jipitec

3 | 2023

Volume 14 (2023)
Issue 3 ISSN 2190-3387

Journal of Intellectual Property, Information Technology, and Electronic Commerce Law

www.jipitec.eu

2023

jipitec

Journal of Intellectual Property, Information Technology and Electronic Commerce Law

Volume 14 Issue 3 Oct 2023
www.jipitec.eu
contact@jipitec.eu

A joint publication of:

Prof. Dr. Thomas Dreier, M. C. J. (NYU)
KIT - Karlsruher Institut für Technologie, Zentrum für Angewandte Rechtswissenschaft (ZAR), Vincenz-Prießnitz-Str. 3, 76131 Karlsruhe Germany

Prof. Dr. Axel Metzger, LL. M. (Harvard)
Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin

Prof. Dr. Gerald Spindler (†)
Dipl.-Ökonom, Georg-August-Universität Göttingen, Platz der Göttinger Sieben 6, 37073 Göttingen

Karlsruhe Institute of Technology, Humboldt-Universität zu Berlin and Georg-August-Universität Göttingen are corporations under public law, and represented by their respective presidents.

ISSN 2190-3387

Funded by
DGRI Deutsche Gesellschaft für Recht und Informatik e.V.

Table Of Contents

Editors' note

Prof. Dr. Gerald Spindler (1960-2023)

This is the first issue that we publish after the untimely and sudden passing of Prof. Dr. Gerald Spindler, initiator and co-founder of our online journal JIPITEC, at only 62, on 11 September 2023.

We are eternally indebted to his tireless commitment to JIPITEC, to his constant efforts in the promotion, coordination and generous support of the journal through his Chair at the Georg-August Universität in Göttingen. The international legal community will miss his immense knowledge and contribution in the broad areas of information technology and commerce law.

Thomas Dreier
Séverine Dusollier
Lucie Guibault
Orla Lynskey
Axel Metzger
Miquel Peguera
Karin Sein
Lars Flamme

Editorial

by **Enguerrand Marique ***

Recommended citation: Enguerrand Marique, Editorial, 14 (2023) JIPITEC 362 para 1.

1 Over the last twenty years, justice systems have increasingly experimented with digitalization. This phenomenon has been praised for its ability to gather and process vast amounts of information, foster innovation, and offer cost-effectiveness. All of these features should improve the smooth functioning of the European market by contributing to secure economic transactions, as they are expected to solve disputes expeditiously and inexpensively. However, digitalization meets with resistance due to its drawbacks (especially, unequal access to digital tools, biases, and replications of past solutions).

2 The European Union has made continuous efforts to modernise judicial procedures in Europe through the advancement of digitalization, as exemplified by the adoption of Regulation 861/2007[1] and Reglation 2020/1783[2]. However, the emergence of new technologies brings with it the need for ongoing assessment of their impact – as evidenced by the cautionary approach taken in the Artificial Intelligence Act proposal, which allows the use of Artificial Intelligence (hereafter 'AI') only when strict conditions are met.[3]

3 Concurrently, private players within the digital industry have pro-gressively assumed roles as dispute resolution agents. Platforms must solve the issues that arise with their users, between their users, as well as between users and third parties (such as copyrights dis-putes). While the e-commerce Directive[4], adopted in 2000, had long been regulating (or exempting from liability) digital intermediaries, it failed to address this particular role of private entities. In 2022, the European Union chose to address the issue and imposed new re-quirements on digital platforms – mostly of a procedural nature, in order to facilitate online dispute resolution, as well as to offer mini-mal safeguards. These issues are now (partially) dealt with in the Digital Services Act (hereafter 'DSA')[5].

4 These various developments have raised a set of questions, which were compiled in a call for papers in June 2022, resulting in the publication of the present special issue on Administration of Justice in

* Enguerrand Marique is an Assistant Professor at Radboud Universiteit and a Guest Lecturer at the UCLouvain.

1 European Parliament and Council Regulation 861/2007 establishing a European Small Claims Procedure, *OJ L* 199, 31 July 2007, p. 1–22.

2 European Parliament and Council Regulation 2020/1783 on cooperation between the courts of the Member States in the taking of evidence in civil or commercial matters (taking of evidence) (recast), *OJ L 405*, 2 December 2020, p. 1–39.

3 European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act, Com/2021/206 Final, 21 April 2021.

4 European Parliament and Council Directive 2000/31/EC on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, *OJ L 178,* 17 July 2000, p. 1–16.

5 European Parliament and Council Regulation 2022/2065 on a Single Market For Digital Services, *OJ L 277*, 27 October 2022, p. 1–102.

the Digital Era. Part 1 of the present issue contains a series of papers directly addressing the digitalization of administration of justice at the European level. Part 2 offers a number of regular JIPITEC articles, not making part of the call for papers, but nonetheless, also prompting contemplation of the evolving role of courts and judges in a digital era.

Part 1. Administration of Justice in a Digital Era

5 Kalliopi Terzidou sets the stage for the use of AI in the European Union for court administration. This contribution reviews the definitions and typologies that have been applied to the concept of AI and how these approaