDaG leads to exclusion from the procedure under § 18(3) UrhDaG.

³⁵⁸ See Art 16 DSA Regulation. The provision applies only to illegal content, not to community standards.

³⁵⁹ Art 16(2) s 1, 2 lit a) DSA Regulation.

³⁶⁰ Art 16(2) s 2 lit b) DSA Regulation.

³⁶¹ Art 16(2) s 2 lit c) DSA Regulation.

³⁶² Art 16(2) s 2 lit d) DSA Regulation. The DSA thus waives the requirement for an affidavit. – Pursuant to Art 6(1) lit b) of the DSA Regulation, such notifications trigger an obligation on the part of the service provider to act promptly after becoming aware of them in order to block access to the illegal content or to remove it. If the providers duly comply with this obligation, they are no longer liable for the stored information.

1.3.2.3.3 Transparency Obligations

- 134 The enforcement of state rights and private standards through platforms turns out to be opaque, complex and dynamically open for development in its prerequisites and implementation. Therefore, it structurally disadvantages the addressees of enforcement. This requires the establishment of transparent platform procedures. Corresponding transparency obligations of the platforms manifest themselves, for example, by means of periodic transparency reports³⁶³ or a right of access for research purposes.³⁶⁴ In essence, they are intended to ensure preventive public scrutiny of platforms – whether by public authorities (to protect competition and consumer interests, for example) or by academia, the media and non-governmental organizations. In addition, clear provisions in the general terms and conditions of the online platforms on the sanctions available should make them predictable for the individual platform user. The factually limited perception of general terms and conditions by users is a general problem that has also been diagnosed and controversially discussed in other fields (such as consumer protection law³⁶⁵),³⁶⁶ and to which the Digital Services Act has already responded in a regulatory way.³⁶⁷ This finding does not, however, fundamentally call into question the usefulness of transparency obligations for platforms.³⁶⁸ One

³⁶³ Art 15 of the DSA Regulation recently introduced transparency reporting obligations for providers of moderation services. The provision obliges the latter to make publicly available at least once a year, in a machine-readable format and in an easily accessible manner, clear, easily understandable reports on the content moderation they have carried out during the period in question (para 1, sentence 1). – The transparency reporting obligations under Art 24, 42 DSA Regulation also apply to providers of online platforms – § 2 of the German NetzDG also stipulates a corresponding reporting obligation. – Cf also the voluntarily prepared transparency reports of major online platforms such as: Meta, ‘Community Standards Enforcement Report’ <https://transparency.fb.com/reports/community-standards-enforcement/> accessed 31 December 2023 or: Google, ‘Transparency Report’ <https://transparencyreport.google.com/> accessed 31 December 2023.

³⁶⁴ Art 40(4) DSA Regulation vis-à-vis providers of very large online platforms.

³⁶⁵ Criticism of the ‘information overload’, the confidence-building effect of such information obligations, and the high information costs of the platforms caused by this: N Maaß, ‘Sorgfaltspflichten der Anbieter von Vermittlungsdiensten’ in Kraul (ed), *Das neue Recht der digitalen Dienste: Digital Services Act* (Nomos 2023) § 4 para 51-53.

³⁶⁶ By compensating for disturbed contractual parity between platforms and their innumerable customers, the law of general terms and conditions aims to prevent incentives on the part of the clause user to exploit the clause opponents’ lack of motivation and capacity to examine the general terms and conditions in detail, cf Mast (n 336), (2023) 78(7) JZ (Juristenzeitung) 287 with further evidence.

³⁶⁷ See Art 14(5) DSA Regulation with its obligation of providers of very large online platforms and of very large online search engines to provide recipients of services with a machine-readable summary of the terms and conditions, including the available remedies and redress mechanisms, in clear and unambiguous language, aimed at reducing complexity.

³⁶⁸ Questioning the usefulness of transparency obligations: Gielen and Uphues (n 84), (2021) 32(14) EuZW (Europäische Zeitschrift für Wirtschaftsrecht) 627, 636. – On the feasibility of transparency obligations: J Drexler, ‘Bedrohung der Meinungsvielfalt durch Algorithmen’ (2017) 61(7) ZUM (Zeitschrift für Urheber- und Medienrecht) 529, 542; cf also K-N Peifer, ‘Die neuen Transparenzregeln im UWG (Bewertungen, Rankings und Influencer)’ (2021) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 1453, 1454; F Hofmann, ‘Die neuen Transparenzvorgaben im UWG 2022 im Kontext lauterkeitsrechtlicher Plattformregulierung’ (2022) 124(11) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 780, 785.

reason being the public control function of such obligations, but also the fact that algorithmic – and in this respect dynamic – decision-making (especially through the use of artificial intelligence in the form of machine learning and deep learning) is practically unpredictable for users.³⁶⁹

- 135 In this light, online platforms should report on the course of the platform procedures themselves. By way of example, information should be provided in advance as to whether automated systems are used,³⁷⁰ in particular whether or in which cases algorithmic decision-making³⁷¹ and upstream content monitoring takes place. The obligation to notify should also extend to the type of sanctioned violations (platform's own standards and/or state law) and to what extent they are actually punished within a certain reporting period.³⁷² Platforms should further provide information on the performance of the algorithms used for decision-making, including their (assessment) evaluation premises.³⁷³ Moreover, whether individual users or user groups are treated differently from the outset, for example with regard to qualitatively graduated access to a platform's internal complaints procedure. The reporting obligation should also include whether and on what basis a platform has taken measures at the instigation of a public authority,³⁷⁴ how the internal complaints procedure is designed³⁷⁵ and how it has been used in detail, especially with regard to the basis and number of complaints as well as the type, number and (median) duration of the platform decisions issued in this regard.³⁷⁶ Finally, a platform's decisions on the removal and blocking of access or downgrading of content deemed to be in violation of the law and standards – whether or not such decisions are made – should be stored in a publicly accessible database.³⁷⁷

³⁶⁹ See already above para 95-98.

³⁷⁰ Art 5 P2B Regulation obliges providers of online intermediary services to present in their general terms and conditions the main parameters determining the ranking and the reasons for the relative weighting of these main parameters compared to other parameters.

³⁷¹ According to Art 14(1) s 2, recital 45 s 2 of the DSA Regulation, providers of intermediary services shall include in their general terms and conditions, among other things, information on all guidelines (including permissible content and the manner in which it is presented; requirements for restrictions), procedures (for the restriction of content as well as for internal complaint management), measures and their and tools used for content moderation (such as blocking of access, restriction of visibility or demotion or removal of content), including algorithmic decision-making. Art 27 DSA Regulation requires the disclosure of the parameters of recommendation systems within the meaning of Art 3 lit s) of the Regulation.

³⁷² Cf, inter alia, Art 15(1) lit b) and c) DSA Regulation.

³⁷³ For example, Art 27, Art 3 lit s) DSA Regulation stipulates an obligation to specify the parameters regarding recommendation systems of online platforms.

³⁷⁴ Cf Art 15(1) lit a) DSA Regulation.

³⁷⁵ See Art 14(1) DSA Regulation; Art 11(3), (4) P2B Regulation; Art 11(1) TCO Regulation.

³⁷⁶ Details: Art 15(1) lit d) DSA Regulation.

³⁷⁷ Cf under the Digital Services Act: Art 24(5) (although not vis-à-vis decisions to the detriment of rightsholders) as in Germany § 3(2) no 4 NetzDG with a too short storage period of only 10 weeks.

136 Regarding the AI Act of the European Union,³⁷⁸ which will apply from 2 August 2026, its Art 13 obliges transparency and the provision of information. This obligation is not aimed at informing platform users whose legal sphere may be restricted by AI decisions. Rather, the transparency obligation benefits persons who use a high-risk AI system under their own responsibility (so-called deployers³⁷⁹). For them, the operation of a high-risk AI system should be sufficiently transparent so that the outputs of the system can be interpreted and used appropriately.³⁸⁰

1.3.2.3.4 Information and Justification Obligations with Regard to Individual Moderation or Enforcement Decisions

137 In contrast, information and justification obligations of platforms relate to individual moderation and enforcement decisions. They are primarily intended to improve the reviewability of such decisions by way of appeal, and also to counteract hidden sanctions by the platforms such as shadow banning or downranking. The aim of information and justification obligations is thus to ensure effective protection of individual rights, which is regularly done downstream.³⁸¹

138 As a consequence, enforcement addressees should have a claim against online platforms to disclose the respective, possibly user-individualized basis for decision-making (such as information pertaining to a user's credit score and the data used to calculate it) – as well as any changes to them.³⁸² This includes disclosing whether automated means were used to process the reports and to make decisions.³⁸³ Above all, platform operators must disclose the function or cause in which they act on a case-by-case basis in individual cases or the reasons for which they act, when enforcing its own standards or third-party rights granted by the state.

³⁷⁸ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act, hereinafter: AI Act). Pursuant to Art 113(2) AI Act, the Regulation applies mainly from 2 August 2026. The AI Act was published on 12 July 2024 and enters into force on the twentieth day following its publication in accordance with Art 113(1).

³⁷⁹ Art 3(4) AI Act.

³⁸⁰ Art 13(1) s 1 AI Act.

³⁸¹ Compare this with regard to the DSA Regulation: B Raue in F Hofmann and B Raue (ed), *Digital Services Act* (Nomos 2023) Art 14 para 3.

³⁸² A corresponding updating obligation provides recital 45 s 2 DSA Regulation. – For the whole, see P Leerssen, 'An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation' (2023) 48 Computer Law & Security Review 105790, 6.

³⁸³ Cf Art 16(6) vis-à-vis the whistleblower ('provide information'); Art 17(3) lit c) DSA Regulation vis-à-vis the users concerned (but only 'where applicable').

- 139 Furthermore, online platforms must promptly inform affected parties – users and rightsholders³⁸⁴ – of enforcement measures taken and provide clear, sufficient and case-specific justification for these measures³⁸⁵ in a way that ensures effective defense of rights.³⁸⁶ The obligation to provide reasons should apply regardless of whether a platform enforces rights granted by the state to third parties or its own standards.³⁸⁷ This obligation should be backed up by the right of data subjects to contest incomplete or otherwise inaccurate information.³⁸⁸ For reasons of effective legal protection, platforms should also have to justify to whistleblowers why they have not enforced the rights of third parties as well as general terms and conditions that protect third parties and have therefore not complied with a report on the matter.³⁸⁹

³⁸⁴ It should be criticized that the Digital Services Act in Art 17(1) lit a) does not stipulate an obligation to give reasons for decisions to affected rightsholders – nor to store them in a publicly accessible database pursuant to Art 24(5) – and thus treats decisions to the detriment of the right of personality differently from those to the detriment of freedom of expression. Similarly: K-H Ladeur, ‘Schutz vor Verletzung von Persönlichkeitsrechten und “Desinformation” in sozialen Medien unter Bedingungen der politischen Polarisierung’ in *Verfassungsblog* <https://verfassungsblog.de/personlichkeitsrecht-soziale-medien/> accessed 31 December 2023.

³⁸⁵ For details, see Art 17(1), (3) DSA Regulation; Art 4(1), (2), (5) P2B Regulation. See further already BGH (Germany), 29 July 2021, III ZR 179/20, (2021) 74(43) NJW (Neue Juristische Wochenschrift) 3179 para 88 f; BGH (Germany), 29 July 2021, III ZR 192/20, (2021) 25(11) ZUM-RD (Zeitschrift für Urheber- und Medienrecht – Rechtsprechungsdienst) 612 para 97 f. Compare to the whole: Nahmias and Perel (n 78), (2021) 58(1) Harvard Journal on Legislation 145, 167 ff; Van Loo (n 100), (2021) 88(4) University of Chicago LRev 829, 841. – See also the right to access meaningful information according to Art 13-25 GDPR: B Casey et al, ‘Rethinking Explainable Machines: The GDPR’s “Right to Explanation” Debate and the Rise of Algorithmic Audits in Enterprise’ (2019) 34(1) BERKELEY TECH LJ 145, 158-162.

³⁸⁶ At the level of the initial decision of a platform, in view of its large number (and corresponding follow-up costs), this does not, a priori, exclude the use of not comprehensively individualized forms of justification.

³⁸⁷ According to its clear wording, this is also the case in Art 17(1) DSA Regulation.

³⁸⁸ D K Citron and F Pasquale, ‘The Scored Society: due process for automated predictions’ (2014) 89(1) WASH L REV 1, 20; Nahmias and Perel (n 78), (2021) 58(1) Harvard Journal on Legislation 145, 167 f; cf also Art 22(3) GDPR.

³⁸⁹ It should therefore be criticized that Art 16(5) DSA Regulation only requires notification of ‘the decision’ and the legal remedies available to the whistleblower, whereas not a statement of the reasons for the decision, which would enable the whistleblower to effectively lodge an appeal under Art 20 DSA Regulation: Principle of effective legal protection based on Art 17(3) lit b), d) and e) DSA Regulation. – In addition, information must be provided on whether a platform has changed its original decision in response to a successful complaint by the user concerned. Similarly: S Gerdemann and G Spindler, ‘Das Gesetz über digitale Dienste (Digital Services Act) (Part 2) – Die Regelungen für Online-Plattformen sowie sehr große Online-Plattformen und -Suchmaschinen’ (2023) 125(3) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 115, 116.

1.3.2.3.5 Access to the Procedure

1.3.2.3.5.1 Access to the Notification and Redress Procedure

140 The requirement of effective legal protection grants, among other things, the right of access to a procedure.³⁹⁰ Transferred to platform-based law enforcement, this means: Individuals whose rights are allegedly violated can claim to effectively participate in a notification and redress procedure to be established by online platforms.³⁹¹ This includes, for one the right to easy ('user-friendly'), but above all uniform access to such procedures.³⁹² Accordingly, online platforms should be required to decide on reported suspected infringements objectively and free of arbitrariness on a uniform basis.³⁹³ In this context, the requirement for a fundamentally rapid response to reports is condensed into an obligation to act immediately if the violation of particularly weighty legal interests (such as the life or physical integrity of a person) is plausibly reported.³⁹⁴

1.3.2.3.5.2 Trusted Whistleblowers (Trusted Flagging)

141 Reports of illegal content by trusted flaggers are treated with priority and without delay.³⁹⁵ Although Art 22 of the DSA Regulation only covers content that violates state law, the Digital Services Act does not prevent online platforms from setting up an independent system in parallel to (privileged) report content that violates a platform's standards.³⁹⁶ The lack of regulation in this area is nevertheless to be criticized, especially in light of the particularly sensitive integration of Internet Referral Units (IRU)³⁹⁷ into the platforms' privileged reporting systems.

142 In addition to a form of civil society supervision,³⁹⁸ trusted whistleblowers represent an instrument to partially compensate for the structural lack of independence of platforms in the enforcement of third-party rights³⁹⁹ and to remedy illegal content more quickly.⁴⁰⁰ According to the DSA regime, the status of trusted whistleblower is granted upon application by so-called coordinators for digital services in the whistleblower's Member State of domicile. The status must be linked to certain conditions, such as, a particular

³⁹⁰ Cf only Art 6(1) ECHR and Art 47(1), (3) EU-GRCh. In general: H-J Blanke in C Calliess and M Ruffert (ed), *EUV/AEUV* (6th edn, Beck 2022) Art 47 EU-GRCh para 9; B Hess, *Europäisches Zivilprozessrecht* (2nd edn, Walter de Gruyter 2021) para 3.67 with further references.

³⁹¹ See Art 16(1) s 1 DSA Regulation.

³⁹² See Art 16(1) s 2 DSA Regulation.

³⁹³ See Art 16(6), recital 52 s 1 DSA Regulation.

³⁹⁴ See recital 52 s 3 DSA Regulation.

³⁹⁵ See Art 22 DSA Regulation.

³⁹⁶ Recital 62, subsec. 2, s 3, 4 DSA Regulation.

³⁹⁷ For this purpose, already above para 78-89.

³⁹⁸ B Raue in F Hofmann and B Raue (ed), *Digital Services Act* (Nomos 2023) Art 22 para 2. The representation of marginalized groups in society is emphasized by N Appelman and P Leerssen, 'On "Trusted" Flaggers' (2022) 24 *Yale Journal of Law & Technology* 452, 469 f.

³⁹⁹ More on this below para 167-170.

⁴⁰⁰ On the acceleration aspect: recital 61 s 1 DSA Regulation.

expertise of the whistleblower and his or her independence from online platforms.⁴⁰¹ With regard to the independence requirement, institutions are generally more suitable for this purpose than individuals.⁴⁰² Otherwise, it is not necessary to act in a strictly altruistic manner to protect collective interests,⁴⁰³ so that the expertise of interest groups of rightsholders can also be used as prioritized whistleblowers. Public institutions such as government agencies⁴⁰⁴ can (and must) also apply for the status of trusted whistleblower under the Digital Services Act model, thus lending (additional) weight to their legal assessments.⁴⁰⁵

- 143 Reporting obligations to be published as well as obligations of the platforms to notify notoriously inaccurate or insufficiently substantiated reports ensure the privileged position of whistleblowers and a certain control over their activities.⁴⁰⁶ The legal consequences of privileged reporting on the (sanction) decision to be taken by the platform vary depending on the codification: While the Digital Services Act leaves this question open, the German Copyright Service Providers Act (UrhDaG) confers a kind of (preliminary) presumption of correctness on notifications from so-called trustworthy rightsholders: corresponding declarations by the rightsholders are capable of rebutting the presumption of legally permitted uses established by § 9(2) UrhDaG. Consequently, the service provider is obligated in accordance with § 14(4) UrhDaG to block content that significantly impairs the economic exploitation of a work on a provisional basis, ie until the internal complaint procedure has been completed. In order to increase the performance of trusted flagging in the context of content moderation, it is proposed to supplement the notice systems, which are still designed for individual cases, with notice and stay down mechanisms with regard to comparable and future content.⁴⁰⁷

⁴⁰¹ Art 22(2) DSA Regulation. The digital services coordinator is equally competent to revoke this status in accordance with paragraph 7. – For more details, see below para 144-147.

⁴⁰² In this sense, the Digital Services Act in recital 61 s 3, which denies trusted whistleblower status to individuals in principle ('should not be granted to individuals'). – On self-regulatory cooperation of platforms with trusted whistleblowers (such as YouTube's Trusted Flaggers program or TikTok Safety Partners) as well as co-regulatory models (such as flagging based on the EU Code of Conduct on Hate Speech), see in more detail Appelman and Leerssen (n 398), (2022) 24 Yale Journal of Law & Technology 452, 454 ff.

⁴⁰³ Accordingly, the Digital Services Act in its final version, notwithstanding recital 61 s 3 (exclusion of individuals). In contrast, § 14(4) of the German UrhDaG explicitly applies to individual, so-called trusted right holders.

⁴⁰⁴ Thus, recital 61 s 4 DSA Regulation mentions Europol in the field of law enforcement.

⁴⁰⁵ K Kaesling, 'Evolution statt Revolution der Plattformregulierung' (2021) 65(3) ZUM (Zeitschrift für Urheber- und Medienrecht) 177, 180; cf also B Raue and H Heesen, 'Der Digital Services Act' (2022) 75(49) NJW (Neue Juristische Wochenschrift) 3537 para 32. On the recognition requirement: B Raue in F Hofmann and B Raue (ed), *Digital Services Act* (Nomos 2023) Art 22 para 38.

⁴⁰⁶ Art 22(3) and (4) DSA Regulation.

⁴⁰⁷ To this end, as to the desideratum of feeding reference files into the platforms' AI moderation system by trusted whistleblowers (similar to the already existing PhotoDNA and Content ID reference databases): Appelman and Leerssen (n 398), (2022) 24 Yale Journal of Law & Technology 452, 473.

1.3.2.3.5.3 Excursus: The Independence of Platforms Vis-À-Vis The State

- 144 In the enforcement of rights and standards, it is also necessary for online platforms to act – in economic, financial and personnel terms – independently of both reporting persons and their users. This applies particularly to the relationship with the state, in order to prevent biased behavior in the form of privileged treatment of state enforcement interests.
- 145 However, the requirement for independence from state actors is not so much to be secured by increasing the responsibility of platforms to provide information about the state's reports and enforcement concerns.⁴⁰⁸ Rather, tighter and at the same time more appropriate limits should be placed on the instrumentalization of platforms in the execution of state interests. Since, it is important to prevent government agencies from acting beyond their legally determined sphere of competence and responsibility by means of private enforcement actors (such as platforms) and thus overstepping the (constitutional) limits of democratic legitimacy and the rule of law – often outside the scope of judicial review. According to the principle of legality (lawfulness of the administration⁴⁰⁹), the state may only intervene in the legal sphere of the citizen on the ground of a formal law or another legal norm based on this (reservation of the law);⁴¹⁰ it must use the instruments and procedures provided for by (administrative) law for its (interventional) actions and, in doing so, take into account the fundamental rights of the addressees of the enforcement (as well as competing guarantees in favor of third parties) affected in the individual case.⁴¹¹

⁴⁰⁸ However, this is the requirement of the Meta Oversight Board in Case 2022-007-IG-MR – *Drill Music*, 21 f ('It is therefore critical that Meta evaluate these requests itself and reach an *independent conclusion*. [...] Independence is crucial, and the evaluation should require specific evidence of how the content cause harm') and 23 ('While there may be good reasons to adopt a prioritization framework that ensures reports from law enforcement are assessed swiftly, that process needs to be designed to ensure that such reports include *sufficient information* to make independent assessment possible, including *seeking further input* from the requesting entity or other parties where necessary').

⁴⁰⁹ In German: 'Grundsatz der Gesetzmäßigkeit der Verwaltung'.

⁴¹⁰ In German: 'Vorbehalt des Gesetzes', see the German Constitutional Court (Bundesverfassungsgericht: BVerfG), 3 February 1959, 2 BvL 10/56, BVerfGE 9, 137 = (1959) 12(21) NJW (Neue Juristische Wochenschrift) 931: 'The principle of the rule of law requires that the administration may only intervene in the legal sphere of the individual if it is authorized to do so by law, and that this authorization is sufficiently determined and limited in terms of content, subject matter, purpose and extent, so that the interventions are measurable and to a certain extent predictable and calculable for the citizen [...]'. Further: BVerfG (Germany), 8 August 1978, 2 BvL 8/77, BVerfGE 49, 89 = (1979) 32(8) NJW 359, 360: 'The same standards are used to assess whether the legislature, as the constitutional reservation of the right to legislate further requires [...], has itself determined the essential normative foundations of the area of law to be regulated with the norm submitted for review and has not left this to the actions of, for example, the administration'.

⁴¹¹ Cf the German Constitutional Court (Bundesverfassungsgericht: BVerfG), 20 April 1982, 2 BvL 26/81, BVerfGE 60, 253 = (1982) 35(43) NJW (Neue Juristische Wochenschrift) 2425: 'Art 19(4), 20(2) s 2 and Section IX of the Grundgesetz [Constitution] prove the rule-of-law idea of binding state power to the

- 146 If follows from this: In case of alleged *illegal* action by platform users, the state is in principle only permitted to sanction this action with the means and in the forms of (police) law, for example by issuing deletion or blocking orders to the respective provider of hosting services. Beyond this legal framework, the state should only be allowed to enforce its enforcement interests to protect private rights by means of decentralized, platform-based reporting and enforcement systems (if only for the purpose of reporting alleged infringements) if such action is expressly permitted by law.⁴¹² In that regard, states are called upon to regulate the reporting process for state whistleblowers:⁴¹³ When using private enforcement agents to (provisionally) sanction rights violations, state actors should, in principle, be subject to the same formal and substantive requirements as within an exclusively state action and enforcement space. If they fail to meet these requirements, the notice and takedown mechanism – underpinned by liability law⁴¹⁴ – should not be allowed to intervene.
- 147 The situation is different when state authorities are solely concerned with using platforms to prevent the behavior of users or certain content created by them which, although presumed to violate a platform's own standards, is nevertheless permitted by law.⁴¹⁵ In this case, the state already exceeds its competence to act with a reference to lawful content directed at the platform, by interfering without a legal basis in areas of freedom that are protected by the law. Such interference should be prohibited by law. If state actors nonetheless make their enforcement concerns known through corresponding notices, platforms would have to document these and publish them – for example, within the framework of transparency reports.⁴¹⁶ Platforms enforcing their own standards in such cases should also be required to disclose any notifications from the state to the addressee and to justify the enforcement solely on the basis of a violation of the platform's own standards. In order to take into consideration the right

law with the establishment of legal protection by independent courts This commitment to the law is indispensable for an order that has placed itself under the claim of the ideas of human dignity, freedom and equality as well as social justice. Freedom requires, above all, the reliability of the legal order. For freedom means above all the possibility of shaping one's own life according to one's own life plans. An essential condition for this is that the circumstances and factors which can have a lasting influence on the possibilities of shaping such life plans and their execution, in particular the state's influence on them, can be assessed as reliably as possible'. Cf also BVerfG (Germany), 15 January 1958, 1 BvR 400/51, BVerfGE 7, 198 = (1958) 11(7) NJW (Neue Juristische Wochenschrift) 257 (on judicial observance of a third-party effect of fundamental rights in private law).

⁴¹² Cf recital 61 s 4 DSA Regulation.

⁴¹³ F Saurwein, 'Regulierung von Internet-Inhalten: Ombudsstellen als Governance-Option an der Schnittstelle von Recht und Ethik', in G Marci-Boehncke, M Rath, M Delere and H Höfer (ed), *Medien – Demokratie – Bildung* (Springer VS 2022) 47-63 https://doi.org/10.1007/978-3-658-36446-5_5 accessed 31 December 2023.

⁴¹⁴ See only Art 16(3) in connection with Art 6(1) lit b) DSA Regulation.

⁴¹⁵ Thus, the case structure in Meta Oversight Board in Case 2022-007-IG-MR – *Drill Music*.

⁴¹⁶ Cf also P Schneiders, 'Hate Speech auf Online-Plattformen: Problematization, Regulation and Evaluation against the Background of the Proposal for a Digital Services Act' (2021) 85(2) UFITA (Archiv für Medienrecht und Medienwissenschaft) 269, 303 ff <https://doi.org/10.5771/2568-9185-2021-2-269> accessed 31 December 2023.

to an effective legal remedy, the internal complaints procedure must be open to the addressee against the enforcement measure without distinction.

1.3.2.3.5.4 Uniform Access to Internal Complaints Procedures

148 Apart from that, all persons affected by measures to enforce rights and a platform's own standards (users, other rightsholders, whistleblowers) should have equal access to internal complaints procedures. At this stage of the procedure, Art 20(4) of the DSA Regulation, for example, requires procedural equality.⁴¹⁷ The principle of equality of legal protection borrowed from the states' codes of (civil) procedure also means that the order and speed of processing must be determined carefully and without discrimination, taking into account the fundamental rights and legitimate interests of the persons concerned.⁴¹⁸ Exceptions to this principle are justified in the case of particular urgency of the decision for the complainant;⁴¹⁹ depending on the individual quality of the complaints, the accuracy, scope and processing time of the complaint decision may vary.⁴²⁰

149 It is not compatible with these requirements if users are not given the opportunity to take action against a platform's decisions by way of an internal complaint within the framework of special procedures (such as Meta's escalation procedure).⁴²¹ The same applies to the lack of legal protection against the application of unwritten exceptions (so-called allowances)⁴²² or against so-called scaling decisions⁴²³.

1.3.2.3.5.5 Uniform Access to the Nature of Decision Making?

150 To be separated from this problem is the question of whether online platforms are already required at the level of legal or standard enforcement to structure the way decisions are made in the same way for all (groups of) enforcement addressees. Beyond explicit legal regulations, the fact that operators are in principle free to structure their platform according to parameters that follow their own economic interests and therefore also to shape the type of decision-making differently depending on the

⁴¹⁷ Similarly, Art 11 subpara 2 s 2 P2B Regulation for the internal complaint management system for business users. On the whole: M Berberich, '§ 5 Sorgfaltspflichten, Moderationsverfahren und prozedurale Fairness' in B Steinrötter (ed), *Europäische Plattformregulierung* (Nomos 2023) para 56; D Holznagel, 'Zu starke Nutzerrechte in Art. 17 und 18 DSA' (2022) 38(9) CR (Computer und Recht) 594, 598.

⁴¹⁸ B Raue in F Hofmann and B Raue (ed), *Digital Services Act* (Nomos 2023) Art 20 para 47.

⁴¹⁹ See recital 52 s 3 DSA Regulation, but only in the case of a platform-based enforcement of state-granted rights; an analogous application of this principle to the enforcement of platform-owned standards suggests itself when the protection of third-party interests is at issue.

⁴²⁰ Cf recital 87 s 7 DSA Regulation for moderation decisions of very large online platforms; also, B Raue in F Hofmann and B Raue (ed), *Digital Services Act* (Nomos 2023) Art 20 para 47.

⁴²¹ So also the Facebook Oversight Board, 'Art. 2 section 1 of the FOB Charter' https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf accessed 31 December 2023.

⁴²² For more details, see below para 153-154.

⁴²³ Citron and Pasquale (n 388), (2014) 89(1) WASH L REV 1, 20.

targeted user (group) speaks against such a requirement for equal treatment (eg, with regard to a human decision-maker).⁴²⁴ Irrespective of an ‘increased commitment of platforms to fundamental rights’,⁴²⁵ this manifests their entrepreneurial freedom (Art 16 EU Charter), including freedom of contract, which is guaranteed by the EU Charter of Fundamental Rights.⁴²⁶ This finding is consistent with the common practice of platforms to base initial decisions, ie the enforcement of standards or third-party rights, vis-à-vis users classified as ‘economically less valuable’, on merely algorithmic decision-making.

- 151 However, one of the limits to the admissibility of this ‘justice by algorithm’ must be seen in the need for effective legal protection. For effective protection of user rights would no longer be guaranteed if algorithmic decision-making were not subject to any human or manual review by the platform.⁴²⁷ As also provided for in the EU’s Digital Services Act,⁴²⁸ platform operators should be required to provide appropriate safeguards as part of internal review procedures.⁴²⁹ This is particularly important in view of the performance limits of algorithmic decision-making systems, which are becoming apparent particularly in the case of a highly context-dependent evaluation of legal issues.⁴³⁰ The use of human decision-makers is also likely to strengthen trust in platform-based legal and contract enforcement overall and contribute to its increased acceptance.⁴³¹

⁴²⁴ Other view: Van Loo (n 182), (2016) 33(2) *Yale Journal on Regulation* 547, 565 f; Van Loo (n 100), (2021) 88(4) *University of Chicago LRev* 829, 876 f. On the similar issue of equal access rights to the legal services market, D Simshaw, ‘Access to A.I. Justice: Avoiding an Inequitable Two-Tiered System of Legal Services’ (2022) 24 *Yale Journal of Law & Technology* 150, 183 ff.

⁴²⁵ So the understanding under German law, see the Federal Court of Justice (Bundesgerichtshof), 29 July 2021, III ZR 192/20, (2021) 25(11) *ZUM-RD* (Zeitschrift für Urheber- und Medienrecht – Rechtsprechungsdienst) 612 para 66, 77, 80; B Raue in F Hofmann and B Raue (ed), *Digital Services Act* (Nomos 2023) Art 14 para 90.

⁴²⁶ This is also explicitly pointed out by recitals 45, 52 s 2 DSA Regulation.

⁴²⁷ Thus Art 14(1) s 2 as well as recitals 45, 58 s 2 DSA Regulation. – Cf also: Van Loo (n 100), (2021) 88(4) *University of Chicago LRev* 829, 876; Lester and Pachamano (n 41), (2017) 24(1) *UCLA Entertainment Law Review* 51, 68.

⁴²⁸ Art 16(6) DSA Regulation permits the use of automated means for processing reports and decision-making in the notification and redress procedure. However, information must be provided about their use (Art 16(6) s 1 DSA Regulation). – In the case of the internal complaints procedure, on the other hand, Art 20(6) DSA Regulation obliges all providers of online platforms to ensure that relevant decisions are made under the supervision of appropriately qualified personnel and not solely by automated means – Cf Maaßen (n 365) § 4 para 71; Spindler (n 338), (2021) 123(4) *GRUR* (Gewerblicher Rechtsschutz und Urheberrecht) 545, 552.

⁴²⁹ Nevertheless, the capacity problem of platforms to maintain sufficient staff to review decisions remains.

⁴³⁰ Of course, there are limits to the context sensitivity of human decision-makers. At least with appropriately qualified personnel, however, they are likely to remain superior to algorithmic decision-making in this respect, given the current state of the art.

⁴³¹ One reason for this may be that the use of human decision-makers can prevent the presumed feeling of not being ‘at the mercy of a machine’, see also R Koulu, ‘Proceduralizing control and discretion:

1.3.2.3.6 Proceduralization of Decision Making

1.3.2.3.6.1 Requirements for the Decision Maker

- 152 To be able to evaluate particularly context-sensitive issues in a normatively correct manner, decisions on platform-based legal and contract enforcement require the deployment of technically qualified personnel with a certain range of language skills.⁴³² Beyond these key suitability requirements, platform-based decision-making structures can be designed in various ways: eg, in the form of collegial bodies,⁴³³ by involving external expertise⁴³⁴ and/or through user participation.⁴³⁵

1.3.2.3.6.2 Decision Parameters and Decision Consistency

- 153 Recently, the decision parameters for (algorithmic) enforcement of third-party rights and platform-specific terms and conditions have become the subject of state regulation. The Digital Services Act of the European Union, for instance, obliges platform operators to act diligently, objectively, non-arbitrarily and proportionately when taking enforcement measures. In particular, legitimate interests, (fundamental) rights and freedoms of all parties involved⁴³⁶ and the freedom and pluralism of the media must be taken into account.⁴³⁷ In the literature, it has also been demanded that the sanctions

Human oversight in artificial intelligence policy' (2020) 27(6) Maastricht Journal of European and Comparative Law 720, 722, 729; about the importance of explained and socially accepted decisions see also F von Ameln, 'Führen und Entscheiden unter Unsicherheit' (2021) 52(4) GIO (Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie) 567, 570.

⁴³² For corresponding transparency obligations of providers of very large online search engines, see Art 42(2) lit a) and b) DSA Regulation.

⁴³³ The Oversight Board of the Meta Group is a paradigm for this. Here, decisions are made by panels consisting of five Board members, see Art 3, in particular Sections 2, 4; Art 7.1 Charter; Art 1, Section 1.1.4.4, 3.1.3 ('Standard Cases'), Art 2, Section 2.1.2 ('Expedited Review') and 2.1.3 ('Summary Decisions') Oversight Board Bylaws.

⁴³⁴ This is the case, for example, in the Network Enforcement Act counter-proceedings (§ 3b NetzDG), but only for illegal content.

⁴³⁵ To this end: S Cooper, C Rule and L Del Duca, 'From Lex Mercatoria to Online Dispute Resolution' (2011) Penn State Legal Studies Research Paper No 09/2011, 13; also, C Busch, 'Mehr Fairness und Transparenz in der Plattformökonomie?' (2019) 121(8) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 788, 796; Mast (n 336), (2023) 78(7) JZ (Juristenzeitung) 287, 291 f; A Ohly in A Ohly and O Sosnitza (ed), *UWG Gesetz gegen den unlauteren Wettbewerb: UWG* (8th edn, Beck 2023) § 8a para 1, 2.

⁴³⁶ R Van Loo argues from a U.S. perspective that enforcement decisions of platforms should at the same time meet the standard of 'reputational accuracy and completeness', in Van Loo (n 100), (2021) 88(4) University of Chicago Law Review 829, 883 f.

⁴³⁷ See Art 14(4) DSA Regulation: 'Providers of intermediary services shall act in a diligent, objective and proportionate manner in applying and enforcing the restrictions referred to in paragraph 1, with due regard to the rights and legitimate interests of all parties involved, including the fundamental rights of the recipients of the service, such as the freedom of expression, freedom and pluralism of the media, and other fundamental rights and freedoms as enshrined in the Charter'. – Art 16(6) s 1 DSA Regulation: 'Providers of hosting services shall process any notices that they receive under the mechanisms referred to in paragraph 1 and take their decisions in respect of the information to which the notices relate, in a timely, diligent, non-arbitrary and objective manner'.

imposed by platforms being ‘perceptible’ for the user acting in violation of pertinent rules (deterrence of sanction).

- 154 In this context, further developments will show to what extent the requirements of objectivity and freedom from arbitrariness (and thus also the fairness characteristics of proportionality and equality) will induce online platforms to gear their enforcement measures more strongly to the ‘decision-making ideal’ of uniformity or consistency and, therefore, less to the (special) interests of individual users or user groups – precisely in the interest of increased predictability and comparability.⁴³⁸ Essentially a platform-based consistency and comparability of decisions is based on two premises: First, that comparable enforcement matters are subsumed under uniform abstract criteria for decision-making;⁴³⁹ second, that this is done in a uniform, permanent manner with regard to the standard of interpretation as well as the depth of review, without regard to the person – be it in the enforcement of rights set by the state or a platform’s own standards, be it in the internal review of enforcement decisions in the complaints procedure.⁴⁴⁰ A problem of lack of consistency arises, not only when one of these conditions is missing, but equally when platforms apply unwritten exceptions, ie, exceptions not specified in their terms and conditions, to forms of conduct that are in themselves sanction-relevant to specific circumstances without disclosing this to the affected users or whistleblowers. The ‘allowances’, which are exclusively used in Meta’s ‘escalation procedure’, serve as a model for this.⁴⁴¹ The Facebook Oversight Board has repeatedly criticized the lack of disclosure of their objective, content-related and temporal application requirements as a violation of the principle of legality (sic!).⁴⁴² There is also a risk of inconsistent decision results because Meta links an in-depth, context-related examination of facts by particularly expert decision-makers to specific

⁴³⁸ On this point, in more detail: Van Loo (n 100), (2021) 88(4) University of Chicago LRev 829, 871.

⁴³⁹ Cf also recital 52 s 1 DSA Regulation: ‘[...] on the basis of rules that are *uniform*, transparent and clear [...]’.

⁴⁴⁰ Cf also T R Tyler, ‘What is procedural justice? Criteria used by citizens to assess the fairness of legal procedures’ (1988) 22(1) Law & Society Review 103, 105: ‘Consistency refers to similarity of treatment and outcomes across people or time or both’.

⁴⁴¹ These are, for example, the facts of the so-called spirit of policy allowance and the newsworthiness allowance; on the latter doctrine, taken from U.S. constitutional law and applied to Facebook’s content moderation, see T Kadri and K Klonick, ‘Facebook v. Sullivan: Public Figures and Newsworthiness in Online Speech Online Speech’ (2019) 93(1) Southern California Law Review 37-99.

⁴⁴² Oversight Board, 9 March 2023, 2022-014-FB-MR – *Sri Lanka Pharmaceuticals*, Part 8.1.I, 14 (‘Secret discretionary exemptions to Meta’s policies are incompatible with the legality standard.’); further: Oversight Board, 27 September 2021, 2021-010-FB-UA – *Colombia Protests*, Part 6 (‘The Board notes that Facebook does not make its criteria for escalation publicly available.’); affirmed in: Oversight Board, 17 June 2022, 2022-001-FB-UA – *Knin Cartoon*, 4, 18; 4: ‘The fact that the content was not sent to Meta’s specialized teams for assessment before it reached the Board shows that the company’s processes for escalating content are not sufficiently clear and effective’; Oversight Board, 14 December 2022, 2022-012-IG-MR – *India Sexual Harassment Video*, under Part 8.1 and 8.3: The Board criticizes the process and, in particular, access to the escalation process as not sufficiently clear and effective.

conditions that are not disclosed to the outside world, such as the fact that government agencies act as reporters.⁴⁴³

1.3.2.3.6.3 'Scaling' of Decisions

- 155 One instrument for increasing a platform-wide formation of decision standards and decision consistency is so-called scaling. This is understood to mean the (technical) possibility for platforms to apply the concrete standards of interpretation obtained in a 'model case' and considered to be of 'high quality' to comparable circumstances.⁴⁴⁴ Scaling thus allows online platforms to quickly adapt and standardize their decision-making practice with regard to new decision-making situations that are similar in a large number of cases and comparable in this respect, not least with the aim of providing more effective protection against systematic human rights violations.⁴⁴⁵ If the legal situation is correctly assessed, this will undoubtedly promote effective legal protection.⁴⁴⁶
- 156 The disadvantages of such scaling processes are obvious: The more context-sensitive disputes are – as for example in the law of statements – the more likely it is that different cases will be treated equally without any objective reason. Distortion effects are more likely to occur if scaling is not strictly party-related but 'supra-individual': This might become relevant if, in the case of defamatory content, a platform were to remove not only those terms or images classified as violating the law or the general terms and conditions that relate to the party in the initiating proceedings, but also, in addition, corresponding content that – in a situational context that is related but not normatively equivalent – relates to other persons. This poses the risk of overenforcement, as well as the opposite effect, if exceptions are scaled excessively ('scaled allowances').⁴⁴⁷

⁴⁴³ Oversight Board, 17 June 2022, 2022-001-FB-UA – *Knin Cartoon*, Oversight Board, 17 June 2022, 2022-001-FB-UA, Part 8.1 (Meta's review process), 17.

⁴⁴⁴ Cf also the 'Media Matching Service Bank' ('escalations bank') used by the Meta Group. Content classified as 'violating' is fed into this database as part of the escalation procedure, and all subsequent identical or 'core-similar' content ('matching content') is automatically sanctioned on this basis. Affected users can file a complaint against this, see for example Oversight Board, 14 December 2022, 2022-011-IG-UA – *Video after Nigeria Church Attack*, About the Case, 2: The FOB reversed Meta's decision to remove a video from Instagram showing the situation immediately following a 5 June 2022 terrorist attack in Nigeria. The panel concluded that by restoring the post with a warning message ('disturbing content'), the privacy of the victims is protected while allowing a discussion of the events that some states may wish to prevent.

⁴⁴⁵ Cf Oversight Board, 9 January 2023, 2022-013-FB-UA – *Iran Protest Slogan*, Parts 6 and 8.1 I. b), 14.

⁴⁴⁶ The entire content moderation system structurally exhibits elements of scaling. In regular moderation, this is evident in the technically simple mechanism of hash matching, for example. Here, too, existing content classified as sanction-worthy – such as words or images – is fed into a database, the corresponding content is detected, and this decision is then replicated (for the technical mode of operation, see above para 30-38).

⁴⁴⁷ Scaled exceptions apply to entire categories of content, not just individual posts – *Sri Lanka Pharmaceuticals*, the Oversight Board (9 March 2023, 2022-014-FB-MR, Part 8.2, 14) upheld Meta's

- 157 As practice shows, scaling can also be based on the decisions of external review bodies. For example, in the case of decisions by the Facebook Oversight Board (FOB) – an external body that also takes into account external standards of interpretation such as human rights when monitoring the platform’s own general terms and conditions – the Meta Group undertakes to examine whether similar situations exist on its platform and whether individual board decisions can be applied to these cases.⁴⁴⁸ In this respect, the Meta Group has a wide scope for estimation.⁴⁴⁹ However, the more individual and context-sensitive the legal disputes are, the smaller the scaling effect of a FOB decision is likely to be.⁴⁵⁰
- 158 Scaling issues also arise at the level of law infringement. In the *Glawischnig* case, for example, the CJEU ruled that hosting providers are obligated to delete content that violates personal rights in parallel cases, provided that the information is identical in wording or – according to the principle – only identical in meaning.⁴⁵¹ In contrast to the scaling of decisions by the FOB on the interpretation of the platform’s own standards, this case law, which is aimed at effectively protecting personal rights, can be based on the authority of a state court decision. Admittedly, the duty of a platform to remove content that violates state law worldwide and thus extraterritorially, or to block access to it, collides with foreign state sovereignty and consequently with the requirement of consideration under international law in view of nationally (strongly) diverging value

decision to allow a Facebook post soliciting drug donations for Sri Lanka to be published when there was a financial crisis there. Meta concluded that under a strict interpretation of the rule, the post violated the Community Standard on Restricted Goods and Services This prohibits content asking for medications However, the FOB applied a scaled ‘within the meaning of the policy’ exception, but in doing so found that undisclosed, arbitrary policy exceptions were inconsistent with Meta’s human rights responsibilities To make ‘in the spirit of the policy’ exceptions more transparent and consistent, the Board made recommendations including that, if applied consistently, they be standardized in the relevant community standards themselves, including the criteria Meta uses to decide whether to scale the exception.

⁴⁴⁸ Art 4 FOB Charter; Art 3, Section 2.3.1 FOB Bylaws For the detailed implementation process of the Board’s decisions by dividing them into the different levels of comparability such as ‘Case Content’, ‘Identical Content with Parallel Context’, and ‘Similar Content’, see Meta, ‘Sharing More Details on How We Will Implement the Oversight Board’s Decisions, Responding to the Oversight Board’s First Decisions’, 28 January 2021 <https://about.fb.com/news/2021/01/responding-to-the-oversight-boards-first-decisions/> accessed 31 December 2023.

⁴⁴⁹ Moreover, there is a corporate reporting requirement on this issue, E Douek, ‘How Much Power Did Facebook Give Its Oversight Board?’, in Lawfare, 25 September 2019 <https://www.lawfaremedia.org/article/how-much-power-did-facebook-give-its-oversight-board> accessed 31 December 2023.

⁴⁵⁰ According to E Douek, even representatives of the Meta Group estimated the number of such identically stored contents to be very small due to the context-specificity. This assessment was made in the context of a discussion between representatives of the Meta Group and ‘stakeholders’ which Douek also attended: ‘The Oversight Board Moment You Should’ve Been Waiting For’, in Lawfare, 26 February 2021 <https://www.lawfaremedia.org/article/oversight-board-moment-you-shouldve-been-waiting-facebook-responds-first-set-decisions> accessed 31 December 2023.

⁴⁵¹ CJEU, 3 October 2019, C-18/18 – *Glawischnig-Piesczek*, ECLI:EU:C:2019:821, para 33 ff.

systems (especially in the law of expression).⁴⁵² Furthermore, if the scaling refers to decisions made by non-governmental dispute resolution bodies (or even platform courts) on the presumed infringement of legal rights, it is to be feared, detached from questions of extraterritorial extension of effect, that in a large number of cases the interpretation and guiding effect of state law beyond the judiciary's responsibility for jurisdiction would be modified, and even undermined in extreme cases.

159 In general, the dangers described highlight the need for transparency and downstream effective legal protection:⁴⁵³ Complaints bodies must be enabled to review facts on the basis of a scaling decision disclosed to the affected parties in a context- and case-specific manner. The more information-intensive the facts and the more context-sensitive the legal situation, the more reluctant platforms should be to use scaling.⁴⁵⁴ Conversely, the acceptance of scaling decisions is likely to be higher if they are based on external objective review standards and independent review bodies.

160 The platforms' reluctance to disclose scaling may be due to the fear that disclosure of scaling will lead to a regional differentiation of a platform's standard terms and conditions, which are often uniform across regions, and that this differentiation will be visible to the outside world. This is precisely the effect that is likely to accompany a corresponding disclosure obligation.

1.3.2.3.7 Excursus: AI Act and Platform Enforcement

161 One question of practical importance is whether the AI Act of the European Union,⁴⁵⁵ which will be applicable from 2 August 2026, will set limits to the AI-based enforcement of standards and third-party rights by online platforms. This is because providers and deployers of AI applications,⁴⁵⁶ which are categorised as high-risk AI systems, are subject to special obligations, including inter alia: Establishing a risk management system (Art 9

⁴⁵² On the instrument of geo-blocking see: R Achleitner, 'The Fight against Geo-Blocking – A Never Ending Story? Policy Paper on Geo-Blocking', 2 February 2021 <https://ssrn.com/abstract=4246896> or <<http://dx.doi.org/10.2139/ssrn.4246896> both accessed 31 December 2023.

⁴⁵³ See below para 167-178.

⁴⁵⁴ The guiding consideration should also be how likely scaling is to cause serious, irreversible damage.

⁴⁵⁵ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act, hereinafter: AI Act). Pursuant to Art 113(2) AI Act, the Regulation applies mainly from 2 August 2026. The AI Act was published on 12 July 2024 and enters into force on the twentieth day following its publication in accordance with Art 113(1).

⁴⁵⁶ Cf Art 2(1) AI Act. According to Art 3(3) AI Act, 'provider' means 'a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has developed an AI system or a general-purpose AI model and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge'.

'Deployer' is defined in accordance with Art 3(4) as 'a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity'.

AI Act); complying with data and data governance requirements, in particular with regard to training, validation and testing data sets (Art 10 AI Act); furthermore, obligations for technical documentation, record-keeping and transparency (Art 11-13 AI Act) and to ensure human oversight (Art 14 AI Act) as well as an appropriate level of accuracy, robustness and cybersecurity over the entire lifetime of AI systems (Art 15 AI Act). In addition, providers of high-risk AI systems must ensure that their systems fulfil the requirements of the Regulation, Art 16(1)(a) AI Act, in particular that they have a quality management system in place, Art 16(1)(b), Art 17 AI Act.

- 162 Whether AI systems are to be classified as high-risk AI systems is determined, among other things, by Art 6(2) in conjunction with Annex III AI Act. In the area of ‘administration of justice and democratic processes’, this applies to ‘AI systems intended to be used by a judicial authority or on their behalf to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts, or to be used in a similar way in alternative dispute resolution’.⁴⁵⁷ The inclusion of such systems in the ADR represents an extension compared to the proposal of the AI Act, which originally only covered their use by judicial authorities.⁴⁵⁸
- 163 This raises the question of whether ‘ADR’ within the meaning of the AI Act also refers to algorithmic rights and standards enforcement through online platforms. This question must be answered by means of an autonomous interpretation of the law of the Union.⁴⁵⁹ Firstly, it should be noted that the AI-based enforcement of rights and standards by online platforms is not ‘purely ancillary administrative activities’ within the meaning of recital 61 s 5 AI Act, which the Regulation excludes from the group of high-risk AI systems. This is because the examples listed there – such as anonymisation or pseudonymisation of judicial decisions, documents, or data – differ categorically in their legal (internal) effects from a platform enforcement aimed at external effects vis-à-vis users or third parties. Furthermore, although the term ‘ADR’ is mentioned in Annex III no 8 lit a) and in recital 61 AI Act, it is not legally defined here or elsewhere in the Regulation. However, its basic conception under EU law is reflected in the ADR Directive.⁴⁶⁰ Without specifying a particular dispute resolution method,⁴⁶¹ the ADR

⁴⁵⁷ Annex III no 8 lit a) AI Act.

⁴⁵⁸ Proposal for a Regulation of the European Parliament and of the Council of 21 April 2021, COM(2021) 206 final, Annex III no 8 lit a).

⁴⁵⁹ In general on the autonomous interpretation of the law of the Union: K Riesenhuber, ‘§ 10 Die Auslegung’ in K Riesenhuber (ed) *Europäische Methodenlehre* (De Gruyter 2021) 285 para 4 ff; R Stotz, ‘Die Rechtsprechung des EuGH’ in K Riesenhuber (ed) *Europäische Methodenlehre* (De Gruyter 2021) 653 para 19.

⁴⁶⁰ Directive 2013/11/EU of the European Parliament and of the Council of 21 May 2013 on alternative dispute resolution for consumer disputes and amending Regulation (EC) No 2006/2004 and Directive 2009/22/EC (Directive on consumer ADR), OJ L 165, 18 June 2013, 63.

⁴⁶¹ G Rühl, ‘The Alternative Dispute Resolution Directive: Handlungsperspektiven und Handlungsoptionen’ (2014) 127(1) ZJP (Zeitschrift für Zivilprozess) 61, 67. So also the implementation of the Directive in the German Verbraucherstreitbeilegungsgesetz (Consumer Dispute Resolution Act), there § 1(1).

Directive defines alternative dispute resolution as ‘a simple, fast and low-cost out-of-court solution to disputes’.⁴⁶² Accordingly, a constitutive feature of out-of-court dispute resolution is the independence or impartiality of the organisation or person appointed to make the decision⁴⁶³ – a principle that the AI Act also expressly assumes in its recital 61 s 1 and 4.⁴⁶⁴ In light of this, the term ‘ADR’, also from the perspective of the AI Act, is to be understood as a form of dispute resolution by independent out-of-court decision-makers which – like ADR entities within the meaning of the ADR Directive⁴⁶⁵ – are similar to state courts in their function and basic structure of independent decision-making.⁴⁶⁶ This deliberately excludes forms of dispute resolution that are operated by companies themselves.⁴⁶⁷

- 164 As a result, there are important reasons against categorizing platform enforcement as an instrument of alternative dispute resolution. This is due to the fact that platform enforcement is not directly aimed at *resolving* disputes in the form of a decision⁴⁶⁸ nor does it provide for the use of *independent* dispute adjudicators. Unlike court proceedings, which typically aim to (provisionally) terminate legal disputes, platform enforcement involves its own concern, namely to unilaterally (if only provisionally) enforce standards or third-party rights on the basis of an independent assessment of the

⁴⁶² Recital 5 s 1 Directive 2013/11/EU, albeit limited to disputes arising from sales contracts or service contracts between consumers and traders. See also Art 2(1) Directive 2013/11/EU. – Even more precise: Commission Recommendations 98/257/EC of 30 March 1998 on the principles applicable to the bodies responsible for out-of-court settlement of consumer disputes, OJ L 115, 17 April 1998, 31, 32: ‘Whereas this recommendation must be limited to procedures which, no matter what they are called, lead to the settling of a dispute through the active intervention of a third party, who proposes or imposes a solution’. This definition was adopted from recitals 3 s 2 and 9 s 1 Commission Recommendation 2001/310/EC of 4 April 2001 on the principles for out-of-court bodies involved in the consensual resolution of consumer disputes (Text with EEA relevance) (notified under document number C(2001) 1016), OJ L 109, 19 April 2001, 56.

⁴⁶³ See recital 32, Art 1 s 1, Art 2(2)(a) and Art 6 Directive 2013/11/EU. Cf also Commission Recommendations 98/257/EC of 30 March 1998 on the principles applicable to the bodies responsible for out-of-court settlement of consumer disputes, OJ L 115, 17 April 1998, 31, 32: English: ‘essential’; German: ‘unerlässliche Voraussetzung’; French: ‘qualités nécessaires’. Also: Commission Recommendation 2001/310/EC (n 462), II.A.

⁴⁶⁴ Recital 61 s 1 (German version): ‘Bestimmte KI-Systeme, die für die Rechtspflege und demokratische Prozesse bestimmt sind, sollten angesichts ihrer möglichen erheblichen Auswirkungen auf die Demokratie, die Rechtsstaatlichkeit, die individuellen Freiheiten sowie das Recht auf einen wirksamen Rechtsbehelf und ein unparteiisches Gericht als hochriskant eingestuft werden’. French version: ‘tribunal impartial’.

Recital 4: ‘The use of AI tools can support the decision-making power of judges or judicial *independence*, but should not replace it’.

⁴⁶⁵ Art 5(1), Art 6 (1), (2) Directive 2013/11/EU.

⁴⁶⁶ The fact that the AI Act restricts the use of high-risk AI systems to functionally judicial proceedings also follows indirectly from its recital 48 s 2: ‘Those rights include [...] the right to an effective remedy and to a fair trial [...]’.

⁴⁶⁷ Art 2(2)(b) Directive 2013/11/EU. See also recital 9 s 4 Commission Recommendation 2001/310/EC (n 462), II.A.

⁴⁶⁸ Cf Art 2(1) Directive 2013/11/EU, German version: ‘Beilegung von [...] Streitigkeiten’; English version: ‘resolution of [...] disputes’; French version: ‘une solution, ou réunit les parties en vue de faciliter la recherche d’une solution amiable’.

legal and factual situation. In that regard, online platforms are not able to guarantee the independence of the decision-maker required by the above regulations, neither when enforcing their own standards nor when enforcing the rights of third parties.⁴⁶⁹ They can therefore be ruled out as institutions for the ‘administration of justice’ according to Annex III no 8 lit a) of the AI Act.

165 A different assessment could only be reached if the ADR concept of the AI Act were to be interpreted broadly, ie teleologically and on the basis of an effects-based approach, in contrast to the understanding of ADR under EU law. However, such an assessment cannot be derived from recital 61 s 3 of the AI Act, which only classifies AI systems used in ADR as high-risk ‘when the outcomes of the alternative dispute resolution proceedings *produce legal effects* for the parties’.⁴⁷⁰ The fact that this can only refer to the binding effect of ADR decisions, but not to any impairment of existing rights and legal interests of platform users or third parties, results on the one hand from the functional-structural proximity of ADR proceedings to court proceedings – a proximity that is also reflected in the reference in recital 61 s 3 AI Act to the legal effect that the results of the alternative dispute resolution produce *for the parties* to the proceedings. A further point of reference arises from Art 86(1) AI Act: The provision confers a right to explanation of individual decision-making if this is based on a high-risk AI system and the decision ‘produces legal effects or similarly significantly affects that person in a way that they consider having an adverse impact on their health, safety or fundamental rights’. Accordingly, the provision presupposes the use of high-risk AI systems within the meaning of Art 6, Annex III AI Act, for example in the course of an ADR decision with binding effect on the parties. This effect must differ from the effects of a decision referred to in Art 86(1) AI Act – ie the legal or factual impairments resulting from it – which would otherwise have no independent normative significance.⁴⁷¹ In other words, a right to an explanation under Art 86(1) of the AI Act does not already follow from the binding effect of state or extrajudicial decisions, but solely from the additional circumstance that such a decision interferes legally or factually with existing rights or legal interests of persons in individual cases.

166 To classify platform enforcement as a suitable field of application for high-risk AI systems within the meaning of Annex III no 8 lit a), a purely teleological interpretation of the provision would have to be adopted, by considering the principle of *effet utile* under EU law. The starting point for such an approach is that the AI Act is designed as a risk-based regulatory system for AI applications.⁴⁷² Accordingly, the riskier the operation of the AI,

⁴⁶⁹ See above para 144-147 with further references.

⁴⁷⁰ German version: ‘die Ergebnisse [...] Rechtswirkung für die Parteien entfalten’; French version: ‘les résultats [...] produisent des effets juridiques pour les parties’.

⁴⁷¹ Arg Art 86(1) AI Act: ‘[...] and which has legal effects or significantly affects them in a similar manner [...]’.

⁴⁷² Thus, expressly recital 26 AI Act.

the more comprehensive the obligations in connection with the operation of the AI.⁴⁷³ This regulatory concept, which is tailored to the intensity and scope of the risks, is manifested in recital 61 s 1 AI Act, among other things. Here, the Union legislator always advocates classifying AI systems as high-risk if they are likely to have a ‘significant impact’⁴⁷⁴ on individual freedoms and the fundamental rights protected by the Charter – such as the right to freedom of expression or the right to an effective remedy⁴⁷⁵. The severity of the harm and its likelihood of occurrence must be taken into account (recitals 48 s 2, 52 s 1 AI Act). The previous explanations have shown⁴⁷⁶ that the potential for errors and the type, scope and probability of imminent risks and damage when AI systems are used in platform enforcement are typically no different or even less pronounced than in the fields of application addressed in Annex III no 8 lit a). In view of the procedural structures proposed for platform enforcement,⁴⁷⁷ it is not apparent that the obligations of providers and operators of high-risk AI systems set out in Art 9-17 AI Act would impose a fundamentally unreasonable hardship on online platforms.

1.3.2.4 Necessary Dovetailing of Internal Complaint Procedures with Out-Of-Court Independent Dispute Resolution

1.3.2.4.1 Structural Bias of Online Platforms

- 167 Online platforms are sometimes described as ‘objective’ dispute decision-makers similar to a judge or even like a judge.⁴⁷⁸ Only recently, such a role description is echoed in the Digital Services Act, when online platforms are required to process complaints submitted

⁴⁷³ M von Welser, ‘Die KI-Verordnung – ein Überblick über das weltweit erste Regelwerk für künstliche Intelligenz’ (2024) 16(15) GRUR-Prax (Gewerblicher Rechtsschutz und Urheberrecht in der Praxis) 485; regarding the Proposal AI Act: I Orsich, ‘Das europäische Konzept für vertrauenswürdige Künstliche Intelligenz’ (2022) 33(6) EuZW (Europäische Zeitschrift für Wirtschaftsrecht) 254, 255.

⁴⁷⁴ Recital 61 S 1 AI Act, German version: ‘erhebliche Auswirkungen’; French version: ‘incidence [...] significative’.

⁴⁷⁵ See recital 48 s 2 AI Act.

⁴⁷⁶ See above 1.2 (Dangers and drawbacks of AI-based law and standards enforcement), para 67 ff.

⁴⁷⁷ See above 1.3.2.3 (Intra-company legal protection proceedings: Basic structures and procedural guarantees), para 127 f.

⁴⁷⁸ Cf the German Federal Court of Justice (Bundesgerichtshof), 25 October 2011, VI ZR 93/10 – *Blögeintrag* (2012) 114(3) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 311 para 25-27 (right of personality). Van Loo (n 182), (2016) 33(2) Yale Journal on Regulation 547, 566 f (‘third-party adjudicator’, ‘network trial’, ‘various court-like roles’), 576 (2016); Van Loo (n 100), (2021) 88(4) University of Chicago LRev 829, 832 (‘most important private judicial system’) 846 (‘quasi-judicial role’ with respect to review and enforcement of the ‘right to be forgotten’ by search engines such as Google), 849, 850 (‘The expanded privatization of U.S. justice through platforms’ internal dispute systems deserves scrutiny’), 865; Haber (n 216), (2016) 40(1) Seattle University Law Review 115, 129 ff; D Holznagel, ‘Melde- und Abhilfverfahren zur Beanstandung rechtswidrig gehosteter Inhalte nach europäischem und deutschem Recht im Vergleich zu gesetzlich geregelten notice and take-down-Verfahren’ (2014) 63(2) GRUR Int (Gewerblicher Rechtsschutz und Urheberrecht Internationaler Teil) 105, 108 (‘Judge’s Role’); see also F Hofmann and T Sprenger, ‘Privatization of Enforcement’ (2021) 85(2) UFITA (Archiv für Medienrecht und Medienwirtschaft) 249, 254 (‘to settle disputes’).

by users as part of a newly established, internal legal protection procedure in an 'objective manner'.⁴⁷⁹

- 168 However, this role model loses sight of the fact that platforms typically act in a liability-based and interest-driven manner and are thus structurally biased.⁴⁸⁰ This becomes obvious when platforms do not enforce third-party rights granted by the state, but rather – often automatically – their own community standards, ie general terms and conditions set by the platform itself and – as in the case of Facebook⁴⁸¹ – applicable worldwide.⁴⁸² In this quantitatively far predominant enforcement situation,⁴⁸³ the platform no longer acts as a mediator between third-party interests, but as an enforcer or – in complaint proceedings – as a judge in its own cause.⁴⁸⁴
- 169 By the same token, it is hardly surprising that platforms have generally shown little inclination so far to remedy complaints against the enforcement of self-imposed standards or – in the absence of liability-reinforcing submission obligations⁴⁸⁵ – to submit them to an independent extrajudicial body.⁴⁸⁶ In Germany, for example, Facebook made use of a corresponding right of referral in only 29 of a total of 123.195 complaints under the Network Enforcement Act (NetzDG) in the 3rd and 4th quarters of

⁴⁷⁹ Cf Art 16(4), (6); Art 14(4), Art 23(3) and recitals 24 s 3, 26 s 2 DSA Regulation.

⁴⁸⁰ See already Laukemann (n 271) 276 f. On the structural bias of the U.S. tourism website Trip-Advisor due to the dependence of its business model on advertising revenue from user-rated companies, see Van Loo (n 100), (2021) 88(4) University of Chicago LRev 829, 869.

⁴⁸¹ Cf Radu (n 261) 179: 'Local values representation is the second point of contention towards Facebook community. The unilateral definition of what is and what is not acceptable online by a company headquartered in the United States is harder to sustain as more than 2 billion people use the platform. Facebook's largest user base at the moment is India, but little of the social and cultural norms there appear to transpire in the global policy of the company'.

⁴⁸² On the legal classification of terms of use and community standards as GTCs, see only German Federal Court of Justice (Bundesgerichtshof), 29 July 2021, III ZR 192/20, (2021) 25(11) ZUM-RD (Zeitschrift für Urheber- und Medienrecht – Rechtsprechungsdienst) 614, para 44 (§§ 305 BGB ff are applicable); further: M Mayer, *Soziale Netzwerke im Internet im Lichte des Vertragsrechts* (Boorberg 2018) 120, 359. – Cf also the broad definitions in Art 3 lit u) DSA Regulation; Art 2 no 10 P2B Regulation and Art 2 no 8 TCO Regulation.

⁴⁸³ 33.700 of all 34.806 pieces of content deleted or blocked by Facebook in the relevant period, ie just under 97%, already (possibly only) violate community standards and are therefore deleted worldwide. Only the remaining share of 3.2% (= 1.106) falls through the 'international grid' and is thus only blocked in Germany, see Meta, 'Facebook Transparency Report of January 2023' 19 327151920_907084790305794_6193992151844220602_n.pdf (fbcdn.net) accessed 28 September 2023.

⁴⁸⁴ On this aspect, see Nahmias and Perel (n 78), (2021) 58(1) Harvard Journal on Legislation 145, 178, referring to Radu (n 261) 179.

⁴⁸⁵ § 3(6) no 3 of the German Network Enforcement Act (NetzDG) links the recognition of an institution as a so-called institution of regulated self-regulation, among other things, to the existence of rules of procedure that regulate submission obligations of the affiliated social networks.

⁴⁸⁶ For Facebook, the remediation rate for the period January to March 2022 is about 8.3% (587.000 appeals to 48.700 remediations); but in some cases also proactive correction due to previous, similar violations.

2022, ie in just around two out of 10.000 cases.⁴⁸⁷ The figures for other platforms are similar – or even more drastic.⁴⁸⁸

- 170 In this context, effective incentives to increase the platforms' willingness to submit documents can also be provided by time limits that are subject to liability or even fines.⁴⁸⁹

1.3.2.4.2 Out-Of-Court Dispute Resolution

- 171 In that regard, internal platform-based complaints procedures should be supplemented by external, out-of-court dispute resolution and these two protection instruments be more closely interlinked.⁴⁹⁰ Only when internal and external complaints procedures are intertwined can an effective legal remedy be guaranteed (see recital 52 p. 2 of the DSA-Regulation). A right to appeal is intended to enable affected users and whistleblowers to address a certified independent dispute resolution body directly (and as an alternative to appealing to state courts⁴⁹¹) after an unsuccessful internal complaint against a decision by the platform about content that is allegedly contrary to law or the terms and

⁴⁸⁷ These complaints were forwarded to the FSM (Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e.V.) in the period between 1 July 2022 and 31 December 2022: Meta, 'Facebook Transparency Report of January 2023 for the 2nd HY 2022' 14 327541832_1414754302684176_3061551644115119140_n.pdf (fbcdn.net) accessed 31 December 2023.

⁴⁸⁸ In the first half of 2023, YouTube submitted 12 of 193.131 reported content: Google, 'Transparency Report for YouTube Platform for January to June 2023', 4, 7 (<https://transparencyreport.google.com/netzdg/youtube?hl=de>). X, formerly Twitter, submitted 66 complaints to German law firms during the same period: Twitter, 'Network Enforcement Report: January-June 2023', 25 <https://transparency.twitter.com/content/dam/transparency-twitter/country-reports/germany/NetzDG-Jan-Jun-2023.pdf> accessed 31 December 2023.

⁴⁸⁹ This is the case if platforms can only avoid a fine if the decision is expected to take longer by submitting the complaint to an out-of-court dispute resolution body before the expiry of a decision deadline (of 7 days). Similarly, the solution of the German NetzDG in § 3(2) s 1 no 3 lit b): 'The procedure must ensure that the social network provider removes or blocks access to any unlawful content without undue delay, usually within seven days of receipt of the complaint; the seven-day period may be exceeded if [...] b) within seven days of receipt of the complaint, the social network provider transfers the decision on the unlawfulness to a body of regulated self-regulation recognized in accordance with paragraphs 6 to 8 and submits to its decision [...]'. However, a violation of the procedural requirements in a specific individual case is not sufficient to warrant a fine (Bundestags-Drucksache 18/12356, 24; M Liesching, *Netzwerkdurchsetzungsgesetz* (Nomos 2018) § 4 para 13. The exact requirements for going beyond the individual case remain unclear. – The Austrian KoPIG does not recognize such an outsourcing of the decision to an independent body.

⁴⁹⁰ See most recently Art 21(1) DSA Regulation, with regard to certified out-of-court dispute resolution bodies, which are to be accredited by Member State coordinators for digital services (Art 21(3) DSA Regulation; revocation of accreditation under paragraph 7).

⁴⁹¹ For both the enforcement addressee and the rightsholder, state legal protection should thus not only be possible downstream, as was recently also the case in Art 21(1) subpara 3 DSA Regulation.

conditions,⁴⁹² even against the will of the platform.⁴⁹³ The decision of the dispute resolution body – which is not to be automated – should in turn be binding for the platform,⁴⁹⁴ at least until a possible state court decision.⁴⁹⁵ It should also be stored anonymously and in an easily accessible manner and published online.⁴⁹⁶

- 172 If, according to the controversial and criticised concept of the DSA, decisions of certified out-of-court dispute settlement bodies do not have any binding effect on the parties,⁴⁹⁷ AI systems that these bodies may use to assist in researching and interpreting facts and the law and in applying the law to a concrete set of facts, are not to be categorized as so-called high-risk AI systems within the meaning of the AI Act of the European Union.⁴⁹⁸ As a consequence, the providers and operators of such high-risk AI systems are not subject to the special obligations under Art 9-17 AI Act.⁴⁹⁹
- 173 In addition, out-of-court dispute resolution (eg, in the context of Art 21 DSA Regulation) opens up an incidence check of the platform general terms and conditions – and the more fluctuating moderation guidelines⁵⁰⁰ that concretize them – with regard to their

⁴⁹² According to the DSA, the relevant criteria for authorization include, in particular, the independence and impartiality of the body, as well as expertise and clear and fair rules that enable an easily accessible procedure aimed at efficient, ie, rapid and cost-effective, dispute resolution. According to the cost regulation of Art 21(5) DSA Regulation, which is advantageous for both users and notifiers, in the event of an unsuccessful complaint, the latter shall in principle only bear their own fees and other reasonable costs, but not those of the online platform. In the opposite case, the online platform bears the full cost burden, including the costs of the prevailing opposing party. – In addition, the DSA stipulates a duty to inform the platforms about the possibility of appealing to such institutions.

⁴⁹³ The dispute resolution body could be called upon in two ways: on the one hand as a kind of ‘complaints authority’ by users or whistleblowers, and on the other, on mandatory submission by the platform in the event of an unclear legal or infringement situation.

⁴⁹⁴ According to Art 21(4) subpara 3 DSA Regulation, a decision should in principle not take longer than 90 calendar days, and in the case of ‘highly complex’ disputes a maximum of 180 days.

⁴⁹⁵ See Hofmann and Specht-Riemenschneider (n 347), (2021) 13(1) ZGE (Zeitschrift für geistiges Eigentum) 48, 103 f, who in this context refer to a study by Fiala and Husivec, according to which the risk of overblocking can be significantly reduced by using out-of-court dispute resolution mechanisms: F Fiala and M Husivec, ‘Using Experimental Evidence to Improve Delegated Enforcement’ (2018) International Review of Law and Economics, Forthcoming TILEC Discussion Paper no 2018-028, 25. – It is therefore to be criticized if Art 21(2) subpara 3 DSA Regulation explicitly denies a corresponding binding effect.

⁴⁹⁶ Art 21(4) and Art 24(5) DSA Regulation refrain from this and only require the bodies to report annually to the Digital Services Coordinator on the number, duration and outcome of disputes the lack of storage of these often-important decisions under the DSA is to be criticized. – As a flanking measure, Art 24(1) lit a) DSA Regulation obliges online platform providers to report on the number of disputes submitted to the out-of-court dispute resolution bodies referred to in Art 21, the results of dispute resolution and mediation duration until the conclusion of dispute resolution proceedings, as well as the proportion of disputes in which the online platform providers implemented the body’s decisions.

⁴⁹⁷ Art 21(2) subpara 3 DSA Regulation. Critical: D Holznapel in R Müller-Terpitz and M Köhler (ed), *Digital Services Act* (Beck 2024) Art 21 para 39-41.

⁴⁹⁸ Art 6(2), Annex III no 8 lit a) AI Act. See in more detail above para 162 ff.

⁴⁹⁹ See in more detail above para 161 ff.

⁵⁰⁰ Cf P McColgan, ‘Das wird man wohl noch löschen dürfen? – Control Standards for Opinion Rules on the Internet’ (2021) 1(12) RDi (Recht Digital) 605, 610 f, 615 f; Mast (n 336), (2023) 78(7) JZ (Juristenzeitung) 287, 292; further Klonick (n 94), (2018) 131(6) Harvard Law Review 1598, 1639.

conformity with the procedural and contractual requirements⁵⁰¹ of pertinent EU regulations (Digital Services Act; P2B Regulation).⁵⁰² In view of the complex and numerous requirements of EU law, from the perspective of platform operators the law on general terms and conditions is turning into a freedom-restricting vehicle for inclusion in private platform usage contracts. Platform general terms and conditions that have been overhauled by state law are taking on a regulatory function. In this way, however, the state instrumentalizes private regulatory potential even in areas that – for reasons of protected spheres of freedom of expression – are typically removed from the state’s legislative competence.⁵⁰³

1.3.2.4.3 The Establishment of ‘Platform Courts’

- 174 Even before the Facebook Oversight Board was set up, there were suggestions that a platform-based enforcement of rights and, above all, of a company’s own standards and terms and conditions should be flanked by external expertise, namely in the form of non-governmental, court-like dispute resolution bodies.⁵⁰⁴ Corresponding ‘platform courts’ will not only have the task of settling individual disputes (usually of particular relevance) with particular proximity to the platform’s subject matters and continuously harmonizing decision-making and interpretation of the community standards of a platform.⁵⁰⁵ Rather, they are also assigned a regulatory function: In implementing what is essentially a ‘procedural concept of (self-)regulation’, platform courts shape a private process of norm-building that is (partly) beyond a state’s control of general terms and conditions and is communicatively linked back to the public.⁵⁰⁶ In this process, the dimensions of mass and time that are typical of platforms (for example, with regard to ephemeral communication and evaluation processes) must be managed in a regulatory manner.⁵⁰⁷ In the interest of a (mostly heterogeneous) mass of platform users, the aim

⁵⁰¹ Mast (n 336), (2023) 78(7) JZ (Juristenzeitung) 287, 292.

⁵⁰² For more information on this and on the linkage of § 307(1) BGB (German Civil Code) to the standards of European Union law, see Mast (n 336), (2023) 78(7) JZ (Juristenzeitung) 287, 290. – In the case of the review of unlawful platform GTCs by state courts – for example in German law on the basis of the Unterlassungsklagegesetz (UKlaG) or supplementary pursuant to Art 14(1) P2B Regulation, court decisions have a broad effect, for example under § 11 UKlaG or by means of a nullity order against GTCs that violate Art 3(1) P2B Regulation (thus Art 3(3) in conjunction with recital 20 of the P2B Regulation).

⁵⁰³ Thus aptly: Mast (n 336), (2023) 78(7) JZ (Juristenzeitung) 287, 289, 292, 295, who in this respect speaks of the ‘normative force of platform GTCs’.

⁵⁰⁴ See initially K-H Ladeur, ‘Neue Institutionen für den Daten- und Persönlichkeitsschutz im Internet: „Cyber-Courts“ für die Blogosphere’ (2012) 36(10) DuD (Datenschutz und Datensicherheit) 711 ff; further: Van Loo (n 100), (2021) 88(4) University of Chicago Law Review 829 ff.

⁵⁰⁵ Cf also Ladeur (n 384).

⁵⁰⁶ On identity and personality formation as a communicative process, see D Wielsch, ‘Medienregulierung durch Persönlichkeits- und Datenschutzrechte’ (2020) 75(3) JZ (Juristenzeitung) 105, 107 ff.

⁵⁰⁷ T Wu argues that in order to cope with the mass problem, also for reasons of optimized resource allocation, simple cases (‘easy cases’: simple facts; clear, little context-dependent legal situation) should be decided by means of automated moderation, while complex case constellations (‘hard cases’:

is twofold: on the one hand, to ensure more effective protection of individual rights in a multitude of (parallel) cases and, on the other, to enable a continuously adapted, by its nature mostly rudimentary-experimental rule formation through and within the framework of procedural structures.

- 175 The conceptual approaches advocated in this context are linked in different ways: either effect-related to a norm-building process through (limited) precedent;⁵⁰⁸ or dogmatically through its foundation on ‘treaty networks’;⁵⁰⁹ or institutionally, for example in the form of an independent, cross-platform ‘court of justice’ for the formation of a (sectorally differentiated) substantive ‘platform common law’.⁵¹⁰ Parallels can also be drawn functionally to the further development of law by the appellate courts or, more concretely, to the functioning of the WTO Appellate Body.⁵¹¹ In the absence of statutory regulation, such projects presuppose that platforms voluntarily submit to private platform jurisdiction – and its indirect regulatory effects – at least within certain business areas. Irrespective of how closely the relevant dispute resolution bodies are organizationally linked to a platform or how their internal structure is differentiated by various normative functionaries, interpretation (dispute-related review), setting and enforcement of private standards ultimately amalgamate into ‘one power’.

context-, consideration-intensive and complex legal situations) the decision should be reserved for humans – especially committees, in: T Wu, ‘Will Artificial Intelligence Eat the Law? The Rise of Hybrid Social-Ordering Systems’ (2019) 119(7) *Columbia Law Review* 2001 ff; see also M Denga, ‘Platform Regulation by European Values: On the Binding of Opinion Platforms to EU Fundamental Rights’ (2021) 56(5) *EuR (Europarecht)* 569, 572. – In contrast, the approach is to reduce the scrutiny of significant communication processes by raising the threshold of infringing conduct within social media. At the same time, the concretization of platform-typical communication customs contributes to the formation of area-specific rules, see on the whole: Ladeur (n 384).

⁵⁰⁸ Van Loo (n 100), (2021) 88(4) *University of Chicago Law Review* 829, 867 ff. However, the link to precedents should not be strict so as not to block innovative adaptations to a rapidly changing online environment.

⁵⁰⁹ Ladeur (n 384), (2012) 36(10) *DuD (Datenschutz und Datensicherheit)* 711, 714.

⁵¹⁰ Comparisons to a cross-provider ‘cyber court of second instance’, with greater consideration of regional characteristics: Ladeur (n 384). – In contrast, R Van Loo does not want to take the word of the idea of a sector-wide uniform private regulatory regime – for example, for social media platforms. Consequently, he rejects the idea of (strict) cross-platform precedence. Instead, he considers taking his cue from the common law model of persuasive authority of foreign court decisions. Similar to the idea of a ‘market for rules’, platforms should compete with each other: Van Loo (n 100), (2021) 88(4) *University of Chicago Law Review* 829, 867 f; also McColgan (n 500), (2021) 1(12) *RD (Recht Digital)* 605, 613. Critical of the idea of competition in view of a monopolization of the platform market, in turn, the Union legislature on the occasion of the Digital Services Act, on this L Kumkar, ‘Plattform-Recht revisited: Umgang mit den Marktordnungen digitaler Plattformen de lege lata et ferenda’ (2022) 30(3) *ZEuP (Zeitschrift für Europäisches Privatrecht)* 530, 551; B Raue, ‘Plattformnutzungsverträge im Lichte der gesteigerten Grundrechtsbindung marktstarker sozialer Netze’ (2022) 75(4) *NJW (Neue Juristische Wochenschrift)* 209 para 4, referring to the lock-in and network effect for users. – M Land also argues in favor of different rules adapted to the practices of different platform types, in: M Land, ‘The Problem of Platform Law: Pluralistic Legal Ordering on Social Media’, in P SA Berman (ed), *The Oxford Handbook of Global Legal Pluralism* (2020) 974.

⁵¹¹ On this parallel: Ladeur (n 384).

2 ARTIFICIAL INTELLIGENCE IN MEDIATION AND CONCILIATION

- 176 This chapter takes a closer look at the use of AI in mediation and conciliation. There is no uniform understanding of the exact terminology used in the two dispute resolution systems. For mediation, the definition of the EU Mediation Directive⁵¹² will be used in the following. It defines mediation⁵¹³ in Art 3 lit a) as a ‘structured process [...] whereby two or more parties to a dispute attempt by themselves, on a voluntary basis, to reach an agreement on the settlement of their dispute with the assistance of a mediator. This process may be initiated by the parties or suggested or ordered by a court or prescribed by the law of a Member State.’
- 177 In contrast, the conciliator, as a neutral intermediary, plays a more active role in the conciliation process by submitting – albeit non-binding – proposals for decisions.⁵¹⁴ Conciliation is therefore a third-party decision procedure.⁵¹⁵
- 178 Then again both procedures pursue the goal of an amicable, in principle non-binding dispute resolution,⁵¹⁶ under the intermediary intervention of a neutral third party, regularly detached from substantive and normative requirements.⁵¹⁷

2.1 Fields of Application of Artificial Intelligence

2.1.1 Fields of Application According to Process Stages

2.1.1.1 Procurement of (Legal) Information by Artificial Intelligence

- 179 Legal chatbots are already able to provide low-threshold information for those seeking legal advice in the run-up to mediation and conciliation proceedings – for example about the course of proceedings or the costs involved.⁵¹⁸ Such as answering simple legal

⁵¹² Directive 2008/52/EC of the European Parliament and of the Council of 21 May 2008 on certain aspects of mediation in civil and commercial matters, Official Journal of the European Union from 24 May 2008, L 136/3 ff.

⁵¹³ See also D Rodi, in *Staudinger-BGB*, Buch 2 (19th rev edn, De Gruyter 2022) Anh. zu §§ 305-610 BGB M 33; M Fehrenbach, in *BeckOGK-BGB* (Beck 1 November 2023) § 307 Schlichtungsklausel para 4; P Röthemeyer, ‘Die Schlichtung’ (2013) 16(2) ZKM (Zeitschrift für Konfliktmanagement) 47, 49.

⁵¹⁴ Röthemeyer (n 513), (2013) 16(2) ZKM (Zeitschrift für Konfliktmanagement) 47, 48; M Fehrenbach, in *BeckOGK-BGB* (Beck 1 November 2023) § 307 Schlichtungsklausel para 4.

⁵¹⁵ M Fehrenbach, in *BeckOGK-BGB* (Beck 1 November 2023) § 307 Schlichtungsklausel para 4; R Greger und C Stubbe, *Schiedsgutachten* (1st edn, Beck 2007) para 27.

⁵¹⁶ See R Greger, ‘D Recht der alternativen Konfliktlösung’ in R Greger, H Unberath and F Steffek (ed), *Recht der alternativen Konfliktlösung* (2nd edn, Beck 2016) para 245.

⁵¹⁷ However, according to § 19 of the German Consumer Dispute Resolution Act (Verbraucherstreitbeilegungsgesetzes: VSBG), the accepted conciliation proposal is legally binding. – In contrast, the result of mediation can also end in the conclusion of a settlement agreement, see F Kreis, ‘KI und ADR-Verfahren’ in M Kaulartz and T Braegelmann (ed), *Rechtshandbuch Artificial Intelligence und Machine Learning* (2020) 633, 638 para 17 f.

⁵¹⁸ S J Heetkamp and C Piroutek, ‘ChatGPT in Mediation und Schlichtung’ (2023) 26(3) ZKM (Zeitschrift für Konfliktmanagement) 80.

questions,⁵¹⁹ whereby, depending on the jurisdiction, professional restrictions must be observed.⁵²⁰ However, the use of such large language models⁵²¹ carries the risk of providing incorrect information (so-called ‘hallucinations’), which are not easily recognizable as such to a layperson.⁵²²

⁵¹⁹ T Deichsel, *Digitalisierung der Streitbeilegung* (1st edn, Nomos 2022) 201 f; C Leeb, *Digitalization, Legal Technology and Innovation* (1st edn, Dr. Otto Schmidt 2019) 238.

⁵²⁰ When using legal advice chatbots in Germany, the limits of the Legal Services Act (Rechtsdienstleistungsgesetz: RDG) must be observed: For example, in the legal services market, non-lawyer legal advice is prohibited as soon as the threshold of independent provision of legal services is exceeded, § 2(1) RDG. – In the case of debt collection services, § 10(1) s 1 no 1 RDG and § 6(1) RDG must be observed, which define the limits of the permissible scope of activities. Accordingly, legal advice (free of charge) provided by private companies such as legal techs is not permitted. A debt collection service provider’s authority to provide advice is limited to the activity of asserting claims – in particular, it is therefore not comprehensive, see in more detail M Hartung, ‘Sonstige Akteure und Rahmenbedingungen’ in M Hartung, M-M Bues and G Halbleib (ed), *Legal Tech: Die Digitalisierung des Rechtsmarkts* (1st edn, Beck 2018) 215 para 1044. The question of whether and under what conditions chatbots can even provide legal services within the meaning of § 2 RDG is still highly controversial. In particular, it is unclear to what extent the use of artificial intelligence constitutes a ‘legal examination of the individual case’ (pursuant to § 2(1) RDG) in view of its technical functionality. This, in turn, is partly denied due to the lack of subsumption. Programming abstract legal decision trees is only an abstract activity, not an examination of the concrete legal situation, cf C Deckenbrock and M Henssler, in C Deckenbrock and M Henssler (ed), *Rechtsdienstleistungsgesetz: RDG* (5th edn, Beck 2021) § 2 para 54g. By contrast, it will be necessary to differentiate: If only abstract information is provided or a chatbot is only used to establish the facts of the case, there is no requirement for a legal review. However, as soon as legal information is provided that is adapted to the data previously entered by the person seeking legal advice, the threshold for a specific examination of the individual case is likely to be exceeded, see also B Brechmann, *Legal Tech und das Anwaltsmonopol* (Mohr Siebeck 2021) 61; F Remmert and M Krenzler, in M Krenzler and F Remmert (ed), *Rechtsdienstleistungsgesetz* (3rd edn, Nomos 2023) § 2 para 71a; see also the final report of the State Working Group: Abschlussbericht der Länderarbeitsgruppe, ‘Legal Tech: Herausforderungen für die Justiz’ (2019) 40 f, https://www.schleswig-holstein.de/DE/landesregierung/ministerien-behoerden/II/Minister/Justizministerkonferenz/Downloads/190605_beschluesse/TOPI_11_Abschlussbericht.pdf?__blob=publicationFile&v=1. – If legal chatbots are used in court-connected dispute resolution, the scope of application of the RDG – which is limited to out-of-court legal services – is already excluded, § 1(1) s 1 RDG. The activities of conciliation boards, arbitrators as well as mediation and any comparable form of alternative dispute resolution – insofar as the activity does not intervene in the discussions of the parties involved by proposing legal regulations – do not constitute legal services within the meaning of § 2(2) no 2, 4 RDG. For the whole, see Leeb (n 519) 75, 280 ff.

⁵²¹ Large language models such as ChatGPT are based on machine learning technology and form neural networks that enable the AI system to answer questions or work assignments (so-called ‘prompts’) posed by the user on the basis of the underlying training data. These are statistical models that do not retrieve ‘knowledge’, but calculate probable word sequences based on an analysis of recognized text patterns and contexts and output them as answers. In addition, large language models are able to contextualize a user’s input information and create new content based on their training data, which can then be restructured or linguistically adapted according to a user’s requirements, see Heetkamp and Piroutek (n 518), (2023) 26(3) ZKM (Zeitschrift für Konfliktmanagement) 80.

⁵²² S Meder, ‘Die Zukunft der juristischen Methode: Rehabilitation durch Chat-GPT?’ (2023) 78(23) JZ (Juristenzeitung) 1041, 1051 at fn 109; Heetkamp and Piroutek (n 518), (2023) 26(3) ZKM (Zeitschrift für Konfliktmanagement) 80.

2.1.1.2 Automated Document Analysis, Fact-Finding and Case Management

180 Similar to arbitration proceedings, the use of automated document analysis also promises to reduce the workload for the parties involved in mediation and conciliation. One possibility is the automated structuring of party submissions, whereby the facts of the case are recorded in advance using input masks.⁵²³ Further, AI-based automation can be used to classify the status of disputes, for instance in the categories of child maintenance or contact issues. The system can then assign the dispute status prepared this way to a human decision-maker, possibly taking into account the urgency of a decision ('triage'), which is also determined by AI.⁵²⁴ In the form of smart assistants or smart scheduling, there is also potential for automation in the further conduct of proceedings. The use of AI-supported speech recognition and translation is also conceivable,⁵²⁵ as are, in principle, potential applications for an automated selection of mediators or conciliators.⁵²⁶

2.1.1.3 AI-Tools for Preparing Settlement or Decision Proposals (Predictive Tools)

181 The use of predictive tools in mediation and conciliation bears potential. In Singapore, for example, there is a procedural simulator intended to predict the outcome of legal – court⁵²⁷ – disputes for the parties and thus increase their willingness to settle.⁵²⁸ The use of such technology is being discussed in proceedings before consumer conciliation boards, too. Its purpose is to encourage consumers to actually submit their complaint.⁵²⁹ Human conciliators are also beneficiaries, as AI systems create draft texts or drafts of conciliation proposals (*document drafting and pre-drafting*).⁵³⁰ Parties also receive support when drafting an agreement in mediation proceedings.⁵³¹ Finally, the proceedings⁵³² – including the decision-making process and in case of conciliation –

⁵²³ T Deichsel, 'Verbraucherschlichtungsstellen – Ein Anwendungsfeld für Legal Tech?' (2020) 35(8) VuR (Verbraucher und Recht) 283, 287.

⁵²⁴ Thus, the proposal for an automated conflict system by H M Anzinger, '10 Jahre Modria – KMS und Online-Mediation auf dem Weg zur Digitalisierung der Justiz – Teil 1' (2021) 24(2) ZKM (Zeitschrift für Konfliktmanagement) 53, 56.

⁵²⁵ Kreis (n 517) 633, 640 para 27 f.

⁵²⁶ In this respect, the problem is likely to be similar to that in arbitration proceedings due to a lack of suitable databases and other information bases, see on automated selections of arbitrators below para 206-207.

⁵²⁷ The Singapore Small Claims Tribunal is responsible for disputes between buyers and sellers in the range of 20.000 to 30.000 Singapore Dollars, see <https://www.judiciary.gov.sg/civil/file-small-claim> accessed 31 December 2023.

⁵²⁸ H M Anzinger, '10 Jahre Modria – KMS und Online-Mediation auf dem Weg zur Digitalisierung der Justiz – Teil 2' (2021) 24(3) ZKM (Zeitschrift für Konfliktmanagement) 84, 87.

⁵²⁹ Deichsel (n 523), (2020) 35(8) VuR (Verbraucher und Recht) 283, 287.

⁵³⁰ Kreis (n 517) 633, 640 para 28.

⁵³¹ Anzinger (n 524), (2021) 24(2) ZKM (Zeitschrift für Konfliktmanagement) 53, 56 f.

⁵³² Anzinger (n 524), (2021) 24(2) ZKM (Zeitschrift für Konfliktmanagement) 53, 57.

could possibly be conducted entirely by machine, although the technical feasibility of this is not yet apparent at present.⁵³³

2.1.2 Sectoral Fields of Application of AI in Mediation and Arbitration

- 182 Platform-based conciliation serves as inspiration for an advanced automation of out-of-court dispute resolution. Large intermediaries not only mediate the exchange of goods, but also take on the role of a third-party adjudicator with respect to the relationship between their customers – eg, between sellers and buyers on a platform (‘third-party-adjudication’).⁵³⁴ By way of example, the chargeback process of the payment service provider PayPal⁵³⁵ can be fully automated with the help of the third-party provider chargeflow: The system automatically recognizes potential chargeback cases, calculates the chances of success, automatically drafts the pleading with corresponding evidence and submits it to PayPal. Further, the appeal against PayPal’s initial decision will be fully automated (so-called charge appeal).⁵³⁶
- 183 In the resolution process of the eBay Resolution Center, an algorithm mediates between the parties to the dispute; 90% of cases are resolved purely automatically in this way.⁵³⁷ As the entire transaction is processed via eBay, the Resolution Center has direct insight into the facts of the case.⁵³⁸

⁵³³ Deichsel (n 523), (2020) 35(8) *VuR* (Verbraucher und Recht) 283, 288.

⁵³⁴ Fundamental in this regard: Van Loo (n 182), (2016) 33(2) *Yale Journal on Regulation* 547, 566 ff.

⁵³⁵ Van Loo (n 182), (2016) 33(2) *Yale Journal on Regulation* 547, 551 f.

⁵³⁶ See also: <https://www.chargeflow.io/paypal-dispute-automation>.

⁵³⁷ W Voß, ‘Gerichtsverbundene Online-Streitbeilegung’ (2020) 84(1) *RabelsZ* (Rabels Zeitschrift für ausländisches und internationales Privatrecht) 62, 65; also G Rühl, ‘Digitale Justiz’ (2020) 75(17) *JZ* (Juristenzeitung) 809, 811 f; A Sela, ‘The Effect of Online Technologies on Dispute Resolution System Design’ (2017) 21(3) *Lewis & Clark Law Review* 633, 662; H Barton, ‘Rebooting Justice’ (2018) 44(4) *Law Practice* 32, 35 f; C Rule, ‘Making Peace on eBay’ (2008) *ACR Resolution* 8, 10.

⁵³⁸ A Sela (n 537), (2017) 21(3) *Lewis & Clark Law Review* 633, 662.

184 In contrast, the degree of automation of state-recognized⁵³⁹ and purely private conciliation bodies as well as ODR providers⁵⁴⁰ is limited to electronic communication.⁵⁴¹ This applies in particular to the area of family mediation.⁵⁴²

185 Despite all these developments, as far as can be seen, AI is currently only used in the procedure simulator in Singapore and in a comparable tool in British Columbia ('Solution Explorer').⁵⁴³ The latter involves the use of a simple, rule-based AI (so-called expert system).⁵⁴⁴

2.2 Use and Dangers of Using Artificial Intelligence

2.2.1 Benefit

186 The (hoped-for) benefits of using AI are obvious in the area of mediation and conciliation as well. On the one hand, procedural efficiency gains are expected, expressed, for instance, in faster processing of applications and similar disputes (eg, before consumer conciliation boards).⁵⁴⁵ This is usually accompanied by a lower cost burden compared to manual processing by humans.⁵⁴⁶ On the other hand, it remains to be seen whether automated or AI-based instruments will equally facilitate access to out-of-court dispute resolution (eg, in consumer matters) and be able to overcome the rational disinterest of

⁵³⁹ For example, certified out-of-court dispute resolution bodies such as 'Der Online-Schlichter'; also: complaint and ombudsman procedures such as the 'Internet Ombudsman' in Austria, see Braegelmann (n 520) 215, 218 para 930. An instructive overview can be found in R Greger (n 516) Part D.

⁵⁴⁰ ODR providers include Modria (<https://www.tylertech.com/products/online-dispute-resolution> accessed 31 December 2023), SquareTrade, Cybersettle (<https://www.cybersettle.com/> accessed 31 December 2023) and Smartsettle (<https://www.smartsettle.com/> accessed 31 December 2023). The NCTDR (The National Center for Technology & Dispute Resolution) provides an overview of private providers at: <https://odr.info/provider-list/> accessed 31 December 2023.

⁵⁴¹ An instructive overview can be found in Anzinger (n 528), (2021) 24(3) ZKM (Zeitschrift für Konfliktmanagement) 84, 85; W Brazil, 'Informalism and Formalism in the History of ADR in the United States' in J Zekoll, M Bälz and I Amelung (ed), *Formalisation and Flexibilisation in Dispute Resolution* (Brill 2014) 250, 280 ff; D Hensler, 'The Private in Public, the Public in Private' in J Zekoll, M Bälz and I Amelung (ed), *Formalisation and Flexibilisation in Dispute Resolution* (Brill 2014) 45, 48 ff, 53-55.

⁵⁴² The dispute resolution of the platform 'uitelkaar.nl' (as successor to the provider 'Rechtwijzer') only takes place online, see Anzinger (n 528), (2021) 24(3) ZKM (Zeitschrift für Konfliktmanagement) 84, 87.

⁵⁴³ See: <https://civilresolutionbc.ca/solution-explorer/> accessed 31 December 2023. The Civil Resolution Tribunal is responsible for car accidents, small claims up to 5.000 Canadian dollars, special tenancy cases (strata property) and proceedings against companies based in British Columbia; see in detail V Tan, 'Online Dispute Resolution for Small Civil Claims in Victoria' (2019) 24 Deakin Law Review 101, 116-118; Anzinger (n 528), (2021) 24(3) ZKM (Zeitschrift für Konfliktmanagement) 84, 86.

⁵⁴⁴ S Salter and D Thompson, 'Public-Centred Civil Justice Redesign' (2016-2017) 3 McGill Journal of Dispute Resolution 113, 129; Tan (n 543), (2019) 24 Deakin Law Review 101, 121.

⁵⁴⁵ Deichsel (n 523), (2020) 35(8) VuR (Verbraucher und Recht) 283, 286 f, 288; S and H Kumar, 'Mediation and Artificial Intelligence' (2021) 4(4) International Journal of Law Management & Humanities 1472, 1477.

⁵⁴⁶ S and H Kumar (n 545), (2021) 4(4) International Journal of Law Management & Humanities 1472, 1476 f; F Specht, 'Chancen und Risiken einer digitalen Justiz für den Zivilprozess' (2019) 22(3) MMR (Multimedia und Recht) 153, 156; Deichsel (n 523), (2020) 35(8) VuR (Verbraucher und Recht) 283, 286 f.

consumers⁵⁴⁷ even in countries where the risks of legal action are better covered by legal aid and legal expenses insurance.⁵⁴⁸ To what extent and in which areas automated online dispute resolution systems are desirable in terms of legal policy is beyond a uniform and – in view of constant further developments – conclusive assessment at the present time. In addition, the attractiveness and actual use of such forms of dispute resolution stands and falls with the relief of the state courts.⁵⁴⁹

2.2.2 Dangers

2.2.2.1 The Black Box Problem and Discrimination

- 187 Similar to platform-based law enforcement and arbitration proceedings, the use of AI harbors the risk of opaque decision-making bases, particularly in the area of dispute resolution (black box problem). Even assuming that modern systems are capable of processing new, yet unknown data after the learning phase, such *modus operandi* inevitably leads to an output becoming independent from the input, so that the individual steps leading to a decision can no longer be reconstructed. This essentially concerns the sources and method of decision-making, ie on the one hand the way in which the multitude of stored information is linked and on the other, and above all, whether the system has merely reproduced a certain normative program or ‘added’ something to it.⁵⁵⁰
- 188 By the same token, it is hardly possible to explain and justify fully automated decisions.⁵⁵¹ Along with this goes a possible loss of legitimacy of AI-based decision-making proposals or suggestions, which may affect the acceptance of such procedures and their decision-makers as a whole.⁵⁵² If AI is not used solely for informational

⁵⁴⁷ On the so-called ‘digital justice gap’, see already: Braegelmann (n 520) 215 para 922; Anzinger (n 528), (2021) 24(3) ZKM (Zeitschrift für Konfliktmanagement) 84, 87 f.

⁵⁴⁸ Critical of the assumption of an extended access: Voß (n 537), (2020) 84(1) RabelsZ (Rabels Zeitschrift für ausländisches und internationales Privatrecht) 62, 64-66. – On the assessment of online dispute resolution mechanisms as a catalyst for effective access to justice, see C Menkel-Meadow, ‘Is ODR ADR? Reflections of an ADR Founder from 15th ODR Conference, The Hague’ (2016) 3(1) IJODR (International Journal on Online Dispute Resolution) 4; O Rabinovich-Einy and E Katsh, ‘The New New Courts’ (2017) 67(1) American University Law Review 165, 169. – On the historical roots of the ADR concept of overcoming structural weaknesses of state court systems and milieu-specific access barriers, see G Wagner, ‘Private Law Enforcement and ADR’ in J Zekoll, M Bälz and I Amelung (ed), *Formalisation and Flexibilisation in Dispute Resolution* (Brill 2014) 369 f; M Wendland, *Mediation und Zivilprozess* (2017), 199 ff.

⁵⁴⁹ See Deichsel (n 523), (2020) 35(8) VuR (Verbraucher und Recht) 283, 286 f.

⁵⁵⁰ Meder (n 522), (2023) 78(23) JZ (Juristenzeitung) 1041, 1047.

⁵⁵¹ For arbitration proceedings: C Sim, ‘Will Artificial Intelligence Take Over Arbitration?’ (2018) 14(1) Asian International Arbitration Journal 1, 8 f; G Vannieuwenhuyse, ‘Arbitration and New Technologies: Mutual Benefits’ (2018) 35(1) Journal of International Arbitration 119, 124.

⁵⁵² S and H Kumar (n 545), (2021) 4(4) International Journal of Law Management & Humanities 1472, 1478 f. On the judicial process: G Rühl, ‘KI in der gerichtlichen Streitbeilegung’ in M Kaulartz and T Braegelmann (ed), *Rechtshandbuch Artificial Intelligence und Machine Learning* (2020) 617, 627 para 20.

purposes in the run-up to mediation or conciliation⁵⁵³ or with the aim of automated document analysis or fact-finding, but (at the same time) with the intention of supporting decision-making in mediative compromise finding or the preparation of conciliation proposals,⁵⁵⁴ there is a risk of discriminating against parties to proceedings (in individual cases) and undermining independent and impartial decision-making with regard to certain facts or groups of persons concerned.

2.2.2.2 ‘Petrification’ of Decision-Making through Strict Reference to the Past

- 189 A key point of criticism of the use of AI in mediation and conciliation, as in arbitration proceedings, is the finding that, for technical reasons, AI is not capable of making cognitive and (spontaneous) value-based decisions⁵⁵⁵ or showing empathy – as a cornerstone of interpersonal communication. AI also lacks self-confidence and the ability to explain its own algorithms.⁵⁵⁶ The reason for this is obvious: Given that training data is only able to represent the past,⁵⁵⁷ AI knows nothing about the current physical world and therefore about *present* life in general. Therefore, AI lacks a *direct* and *immediate* view of this; judgment, intuition, (legal) feeling, *aequitas* and tact are necessarily alien to it.⁵⁵⁸ This disadvantage always becomes apparent when improbable events occur or the existing body of norms mapped in the algorithm is not sufficient and must therefore be supplemented by interpretation or further development of the law.⁵⁵⁹
- 190 Rather, by referring to past case material, AI makes ‘conservative’ decisions and is therefore not technically capable of developing the law, especially in the case of legal situations strongly depending on values and individual circumstances.⁵⁶⁰ Past-oriented decision-making is particularly problematic in the event of changes to the law. The reason is that decisions made on the basis of an old legal situation become incorrect as

⁵⁵³ In addition, Heetkamp and Piroutek (n 518), (2023) 26(3) ZKM (Zeitschrift für Konfliktmanagement) 80.

⁵⁵⁴ Anzinger (n 524), (2021) 24(2) ZKM (Zeitschrift für Konfliktmanagement) 53, 57: ‘Models can be found in systems such as A2JAuthor [<https://www.a2jauthor.org/>] or Law Lift [<https://de.lawlift.com/>>] and Smart Law [<https://www.smartlaw.de/>].’ – On consumer arbitration: Deichsel (n 523), (2020) 35(8) VuR (Verbraucher und Recht) 283, 288.

⁵⁵⁵ See below para 212-214.

⁵⁵⁶ See commentary to Guideline 1 of the Silicon Valley Arbitration & Mediation Center Guidelines, Draft of August 31, 2023 <https://thearbitration.org/wp-content/uploads/2023/08/SVAMC-AI-Guidelines-CONSULTATION-DRAFT-31-August-2023-1.pdf> accessed 31 December 2023.

⁵⁵⁷ Rühl (n 552) 617, 624 f para 15 f; M Scherer, ‘Artificial Intelligence and Legal Decision-Making’ (2019) 36(5) Journal of International Arbitration 539, 557.

⁵⁵⁸ In other words, simultaneous updating of training data is impossible, cf Meder (n 522), (2023) 78(23) JZ (Juristenzeitung) 1041, 1042, 1050.

⁵⁵⁹ Meder (n 522), (2023) 78(23) JZ (Juristenzeitung) 1041, 1047.

⁵⁶⁰ For the arbitration proceedings: D Lindquist and Y Dautaj, ‘AI in International Arbitration’ (2021) (1) Journal of Dispute Resolution 39, 54 f; M Fries, ‘Legal Tech im Schiedsverfahren’ in R Wilhelmi and M Stürner (ed), *Mehrparteienschiedsverfahren* (Springer 2021) 85, 93; see also D Nink, *Justice and Algorithms* (Duncker & Humoldt 2021), 230.

training data and therefore can no longer be used.⁵⁶¹ The ability of artificial intelligence to predict and adapt cannot compensate for the underlying technical problem,⁵⁶² although the low publication density in the area of conciliation reduces the severity of the problem.⁵⁶³

2.2.2.3 'Legal Remoteness' Decision-Making Standards and the Accumulation of Private Decision-Making Power

191 In light of this, trust and willingness to (fully) rely on the results of a machine in matters of dispute resolution are generally (still) not very pronounced.⁵⁶⁴ However, this limits the use of AI from the outset to disputes with more similar conflicts of interest that are easier to map technically in terms of the underlying evaluations.⁵⁶⁵ In addition, especially in the area of platform-based conciliation, decisions are often not made (strictly) on the basis of (and to enforce) substantive law:⁵⁶⁶ Normative standards⁵⁶⁷ are replaced by a company's own standards and general terms and conditions and thus rules being – with a view to the economic interests of the respective user – more focused on fulfilling customer satisfaction than on a finely balanced resolution of conflicts of interest through legal principles or on the execution of legally recognized needs.⁵⁶⁸ As a prominent example of this serves the platform-based dispute resolution of the payment service provider PayPal.⁵⁶⁹

⁵⁶¹ This problem arises both in continental European law and in case law, see Scherer (n 557), (2019) 36(5) *Journal of International Arbitration* 539, 557.

⁵⁶² Scherer (n 557), (2019) 36(5) *Journal of International Arbitration* 539, 557. An overview of the various models of artificial intelligence in the context of legal decision-making processes can be found in: *ibid*, 546 ff.

⁵⁶³ For the parallel issue of arbitration, see below para 215 and para 225.

⁵⁶⁴ S and H Kumar (n 545), (2021) 4(4) *International Journal of Law Management & Humanities* 1472, 1478.

⁵⁶⁵ Heetkamp and Piroutek (n 518), (2023) 26(3) *ZKM (Zeitschrift für Konfliktmanagement)* 80, 81.

⁵⁶⁶ B Gsell, 'Die Umsetzung der Richtlinie über alternative Streitbeilegung Juristisches Fachwissen der streitbeilegenden Personen und Rechtstreue des Verfahrensergebnisses' (2015) 128(2) *ZZP (Zeitschrift für Zivilprozess)* 189, 199 f; M Fries, *Verbraucherrechtsdurchsetzung* (Mohr Siebeck 2016), 245.

⁵⁶⁷ However, cf § 19 of the German Consumer Dispute Resolution Act (*Verbraucherstreitbeilegungsgesetzes: VSBG*): 'The conciliation proposal shall be based on the applicable law and shall in particular comply with the mandatory consumer protection laws'. See §§ 16, 17 of the German

⁵⁶⁸ W Voß, for example, attests to a considerable simplification of state law on service disruption and consumer protection, in Voß (n 537), (2020) 84(1) *RabelsZ (Rabels Zeitschrift für ausländisches und internationales Privatrecht)* 62, 66 f; see also Specht (n 546), (2019) 22(3) *MMR (Multimedia und Recht)* 153, 155; Anzinger (n 524), (2021) 24(2) *ZKM (Zeitschrift für Konfliktmanagement)* 53, 56.

⁵⁶⁹ If a conflict arises between merchant and customer, the employees of the PayPal group essentially 'decide' on the basis that money and goods must not be with the same person, see M Fries, 'PayPal Law und Legal Tech – Was macht die Digitalisierung mit dem Privatrecht?' (2016) 69(39) *NJW (Neue Juristische Wochenschrift)* 2860, 2861 f; also: J Adolphsen, 'Der Zivilprozess im Wettbewerb der Methoden' (2017) 48(4) *BRÄK Mitteilungen* 147, 149; Rühl (n 537), (2020) 75(17) *JZ (Juristenzeitung)* 809, 812. See, however, the decision of the German Federal Court of Justice (*Bundesgerichtshof*) on

- 192 Indeed, the findings of a certain ‘legal remoteness’,⁵⁷⁰ the simplicity of the applicable rules and regulations⁵⁷¹ as well as a ‘privatization of civil law’⁵⁷² (and its procedural implementation) are conceptually inherent to the dispute resolution forms of conciliation and – in particular – mediation and have therefore been known for a long time.⁵⁷³ However, the increased use of AI is accompanied by a shift in emphasis: away from consensual, individualized, case-by-case dispute resolution⁵⁷⁴ towards a more standardized, rule-based, scalable decision-making process in the context of a large number of comparable conflict situations. This, in turn, is associated with the risk of a ‘fossilization’ of dispute resolution, which pushes back the fundamentally necessary possibility of (spontaneous) legal development.⁵⁷⁵
- 193 Last but not least, the progressive absorption of proceedings by intermediaries acting as third-party adjudicators (PayPal, Amazon Resolution Center, etc) is leading to an accumulation of private decision-making power in an imperfectly regulated area of law.⁵⁷⁶ All of these developments threaten to weaken the acceptance of mediation and conciliation as a whole.

2.3 Procedural Answers

- 194 Setting a regulatory framework for the beneficial and risky use of AI is a task of the law that arises in the area of mediation and dispute resolution, too. In terms of implementation, it is (largely) similar to the approaches outlined in the platform-based dispute resolution and arbitration.

the PayPal Buyer Protection Directive: BGH (Germany) 22 November 2017, VIII ZR 83/16, (2018) 71(8) NJW (Neue Juristische Wochenschrift) 537 and VIII ZR 213/16, (2018) 21(3) MMR (Multimedia und Recht) 156.

⁵⁷⁰ See Wendland (n 548) 192-213.

⁵⁷¹ J Adolphsen even speaks of a ‘primitive legal system’: Adolphsen (n 569), (2017) 48(4) BRAK Mitteilungen 147, 149; of a ‘banalization of private law’: C Althammer, ‘Alternative Streitbeilegung im Internet’ in F Faust and H-B Schäfer (ed), *Zivilrechtliche und rechtsökonomische Probleme des Internet und der künstlichen Intelligenz* (Mohr Siebeck 2019), 249, 266.

⁵⁷² M Fries chooses the term ‘de facto privatization of civil law’, in Fries (n 569), (2016) 69(39) NJW (Neue Juristische Wochenschrift) 2860, 2860 f.

⁵⁷³ To this end: Althammer (n 571) 249, 260.

⁵⁷⁴ Thus, according to M Wendland, the result of mediation (as conventionally understood) is a singular product of the individual case and not a ‘party contract law’ or an autonomous private order, see Wendland (n 548) 196.

⁵⁷⁵ See H Prütting, ‘Das neue Verbraucherstreitbeilegungsgesetz: Was sich ändert – und was bleiben wird’ (2016) (3) AnwBl (Anwaltsblatt) 190, 192 f; Althammer (n 571) 249, 260 f. – The phenomenon of the ‘conservatism’ of machine-based decision-making cannot be addressed by publishing paradigmatic conflict cases. On the benefits of corresponding publications outside of AI-based decision-making, see Hess (n 329), (2015) 70(11) JZ (Juristenzeitung) 548, 553.

⁵⁷⁶ U Gläßer, ‘Mediation und Digitalisierung’ in T Riehm and S Dörr (ed), *Digitalisierung und Zivilverfahren* (De Gruyter 2023) 529, 531. On the power of platforms to exclude companies as well: Althammer (n 571), 249, 262; M Fries, ‘Erfüllung von Geldschulden über eigenwillige Zahlungsdienstleister’ (2018) 33(4) VuR (Verbraucher und Recht) 123, 124.

- 195 For one thing, this is linked to the need to create transparency with regard to the use of AI.⁵⁷⁷ The Artificial Intelligence Act recently adopted by the European Union⁵⁷⁸ sets out transparency rules for certain AI systems that are particularly susceptible to manipulation.⁵⁷⁹ These include, for example, AI systems intended for interaction with natural persons, such as chatbots in particular.⁵⁸⁰ The AI Act also standardizes prohibited practices, which may include nudging, dark patterns and behavioral microtargeting.⁵⁸¹ In this regard, certain obligations are imposed on providers of high-risk systems,⁵⁸² such as the creation of a quality management system or technical documentation for the high-risk AI system.⁵⁸³ According to the text of the regulation as amended by the EU Parliament, the information obligations extend to the legal protection afforded by procedures when using AI systems.⁵⁸⁴
- 196 For another thing, the establishment of ‘hybrid decision-making’ is being discussed: AI-based proposals for decisions – such as an out-of-court dispute resolution body in accordance with the DSA⁵⁸⁵ or the German UrhDaG⁵⁸⁶ – should never have binding effect without the consent of the parties to the proceedings. If the parties refuse to give their consent, it would be up to the individuals to make a binding decision on the subject

⁵⁷⁷ Heetkamp and Piroutek (n 518), (2023) 26(3) ZKM (Zeitschrift für Konfliktmanagement) 80, 81. For example, the Digital Services Act requires disclosure of (AI-based) decision parameters and statistical bases of dispute resolution, cf. on the one hand Art 17(3) lit d) DSA Regulation (reference to the legal basis and explanations as to why the information is considered unlawful content on this basis); according to lit c), the hosting service provider’s decision to impose usage restrictions must also indicate whether automated means were used to make the decision, including whether the decision was made in relation to content that was identified or detected by automated means. – Secondly, Art 24(1) DSA Regulation formulates periodic reporting obligations for online platform providers, including on the number, outcome and duration of disputes and dispute resolution.

⁵⁷⁸ See Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 12 July 2024. According to this, the AI Act is to apply alongside the Digital Services Act (DSA), see Art 2(5) AI Act.

⁵⁷⁹ Art 50 AI Act.

⁵⁸⁰ Also covered – although less relevant for the types of procedure examined here – are systems for biometric categorization, Art 50(3) AI Act and AI systems in connection with deep fakes, Art 50(2) AI Act.

⁵⁸¹ P Richter and J Mendelsohn, ‘§ 21 Plattformspezifische Vorgaben des Data Acts’ in B Steinrötter (ed), *Europäische Plattformregulierung* (Nomos 2023) para 17 ff (on the Proposal AI Act).

⁵⁸² Richter and Mendelsohn (n 581) para 31.

⁵⁸³ Art 16 AI Act.

⁵⁸⁴ Art 52(1), subpara 2 of the proposal amended by the EU Parliament: ‘Where appropriate and relevant, this information shall also include which functions are AI enabled, if there is human oversight, and who is responsible for the decision-making process, as well as the existing rights and processes that, according to Union and national law, allow natural persons or their representatives to object against the application of such systems to them and to seek judicial redress against decisions taken by or harm caused by AI systems, including their right to seek an explanation’. See Synopsis AI Act, Commission-Parliament.pdf (P9_TA(2023)0236), https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf accessed 31 December 2023.

⁵⁸⁵ Art 21 DSA Regulation.

⁵⁸⁶ See §§ 16, 17 of the German Copyright Service Provider Act (Urheberrechts-Diensteanbieter-Gesetz: UrhDaG) for out-of-court dispute resolution by private and official conciliation boards.

matter of the dispute.⁵⁸⁷ In the best-case scenario, this model promises to combine the advantages of human decision-making (especially in disputes with a strong focus on normative values and individual circumstances) with the (potential) cost and resource-related efficiency of automated decision-making systems – a pledge for efficiency that has to prove itself in particular when dealing with a large number of similar disputes (scaling of decisions).

- 197 Regarding the need for procedural minimum standards in the area of mediation and conciliation, please refer to the comments on platform-based dispute resolution.⁵⁸⁸
- 198 Both conciliation proceedings, in which a conciliator submits non-binding proposals for decisions to the parties, and mediation, which leaves the parties to resolve the conflict on their own responsibility, do not fulfil the definition of ADR proceedings that AI Act of the European Union establishes⁵⁸⁹ – due to the lack of binding effect of their decisions on the parties. Consequently, AI systems that a conciliator or mediator may use to assist in researching and interpreting facts and the law and in applying the law to a concrete set of facts are not to be classified as so-called high-risk AI systems within the meaning of the EU AI Act. Consequently, the providers and operators of such high-risk AI systems are not subject to the special obligations under Art 9-17 AI Act.⁵⁹⁰

3 ARBITRATION AND ARTIFICIAL INTELLIGENCE

3.1 Fields of Application of Artificial Intelligence in Arbitration Proceedings

- 199 Although the use of artificial intelligence in arbitration proceedings is still limited, the technology has been gaining importance in recent years.⁵⁹¹ Mainly because such use promises an increase in procedural efficiency, in particular a reduction in costs.⁵⁹² The main current fields of application of AI are the following:

⁵⁸⁷ Heetkamp and Piroutek (n 518), (2023) 26(3) ZKM (Zeitschrift für Konfliktmanagement) 80, 81, who consider forms of hybrid decision-making to be particularly suitable in the area of e-commerce. The reason for this is the high rate of automated dispute resolution that can be found there anyway.

⁵⁸⁸ See above para 108-175.

⁵⁸⁹ Art 6(2), Annex III no 8 lit a) AI Act. See in more detail above para 162 ff.

⁵⁹⁰ See in more detail above para 161 ff.

⁵⁹¹ In 2021, for example, 15% of respondents to the International Arbitration Survey stated that they regularly or frequently use artificial intelligence; Queen Mary University and White & Case, 'International Arbitration Survey: Adapting Arbitration to a Changing World' (2021) 21, https://www.qmul.ac.uk/arbitration/media/arbitration/docs/LON0320037-QMUL-International-Arbitration-Survey-2021_19_WEB.pdf accessed 31 December 2023.

⁵⁹² J Rajendra and A Thuraisingam, 'The deployment of artificial intelligence in alternative dispute resolution, the AI augmented arbitrator' (2022) 31(2) Information & Communications Technology Law 176-193.

3.1.1 Document Analysis and Case Management

3.1.1.1 Automated Document Analysis

- 200 Reflecting a general trend, AI is currently predominantly used in the field of automated document analysis.⁵⁹³ The focus lies on technologies for data evaluation and structuring.⁵⁹⁴ The fields of application are diverse: In addition to intelligent document searches (so-called smart searches)⁵⁹⁵ and the automated creation of a process history,⁵⁹⁶ this includes the structuring of party submissions, evidence and presentations of evidence⁵⁹⁷ as well as automated e-discovery.⁵⁹⁸ Use is also made of forecasting tools for the automated collection and analysis of market data (eg, in the energy sector) being relevant to the outcome of proceedings.⁵⁹⁹
- 201 Speech recognition and translation are further fields of application. With the help of AI tools, transcripts of verbal negotiations can be created automatically and in real time⁶⁰⁰ or translations be produced.⁶⁰¹
- 202 In regard, natural language processing (NLP),⁶⁰² optical character recognition (OCR)⁶⁰³ or predictive coding⁶⁰⁴ act as technical tools, although the use of AI is by no means

⁵⁹³ Queen Mary University and White & Case (n 591) 22.

⁵⁹⁴ M Scherer and O Jensen, 'Die Digitalisierung der Schiedsgerichtsbarkeit' in T Riehm and S Dörr (ed), *Digitalisierung und Zivilverfahren* (De Gruyter 2023) 591, 604 para 34. On such software solutions, see Rühl (n 552) 617, 618.

⁵⁹⁵ E Zorrilla, 'Towards a Credible Future' (2018) 16(2) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 106, 113; G Zekos, *Advanced Artificial Intelligence and Robo-Justice* (Springer 2022), 328. One provider of such software is DISCO <https://csdisco.com/offerings/review>.

⁵⁹⁶ Kreis (n 517) 633, 639 para 22.

⁵⁹⁷ Zorrilla (n 595), (2018) 16(2) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 106, 113; Zekos (n 595) 328; Eidenmüller and Wagner (n 194) 192.

⁵⁹⁸ L Bizikova, P Hancock, D Jewell and I Sherr, 'IA Meets AI' (2 October 2023) <https://dailyjus.com/legal-tech/2023/10/ia-meets-ai-rise-of-the-machines> accessed 31 December 2023; Eidenmüller and Wagner (n 194) 192 f; Zekos (n 595) 326. Existing software is, for instance, the e-discovery tool eBravia (<https://www.dfinsolutions.com/products/ebrevia>).

⁵⁹⁹ Zorrilla (n 595), (2018) 16(2) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 106, 113.

⁶⁰⁰ Existing tools include Trint (<https://trint.com/>), Fireflies (<https://firefliesai/>) and Otter (<https://otter.ai>).

⁶⁰¹ Kreis (n 517) 633, 639 para 22; Zekos (n 595) 326; Bizikova, Hancock, Jewell and Sherr (n 598), 'IA Meets AI' (2 October 2023).

⁶⁰² The research field of Natural Language Processing deals with the algorithmic processing of spoken and written language. This can be implemented using artificial intelligence, among other things, see F Deusch and T Eggendorfer, 'IT-Sicherheit' in J Taeger and J Pohle (ed), *Computerrechts-Handbuch* (38th edn, Beck 8/2023) para 232n.

⁶⁰³ This technology is used to recognize and convert text – such as a scan or photo of a typed or handwritten text – into a machine-readable format, see P von Bünnau, 'Künstliche Intelligenz im Recht' in S Breidenbach and F Glatz (ed), *Rechtshandbuch Legal Tech* (2nd edn, Beck/Manz 2021) 71 para 18; regarding arbitration: Eidenmüller and Wagner (n 194) 192.

⁶⁰⁴ Predictive coding – also known as technology-assisted review – involves (technically simplified) complex algorithms that are able to search and analyze large volumes of documents. This technology

mandatory. In individual cases, less technically complex software is also suitable, for example for automated keyword searches.⁶⁰⁵

3.1.1.2 Automated Document Creation and Advanced Analysis Tools

203 As in the legal services market,⁶⁰⁶ AI-driven research and analysis of case law may also take place in arbitration proceedings.⁶⁰⁷ For example, tools are already being used to create summaries of arbitral awards and court judgments based on a database of international law and arbitration law documents.⁶⁰⁸ Another field of application is the automated creation of drafts, such as pleadings (pre-drafting).⁶⁰⁹ Similarly, artificial intelligence can be used to review documents in order to avoid legal and factual errors.⁶¹⁰

3.1.1.3 (Other) AI-Supported Case Management

204 It is also possible to outsource parts of the organization of the procedure to artificial intelligence. For instance, AI systems are capable of checking deadlines⁶¹¹ or schedule meetings and hearings automatically (smart scheduling).⁶¹² The same applies to checks on whether an arbitration award meets certain formal requirements, such as the signature of all arbitrators.⁶¹³ In addition, the use of so-called *smart personal assistants*

is primarily used in e-discovery, see in detail C Yablon and N Landsman-Ross, 'Predictive Coding' (2013) 64(3) South Carolina Law Review 633, 643, 638; regarding arbitration: Zorrilla (n 595), (2018) 16(2) SchiedsVZ (Zeitschrift für Schiedsverfahren) 106, 111; corresponding technology is used, for example, by the service provider DISCO (<https://www.csdisco.com/offers/ediscovery/features-ai>).

⁶⁰⁵ Scherer and Jensen (n 594) 591, 604 para 34; Rühl (n 552) 617, 620 para 6.

⁶⁰⁶ Examples include Westlaw Edge <https://legal.thomsonreuters.com/en/products/westlaw-edge> and Lexis+ <https://www.lexisnexis.com/en-us/products/lexis-pluspage>.

⁶⁰⁷ S Barona Vilar, 'Effizienzsteigerung und Suche nach Beschleunigung von Schiedsverfahren im Spannungsfeld von Mythos, Sublimierung und Vierter Industrieller Revolution (4.0)' (2019) 23 ZZPInt (Zeitschrift für Zivilprozess International) 295, 313; Bizikova, Hancock, Jewell and Sherr (n 598), 'IA Meets AI' (2 October 2023).

⁶⁰⁸ Eg, the 'Jus-AI' of the provider JusMundi <https://jusmundi.com/en>; Daily Jus, 'Jus Mundi Introduces Jus-AI' (29 June 2023) <https://dailyjus.com/news/2023/06/jus-mundi-introduces-jus-ai-a-game-changing-gpt-powered-ai-solution-for-the-arbitration-community> accessed 31 December 2023; overview in Bizikova, Hancock, Jewell and Sherr (n 598), 'IA Meets AI' (2 October 2023).

⁶⁰⁹ Zekos (n 595) 330. – On the pre-drafting of court decisions, see Fries (n 569), (2016) 69(39) NJW (Neue Juristische Wochenschrift) 2860, 2864.

⁶¹⁰ Zorrilla (n 595), (2018) 16(2) SchiedsVZ (Zeitschrift für Schiedsverfahren) 106, 113. Specific providers for the review of documents in arbitration proceedings cannot be found. Technically comparable software for contract drafting is, for example, Luminance (<https://www.luminance.com/overview.html>).

⁶¹¹ Kreis (n 517) 633, 639 para 23.

⁶¹² Rajendra and Thuraisingam (n 592), (2022) 31(2) Information & Communications Technology Law 176, 183 f; Zekos (n 595) 325 f.

⁶¹³ Zorrilla (n 595), (2018) 16(2) SchiedsVZ (Zeitschrift für Schiedsverfahren) 106, 116; Kreis (n 517) 633, 639 para 23.

is conceivable, which are tailored to the respective circumstances of the individual arbitration proceedings.⁶¹⁴

3.1.1.4 Legal Admissibility

205 Asking about the admissibility of such use of AI,⁶¹⁵ the supreme personality of the arbitrator's mandate⁶¹⁶ must be observed when making use of advanced software. As a consequence, the outsourcing of an arbitrator's core tasks without the consent of the parties proves to be inadmissible.⁶¹⁷ A problem arises in particular for auxiliary activities such as the summarization of factual sequences and expert opinions, the review of other documents or the examination of evidence, which arises in a comparable way when deploying tribunal secretaries in arbitration proceedings.⁶¹⁸ In order to avoid an arbitral award being set aside, it is advisable to contractually regulate the involvement of AI in parallel to the involvement of tribunal secretaries.⁶¹⁹

3.1.2 Selection of the Referee Using AI

206 Another auxiliary function of AI is the selection of referees. This is conceivable on the basis of an analysis of specialist databases⁶²⁰ or a broader internet search,⁶²¹ combined with the creation of a ranking list of the candidates found. In that regard, the AI is technically able to consider the parameters specified by the parties and, if necessary, considering conflicting criteria, eg, by implementing an expert system.⁶²²

⁶¹⁴ Eidenmüller and Wagner (n 194) 191 f.

⁶¹⁵ C Aschauer, 'Automated Decision-Making and Artificial Intelligence (AI) in Arbitration' in C Leyens, I Eisenberger and R Niemann (ed), *Smart Regulation* (Mohr Siebeck 2021) 130, 133.

⁶¹⁶ See H Prütting, 'Die rechtliche Stellung des Schiedsrichters' (2011) 9(5) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 233, 235.

⁶¹⁷ Specifically on the use of AI: Kreis (n 517) 633, 644 para 46-50 in relation to 'producing', not merely 'examining tasks' of the arbitrator. – On the supreme personal nature of the mandate: Prütting (n 616), (2011) 9(5) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 233, 235 – On the admissibility of delegation to auxiliary persons: O Jensen, *Tribunal Secretaries in International Arbitration* (Oxford University Press 2019) para 805 ff.

⁶¹⁸ Zorrilla (n 595), (2018) 16(2) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 106, 113; Kreis (n 517) 633, 644 para 46-50. – On the use of tribunal secretaries, see Sim (n 551), (2018) 14(1) *Asian International Arbitration Journal* 1, 6; Jensen (n 617) para 805 ff; M Polkinghorne, 'Different Strokes for Different Folks?' *Kluwer Arbitration Blog* (16 May 2014), <https://arbitrationblog.kluwerarbitration.com/2014/05/17/different-strokes-for-different-folks-the-role-of-the-tribunal-secretary-2/> accessed 31 December 2023.

⁶¹⁹ Kreis (n 517) 633, 644 para 50.

⁶²⁰ One example is *Kluwer Arbitration* <https://www.kluwerarbitration.com/> accessed 31 December 2023.

⁶²¹ The subject of research is, for example, data that arbitrators voluntarily or involuntarily leave behind in social media, for example on family circumstances, political inclinations or general sensitivities, see in more detail Aschauer (n 615) 130, 135 f.

⁶²² Y Rhim and K Park, 'The Artificial Intelligence in International Law' in E Y J Lee (ed), *Revolutionary Approach to international Law: The Role of international Lawyer in Asia* (Springer 2023) 215, 224 f.

207 However, implementation problems arise in view of the extremely incomplete data basis.⁶²³ Although there are some databases specializing in the evaluation of arbitrators and arbitration institutions,⁶²⁴ their analytical capabilities remain deficient as only providing an incomplete and generally unreliable basis for suitable use as training data.

3.1.3 Technologies for Risk Assessment and Prediction of Process Outcomes (Predictive Analytics)

208 Another area of application for AI is risk assessment aiming at a prediction of the outcome of proceedings⁶²⁵ or even the behavior of individual arbitrators⁶²⁶ (predictive analytics). This puts the parties in a position to assess their chances of success (supposedly more reliably), or at least use the AI prediction as a basis for settlement negotiations.⁶²⁷ This technology shows to be of great interest for the rapidly growing market of litigation financing in arbitration proceedings, hoping for a statistically improved prediction of the outcome of proceedings and litigation risks.⁶²⁸

209 The prediction process is of a statistical nature, ie, not a cognitive process involving the application and analysis of relevant legal rules.⁶²⁹ Therefore, the decisions are obtained inductively from a large amount of historical data on decisions and cases. The AI system searches for patterns or correlations relevant to the outcome of the decision. Suitable input data is essentially metadata (metadata analysis)⁶³⁰ and factual data⁶³¹ (factual

⁶²³ The lack of publication of arbitration awards makes it difficult to draw conclusions about the working methods and 'habitus' of the arbitrators.

⁶²⁴ The Arbitrator Intelligence project, for example, has created an evaluation database for arbitrators (based on questionnaires and crowd-sourced arbitration awards). However, the initiator of the project, Professor Catherine Rogers, herself pointed out that the project does not yet have a sufficient data basis for processing by means of machine learning; see in detail Zorrilla (n 595), (2018) 16(2) *SchiedsVZ* (*Zeitschrift für Schiedsverfahren*) 106, 111. – Other databases are: Global Arbitration Review Arbitrator GAR ART, <https://globalarbitrationreview.com/tools/arbitrator-research-tool> accessed 31 December 2023 or Jus mundi, <https://jusmundi.com/en> accessed 31 December 2023, see Aschauer (n 615) 130.

⁶²⁵ Scherer and Jensen (n 594) 591, 615 para 61; Rühl (n 552) 617, 619, para 6-11; Zekos (n 595) 329 f; L Bull and F Steffek, 'The Decoding of Legal Conflicts' (2018) 21(5) *ZKM* (*Zeitschrift für Konfliktmanagement*) 165, 166.

⁶²⁶ Bizikova, Hancock, Jewell and Sherr (n 598), 'IA Meets AI' (2 October 2023); Aschauer (n 615) 130, 135; Barona Vilar (n 607), (2019) 23 *ZZPInt* (*Zeitschrift für Zivilprozess International*) 295, 314.

⁶²⁷ Scherer and Jensen (n 594) 591, 615 para 61.

⁶²⁸ On the increasing importance of litigation funding in arbitration proceedings, see only S Wilske, L Markert and B Ebert, 'Entwicklungen in der internationalen Schiedsgerichtsbarkeit im Jahr 2022 und Ausblick auf 2023' (2023) 21(3) *SchiedsVZ* (*Zeitschrift für Schiedsverfahren*) 121, 125.

⁶²⁹ Scherer and Jensen (n 594) 591, 615 para 61.

⁶³⁰ Predictive tools that work on the basis of metadata analysis include: LexMachina of patent disputes <https://lexmachina.com/> and Predictice <https://predictice.com>, all accessed 31 December 2023. Another well-known example is the study on the prediction of decisions of the US Supreme Court, see T Ruger, P Kim, A Martin and K Quinn, 'The Supreme Court Forecasting Project' (2004) 104(4) *Colum. L. Rev.* 1150, 1163 ff.

⁶³¹ The factual data analysis is based on a comparison of the facts of the case with the facts of relevant preliminary decisions, see Rühl (n 552) 617, 621 f para 9 f.

data analysis),⁶³² being technically implemented by using the random forest method⁶³³ or natural language processing.⁶³⁴

210 From a legal perspective, the use of such predictive analytics – still being unregulated in arbitration proceedings⁶³⁵ – is not uncontroversial. Nevertheless, the use of such tools raises concerns regarding the independence and impartiality of arbitrators, especially when an arbitrator is appointed on the basis of decision predictions.⁶³⁶ This raises the intricate question of whether any duty of investigation on the part of the arbitrator – resulting from an obligation of disclosure⁶³⁷ – with regard to conflicts of interest⁶³⁸ also extends to an investigation into the use of such predictive tools by the nominating party or a third-party financier.⁶³⁹ It is equally questionable whether an arbitrator can be challenged because the prognosis shows with overwhelming probability that the case can only be decided by him with a certain procedural outcome. In line with the ‘judge scoring’ now prohibited by law in France,⁶⁴⁰ there are some calls for a complete ban on predictive analytics in arbitration proceedings.⁶⁴¹

3.1.4 Fully Automated Process Management and Decision-Making (‘Robo Arbitrator’)

211 Fully automated arbitration procedures in which decision-making is carried out exclusively by AI are still far beyond the current state of the Art Nevertheless, a brief analysis is already worthwhile at this stage, not least due to the rapid pace of

⁶³² Rühl (n 552) 617, 620 para 6. On predictive tools that work on the basis of factual data analysis, see also the study on the prediction of decisions of the ECtHR: M Medvedeva, M Vols and M Wieling, ‘Using machine learning to predict decisions of the European Court of Human Rights’ (2020) 28(2) *Artificial Intelligence and Law* 237, 266; see also the study on predicting decisions of the Financial Ombudsman in the UK (Case Cruncher Alpha): R Cellan-Jones, ‘The robot lawyers are here’ (1 November 2017) BBC News, <https://www.bbc.com/news/technology-41829534> accessed 31 December 2023.

⁶³³ The random forest method is used in the context of machine learning. This is a combination of decision trees (random forest): Each tree depends on the values of a random vector that is determined independently and with the same distribution for all decision trees in the forest, see L Breiman, ‘Random Forests’ (2001) 45(1) *Machine Learning* 5 ff.

⁶³⁴ Deichsel (n 519), 98 f.

⁶³⁵ This applies both to the use by parties and litigation funders as well as by the arbitrator himself, see S Marmont, ‘Keeping Up with Legal Technology’ (2019) 1(2) *ITA in Review* 37, 41, 48.

⁶³⁶ Aschauer (n 615) 130, 137.

⁶³⁷ See J Lew, L Mistelis and S Kröll, *Comparative International Commercial Arbitration* (Kluwer Law International 2003) 265.

⁶³⁸ On the existence and scope of this obligation to investigate and the consequences of its violation, see Lew, Mistelis and Kröll (n 637) 269.

⁶³⁹ The scope of this obligation to investigate is limited by the principle of reasonableness. For the whole, see S Marmont (n 635), (2019) 1(2) *ITA in Review* 37, 41, 47 f.

⁶⁴⁰ Art 33 Law of March 23, 2019 (Loi de programmation 2018-2022 et réforme pour la justice) no 2019/222, available at <https://www.legifrance.gouv.fr/> accessed 31 December 2023. In that regard, the French legislator has generally prohibited the use of data on the identity of judges for the evaluation, analysis, comparison or prediction of their decisions.

⁶⁴¹ Aschauer (n 615) 130, 137.

technological progress.⁶⁴² Fully automated decision-making raises both technical and legal concerns:

3.1.4.1 Computability of Law?

- 212 Given that decision-making by AI is purely based on pattern recognition and statistical methods, rather than of a legal methodological subsumption act,⁶⁴³ AI is not capable of examining and deciding complex legal situations – which are often the subject of arbitration proceedings.⁶⁴⁴ As a result, a purely AI-based decision-making process would only be suitable for disputes that are uniform in nature and based on easily verifiable facts – often (but not necessarily) with a low value in dispute (small claims). Examples include flight delays⁶⁴⁵ or consumer arbitration, being widespread in the USA.
- 213 The already established lack of AI systems to make human assessments⁶⁴⁶ and show emotions and empathy⁶⁴⁷ is even more pronounced in arbitration proceedings, where there is a much greater need for individual case and fairness decisions.⁶⁴⁸ Albeit existing AI being capable of recognizing emotions based on facial expressions and gestures, this technology is still at a rudimentary level. In addition, these parameters would have to be included in the machine decision-making process, thus posing considerable technical difficulties.⁶⁴⁹
- 214 In this light, the replacement of referees by AI is not even remotely foreseeable at the present time.

⁶⁴² This is also the case with Kreis (n 517) 633, 647 para 62.

⁶⁴³ J Schwartz, 'Artificial Arbitration?' in R Wilhelmi and M Stürner (ed), *Mehrparteien-Schiedsverfahren: Unter besonderer Berücksichtigung gesellschaftsrechtlicher Streitigkeiten* (Springer 2021) 95, 120 f.

⁶⁴⁴ Bizikova, Hancock, Jewell and Sherr (n 598), 'IA Meets AI' (2 October 2023); different, albeit without further justification: Eidenmüller and Wagner (n 194), 203.

⁶⁴⁵ M Kaulartz, 'Smart Contract Dispute Resolution' in M Fries and B-P Paal (ed), *Smart Contracts* (Mohr Siebeck 2019) 73, 80 f.

⁶⁴⁶ Interesting here is ChatGPT's answer to the question 'Can you act as an arbitrator in an arbitration?': 'No, I cannot act as an arbitrator in an arbitration. The role of an arbitrator requires a specific set of skills, experience and qualifications that include human characteristics and judgment. As an AI model, I lack these qualifications and the ability to make human decisions' (as of 8 December 2023).

⁶⁴⁷ See above para 189-190. On arbitration proceedings: Lindquist and Dautaj (n 560), (2021) (1) *Journal of Dispute Resolution* 39, 48 f, 51 ff; Vannieuwenhuysen (n 551), (2018) 35(1) *Journal of International Arbitration* 119, 124; G Halis Kasap, 'Can Artificial Intelligence ("AI") Replace Human Arbitrators?' (2021) (2) *Journal of Dispute Resolution* 209, 232 ff; Nink (n 560), 231 ff; Schwartz (n 643), 95, 122 f; Bizikova, Hancock, Jewell and Sherr (n 598), 'IA Meets AI' (2 October 2023); Aschauer (n 615) 130, 134; Kreis (n 517) 633, 648 para 67 ff.

⁶⁴⁸ Schwartz (n 643) 95, 122 f; Aschauer (n 615) 130, 134; Lindquist and Dautaj (n 560), (2021) (1) *Journal of Dispute Resolution* 39, 48 f, 51 ff.

⁶⁴⁹ Lindquist and Dautaj (n 560), (2021) (1) *Journal of Dispute Resolution* 39, 49.

3.1.4.2 The problem of obtaining training data

215 Another technical barrier is the acquisition of training data. Due to the low publication density of arbitration awards, there is little case material available.⁶⁵⁰ Indeed, publicly accessible data sets already exist, including, *inter alia*, the decisions of the International Center for Settlement of Investment Disputes (ICSID), the Society of Maritime Lawyers (SMA) and the Court of Arbitration for Sport (CAS).⁶⁵¹ By the same token, organizations have developed databases containing decisions and other data sets (such as interviews with the parties to the arbitration proceedings).⁶⁵² Regardless of whether a trend towards the increasing publication of arbitration awards can actually be observed,⁶⁵³ existing data is only suitable for AI-supported analysis to a limited extent due to its incompleteness, eg, with respect to a lack of (exhaustive) reasons and the names of the parties involved, including the arbitrator.⁶⁵⁴

3.1.4.3 Legal assessment of autonomous decision-making by AI

216 A fully AI-driven decision-making process also raises considerable objections from a legal perspective. First of all, there is no question that outsourcing decision-making authority to AI systems as part of an arbitrator's personal mandate would require consent by party agreement.⁶⁵⁵ However, even beyond this consent requirement, the question of legal conformity of extensively automated arbitral awards arises. Both international arbitration law and the vast majority of national legal systems assume the requirement that only natural persons or persons with legal personality are suitable as arbitrators. Furthermore, concerns are raised regarding fundamental procedural guarantees – namely procedural equality of arms, the granting of a fair hearing as well as the independence of arbitrators and, not least, data protection.

⁶⁵⁰ Rhim and Park (n 622) 215, 225; Zekos (n 595) 329; Schwartz (n 643) 95, 123; Halis Kasap (n 647), (2021) 2 Journal of Dispute Resolution 209, 221 ff.

⁶⁵¹ J Münch in *Münchener Kommentar zur ZPO* (6th edn, Beck 2022) vor § 1025, para 5, 127; Rhim and Park (n 622) 215, 225; K Paisley and E Sussman, 'Artificial Intelligence Challenges and Opportunities for International Arbitration' (2018) 11(1) New York Dispute Resolution Lawyer 35, 37.

⁶⁵² Cf for example: Arbitrator Intelligence <https://arbitratorintelligence.vercel.app/> accessed 31 December 2023, see P Shaughnessy and C Rogers, 'Arbitrator Intelligence – An Interview with its Founder and Director, Professor Catherine Rogers' (2015) 87 Journal on Technology in International Arbitration 87, 96; Dispute Resolution Data <https://www.disputeresolutiondata.com/> accessed 31 December 2023; Global Arbitration Review Arbitrator Research Tool (GAR ART, <https://globalarbitrationreview.com/tools/arbitrator-research-tool> accessed 31 December 2023). On the whole: Paisley and Sussman (n 651), (2018) 11(1) New York Dispute Resolution Lawyer 35, 38; Rhim and Park (n 622) 215, 225 f.

⁶⁵³ Eidenmüller and Wagner (n 194) 202 f.

⁶⁵⁴ Rhim and Park (n 622) 215, 226. – Regarding data protection concerns, cf Paisley and Sussman (n 651), (2018) 11(1) New York Dispute Resolution Lawyer 35, 38.

⁶⁵⁵ Kreis (n 517) 633, 646-648 para 61 ff, 76; Kaulartz (n 645) 73, 80; Halis Kasap (n 647), (2021) 2 Journal of Dispute Resolution 209, 237 ff.

3.1.4.3.1 International arbitration law: UNCITRAL Model Law and New York Convention

217 While the UNCITRAL Model Law (arguably) assumes that the arbitrator must be a natural person,⁶⁵⁶ there is disagreement about the admissibility of fully automated decision-making systems under the New York Convention on the Recognition and Enforcement of Foreign Arbitral Awards. The wording of the Convention (Art I no 2 NYC) only refers to ‘arbitrators’ as such.⁶⁵⁷ It has occasionally been deduced from this that an arbitrator does not necessarily have to be a natural person, thereby alluding to the purpose of the Convention, ie the cross-border recognition and enforcement of arbitral awards.⁶⁵⁸ With reference to Art IV no 1 lit a) NYC), however, this view is rejected.⁶⁵⁹

3.1.4.3.2 National arbitration law

218 Some legal systems expressly ask for a natural person as arbitrator: for example, French,⁶⁶⁰ Dutch,⁶⁶¹ Spanish⁶⁶² and Turkish⁶⁶³ arbitration law, as well as Peru, Brazil and Ecuador.⁶⁶⁴ Based on the personal mandate of arbitrators, this also corresponds to the prevailing view in German literature.⁶⁶⁵ Correspondingly, some arbitration rules of the major arbitration institutions assume that a natural person always conducts and decides

⁶⁵⁶ Cf Art. 11, 12(1) UNCITRAL Model Law. Agreeing: Rhim and Park (n 622) 215, 225; Eidenmüller and Wagner (n 194) 215.

⁶⁵⁷ Zekos (n 595) 340; H Snijders, *Arbitration and AI*, *Arbitration* (1st edn, Wolters Kluwer 2023) 224, 234 ff.

⁶⁵⁸ Eidenmüller and Wagner (n 194) 209 f; Rhim and Park (n 622) 215, 225; Zekos (n 595) 381 f; Halis Kasap (n 647), (2021) (2) *Journal of Dispute Resolution* 209, 237.

⁶⁵⁹ Scherer and Jensen (n 594) 591, 617 f.

⁶⁶⁰ Art 1450(1) Code de procédure civile: ‘La mission d’arbitre ne peut être exercée que par une personne physique jouissant du plein exercice de ses droits’. See M Scherer, ‘International Arbitration 3.0. How Artificial Intelligence Will Change Dispute Resolution’ in C Klausegger et al (ed), *Austrian Yearbook on International Arbitration* (1st edn, Beck 2019) 503, 512 fn 22; Bizikova, Hancock, Jewell and Sherr (n 598), ‘IA Meets AI’ (2 October 2023).

⁶⁶¹ Art 11.1 s 1 NAI-SchO (Schiedsordnng Nederlands Arbitrage Instituut: Arbitration Rules of the Netherlands Arbitration Institute): ‘Any natural person [‘natuurlijke persoon’] of legal capacity may be appointed as arbitrator’.

⁶⁶² Art 13 Spanish Arbitration Act: ‘All natural persons in full possession of their civil rights may act as arbitrators, provided that they are not restricted by the legislation applicable to them in the exercise of their profession’.

⁶⁶³ Turkish International Arbitration Law, Article 7(B)(I): ‘Only natural persons can be selected as arbitrators’.

⁶⁶⁴ G Maxwell and G Vannieuwenhuysse, ‘Robots Replacing Arbitrators: Smart Contract Arbitration’ (2018) (1) *ICC Dispute Resolution Bulletin* 24, 31.

⁶⁶⁵ J Münch in *Münchener Kommentar zur ZPO* (6th edn, Beck 2022) vor § 1034 para 18-21; vor § 1025 para 5; § 1025 para 10. – The opposing view, which allows the parties to appoint AI as arbitrator, also places party autonomy under the proviso that basic procedural guarantees are preserved. It is based on state protection obligations and barriers that claim to be valid both in the interests of the parties and the general public, see Kaulartz (n 645) 73, 80 f; J Münch in *Münchener Kommentar zur ZPO* (6th edn, Beck 2022) vor § 1025 para 6.

the proceedings.⁶⁶⁶ Other legal systems explicitly require the legal capacity⁶⁶⁷ of arbitrators, including Sweden,⁶⁶⁸ Italy⁶⁶⁹ and England⁶⁷⁰.

3.1.4.3.3 Formal and data protection hurdles

- 219 For recognition, Art IV no 1 lit a) NYC requires a certified original or certified copy of the arbitral award. Whether electronic documents are sufficient in this respect is governed by national law.⁶⁷¹ In this light, there is legal uncertainty as to whether electronic arbitral awards may be subject to recognition at all.⁶⁷² German arbitration law, for example, requires arbitral awards to be signed.⁶⁷³ Whether this written form requirement can be satisfied electronically and therefore also by a ‘robo-arbitrator’ is certainly doubtful, given the different regulatory functions of the form requirement.⁶⁷⁴
- 220 The requirements of the General Data Protection Regulation (GDPR) also apply to the use of AI in arbitration proceedings.⁶⁷⁵ According to Art 22(1) GDPR, fully automated decisions are prohibited in arbitration proceedings unless the data subject has given their explicit consent.⁶⁷⁶ From a data protection perspective, it would therefore make sense to include the use of fully automated decision-making systems in a party agreement.

⁶⁶⁶ Art 13(1) ICC Arbitration Rules 2021, referring to the nationality of the arbitrator; furthermore, Art. 16(1) of the Vienna Rules 2018, which refers to legal capacity, see Aschauer (n 615) 130, 133.

⁶⁶⁷ With regard to a possible legal capacity of the arbitrator, it is proposed, in line with the discussion held in the EU Parliament in 2017, to provide automated systems with legal capacity (‘e-personality’) or to allow the fully automated management of legal entities (‘self-driving corporation’), so: Eidenmüller and Wagner (n 194) 201 f, 157 ff.

⁶⁶⁸ Section 7 LoS: ‘Var och en som råder över sig själv och sin egendom’ [Anyone who has legal capacity]), cf J Münch in *Münchener Kommentar zur ZPO* (6th edn, München 2022) vor § 1025 para 5 fn 13.

⁶⁶⁹ Art 812(1) Codice di procedura civile: ‘La norma in analisi indica il requisito fondamentale di capacità degli arbitri, ovvero il pieno possesso della piena capacità legale di agire’.

⁶⁷⁰ Section 26(1) of the English Arbitration Act 1996: ‘The authority of the arbitrator is personal and ceases on his death’.

⁶⁷¹ Kreis (n 517) 633, 694 para 72.

⁶⁷² Scherer and Jensen (n 594) 591, 618.

⁶⁷³ § 1054(1) ZPO (German Code of Civil Procedure).

⁶⁷⁴ Such as a clarification, evidence, authentication and concluding function. – Affirmative: Kreis (n 517) 633, 648 f para 70 f.

⁶⁷⁵ Maxwell and Vannieuwenhuyse (n 664), (2018) (1) ICC Dispute Resolution Bulletin 24, 30; generally on data protection in arbitration, see A Cervenka and P Schwarz, ‘Data Protection in Arbitration Proceedings’ (2020) 18(2) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 78, 79 f; G Fritz, D Prantl, N Leinwather and M Hofer, ‘Data Protection in International Arbitration Proceedings’ (2019) 17(6) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 301, 302 f; C Boll-Kempelmann, ‘Data protection and the evidence procedure in arbitration proceedings’ (2022) 20(5) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 241.

⁶⁷⁶ Art 22(2) lit c) GDPR. On the judicial procedure, see Nink (n 560) 251 ff.

3.1.4.3.4 Outlook: Blockchain arbitration

- 221 An alternative enforcement regime⁶⁷⁷ is blockchain arbitration.⁶⁷⁸ Such decentralized dispute resolution procedures⁶⁷⁹ are characterized by the fact that the contracting parties register the contract in dispute on a blockchain-based platform and deposit a security amount there, which corresponds to the amount in dispute.⁶⁸⁰ The dispute is decided according to the majority principle by a collective of human jurors selected by algorithm, but without them communicating with each other or with the contracting parties. Subsequently, only the jurors who decide in accordance with the majority principle receive remuneration and the decision is implemented immediately through the automated distribution of the deposited contributions.
- 222 Such dispute resolution mechanisms are often referred to as ‘arbitration’. However, it is very doubtful whether this classification is correct from the perspective of state law and whether it could therefore have binding and blocking effects on state court proceedings. One argument against this is that the current procedures probably do not meet the minimum legal requirements in several aspects.⁶⁸¹ Therefore, such proceedings may not replace but rather complement arbitration proceedings.⁶⁸²

3.2 Risks Associated with the Use of Artificial Intelligence in Arbitration Proceedings

3.2.1 Unconscious Bias and Risk of Discrimination

- 223 Using AI merely in a supportive way harbors the risk of unconsciously influencing human arbitrators. Such a scenario is conceivable in the form of a cognitive bias to follow the

⁶⁷⁷ Explicitly Zekos (n 595) 331 ff.

⁶⁷⁸ In this context, enforcement is self-executing, meaning that legal conformity with the New York Convention and other arbitration law is irrelevant for the enforcement of the arbitral award. Generally with regard to blockchain arbitration: T Kindt, ‘Blockchainbasierte dezentrale Streitbeilegungsverfahren und ihr Verhältnis zur Schiedsgerichtsbarkeit’ (2023) 21(5) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 241 ff.; G Wagner, *Legal Tech und Legal Robots* (2nd edn, Springer 2020) 34 f.

⁶⁷⁹ One provider of such procedures is the company Kleros, offering peer-to-peer arbitration proceedings for small claims, as well as in e-commerce, IP law and insurance law <https://kleros.io/en/> accessed 31 December 2023; another one is Aragon (‘Aragon Court’).

⁶⁸⁰ On this and in the following: Kindt (n 678), (2023) 21(5) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 241, 245 f.

⁶⁸¹ First of all, judicial independence and impartiality are questionable, as there is in fact no possibility of reviewing the anonymous jurors. Furthermore, their remuneration is directly linked to the outcome of the proceedings. Also, the blockchain-based procedure probably does not meet the right to a fair hearing due to lacking opportunities for the parties to express themselves after the proceedings have been initiated. Finally, doubts about the legal form of the decision arise in view of a lack of uniform decision-making standards. To all these aspects, see Kindt (n 678), (2023) 21(5) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 241, 246, 248-251.

⁶⁸² Such blockchain procedures may complement arbitration proceedings by either preceding them (eg within the framework of an escalation clause) or being integrated into them, see Kindt (n 678), (2023) 21(5) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 241, 252 f.

results of automated reviews ('anchor-effect').⁶⁸³ In addition, there is a risk of algorithmic discrimination in arbitration proceedings, too, eg, on the basis of gender, ethnicity or age.⁶⁸⁴ The technical cause of this can be the training data, actions of the self-learning system in its environment, but also deliberate manipulation.⁶⁸⁵ The frequently criticized investor-friendliness of arbitration tribunals may serve as a concrete example in the area of investment arbitration. Such a bias in arbitral awards threaten to continue in their future use as training data in the AI model.⁶⁸⁶ Both risks, anchor effects as well as discrimination, structurally impair the independence of human arbitrators.⁶⁸⁷

3.2.2 The Black Box Problem

224 A lack of transparency and traceability of automated decision-making become no less virulent in arbitration proceedings. As a matter of principle, the operation and results of AI systems are not comprehensible to users, nor (usually) to developers of the system.⁶⁸⁸ However, if the black box problem prevents sufficient explanation and justification of decisions,⁶⁸⁹ the acceptance of arbitral awards dwindles,⁶⁹⁰ possibly the legitimacy of the decision-making body as such,⁶⁹¹ but also the behavior-guiding effect of decisions on the parties.⁶⁹² The relevance of a statement of reasons for decisions follows from the fact that parties are, statistically speaking, rarely willing to dispense a statement of reasons or to allow an arbitrator to decide *ex aequo et bono*.⁶⁹³ Furthermore, due to the lack of normative reasoning, AI-based decisions are not suitable for use as a reference by other (arbitration) courts.⁶⁹⁴ In contrast to decisions by state courts, however, this problem is likely to arise less frequently.

⁶⁸³ Generally: D Kahneman, *Thinking, Fast and Slow* (1st edn, Penguin 2013) 119; Scherer (n 557), (2019) 36(5) *Journal of International Arbitration* 539, 557-562. In the context of arbitration: Zorrilla (n 595), (2018) 16(2) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 106, 113; on this point.

⁶⁸⁴ Rhim and Park (n 622) 215, 225. – On the technical foundations of the discrimination problem, see Ebers (n 191) 75 para 101-138; Scherer (n 660) 503, 510; Sim (n 551), (2018) 14(1) *Asian International Arbitration Journal* 1, 7 ff; Halis Kasap (n 647), (2021) (2) *Journal of Dispute Resolution* 209, 225 ff.

⁶⁸⁵ Scherer (n 557), (2019) 36(5) *Journal of International Arbitration* 539, 561.

⁶⁸⁶ The empirical situation is not clear. On the whole: Scherer (n 557), (2019) 36(1) *Journal of International Arbitration* 539, 559-561.

⁶⁸⁷ On the neutrality requirement: Kreis (n 517) 633, 645 f para 54-57.

⁶⁸⁸ Zorrilla (n 595), (2018) 16(2) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 106, 112 ff; Scherer (n 660) 503, 511; Halis Kasap (n 647), (2021) (2) *Journal of Dispute Resolution* 209, 229 ff.

⁶⁸⁹ Sim (n 551), (2018) 14(1) *Asian International Arbitration Journal* 1, 8 f; Vannieuwenhuyse (n 551), (2018) 35(1) *Journal of International Arbitration* 119, 124.

⁶⁹⁰ On judicial proceedings: Rühl (n 552) 617, 627 para 20; Scherer (n 660) 503, 511; Scherer (n 557), (2019) 36(5) *Journal of International Arbitration* 539, 562.

⁶⁹¹ It is well known that trust in the expertise, reputation and personality of an arbitrator is particularly relevant. On the whole: Scherer (n 557), (2019) 36(5) *Journal of International Arbitration* 539, 565; Maxwell and Vannieuwenhuyse (n 664), (2018) (1) *ICC Dispute Resolution Bulletin* 24, 32; Halis Kasap (n 647), (2021) (2) *Journal of Dispute Resolution* 209, 230.

⁶⁹² Scherer (n 660) 503, 512; Scherer (n 557), (2019) 36(5) *Journal of International Arbitration* 539, 562.

⁶⁹³ Sim (n 551), (2018) 14(1) *Asian International Arbitration Journal* 1, 8 f.

⁶⁹⁴ Scherer (n 660) 503, 512; Scherer (n 557), (2019) 36(5) *Journal of International Arbitration* 539, 562.

3.2.3 Danger of 'Petrification'

- 225 As artificial intelligence makes decisions on the basis of training data from the past,⁶⁹⁵ the phenomenon of conservative decision-making closed to a development of the law also occurs in arbitration proceedings. This procedural area is by no means alien to a development of the law,⁶⁹⁶ even though the problem described is much less pronounced than in state court proceedings due to the low publication density of arbitration awards.

3.2.4 Failure to Grant the Right to Be Heard

- 226 It is true that complete automation of proceedings as such does not constitute a violation of the right to be heard. Yet, such a violation must be assumed if significant party submissions cannot be considered in the arbitral decision due to an error in the AI system.⁶⁹⁷

3.3 Procedural Answers

- 227 Notwithstanding all technical limitations, AI-based arbitration awards run – at least in some jurisdictions – the risk of being set aside or unenforceable.⁶⁹⁸ In view of the risks associated with the use of AI in arbitration proceedings, regulatory requirements become necessary with the aim of ensuring legal certainty and protect the legal interests of both the parties and the general public. In this context, the principles of transparency and fairness of proceedings as well as the protection of trust in the accuracy of AI systems and the integrity of arbitration proceedings as such should be mentioned in particular, eg, as concerns an AI-based reproduction or summary of the facts, the legal situation or the evidence in a specific proceeding. In that regard, the draft of the Guidelines on the Use of Artificial Intelligence in Arbitration,⁶⁹⁹ recently published by the Silicon Valley Arbitration & Mediation Center, provides useful indications on the design of arbitration proceedings using AI.⁷⁰⁰ In this light, the following regulatory approaches, which are by no means exhaustive, should be considered:

⁶⁹⁵ See above para 189-190.

⁶⁹⁶ In some cases, disputes relating to entire areas of law are mainly settled in arbitration proceedings.

⁶⁹⁷ Kreis (n 517) 633, 645 para 51 f.

⁶⁹⁸ Bizikova, Hancock, Jewell and Sherr (n 598), 'IA Meets AI' (2 October 2023).

⁶⁹⁹ Silicon Valley Arbitration & Mediation Center, SVAMC Guidelines on the Use of Artificial Intelligence in Arbitration, Draft of 31 August 2023 <https://the Arbitration.org/wp-content/uploads/2023/08/SVAMC-AI-Guidelines-CONSULTATION-DRAFT-31-August-2023-1.pdf> accessed 31 December 2023.

⁷⁰⁰ The purpose of the guidelines (*reference framework*) is described as follows (SVAMC Guidelines, 3): 'The Guidelines seek to establish a set of general principles for the use of AI in arbitration. Intended to guide rather than dictate, they are meant to accommodate case-specific circumstances and technological developments, promoting fairness, efficiency, and transparency in arbitral proceedings' – The term AI is defined as follows, SVAMC Guidelines, 3: '[...] the term 'AI' refers to computer systems that perform tasks commonly associated with human cognition, such as understanding natural language, recognizing complex semantic patterns, and generating human-like outputs'.

- 228 In view of an arbitrator's personal responsibility for the decision-making process, any transfer of his personal mandate to AI systems – even if only on a pro rata basis – should be excreted.⁷⁰¹ This is the only way to ensure that the right to be heard and the principles of fairness and integrity of arbitration proceedings are adequately considered.
- 229 In addition, every output from AI systems on which a decision is based should be checked by a human.⁷⁰² This would allow the advantages of machine and human decision-making to be combined in individual cases.⁷⁰³ Naturally speaking, this does not rule out the possibility of using AI tools downstream, for example following a human search.⁷⁰⁴
- 230 Moreover, it is particularly important to prematurely inform the (opposing) parties in a comprehensible and comprehensive manner about the use of AI tools, including its nature and scope. This should encompass the function as well as the intended type of use of the respective AI tool and concrete (significant) effects on the proceedings.⁷⁰⁵ Nothing else should apply to information about the complete prompt and the associated output of an AI processing.⁷⁰⁶ In particular, decision parameters and statistical bases should generally be disclosed, as well as the decision-relevant use of AI results outside a protocol that (may) actually influence an arbitrator's understanding.
- 231 In accordance with the SVAMC guidelines, the use of 'explainable AI' functions is recommended.⁷⁰⁷ The aim is to make it comprehensible to human users how an AI system arrives at a certain result based on certain inputs. Such explanations are owed irrespective of any technical and cost limitations of explaining how complex AI systems work or of the general need for all participants in arbitration to exercise their own judgment independently and to be aware of potential bias that may be inherent in the outcome of AI systems.⁷⁰⁸ Accordingly, the due process principle requires arbitrators to independently and critically assess the reliability of AI information.⁷⁰⁹
- 232 The guidelines also include obligations for the parties and their representatives: These include upstream due diligence obligations when using AI tools (eg, to check them in

⁷⁰¹ See also SVAMC Guidelines (n 699) 17 (Commentary to Guideline 6).

⁷⁰² See also SVAMC Guidelines (n 699) Guideline 7. – The degree of depth of review must, of course, be weighed in each individual case against the cost and time savings hoped for (and achieved) through the use of AI.

⁷⁰³ Schwartz (n 643) 95, 124 f.

⁷⁰⁴ Similarly, Cohen, who would like to use AI to correct a bias in human arbitrators: P Cohen, 'Bytes and Prejudice' (2015) 1(1) *Journal of Technology in International Arbitration* 57, 66. This, in turn, could avoid anchor effects due to an upstream machine decision.

⁷⁰⁵ In individual cases, this may also relate to the question of the extent to which the use of AI or a specific AI tool is preferable to the use of *primary source material*.

⁷⁰⁶ See also SVAMC Guidelines (n 699) 9 ff, 13 (Commentary to Guideline 3), with the additional concern that innocuous and uncontroversial uses of AI should not be prevented by overly strict procedural requirements.

⁷⁰⁷ See also SVAMC Guidelines (n 699) 6 f (Commentary to Guideline 1).

⁷⁰⁸ SVAMC Guidelines (n 699) 7 f (Commentary to Guideline 1).

⁷⁰⁹ SVAMC Guidelines (n 699) 18 f (Commentary to Guideline 7).

advance for application errors such as hallucinations)⁷¹⁰ or the obligation to refrain from using (generative) AI if this is likely to jeopardize the integrity of the proceedings or the authenticity of evidence (eg, by means of deep fakes).⁷¹¹

233 Finally, it must be ensured that the use of AI tools complies with the (legal) obligations to protect confidential information.⁷¹²

234 If AI is used to support arbitration, be it in researching and interpreting facts and the law or in applying the law to a concrete set of facts, these are to be classified as so-called high-risk AI systems within the meaning of Art 6(2), Annex III no 8 lit a) of the AI Act of the European Union.⁷¹³ This follows from the fact that decisions of arbitration tribunals are binding on the parties. Consequently, the providers and operators of such high-risk AI systems are subject to the special obligations of Art 9-17 AI Act.⁷¹⁴

4 Summary of the Main Results

4.1 Platform-Based Enforcement of Rights and Standards

4.1.1 Economization of Platform Procedures

235 Online platforms (provisionally) enforce state law as well as their own standards. By providing independent complaint procedures for arising disputes, usually in an easily accessible, cost- and time-saving manner, such disputes almost invariably no longer reach the state courts. This ‘absorption’ effect stems from the privatization of judicial, enforcement and regulatory tasks by private actors. What is more, the business model tailored by a platform to specific user (groups) is often continued in an ‘economization’ of internal procedural structures: Not infrequently, the result is a different treatment of (sometimes segregated) user groups: be it in terms of access to internal complaint procedures, be it with regard to the person or the technical (algorithmic) medium designated by a platform for making the decision. As a result, the principle of equal treatment of parties or participants with equivalent procedural roles, which is derived from national codes of (civil) procedure, is partly put into perspective. At the same time, the economization of platform proceedings is accompanied by adapted patterns of behavior on the part of rightsholders. In the context of copyright law, such adaptation processes are expressed, for example, in an increasing monetization of protected content, thereby replacing reactive law enforcement usually aimed at deleting infringing content. In this respect, the (procedural) role of online platforms is advancing from ‘mere’ mediators of sanctions to mediators of economic revenue growth.

⁷¹⁰ SVAMC Guidelines (n 699) 14 (Commentary to Guideline 4).

⁷¹¹ SVAMC Guidelines (n 699) 16 (Commentary to Guideline 5).

⁷¹² SVAMC Guidelines (n 699) 9 (Commentary to Guideline 2).

⁷¹³ See in more detail above para 162 ff.

⁷¹⁴ See in more detail above para 161 ff.

4.1.2 *Merging of State and Private Enforcement Interests*

236 Another line of development emerges in parallel: the merging of state and private enforcement interests. Which proves to be problematic whenever (and due to) online platforms not (clearly) separating the enforcement of state-granted rights on the one hand and their own private standards on the other. By resorting to platforms in order to ostensibly enforce the standards of those private actors without disclosing and justifying the actual public enforcement interests, state authorities run the risk of circumventing the rule of law. In addition, private platform standards are often interpreted in line with a specific pre-understanding or regulatory calculation of a state. As a consequence, the guiding function and recognizability of state law will be weakened. Compared to this, internal complaint procedures offer little effective protection, especially as online platforms do not (yet) provide any specific procedures or procedural rules against forms of state involvement in a platform-based enforcement of rights and standards. Last but not least, by instrumentalizing large international online platforms, government agencies hope to enforce national law *de facto* with transnational effect, as illustrated by the global application of the fair use principle inherent in US copyright law.

4.1.3 *Potential Effects of the Use of Artificial Intelligence*

237 The use of AI supports and reinforces such effects. In platform-based dispute resolution, as in arbitration or conciliation, it harbors the risk of opaque decision bases (the so-called black box problem). Thus, AI regularly prevents individual steps in the decision-making process from being reconstructed. It is the independence of the output from the input, often in conjunction with feedback loops, that favors the phenomenon of algorithmic discrimination. Furthermore, AI is technically incapable of showing empathy – the cornerstone of interpersonal communication – nor can it make cognitive or (spontaneous) judgment-based or value-related decisions. The same applies to the (lacking) ability to grasp highly context-sensitive content in a legally correct manner. In this light, it is hardly possible to justify (enforcement) decisions at all. In addition, AI's strict reference to the past makes it impossible to develop the law in individual cases. All of these potential consequences considerably affect the effective legal defense of both platform users and rightsholders. At the same time, rapid technical progress offers the possibility of ever greater flexibility and individualization of platform procedures.

4.1.4 *Proceduralization of Private Rights and Standards Enforcement*

238 To such dangers the (increased) proceduralization of private rights and standards enforcement offers a solution. This is because procedural structures fulfil an independent and necessary function in order to ensure effective legal protection in the context of private, platform-based rights enforcement: Only then do they enable the addressee of enforcement, but also affected rightsholders, especially in particularly dangerous situations, to effectively realize their rights and thus to provisionally balance them in the context of uncertain legal and factual situations in a less faulty manner. The

inclusion of platforms under liability law alone does not do justice to this regulatory objective: it is neither geared towards such a balancing of interests nor does it ask about the procedural conditions of possibility and the implementation of duties of disclosure.

- 239 Recently, state regulatory acts have taken a first, corresponding approach. The European Union's Digital Services Act serves as the inspiration. Hereby procedural (minimum) standards, essentially borrowed from the guarantees of state (civil) procedure, are being put in place. These include, *inter alia*, the right to equal access to proceedings, the obligation of a platform operator to justify (enforcement) decisions as well as the right to an effective complaint. In the interests of platform users, these rights particularly aim to guarantee the right to be heard and an effective legal defense and, therefore, to balance structural power and information imbalances within platform structures. Above this, a multipolar, more differentiated conflict resolution system is emerging that does not only include online platforms, rightsholders and users: In order to promote effective law enforcement, platform procedures also involve third parties such as (trusted) whistleblowers.
- 240 Nevertheless, there is still a need for further regulation: For instance, the objective of more closely dovetailing a platform's internal complaint procedures with procedures for independent out-of-court dispute resolution has not yet been sufficiently implemented. This applies all the more in light of a structural bias of platforms, especially when enforcing their own standards. In order to compensate for the limited redeemability of substantively correct private legal enforcement, internal complaint procedures – all the more in view of (automated) enforcement of a platform's own standards – should be supplemented by an out-of-court dispute settlement subject to a subsequent state court decision. The procedure should be freely accessible to affected users, rightsholders and whistleblowers and binding for platforms.
- 241 The model of provisional legal protection as provided for by the national codes of civil procedure is a suitable starting point for proceduralizing platform-based legal enforcement: In particular, this applies with regard to a platform's reduced duties to clarify the facts and assumptions of probability underlying a platform's decision, but also to a hearing of addressees downstream of urgent enforcement measures. In order to assess a platform's liability risks more reliably, clearly defined procedural (reaction) obligations should take the place of case-by-case duties of disclosure; the substantive standard of review should be simplified and, for example, substantiated by legal presumptions or a typification of (un)permitted user behavior.
- 242 Equally a stronger proceduralization of platform-based decision-making has only been implemented to a limited extent through government regulatory acts. In addition to requirements for (better) qualification of platform decision-makers and fundamental decision-making standards under the rule of law for platform-based legal enforcement (such as the obligation to act carefully, objectively, without arbitrariness and proportionately), this also concerns issues of consistency and comparability of platform

decisions, which are often made more difficult, if not prevented, by non-transparent decision-making standards. Rules on the scaling of decisions, ie, when platforms transfer specific interpretation standards derived from ‘model cases’ and deemed to be of ‘high quality’ to comparable situations are also required. Better protection against distortion effects, which are particularly associated with context-sensitive disputes and often have a transnational reach, is promised by, among other things, increased transparency obligations, further a stronger link between scaling decisions and external objective review standards and independent review bodies, as well as effective downstream legal protection.

- 243 Transparency obligations of platforms are a key component of proceduralization. In this context, they fulfill both an auxiliary procedural function and a function of institutional control. By the means of periodic transparency reports or a right of access for research purposes, they are intended to ensure preventive public control of platforms – whether by authorities (eg, to protect competition and consumer interests) or by academia, the media and non-governmental organizations. Then again, clear provisions in the general terms and conditions of online platforms regarding the intended forms of sanctions should make these predictable for individual platform users, especially when using AI: Corresponding information obligations should cover, among other things, the course of platform procedures, whether automated systems are used or whether and on what basis platforms have taken measures at the instigation of a state authority. Last but not least, specific decisions made by online platforms, for instance, on the removal and blocking of access or the downgrading of content, should be stored in a publicly accessible database. Ensuring procedural transparency is therefore a central procedural principle of platform-based law enforcement.

4.1.5 The Regulatory Dimension: Interaction Between (Individual) Rule Enforcement and (Collective) Setting of Rules

- 244 Standardization of platform-based rule enforcement can also be observed. This is the result of feedback from individual platform decisions to a platform’s AI systems. At the same time, the ongoing implementation of platform-specific standards influences their dynamic, regulatory development.
- 245 This ‘collectivizing’ effect is reinforced by so-called platform courts, ie, non-governmental, court-like dispute resolution bodies such as the Facebook Oversight Board. Such courts will not only have the task of settling individual disputes (usually of particular relevance) and continuously harmonizing decision-making and interpretation of a platform’s community standards. Platform courts at once shape a private process of rule-setting, essentially following a ‘procedural concept of (self-)regulation’. This process of norm-building is (partly) beyond state control of the framework conditions, while at the same time being communicatively tied back to the public. The aim is twofold: on the one hand, to ensure more effective protection of individual rights in a multitude of (parallel) cases and, on the other, to enable a continuously adapted, by its nature mostly

rudimentary-experimental rule building through and within the framework of procedural structures. In this way, interpretation (contentious review), setting and enforcement of private standards ultimately amalgamate into 'one power'.

246 Such a concentration of power of online platforms can be countered from the perspective of various areas of law. As far as procedural solutions are concerned, out-of-court dispute settlement (cf Art 21 of the EU DSA Regulation) opens up an incidence check of a platform's T&Cs and its moderation guidelines concretizing them. Reference point of control is their conformity with the procedural and contractual requirements of relevant EU legal acts (cf Digital Services Act; P2B Regulation). From the perspective of such legislative overlay, platform T&Cs thus (de facto) assume a regulatory function. This indirect control of a platform's own standards is flanked by opportunities for direct state influence, primarily in the form of regulating internal platform procedures. This approach finds paradigmatic expression in the Digital Services Act in the binding specification of procedural (minimum) requirements.

4.2 Use of Artificial Intelligence in Mediation, Dispute Resolution and Arbitration Proceedings

247 In contrast to a platform-based enforcement of rights and standards, the use of AI in mediation, dispute resolution and arbitration proceedings has so far only fulfilled an upstream auxiliary function. It is primarily aimed at increasing procedural efficiency, ie, essentially at saving time and costs in the interests of the parties to the dispute and the independent arbitrator, conciliator, or mediator. The scope of application of AI ranges from providing procedural information, including simple legal information, automated document analyses and other forms of case management to the preparation of settlement and decision proposals (predictive tools). A 'supra-individual' standardization of dispute resolution is beginning to emerge, insofar as companies geared towards dispute resolution store the data of all disputes, which are then available to them as data sets for each new legal dispute pending before such companies.

248 The risks associated with these areas of AI application are similar to those of platform-based law and standards enforcement, for example in the form of the black box problem or algorithmic discrimination. However, these risks have a less serious impact on arbitration proceedings as well as in mediation and dispute resolution. This is because fully automated or AI-based proceedings without a human decision-maker are currently neither technically possible nor desirable in terms of legal policy. Only forms of 'hybrid decision-making' are already in use. The de facto settlement of disputes, which is characteristic of a platform-based enforcement of rights and standards, does not occur in mediation, dispute resolution or arbitration proceedings either.

ABBREVIATIONS AND ACRONYMS

ADR	Alternative Dispute Resolution
Art	Article/Articles
BGB	Bürgerliches Gesetzbuch (Civil Code) [Germany]
BGH	Bundesgerichtshof (Federal Court of Justice) [Germany]
BVerfG	Bundesverfassungsgericht (Federal Constitutional Court) [Germany]
cf	confer (compare)
ch	chapter
CJEU	Court of Justice of the European Union
ECLI	European Case Law Identifier
ECtHR	European Court of Human Rights
DMCA	Digital Millennium Copyright Act [USA]
DSA or DSA Regulation	Digital Services Act [European Union]
ECG	E-Commerce-Gesetz (E-Commerce Act) [Austria]
ed	editor/editors
edn	edition/editions
eg	exempli gratia (for example)
etc	et cetera
EU	European Union
EUR	Euro
f/ff	following
fn	footnote (external, ie, in other chapters or in citations)
GSR	General Secondary Response
GG	Grundgesetz (Fundamental Law) [Germany]
GDPR	General Data Protection Regulation [European Union]
GTC	General terms and conditions
ibid	ibidem (in the same place)
ie	id est (that is)
KoPIG	Kommunikationsplattformengesetz (Communications Platforms Act) [Austria]
n	footnote (internal, ie, within the same chapter)
NetzDG	Netzwerkdurchsetzungsgesetz (Network Enforcement Act) [Germany]
no	number/numbers
öABGB	Allgemeines Bürgerliches Gesetzbuch (General Civil Code) [Austria]
öUrhG	Urheberrechtsgesetz (Copyright Act) [Austria]

P2B Regulation	Regulation (EU) 2019/1150 of 20 June 2019 on promoting fairness and transparency for business users of online intermediary services
PACT or PACT Act	Platform Accountability and Consumer Transparency Act [USA]
para	paragraph/paragraphs
pt	part
RDG	Rechtsdienstleistungsgesetz (Legal Services Act) [Germany]
S.D. Cal.	District Court for the Southern District of California [USA]
sec	Section/Sections
supp	supplement/supplements
TMG	Telemediengesetz (Telemedia Act) [Germany]
trans/tr	translated, translation/translator
UGC	User-generated content
UK	United Kingdom
UNIDROIT	Institut international pour l'unification du droit privé (International Institute for the Unification of Private Law)
UrhDaG	Urheberrechts-Diensteanbieter-Gesetz (Act on Copyright Content Sharing Service Providers) [Germany]
UrhG	Urheberrechtsgesetz (Copyright Act) [Germany]
US/USA	United States of America
U.S.C.	United States Code
USD	United States Dollar
v	versus
vol	volume/volumes
VSBG	Verbraucherstreitbeilegungsgesetz (Consumer Dispute Resolution Act) [Germany]
ZPO	Zivilprozessordnung (Code of Civil Procedure) [Germany]
***	***

LEGISLATION

International/Supranational

Charter of Fundamental Rights of the European Union, OJ C 364, 18 December 2000, 1.

Comission and Parliament (EU), Synopsis AI Act, P9_TA(2023)0236, https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf accessed 31 December 2023.

Commission Recommendation (EU) 2018/334 of 1 March 2018 on measures to effectively tackle illegal content online, OJ L 62, 6 March 2018, 50.

Directive 2000/31/EC of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'), OJ L 178, 17 July 2000, 1.

Directive 2001/29/EC of 22 May 2001 on the harmonization of certain aspects of copyright and related rights in the information society, OJ L 167, 22 June 2001, 10.

Directive 2004/48/EC of 29 April 2004 on the enforcement of intellectual property rights, OJ L 157, 30 April 2004, 45.

Directive 2008/52/EC of the European Parliament and of the Council of 21 May 2008 on certain aspects of mediation in civil and commercial matters, OJ L 136, 24 May 2008, 3.

Directive 2013/11/EU of 21 May 2013 on alternative dispute resolution for consumer disputes and amending Regulation (EC) No 2006/2004 and Directive 2009/22/EC (Directive on consumer ADR), OJ L 165, 18 June 2013, 63.

Directive 2019/790/EU of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, OJ L 130, 17 May 2019, 92.

ICC Arbitration Rules 2021.

Regulation (EU) 2019/1150 of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services, OJ L 186, 11 July 2019, 57.

Regulation (EU) 2021/784 of 29 April 2021 on addressing the dissemination of terrorist content online, OJ L 172, 17 May 2021, 79.

Regulation (EU) 2022/2065 of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act; hereinafter: DSA Regulation), OJ L 277, 27 October 2022, 1.

Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives

2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 12 July 2024.

UNCITRAL Model Law on International Commercial Arbitration (1985).

National

Arbitration Rules of the Netherlands Arbitration Institute (Schiedsordnung Nederlands Arbitrage Instituut).

Austrian Communications Platforms Act (Kommunikationsplattformengesetz: KoPlG).

Austrian Copyright Act (Urheberrechtsgesetz: öUrhG).

Austrian General Civil Code (Allgemeines Bürgerliches Gesetzbuch: öABGB).

Austrian E-Commerce Act (E-Commerce-Gesetz: ECG).

English Arbitration Act 1996.

French Code of Civil Procedure (Code de procédure civile).

French Law of 23 March 2019 (Loi de programmation 2018-2022 et réforme pour la justice) No 2019/222.

German Act on Copyright Content Sharing Service Providers (Urheberrechts-Diensteanbieter-Gesetz: UrhDaG).

German Fundamental Law (Grundgesetz: GG).

German Civil Code (Bürgerliches Gesetzbuch: BGB).

German Code of Civil Procedure (Zivilprozessordnung: ZPO).

German Consumer Dispute Resolution Act (Verbraucherstreitbeilegungsgesetz: VSBG).

German Copyright Act (Urheberrechtsgesetz: UrhG).

German Legal Services Act (Rechtsdienstleistungsgesetz: RDG).

German Network Enforcement Act (Netzwerkdurchsetzungsgesetz: NetzDG).

German Telemedia Act (Telemediengesetz: TMG).

Italian Code of Civil Procedure (Codice di procedura civile: CPC).

Spanish Arbitration Act (Lex 60/2003 de 23 de diciembre, de Arbitraje).

Title 17 of the United States Code – Copyright Law (17 U.S.C.).

Title 47 of the United States Code – Telecommunications (47 U.S.C.).

Turkish International Arbitration Law (Uluslararası Tahkim Kanunu).

United States of America Platform Accountability and Consumer Transparency Act,
116th Cong. § 5(2) (2020) (PACT Act).

United States of America Digital Millennium Copyright Act (DMCA).

CASES

International/Supranational

CJEU, 23 March 2010, C-236/08 and C-238/08 – *Google France*, ECLI:EU:C:2010:159.

CJEU, 12 July 2011, C-324/09 – *L'Oréal v eBay*, ECLI:EU:C:2011:474.

CJEU (Grand Chamber), 24 September 2019, C-507/17 – *Google (Spatial scope of delisting)*, ECLI:EU:C:2019:772.

CJEU, 3 October 2019, C-18/18 – *Glawischnig-Piesczek*, ECLI:EU:C:2019:821.

CJEU, 26 April 2022, C-401/19 – *Poland v Parliament and Council*, ECLI:EU:C:2022:297.

European General Court, 27 September 2023, T-367/23 – *Amazon Services Europe v Commission*, ECLI:EU:T:2023:589.

Oversight Board, 28 January 2021, 2020-005-FB-UA – *Nazi Quote*.

Oversight Board, 5 May 2021, 2021-001-FB-FBR – *Former President Trump's suspension*.

Oversight Board, 8 July 2021, 2021-006-IG-UA – *Ocalan's Isolation*.

Oversight Board, 14 September 2021, 2021-009-FB-UA – *Shared Al Jazeera Post*.

Oversight Board, 27 September 2021, 2021-010-FB-UA – *Colombia Protests*.

Oversight Board, 17 June 2022, 2022-001-FB-UA – *Knin Cartoon*.

Oversight Board, 15 September 2022, 2022-005-FB-UA – *Mention of the Taliban in News Reporting*.

Oversight Board, 22 November 2022, 2022-007-IG-MR – *UK Drill Music*.

Oversight Board, 6 December 2022, 2021-002-FB-PAO – *Meta's Cross-Check Program*.

Oversight Board, 14 December 2022, 2022-012-IG-MR – *India Sexual Harassment Video*.

Oversight Board, 14 December 2022, 2022-011-IG-UA – *Video after Nigeria Church Attack*.

Oversight Board, 9 January 2023, 2022-013-FB-UA – *Iran Protest Slogan*.

Oversight Board, 9 March 2023, 2022-014-FB-MR – *Sri Lanka Pharmaceuticals*.

Oversight Board, 18 December 2023, 2023-054-FB-UA, 2023-055-FB-UA, 2023-056-FB-UA, 2023-057-FB-UA – *Goebbels Quote*.

National

BGH (Germany), 17 August 2011, I ZR 57/09 – *Stiftparfüm*, BGHZ 191, 19 = (2011) 113(11) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 1038.

BGH (Germany), 25 October 2011, VI ZR 93/10 – *Blog Eintrag*, BGHZ 191, 219 = (2012) 65(3) NJW (Neue Juristische Wochenschrift) 148 = (2012) 114(3) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) GRUR 2012, 311.

BGH (Germany), 27 October 2011, I ZR 131/10 – *regierung-oberfranken.de*, (2012) 65(31) NJW (Neue Juristische Wochenschrift) 2279.

BGH (Germany), 12 July 12, 2012, I ZR 18/11 – *Alone in the dark*, BGHZ 194, 339 = (2013) 66(11) NJW (Neue Juristische Wochenschrift) 784.

BGH (Germany), 18 June 2015, I ZR 74/1 – *Liability for Hyperlink*, BGHZ 206, 103 = (2016) 69(11) NJW (Neue Juristische Wochenschrift) 804.

BGH (Germany), 1 March, 2016, VI ZR 34/15 – *Ärztebewertungsportal III (jameda.de)*, BGHZ 209, 139 = (2016) 65(3) NJW (Neue Juristische Wochenschrift) 2106.

BGH (Germany), 22 November 2017, VIII ZR 83/16, (2018) 71(8) NJW (Neue Juristische Wochenschrift) 537.

BGH (Germany), 22 November 2017, VIII ZR 213/16, (2018) 21(3) MMR (Multimedia und Recht) 156.

BGH (Germany), 29 July 2021, III ZR 179/20, (2021) 74(43) NJW (Neue Juristische Wochenschrift) 3179.

BGH (Germany), 29 July 2021, III ZR 192/20, (2021) 25(11) ZUM-RD (Zeitschrift für Urheber- und Medienrecht – Rechtsprechungsdienst) 612.

BVerfG (Germany), 3 February 1959, 2 BvL 10/56, BVerfGE 9, 137 = (1959) 12(21) NJW (Neue Juristische Wochenschrift) 931.

BVerfG (Germany), 8 August 1978, 2 BvL 8/77, BVerfGE 49, 89 = (1979) 32(8) NJW (Neue Juristische Wochenschrift) 359.

BVerfG (Germany), 20 April 1982, 2 BvL 26/81, BVerfGE 60, 253 = (1982) 35(43) NJW (Neue Juristische Wochenschrift) 1982, 2425.

District Court for the Southern District of California (USA), *United States v Green*, 857 F. Supp. 2d 1015, 1018 (S.D. Cal. 2012).

BIBLIOGRAPHY

Achleitner R, 'The Fight against Geo-Blocking – A Never Ending Story? Policy Paper on Geo-Blocking' <https://ssrn.com/abstract=4246896> or <http://dx.doi.org/10.2139/ssrn.4246896> accessed 31 December 2023.

Adolphsen J, 'Der Zivilprozess im Wettbewerb der Methoden' (2017) 48(4) BRAK Mitteilungen 147.

Althammer C, 'Alternative Streitbeilegung im Internet' in F Faust and H-B Schäfer (ed), *Zivilrechtliche und rechtsökonomische Probleme des Internet und der künstlichen Intelligenz* (Mohr Siebeck 2019) 249.

Ameln F von, 'Führen und Entscheiden unter Unsicherheit' (2021) 52(4) GIO (Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie) 567.

Andrea R, 'No Safe Harbor: YouTube's Content ID and Fair Use' (2020) Boston College Intellectual Property & Technology Forum 1.

Anzinger H M, '10 Jahre Modria – KMS und Online-Mediation auf dem Weg zur Digitalisierung der Justiz – Teil 1' (2021) 24(2) ZKM (Zeitschrift für Konfliktmanagement) 53.

Anzinger H M, '10 Jahre Modria – KMS und Online-Mediation auf dem Weg zur Digitalisierung der Justiz – Teil 2' (2021) 24(3) ZKM (Zeitschrift für Konfliktmanagement) 84.

AIRBNB, 'Scoring the user to prevent "suspicious" activity before it occurs: What Does It Mean When Someone's ID Has Been Checked?', <https://www.airbnb.com/help/article/2356/what-does-it-mean-when-someones-id-has-been-checked> accessed 31 December 2023.

Appelman N and Leerssen P, 'On "Trusted" Flaggers' (2022) 24 Yale Journal of Law & Technology 452.

Arbitrator Intelligence, 'Arbitrator Intelligence Database', <https://arbitratorintelligence.vercel.app/> accessed 31 December 2023.

Aschauer C, 'Automated Decision-Making and Artificial Intelligence (AI) in Arbitration' in C Leyens, I Eisenberger and R Niemann (ed), *Smart Regulation* (Mohr Siebeck 2021) 130.

Askani A, *Private Rechtsdurchsetzung bei Urheberrechtsverletzungen im Internet* (Nomos 2021).

Bambauer D E, 'Against Jawboning' (2015) 100(1) MINN. L. REV. 51.

Barocas S and Self A D, 'Big Data's Disparate Impact' (2016) 104(3) California Law Review 671 (2016).

Barona Vilar S, 'Effizienzsteigerung und Suche nach Beschleunigung von Schiedsverfahren im Spannungsfeld von Mythos. Sublimierung und Vierter Industrieller Revolution (4.0)' (2018) 23 ZZPInt (Zeitschrift für Zivilprozess International) 295.

Barton H, 'Rebooting Justice' (2018) 44(4) Law Practice 32.

Bar-Ziv S and Elkin-Koren N, 'Behind the Scenes of Online Copyright Enforcement: Empirical Evidence on Notice & Takedown' (2018) 50(2) Connecticut Law Review 339.

Berberich M and Conrad A, '§ 30 Plattformen und KI' in M Ebers, CA Heinze and B Steinrötter (ed), *Künstliche Intelligenz und Robotik* (Beck 2020) 930.

Berberich M, '§ 5 Sorgfaltspflichten, Moderationsverfahren und prozedurale Fairness' in B Steinrötter (ed), *Europäische Plattformregulierung* (Nomos 2023) 126.

Bizikova L, Hancock P, Jewell D and Sherr I, 'IA Meets AI – Rise of the Machines', <<https://dailyjus.com/legal-tech/2023/10/ia-meets-ai-rise-of-the-machines>> accessed 31 December 2023.

Bloch-Wehba H, 'Automation in Moderation' (2020) 53(1) Cornell International Law Journal 41.

Bloch-Wehba H, 'Global Platform Governance: Private Power in the Shadow of the State' (2019) 72(1) SMU Law Review 27.

Boll-Kempelmann C, 'Datenschutz und das Beweisverfahren im Schiedsverfahren' (2022) 20(5) SchiedsVZ (Zeitschrift für Schiedsverfahren) 241.

Bomhard D and Merkle M, 'Regulation of Artificial Intelligence' 2021 (6) EuCML (Journal of European Consumer and Market Law) 257.

Brazil W, 'Informalism and Formalism in the History of ADR in the United States' in J Zekoll, M Bälz and I Amelung (ed), *Formalisation and Flexibilisation in Dispute Resolution* (Brill 2014) 250.

Brechmann B, *Legal Tech und das Anwaltsmonopol* (Mohr Siebeck 2021).

Breiman L, 'Random Forests' (2001) 45(1) Machine Learning 5.

Bridy A, 'Intellectual Property' in Keller D (ed), *Law, Borders, and Speech: Proceedings and Materials* (The Center for Internet and Society 2017) 9.

Bünnau P von, 'Künstliche Intelligenz im Recht' in S Breidenbach and F Glatz (ed), *Rechtshandbuch Legal Tech* (2nd edn, Beck/Manz 2021) 71.

Bull L and Steffek F, 'The Decoding of Legal Conflicts' (2018) 21(5) ZKM (Zeitschrift für Konfliktmanagement) 165.

Burk D L, 'Algorithmic Fair Use' (2019) 86(2) University of Chicago Law Review 283.

Burrell J, 'How the machine "thinks": Understanding opacity in machine learning algorithms' (2016) (1) Big Data & Society 3, <https://doi.org/10.1177/2053951715622512> accessed 31 December 2023.

Busch C, 'Mehr Fairness und Transparenz in der Plattformökonomie?' (2019) 121(8) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 788.

Calliess C and Ruffert M (ed), *EUV/AEUV* (6th edn, Beck 2022).

Casey B et al, 'Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise' (2019) 34(1) BERKELEY TECH. L.J. 145.

Castets-Renard C, 'Algorithmic Content Moderation on Social Media in EU Law: Illusion of Perfect Enforcement' (2020) (2) University of Illinois Journal of Law, Technology & Policy 283.

Cellan-Jones R, 'The robot lawyers are here', <<https://www.bbc.com/news/technology-41829534>> accessed 31 December 2023.

Cervenka A and Schwarz P, 'Data Protection in Arbitration Proceedings' (2020) 18(2) SchiedsVZ (Zeitschrift für Schiedsverfahren) 78.

Chang B, 'From Internet Referral Units to International Agreements: Censorship of the Internet by the UK and EU' (2018) 49(2) Columbia Human Rights Law Review 114.

Chander A, 'The Racist Algorithm?' (2017) 115(6) Michigan Law Review 1023.

Chargeflow, 'Paypal Dispute automation', <https://www.chargeflow.io/paypal-dispute-automation> accessed 31 December 2023.

Citron D K, 'Technological due process' (2008) 85(6) Washington University Law Review 1249.

Citron D K and Pasquale F, 'The Scored Society: due process for automated predictions' (2014) 89(1) WASH. L. REV. 1.

Civil resolutional tribunal, 'Civil resolutional tribunal', <https://civilresolutionbc.ca/solution-explorer/> accessed 31 December 2023.

Cohen J E, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (Oxford University Press 2019).

Cohen P, 'Bytes and Prejudice' (2015) 1(1) Journal of Technology in International Arbitration 57.

Conrad A and Nolte G, 'Schrankenbestimmungen im Anwendungsbereich des UrhDaG' (2021) 65 ZUM (Zeitschrift für Urheber- und Medienrecht) 111.

Cooper S, Rule C and Del Duca L, 'From Lex Mercatoria to Online Dispute Resolution' (2011) 43 Uniform Commercial Code Law Journal, Penn State Legal Studies Research Paper No 09/2011.

- Deckenbrock C and Henssler M, *Rechtsdienstleistungsgesetz* (5th edn, Beck 2021).
- Deichsel T, *Digitalisierung der Streitbeilegung* (Nomos 2022).
- Deichsel T, 'Verbraucherschlichtungsstellen – Ein Anwendungsfeld für Legal Tech?' (2020) 35(8) VuR (Verbraucher und Recht) 283.
- Denga M, 'Platform Regulation by European Values: On the Binding of Opinion Platforms to EU Fundamental Rights' (2021) 56(5) EuR (Europarecht) 569.
- Deusch F and Eggendorfer T, 'IT-Sicherheit' in J Taeger and J Pohle (ed), *Computerrechts-Handbuch* (38th edn, Beck 2023) ch 50.1.
- Dewey C, '98 personal data points that Facebook uses to target ads to you', The Washington Post, 19 August 2016, <https://www.washingtonpost.com/news/the-intersect/wp/2016/08/19/98-personal-data-points-that-facebook-uses-to-target-ads-to-you/> accessed 31 December 2023.
- Dispute Resolution Data, 'Dispute Resolution Database', <https://www.disputeresolutiondata.com/> accessed 31 December 2023.
- Disco, 'Disco Powerful AI and Analytics', <https://www.csdisco.com/offering/ediscovery/features-ai> accessed 31 December 2023.
- Douek E, 'Content Moderation as Systems Thinking' (2022) 136(2) Harvard Law Review 528.
- Douek E, 'How Much Power Did Facebook Give Its Oversight Board?', <<https://www.lawfaremedia.org/article/how-much-power-did-facebook-give-its-oversight-board>> accessed 31 December 2023.
- Douek E, 'The Oversight Board Moment You Should've Been Waiting For, Lawfare', <<https://www.lawfaremedia.org/article/oversight-board-moment-you-shouldve-been-waiting-facebook-responds-first-set-decisions>> accessed 31 December 2023.
- Drexl J, 'Bedrohung der Meinungsvielfalt durch Algorithmen' (2017) 61(7) ZUM (Zeitschrift für Urheber- und Medienrecht) 529.
- Drupal, 'a2J Author', <https://www.a2jauthor.org/> accessed 31 December 2023.
- Dubois E and Blank G, 'The echo chamber is overstated: the moderating effect of political interest and diverse media' (2018) 21(5) Information, Communication & Society 729.
- Ebrevia, 'DraftPro', <https://www.dfinsolutions.com/products/ebrevia> accessed 31 December 2023.
- Eghbariah R and Metwally A, 'Informal Governance: Internet Referral Units and the Rise of State Interpretation of Terms of Service' (2021) 23 Yale J.L. & Tech. 545.
- Eidenmüller H and Wagner G, *Law by Algorithm* (Mohr Siebeck 2021).

Eifert M et al, *Netzwerkdurchsetzungsgesetz in der Bewährung* (Nomos 2020).

Elkin-Koren N, 'After twenty years: revisiting copyright liability of online intermediaries' in S Frankel and D Gervais (ed), *The Evolution and Equilibrium of Copyright in the Digital Age* (Cambridge University Press 2014) 29.

Elkin-Koren N, 'Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence' (2020) 7(2) Big Data & Society 1, <https://journals.sagepub.com/doi/epub/10.1177/2053951720932296> accessed 31 December 2023.

Elkin-Koren N and Perel M, 'Separation of Functions for AI: Restraining Speech Regulation by Online Platforms' (2020) 24(3) Lewis & CLARK L. REV. 857.

Engert A, 'Digitale Plattformen' (2018) 218(2-4) AcP (Archiv für die civilistische Praxis) 304.

Engstrom E and Feamster N, *The Limits of Filtering: A Look at the Functionality & Shortcomings of Content Detection Tools* (Engine 2017), <https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/58d058712994ca536bbfa47a/1490049138881/FilteringPaperWebsite.pdf> accessed 31 December 2023.

European Commission, 'Staff Working Document Impact Assessment, Accompanying the document Pro-posal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online' (12 September 2018) SWD (2018), 408 final.

European Commission, 'Tackling online disinformation: Commission proposes an EU-wide Code of Practice' (26 April 2018), https://ec.europa.eu/commission/presscorner/detail/en/IP_18_3370 accessed 31 December 2023.

Europol, '2018 Consolidated Annual Activity 44' (23 May 2019), https://www.europol.europa.eu/cms/sites/default/files/documents/consolidated_annual_activity_report_2018.pdf accessed 23 December 2023.

Europol, 'EU Internet Referral Unit – EU IRU', <https://www.europol.europa.eu/about-europol/european-counter-terrorism-centre-ectc/eu-internet-referral-unit-eu-iru> accessed 31 December 2023.

Facebook Oversight Board, 'Oversight Board Bylaws', <https://www.oversightboard.com/wp-content/uploads/2024/03/Oversight-Board-Bylaws.pdf> accessed 31 December 2023.

Facebook Oversight Board, 'Oversight Board Charter', https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf accessed 31 December 2023.

Fiala F and Husovec M, 'Using Experimental Evidence to Improve Delegated Enforcement' (2022) 71 International Review of Law and Economics 106079 and (2018) 28 Forthcoming TILEC Discussion Paper no 2018-028.

Fireflies, 'Fireflies AI', <https://firefliesai/> accessed 31 December 2023.

Fries M, 'Erfüllung von Geldschulden über eigenwillige Zahlungsdienstleister' (2018) 33(4) VuR (Verbraucher und Recht) 123.

Fries M, 'Legal Tech im Schiedsverfahren' in R Wilhelmi and M Stürner (ed), *Mehrparteienschiedsverfahren. Unter besonderer Berücksichtigung gesellschaftsrechtlicher Streitigkeiten* (Springer 2021) 85.

Fries M, 'PayPal Law und Legal Tech – Was macht die Digitalisierung mit dem Privatrecht?' (2016) 69(39) NJW (Neue Juristische Wochenschrift) 2860.

Fries M, *Verbraucherrechtsdurchsetzung* (Mohr Siebeck 2016).

Fritz G, Prantl D, Leinwather N and Hofer M, 'Data Protection in International Arbitration Proceedings' (2019) 17(6) SchiedsVZ (Zeitschrift für Schiedsverfahren) 301.

Frosio G, 'Algorithmic Enforcement Online' in P Torremans (ed), *Intellectual Property and Human Rights* (4th edn, Kluwer Law International 2020) 24.

Gerdemann S and Spindler G, 'Das Gesetz über digitale Dienste (Digital Services Act) (Part 2) – Die Regelungen für Online-Plattformen sowie sehr große Online-Plattformen und Suchmaschinen' (2023) 125(3) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 115.

Gielen N and Uphues S, 'Digital Markets Act und Digital Services Act' (2021) 32(14) EuZW (Europäische Zeitschrift für Wirtschaftsrecht) 627.

Gillespie T, 'The Relevance of Algorithms' in T Gillespie, P J Boczkowski and K A Foot (ed), *Media Technologies* (The MIT Press 2014) 188.

Gläßer U, 'Mediation und Digitalisierung' in T Riehm and S Dörr (ed), *Digitalisierung und Zivilverfahren* (De Gruyter 2023) 529.

Global Arbitration Review (GAR), 'Database', <https://globalarbitrationreview.com/tools/arbitrator-research-tool> accessed 31 December 2023.

Glogowski M, *Plattformbedingungen* (Mohr Siebeck 2022).

Google, 'Asset erstellen', https://support.google.com/youtube/answer/3011552?hl=de&ref_topic=3011550 accessed 31 December 2023.

Google, 'Transparency Reports', <<https://transparencyreport.google.com/>> accessed 31 December 2023.

Google, 'Transparency Report for YouTube Platform for January to June 2023' <<https://transparencyreport.google.com/netzdg/youtube?hl=de>> accessed 31 December 2023.

Google, 'How Google Fights Piracy', <https://www.youtube.com/watch?v=QuOn93KPlr4> accessed 31 December 2023.

Google, 'Counter notification' https://support.google.com/youtube/answer/6005919?hl=de&ref_topic=9282678 accessed 31 December 2023.

Gorwa R, Binns R and Katzenbach C, 'Algorithmic content moderation' (2020) 7(1) Big Data & Society 1, <https://journals.sagepub.com/doi/epub/10.1177/2053951719897945> accessed 31 December 2023.

Gray J E and Suzor N P, 'Playing with machines: using machine learning to understand automated copyright enforcement at scale' (2020) 7(1) Big Data & Society 1, <https://journals.sagepub.com/doi/epub/10.1177/2053951720919963> accessed 31 December 2023.

Greger R and Stubbe C, *Schiedsgutachten* (Beck 2007).

Greger R, 'Recht der alternativen Konfliktlösung' in R Greger, H Unberath and F Steffek (ed), *Recht der alternativen Konfliktlösung* (2nd edn, Beck 2016) 270.

Grimmelmann J, 'The Virtues of Moderation' (2015) 17 YALE J.L. & TECH. 42.

Grosse Ruse-Kahn H, 'Automated Copyright Enforcement Online: From Blocking to Monetization of User-Generated Content' PIJIP Research Paper Series 51, <https://digitalcommons.wcl.american.edu/research/51> accessed 31 December 2023.

Grosse Ruse-Kahn H, 'Global Content Protection through Automation' (2018) 49(9) IIC (International Review of Intellectual Property and Competition Law) 1017.

Gsell B, 'Die Umsetzung der Richtlinie über alternative Streitbeilegung Juristisches Fachwissen der Streitbeilegenden Personen und Rechtstreue des Verfahrensergebnisses' (2015) 128(2) ZJP (Zeitschrift für Zivilprozess) 189.

Gsell B, Krüger W, Lorenz S and Reymann C (ed), *Beck-Online Grosskommentar zum BGB* (Beck 2023).

Gunning D et al., XAI-Explainable artificial intelligence, 4(37) Science Robotics (2019), <https://www.science.org/doi/10.1126/scirobotics.aay7120> accessed 31 December 2023.

Haber E, 'Privatization of the Judiciary' (2016) 40(1) Seattle University Law Review 115.

Halis Kasap G, 'Can Artificial Intelligence ("AI") Replace Human Arbitrators?' (2021) (2) Journal of Dispute Resolution 209.

Hartung M, 'Sonstige Akteure und Rahmenbedingungen' in M Hartung, M-M Bues and G Halbleib (ed), *Legal Tech: Die Digitalisierung des Rechtsmarkts* (Beck 2018) 215.

Heetkamp S J and Piroutek C, 'ChatGPT in Mediation und Schlichtung' (2023) 26(3) ZKM (Zeitschrift für Konflikt-management) 80.

Heiss S, 'Artificial Intelligence Meets European Union Law' (2021) 10(6) EuCML (Journal of European Consumer and Market Law) 252.

Heldt A, *Intensivere Drittwirkung* (Mohr Siebeck 2023).

Helfer L and Land M K, 'The Meta Oversight Board's Human Rights Future' (2023) 44(106) Cardozo Law Review 2233.

Helmond A, 'The Platformization of the Web: Making Web Data Platform Ready' (2015) Social Media and Society 1.

Hess B, *Europäisches Zivilprozessrecht* (2nd edn, De Gruyter 2021).

Hess B, 'Prozessuale Mindestgarantien in der Verbraucherschlichtung' (2015) 70(11) JZ (Juristenzeitung) 548.

Hess T and Waltermann H, 'Upload-Filter für Content' (2019) 16(2) MedienWirtschaft 19.

Hoeren T, 'Sperrpflichten eines Hosting-Anbieters bei rechtswidrigen Informationen sowie wort- und sinngleichen Inhalten' (2020) (4) LMK (Leitsätze mit Kommentierung) 425949.

Hofmann F, 'Die neuen Transparenzvorgaben im UWG 2022 im Kontext lauterkeitsrechtlicher Plattformregulierung' (2022) 124(11) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 780.

Hofmann F, 'Fünfzehn Thesen zur Plattformhaftung nach Art 17 DSM-RL' (2019) 121(12) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 1219.

Hofmann F, 'Mittelbare Verantwortlichkeit im Internet' (2017) 57(8) JuS (Juristische Schulung) 713.

Hofmann F, 'Prozeduralisierung der Haftungsvoraussetzungen im Medienrecht – Vorbild für die Intermediärhaftung' (2017) 61(2) ZUM (Zeitschrift für Urheber- und Medienrecht) 102.

Hofmann F and Raue B (ed), *Digital Services Act* (Nomos 2023).

Hofmann F and Specht-Riemenschneider L, 'Verantwortung von Online-Plattformen (Responsibility of Online Platforms)' (2021) 13(1) ZGE (Zeitschrift für geistiges Eigentum) 48.

Hofmann F and Sprenger T, 'Privatization of Enforcement' (2021) 85(2) UFITA (Archiv für Medienrecht und Medienwissenschaft) 249.

Holznagel D, 'Melde- und Abhilfeverfahren zur Beanstandung rechtswidrig gehosteter Inhalte nach europäischem und deutschem Recht im Vergleich zu gesetzlich geregelten notice and take-down-Verfahren' (2014) 63(2) GRUR Int. (Gewerblicher Rechtsschutz und Urheberrecht Internationaler Teil) 105.

Holznagel D, *Notice and Takedown as Part of Provider Liability* (Mohr Siebeck 2013).

Holznagel D, 'Nutzerrechte bei Facebook: Klärung durch den BGH und bevorstehende Irrwege des EU-Gesetzgebers' (2021) 37(11) CR (Computer und Recht) 733.

Holznagel D, 'Zu starke Nutzerrechte in Art. 17 und 18 DSA' (2022) 38(9) CR (Computer und Recht) 594.

Horwitz J, 'Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt', (2021) Wall Street Journal, <https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353> accessed 31 December 2023.

Hörnle J, *Internet Jurisdiction* (Oxford University Press 2021).

Jacques S, Garstka K, Hviid M and Street J, 'An empirical study of the use of automated anti-piracy systems and their consequences for cultural diversity' (2018) 15(2) Script-Ed 277.

Janal R, 'Haftung und Verantwortung im Entwurf des Digital Services Acts' (2021) 29(2) ZEuP (Zeitschrift für Europäisches Privatrecht) 227.

Jensen O, *Tribunal secretaries in international arbitration* (Oxford University Press 2019).

Jus Mundi, 'AI-Powered Search for International Law and Arbitration', <https://jusmundi.com/en> accessed 31 December 2023.

Jus Mundi, 'Jus Mundi Introduces Jus AI: A Game-Changing GPT-Powered AI Solution for the Arbitration Community' <https://dailyjus.com/news/2023/06/jus-mundi-introduces-jus-ai-a-game-changing-gpt-powered-ai-solution-for-the-arbitration-community> accessed 31 December 2023.

Kadri T and Klonick K, 'Facebook v. Sullivan: Public Figures and Newsworthiness in Online Speech Online Speech' (2019) 93(1) Southern California Law Review 37.

Kaesling K, 'Evolution statt Revolution der Plattformregulierung' (2021) 65(3) ZUM (Zeitschrift für Urheber- und Medienrecht) 177.

Kaesling K, 'Privatising Law Enforcement in Social Networks: A Comparative Model Analysis' (2018) 11(3) Erasmus Law Review 151.

Kahneman D, *Thinking, Fast and Slow* (Farrar, Straus and Giroux 2013).

Kalbhenn J-C, 'Design Specifications for Chatbots, Deepfakes, and Emotion Recognition Systems' (2021) 65(8/9) ZUM (Zeitschrift für Urheber- und Medienrecht) 663.

Kaminski M E, 'Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability' (2019) 92(6) S. CAL. L. REV. 1529.

Katsh E and Rabinovich-Einy O, *Digital Justice* (Oxford University Press 2017).

Katyal S, 'Private Accountability in the Age of Artificial Intelligence' (2019) 66(1) UCLA Law Review 55.

Katzenbach C, 'The "Alghorithmic turn" in platform governance' (2022) 74(1 supp) Cologne Journal of Sociology and Social Psychology 283.

Kaulartz M, 'Smart Contract Dispute Resolution' in M Fries and B-P Paal (ed), *Smart Contracts* (Mohr Siebeck 2019) 73.

Kaye D, *Speech Police* (Columbia Global Reports 2019).

Keller D, 'Internet Platforms: Observations on Speech, Danger, and Money' (2018) Hoover Inst. Aegis Paper Series No 1807.

Kindt T, 'Blockchainbasierte dezentrale Streitbeilegungsverfahren und ihr Verhältnis zur Schiedsgerichtsbarkeit' (2023) 21(5) SchiedsVZ (Zeitschrift für Schiedsverfahren) 241.

Kleros, 'Arbitration Service', <https://kleros.io/en/> accessed 31 December 2023.

Klonick K, 'The New Governors: The People, Rules and Process Governing Online Speech' (2018) 131(6) Harvard Law Review 1598.

Klonick K, 'Why the History of Content Moderation Matters' (2018) TECHDIRT, <<https://www.techdirt.com/articles/20180129/21074939116/whyhistory-content-moderation-matters.shtml>> accessed 31 December 2023.

Kosseff J, 'Private Computer Searches and the Fourth Amendment' (2018) 14(2) I/S A Journal of Law and Policy 187.

Koulu R, 'Proceduralizing control and discretion: Human oversight in artificial intelligence policy' (2020) 27(6) Maastricht Journal of European and Comparative Law 720.

Kreis F, 'KI und ADR-Verfahren' in Kaulartz M and Braegelmann T (ed), *Rechtshandbuch Artificial Intelligence und Machine Learning* (Beck and Vahlen 2020) 633.

Krenzler M and Remmert F R (ed), *Rechtsdienstleistungsgesetz* (3rd edn, Nomos 2023).

Krüger W and Rauscher T (ed), *Münchener Kommentar zur ZPO* (6th edn, Beck 2022).

Kumar S and Kumar H, 'Mediation and Artificial Intelligence' (2021) 4(4) International Journal of Law Management & Humanities 1472.

Kumkar L, 'Plattform-Recht revisited: Umgang mit den Marktordnungen digitaler Plattformen de lege lata et ferenda' (2022) 30(3) ZEuP (Zeitschrift für Europäisches Privatrecht) 530.

Kuczerawy A, 'The Good Samaritan that wasn't: voluntary monitoring under the (draft) Digital Services Act' (2021) Verfassungsblog, <https://verfassungsblog.de/good-samaritan-dsa/> accessed 31 December 2023.

Kraul T (ed), *Das neue Recht der digitalen Dienste – Digital Services Act (DSA)* (Nomos 2023).

Ladeur K-H, 'Neue Institutionen für den Daten- und Persönlichkeitsschutz im Internet: „Cyber-Courts“ für die Blogosphere' (2012) 36(10) DuD (Datenschutz und Datensicherheit) 711.

Ladeur K-H, 'Schutz vor Verletzung von Persönlichkeitsrechten und "Desinformation" in sozialen Medien unter Bedingungen der politischen Polarisierung' <<https://verfassungsblog.de/personlichkeitsrecht-soziale-medien/>> accessed 31 December 2023.

Länderarbeitsgruppe, 'Abschlussbericht der Länderarbeitsgruppe, "Legal Tech: Herausforderungen für die Justiz"' (2019), https://www.schleswig-holstein.de/DE/landesregierung/ministerienbehoerden/II/Minister/Justizministerkonferenz/Downloads/190605_beschluesse/TOPI_11_Abschlussbericht.pdf?blob=publicationFile&v=1 accessed 31 December 2023.

Land M, 'The Problem of Platform Law: Pluralistic Legal Ordering on Social Media' in P SA Berman (ed), *The Oxford Handbook of Global Legal Pluralism* (Oxford University Press 2020) 974.

Laukemann B, 'Private law enforcement and intellectual property: Regulatory challenges in a digital era' in B Hess, E Jayme and H-P Mansel (ed), *Europa als Rechts- und Lebensraum: Liber Amicorum für Christian Kohler zum 75. Geburtstag* (Gieseking 2018) 357.

Laukemann B, 'Private Rechtsdurchsetzung zwischen (digitaler Selbsthilfe) und gerichtlichem Rechtsschutz' (2022) 8(3) ZfPW (Zeitschrift für die gesamte Privatrechtswissenschaft) 357.

Lawlift GmbH, 'Lawlift', <https://de.lawlift.com/> accessed 31 December 2023.

Ledwich M and Zautsev A, 'Algorithmic Extremism: Examining YouTube's Rabbit Hole of Radicalization', Cornell University (2019), <https://arxiv.org/abs/1912.11211> accessed 31 December 2023.

Leeb C-M, *Digitalisierung, Legal Technology und Innovation* (Duncker & Humboldt 2019).

Legal Grade AI, 'Luminace', <https://www.luminance.com/overview.html> accessed 31 December 2023.

Leenheer Zimmerman D, 'A Tale of Legislative Abdication' (2014) 35(1) Pace Law Review 260.

Leerssen P, 'An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation' (2023) 48 Computer Law & Security Review 6.

Lennartz J and Kraetzig V, 'Filtering fundamental Rights' (2022) Verfassungsblog, <https://verfassungsblog.de/filtering-fundamental-rights/> accessed 31 December 2023.

Lessig L, *Code and Other Laws of Cyberspace* (Basic Books 1999).

Lester T and Pachamano D, 'The Dilemma of False Positives: Making Content ID Algorithms More Conducive to Fostering Innovative Fair Use in Music Creation' (2017) 24(1) UCLA Entertainment Law Review 51.

Lev-Aretz Y, 'Second Level Agreements' (2012) 45(1) AKRON L. REV. 137.

Lew J, Mistelis L and Kröll S, *Comparative International Commercial Arbitration* (Wolters Kluwer 2003).

Lex Machina, 'Lex Machina Legal Analytics', <https://lexmachina.com/> accessed 31 December 2023.

Liesching M (ed), *Netzwerkdurchsetzungsgesetz* (Nomos 2018).

Lindquist D and Dautaj Y, 'AI in International Arbitration' (2021) (1) Journal of Dispute Resolution 39.

Loo R Van, 'Federal Rules of Platform Procedure' (2021) 88(4) University of Chicago Law Review 829.

Loo R Van, 'The Corporation as Courthouse' (2016) 33(2) Yale Journal on Regulation 547.

Luhmann N, *Legitimation durch Verfahren* (11th edn, Suhrkamp 2019).

Lüdemann J, 'Privatisierung der Rechtsdurchsetzung in sozialen Netzwerken?' in M Eifert and T Gostomzyk (ed), *Netzwerkrecht* (2018) 165.

Lüdemann J, 'Warum und wie reguliert man digitale Informationsintermediäre?' in J Lüdemann and Y Hermstrüwer (ed), *Der Schutz der Meinungsbildung im digitalen Zeitalter* (Mohr Siebeck 2021) 15.

Maier H, *Remixes on Hosting Platforms* (Mohr Siebeck 2018).

Marmont S, 'Keeping Up with Legal Technology' (2019) 1(2) ITA in Review 37.

Mast T, 'AGB-Recht als Regulierungsrecht' (2023) 78(7) JZ (Juristenzeitung) 287.

Maxwell G and Vannieuwenhuysen G, 'Robots Replacing Arbitrators: Smart Contract Arbitration' (2018) (1) ICC Dispute Resolution Bulletin 24.

Mayer M, *Soziale Netzwerke im Internet im Lichte des Vertragsrechts* (Richard Boorberg 2018).

McColgan P, 'Das wird man wohl noch löschen dürfen? – Control Standards for Opinion Rules on the Internet' (2021) 1(12) RDi (Recht Digital) 605.

Meder S, 'Die Zukunft der juristischen Methode: Rehabilitierung durch Chat-GPT?' (2023) 78(23) JZ (Juristenzeitung) 1041.

Medvedeva M, Vols M and Wieling M, 'Using machine learning to predict decisions of the European Court of Human Rights' (2020) 28(2) Artificial Intelligence and Law 237.

Mendelsohn J K, 'Die "normative Macht" der Plattformen – Gegenstand der zukünftigen Digitalregulierung' (2021) 24(11) MMR (Multimedia und Recht) 857.

Menkel-Meadow C, 'Is ODR ADR? Reflections of an ADR Founder from 15th ODR Conference, the Hague, the Netherlands, 22-23 May 2016' (2016) 3(1) IJODR (International Journal on Online Dispute Resolution) 4.

Meta, 'Community Standards Enforcement Report: Child Endangerment: Nudity and Physical Abuse and Child Sexual Exploitation', <https://transparency.fb.com/data/community-standards-enforcement/child-nudity-and-sexual-exploitation/facebook/> accessed December 2023.

Meta, 'Community Standards Enforcement Report', <https://transparency.fb.com/reports/community-standards-enforcement/> accessed 31 December 2023.

Meta, 'How Facebook uses super-efficient AI models to detect hate speech, 19 November 2020', <https://ai.facebook.com/blog/how-facebook-uses-super-efficient-ai-models-to-detect-hate-speech/> accessed 31 December 2023.

Meta, 'How we review content', 11 August 2020, <https://about.fb.com/news/2020/08/how-we-review-content/> accessed 31 December 2023.

Meta, 'How we review Content – Prioritization', 11 August 2020, <https://about.fb.com/news/2020/08/how-we-review-content/> accessed 31 December 2023.

Meta, 'Sharing More Details on How We Will Implement the Oversight Board's Decisions, Responding to the Oversight Board's First Decisions', 28 January 2021, <https://about.fb.com/news/2021/01/responding-to-the-oversight-boards-first-decisions/> accessed 31 December 2023.

Meta, 'Facebook Transparency Report of January 2023', <327151920_907084790305794_6193992151844220602_n.pdf> accessed 31 December 2023.

Meta Transparency Center, 'Global restrictions', <https://transparency.fb.com/data/content-restrictions/> accessed 31 December 2023.

Meta Transparency Center, 'How technology detects violations', 18 October 2023, <https://transparency.meta.com/de-de/enforcement/detecting-violations/technology-detects-violations/> accessed 31 December 2023.

Meta Transparency Center, 'How we assess reports of content violating local law', <https://transparency.fb.com/data/content-restrictions/> accessed 31 December 2023.

Meta Transparency Center, 'Reviewing high-impact content accurately via our cross-check system', 12 May 2023, <https://transparency.fb.com/enforcement/detecting-violations/reviewing-high-visibility-content-accurately/> accessed 31 December 2023.

Metzger A and Senftleben M, 'Selected Aspects of Implementing Article 17 of the Directive on Copyright in the Digital Single Market into National Law – Comment of the European Copyright Society' (2020) 11(2) *Journal of Intellectual Property, Information Technology and E-Commerce Law* 1.

Milano D, *Content control: Digital Watermarking and Fingerprinting* (Rhozet 2013).

Monroy M, 'EU-Kommission droht mit "gesetzgeberischen Maßnahmen" zur Entfernung von Internetinhalten', <https://netzpolitik.org/2018/eu-kommission-droht-mit-gesetzgeberischen-massnahmen-zur-entfernung-von-internetinhalten/> accessed 31 December 2023.

Monroy M, 'EU-Internetforum: Viele Inhalte zu "Extremismus" werden mit Künstlicher Intelligenz aufgespürt', <https://netzpolitik.org/2018/eu-kommission-droht-mit-gesetzgeberischen-massnahmen-zur-entfernung-von-internetinhalten/> accessed 31 December 2023.

Mostert F, 'Free Speech and Internet Regulation' (2019) 14(8) *Journal of Intellectual Property Law & Practice* 607.

Müller-Terpitz R and Köhler M (ed), *Digital Services Act* (Beck 2024).

Muller C, 'Setting the Record Straight', <https://blog.youtube/news-and-events/setting-record-straight/> accessed 31 December 2023.

Murgia M, Warell H and Bond D, 'YouTube revenues under threat over ads alongside extremist videos' (2017) *Financial Times*, <https://www.ft.com/content/04f8bf56-0b12-11e7-97d1-5e720a26771b> accessed 31 December 2023.

Musa S and Bendett S, 'Islamic Radicalization in the United States – New Trends and a Proposed Methodology for Disruption' (2010) Washington D.C.: National Defense

University, Washington DC Center for Technology and National Security Policy, <https://apps.dtic.mil/sti/pdfs/ADA532696.pdf> accessed 31 December 2023.

Nahmias Y and Perel M, 'The oversight of content moderation by AI: Impact assessment and their limitations' (2021) 58(1) *Harvard Journal on Legislation* 145.

Nathenson I S, 'The Procedural Foundations of Information Regulation' (2020) 24(1) *Lewis & Clark Law Review* 109.

Nink D, *Justiz und Algorithmen* (Duncker & Humboldt 2021).

Nolte G, 'Three Theses on the Current Debate on Liability and Distributive Justice in Hosting Services with User-Generated Content (the so-called "Value Gap" Debate)' (2017) 61(4) *ZUM (Zeitschrift für Urheber- und Medienrecht)* 304.

Nunziato D C, 'The Beginning of the End of Internet Freedom' (2014) 45(2) *Georgetown Journal of International Law* 383.

Ohly A and Sosnitza O (ed), *Gesetz Gegen den unlauteren Wettbewerb: UWG* (8th edn, Beck 2023).

Orsich I, 'Das europäische Konzept für vertrauenswürdige Künstliche Intelligenz' (2022) 33(6) *EuZW (Europäische Zeitschrift für Wirtschaftsrecht)* 254.

Otter, 'Otter AI', <https://otter.ai> accessed 31 December 2023.

Paisley K and Sussman E, 'Artificial Intelligence Challenges and Opportunities for International Arbitration' (2018) 11(1) *New York Dispute Resolution Lawyer* 35.

Pasquale F, *The Black Box Society* (Harvard University Press 2015).

Peifer K-N, 'Die neuen Transparenzregeln im UWG (Bewertungen, Rankings und Influencer)' (2021) 123(12) *GRUR (Gewerblicher Rechtsschutz und Urheberrecht)* 1453.

Perel M and Elkin-Koren N, 'Accountability in Algorithmic Copyright Enforcement' (2016) 19(3) *Stan. Tech. L. Rev.* 473.

Plantin J-C, Lagoze C, Edwards P N and Sandvig C, 'Infrastructure Studies Meet Platform Studies in the Age of Google and Face-book' (2016) 20(1) *New MEDIA & Society* 293.

Polkinghorne M, 'Different Strokes for Different Folks?', <<https://arbitrationblog.kluwerarbitration.com/2014/05/17/different-strokes-for-different-folks-the-role-of-the-tribunal-secretary-2/>> accessed 31 December 2023.

Prütting H, 'Das neue Verbraucherstreitbeilegungsgesetz: Was sich ändert – und was bleiben wird' (2016) (3) *AnwBl (Anwaltsblatt)* 190.

Prütting H, 'Die rechtliche Stellung des Schiedsrichters' (2011) 9(5) *SchiedsVZ (Zeitschrift für Schiedsverfahren)* 233.

Queen Mary University and White & Case, '2021 International Arbitration Survey: Adapting Arbitration to a Changing World', <https://arbitration.qmul.ac.uk/research/2021-international-arbitration-survey/> accessed 31 December 2023.

Rabinovich-Einy O and Katsh E, 'The New New Courts' (2017) 67(1) Am. U. L. Rev. 165.

Radu R, *Negotiating Internet Governance* (Oxford University Press 2019).

Rajendra J B and Thuraisingam A S, 'The deployment of artificial intelligence in alternative dispute resolution: the AI augmented arbitrator' (2022) 31(2) Information & Communications Technology Law 176.

Raue B, 'Plattformnutzungsverträge im Lichte der gesteigerten Grundrechtsbindung marktstarker sozialer Netze' (2022) 75(2) NJW (Neue Juristische Wochenschrift) 209.

Raue B and Heesen H, 'Der Digital Services Act' (2022) 75(49) NJW (Neue Juristische Wochenschrift) 3537.

Reuter M, 'Facebook Knew What All Was Going Wrong' (2021) Netzpolitik.org, <https://netzpolitik.org/2021/facebook-files-facebook-wusste-was-alles-schieflaeuft/> accessed 31 December 2023.

Riesenhuber K, '§ 10 Die Auslegung' in K Riesenhuber (ed), *Europäische Methodenlehre* (De Gruyter 2021) 285.

Richter P and Mendelsohn J, '§ 21 Plattformspezifische Vorgaben des Data Acts' in B Steinrötter (ed), *Europäische Plattformregulierung* (Nomos 2023) 564.

Rhim Y and Park K, 'The Artificial Intelligence in International Law' in E Y J Lee (ed), *Revolutionary Approach to international Law: The Role of international Lawyer in Asia* (Springer 2023) 215.

Röthemeyer P, 'Die Schlichtung' (2013) 16(2) ZKM (Zeitschrift für Konfliktmanagement) 47.

Ruger T, Kim P, Martin A and Quinn K, 'The Supreme Court Forecasting Project' (2004) 104(4) Colum. L. Rev. 1150.

Rule C, 'Making Peace on eBay' (2008) ACR Resolution 8.

Rule C, 'Quantifying the Economic Benefits of Effective Redress: Large ECommerce Data Sets and the Cost-Benefit Case for Investing in Dispute Resolution' (2012) 34(4) U. ARK. LITTLE ROCK L. REV. 767.

Rühl G, 'Die Richtlinie über alternative Streitbeilegung: Handlungsperspektiven und Handlungsoptionen' (2014) 127(1) ZZP (Zeitschrift für Zivilprozess) 61.

Rühl G, 'Digitale Justiz' (2020) 75(17) JZ (Juristenzeitung) 809.

Rühl G, 'KI in der gerichtlichen Streitbeilegung' in M Kaulartz and T Braegelmann (ed), *Rechtshandbuch Artificial Intelligence und Machine Learning* (Beck 2020) 617.

Salter S and Thompson D, 'Public-Centred Civil Justice Redesign' (2016-2017) 3 McGill Journal of Dispute Resolution 113.

Saurwein F, 'Regulierung von Internet-Inhalten: Ombudsstellen als Governance-Option an der Schnittstelle von Recht und Ethik' in G Marci-Boehncke, M Rath, M Delere and H Höfer (ed), *Medien – Demokratie – Bildung* (Springer 2022) 47 https://doi.org/10.1007/978-3-658-36446-5_5 accessed 31 December 2023.

Schillmöller J and Doseva S, '"Chilling effects" durch YouTubes Content ID?' (2022) 25(3) MMR (Multimedia und Recht) 181.

Schneiders P, 'Hate Speech auf Online-Plattformen: Problematization, Regulation and Evaluation against the Background of the Proposal for a Digital Services Act' (2021) 85(2) UFITA (Archiv für Medienrecht und Medienwissenschaft), <https://doi.org/10.5771/2568-9185-2021-2-269> accessed 31 December 2023 accessed 31 December 2023.

Scherer M, 'Artificial Intelligence and Legal Decision-Making' (2019) 36(5) Journal of International Arbitration 539.

Scherer M, 'International Arbitration 3.0. How Artificial Intelligence Will Change Dispute Resolution' in C Klausegger et al (ed), *Austrian Yearbook on International Arbitration* (Beck 2019) 503.

Scherer M and Jensen O, 'Die Digitalisierung der Schiedsgerichtsbarkeit' in T Riehm and S Dörr (ed), *Digitalisierung und Zivilverfahren* (De Gruyter 2023), 591.

Schwartz J, 'Artificial Arbitration?' in R Wilhelmi and M Stürner (ed), *Mehrparteien-Schiedsverfahren: Unter besonderer Berücksichtigung gesellschaftsrechtlicher Streitigkeiten* (Springer 2021) 95.

Seetharaman D, Horwitz J and Scheck J, 'Facebook Says AI Can Enforce Its Rules, but the Company's Own Engineers Are Doubtful' (2021) Wall Street Journal, <https://www.wsj.com/articles/facebook-ai-enforce-rules-engineers-doubtful-artificial-intelligence-11634338184> accessed 31 December 2023.

Sela A, 'The Effect of Online Technologies on Dispute Resolution System Design' (2017) 21(3) Lewis & Clark Law Review 633.

Senftleben M, 'Institutionalized Algorithmic Enforcement – The Pros and Cons of the EU Approach to UGC Platform Liability' (2020) 14(2) FIU Law Review 299.

Shaughnessy P and Rogers C, 'Arbitrator Intelligence - An Interview with its Founder and Director, Professor Catherine Rogers' (2015) 1(1) Journal on Technology in International Arbitration 87.

Shinn L D, 'YouTube's Content ID as a Case Study of Private Copyright Enforcement Systems' (2015) 43(2/3) AIPLA Quarterly Journal 359.

Silicon Valley Arbitration and Mediation Center, 'Silicon Valley Arbitration & Mediation Center Guidelines, Draft of August 31, 2023', <https://thearbitration.org/wp-content/uploads/2023/08/SVAMC-AI-Guidelines-CONSULTATION-DRAFT-31-August-2023-1.pdf> accessed 31 December 2023.

Sim C, 'Will Artificial Intelligence Take Over Arbitration?' (2018) 14(1) Asian International Arbitration Journal 1.

Simshaw D, 'Access to A.I. Justice: Avoiding an Inequitable Two-Tiered System of Legal Services' (2022) 24 Yale Journal of Law & Technology 150.

Snijders H, 'Arbitration and AI, Arbitration 2023' in H Snijders (ed), *Arbitration and AI, Arbitration* (Wolters Kluwer 2023) 224.

Solomon L, 'Fair users or content abusers' (2015) 44(1) Hofstra L. Rev. 237.

Specht F, 'Chancen und Risiken einer digitalen Justiz für den Zivilprozess' (2019) 22(3) MMR (Multimedia und Recht) 153.

Spindler G, 'Der Vorschlag für ein neues Haftungsregime für Internetprovider – der EU-Digital Services Act (Teil 1)' (2021) 123(4) GRUR (Gewerblicher Rechtsschutz und Urheberrecht) 545.

Spoerri T, 'On Upload Filters and other Competitive Advantages for Big Tech Companies under Article 17 of the Directive on Copyright in the Digital Single Market' (2019) 10(2) Journal of Intellectual Property, Information Technology and Electronic Commerce Law 173.

Staudinger J von von (ed), *Kommentar zum BGB*, Buch 2 (19th rev edn, De Gruyter 2022).

Stotz R, 'Die Rechtsprechung des EuGH' in K Riesenhuber (ed) *Europäische Methodenlehre* (De Gruyter 2021) 653.

Sunstein C R, *Republic – Divided Democracy in the Age of Social Media* (Princeton University Press 2017).

Suzor N P, *Lawless: The Secret Rules that Govern our Digital Lives* (Cambridge University Press 2019).

Suzor N P and Gray J E, 'Playing with machines: Using machine learning to understand automated copyright enforcement at scale' (2020) 7(1) Big Data & Society, <https://doi.org/10.1177/2053951720919963> accessed 31 December 2023.

Sweeney L, 'Discrimination in Online Ad Delivery' May 2013 Comm. ACM 44, <http://cacm.acm.org/magazines/2013/5/163753-discrimination-in-online-ad-delivery/> accessed 31 December 2023.

Taeger J and Kremer S, *Recht im E-Commerce und Internet* (Beck 2021).

Tan V, 'Online Dispute Resolution for Small Civil Claims in Victoria' (2019) 24 Deakin Law Review 101.

Titlow J P, 'Youtube is using AI to police copyright to the tune of \$2 billion in payouts', 31 July 2016, <https://www.fastcompany.com/4013603/youtube-is-using-ai-to-police-copyright-to-the-tune-of-2-billion-in-payouts> accessed 31 December 2023.

Trint, 'Trint', <https://trint.com/> accessed 31 December 2023.

Tushnet R, 'All of this has happened before and all of this will happen again: Innovation in copyright licensing' (2014) 29(3) Berkeley Technology Law Journal 1447.

Twitter, 'Network Enforcement Report: January-June 2023', <https://transparency.twitter.com/content/dam/transparency-twitter/country-reports/germany/NetzDG-Jan-Jun-2023.pdf> accessed 31 December 2023.

Tyler T R, 'What is procedural justice?: Criteria used by citizens to assess the fairness of legal procedures' (1988) 22(1) Law & Society Review 103.

Tyler T R, *Why people obey the law* (Princeton University Press 2006).

Urban J, Karagani J and Schofield B, 'Notice and Takedown in Everyday Practice' (2016) UC Berkeley Public Law Research Paper No 2755628, <https://ssrn.com/abstract=2755628> or <http://dx.doi.org/10.2139/ssrn.2755628>.

Valkanova M, 'Trainieren von KI-Modellen' in M Kaulartz and T Braegelmann (ed), *Rechtshandbuch Artificial Intelligence und Machine Learning* (Beck 2020) 336.

Vannieuwenhuysse G, 'Arbitration and New Technologies: Mutual Benefits' (2018) 35(1) Journal of International Arbitration 119.

Voß W, 'Gerichtsverbundene Online-Streitbeilegung' (2020) 84(1) RabelsZ (Rabels Zeitschrift für ausländisches und internationales Privatrecht) 62.

Wagner B, *Global Free Expression – Governing the Boundaries of Internet Content* (Springer 2016).

Wagner G, 'Haftung von Plattformen für Rechtsverletzungen (Teil 2)' (2020) 122(5) GRUR (Gewerblicher Rechts-schutz und Urheberrecht) 447.

Wagner G, 'Private Law Enforcement and ADR' in J Zekoll, M Bälz and I Amelung (ed), *Formalisation and Flexibilisation in Dispute Resolution* (Brill 2014) 369.

Wagner J, *Legal Tech und Legal Robots* (2nd edn, Springer Gabler 2020).

Walker K, 'Four ways Google will help to tackle extremism' (18 June 2017) Financial Times, <https://www.ft.com/content/ac7ef18c-52bb-11e7-a1f2-db19572361bb> accessed 31 December 2023.

Wall Street Journal, 'The Facebook Files' (13 September 2021), <https://www.wsj.com/articles/the-facebook-files-11631713039> accessed 31 December 2023.

Welser M von, 'Die KI-Verordnung – ein Überblick über das weltweit erste Regelwerk für künstliche Intelligenz' (2024) 16(15) GRUR-Prax (Gewerblicher Rechtsschutz und Urheberrecht in der Praxis) 485.

Wendland M, *Mediation und Zivilprozess* (Mohr Siebeck 2017).

Wielsch D, 'Die Ordnungen der Netzwerke, AGB – Code – Community Standards' in M Eifert and T Gostomzyk (ed), *Netzwerkrecht* (Nomos 2018) 61.

Wielsch D, 'Medienregulierung durch Persönlichkeits- und Datenschutzrechte' (2020) 75(3) JZ (Juristenzeitung) 105.

Wilske S, Markert L and Ebert B, 'Entwicklungen in der internationalen Schiedsgerichtsbarkeit im Jahr 2022 und Ausblick auf 2023' (2023) 21(3) SchiedsVZ (Zeitschrift für Schiedsverfahren) 121.

Wolfowicz M, Weisburd D and Hasisi B, Examining the interactive effects of the filter bubble and the echo chamber on radicalization, (2023) 19(5) Journal of Experimental Criminology 119.

Wolters Kluwer, 'Arbitration Database', <https://www.kluwerarbitration.com/> accessed 31 December 2023.

Wolters Kluwer, 'smartlaw', <https://www.smartlaw.de/> accessed 31 December 2023.

Wu T, 'Will Artificial Intelligence Eat the Law? The Rise of Hybrid Social-Ordering Systems' (2019) 119(7) Columbia Law Review 2001.

Yeung K, "'Hybernudge": Big Data as a mode of regulation by design, Information' (2017) 20(1) Communication & Society 118.

Yablon Y and Landsman-Ross N, 'Predictive Coding' (2013) 64(3) South Carolina Law Review 633.

YouTube, 'Accelerated Content ID & Complaint Process', <https://support.google.com/youtube/thread/171619847> accessed 31 December 2023.

YouTube, 'Answers to common questions about Copyright claims on YouTube', <https://support.google.com/youtube/thread/1281991> accessed 31 December 2023.

YouTube, 'Appeal a Content ID claim', <https://support.google.com/youtube/answer/12104471> accessed 31 December 2023.

YouTube, 'Best practices for claims', <https://support.google.com/youtube/answer/4352063> accessed 31 December 2023.

YouTube, 'Best practices for references', <https://support.google.com/youtube/answer/107008> accessed: 31 December 2023.

YouTube, 'Content eligible for Content ID',
<https://support.google.com/youtube/answer/2605065?hl=en> accessed 31 December 2023.

Youtube 'Copyright Strike Basics'
<https://support.google.com/youtube/answer/2814000#zippy=%2Cfolgen-einer-urheberrechtsverwarnung> accessed 31 December 2023.

YouTube, 'Copyright Transparency Report H2 2021',
https://storage.googleapis.com/transparencyreport/report-downloads/pdf-report-22_2021-7-1_2021-12-31_en_v1.pdf accessed 31 December 2023.

YouTube, 'Create an asset',
https://support.google.com/youtube/answer/3011552?hl=en&ref_topic=3011550 accessed 31 December 2023.

YouTube, 'Deliver content using spreadsheet templates',
<https://support.google.com/youtube/answer/6066171> accessed 31 December 2023.

YouTube, 'Dispute a Content ID claim',
https://support.google.com/youtube/answer/2797454?hl=en&ref_topic=9282678#zippy=%2Coptionen-f%C3%BCr-den-anspruchsteller accessed 31 December 2023.

YouTube, 'Fix reference overlaps',
https://support.google.com/youtube/answer/3022604?hl=en&ref_topic=3013248 accessed 31 December 2023.

YouTube, 'Frequently asked questions about copyright',
<https://support.google.com/youtube/answer/2797449?hl=en> accessed 31 December 2023.

YouTube, 'Frequently asked questions about fair use',
<https://support.google.com/youtube/answer/6396261#zippy=%2Ci-posted-a-disclaimer-on-my-video%2Ci-gave-credit-to-the-copyright-owner%2Cim-using-the-content-for-entertainment-or-non-profit-uses%2Cwhen-does-fair-use-apply%2Cwhat-constitutes-fair-use%2Chow-does-fair-use-work%2Chow-does-content-id-work-with-fair-use> accessed 31 December 2023.

YouTube, 'How policies are applied',
<https://support.google.com/youtube/answer/3369929> accessed 31 December 2023.

YouTube, 'Monetization during Content ID disputes',
https://support.google.com/youtube/answer/7000961?hl=en&ref_topic=9282678 accessed 31 December 2023.

YouTube, 'Qualify for Content ID',
<https://support.google.com/youtube/answer/1311402> accessed 31 December 2023.

YouTube, 'Requirements for counter notifications', https://support.google.com/youtube/answer/6005919?hl=en&ref_topic=9282678 accessed 31 December 2023.

YouTube, 'Review potentially invalid references', <https://support.google.com/youtube/answer/6013183> accessed 31 December 2023.

YouTube, 'Update: Improving Content ID for creators', <https://blog.youtube/news-and-events/update-improving-content-id-for-creators/> accessed 31 December 2023.

YouTube, 'Use Content ID matching on live streams', <https://support.google.com/youtube/answer/9896248?hl=en> accessed 31 December 2023.

YouTube, 'Using Content ID', <https://support.google.com/youtube/answer/3244015?hl=en> accessed 31 December 2023.

YouTube, 'Using the YouTube DDEX feed', <https://support.google.com/youtube/topic/3505247> accessed 31 December 2023.

YouTube, 'What are policies?', https://support.google.com/youtube/answer/107383?hl=en&ref_topic=24332 accessed 31 December 2023.

YouTube, 'What Does Fair Use Mean', <https://support.google.com/youtube/answer/9783148?hl=de> accessed 31 December 2023.

YouTube, 'What is copyright?', <https://support.google.com/youtube/answer/2797466#refrained&zipppy=%2Cmissvest%3%A4ndnis-nr-if-you-angive-that-your-content-does-not-serve-commercial-purposes-you-can-use-any-content> accessed 31 December 2023.

YouTube, 'What is a reference?' <https://support.google.com/youtube/answer/107004?hl=en> accessed 31 December 2023.

YouTube, 'Where can I get more information about copyright outside the U.S.?' https://support.google.com/youtube/answer/2797449?hl=de&ref_topic=2778546#zipppy=%2Cwo-erhalte-ich-weitere-informationen-zum-urheberrecht-außerhalb-der-usa accessed 31 December 2023.

YouTube, 'YouTube Partner Program overview & eligibility' <https://support.google.com/youtube/answer/72851> accessed 31 December 2023.

Zawada K L, 'The Emergence and Development of Content ID in Light of User-generated Law, in How Deep is your Law?' (2017) 5th International Conference of PhD Students and Young Researchers Conference Papers 483.

Zekos G, *Advanced Artificial Intelligence and Robo-Justice* (Springer 2022).

Zhou T, Postmortem: 'Every Frame a Painting (2017)' <https://perma.cc/U5WU-M6ZZ> accessed 31 December 2023.

Zorrilla E, 'Towards a Credible Future' (2018) 16(2) *SchiedsVZ* (Zeitschrift für Schiedsverfahren) 106.
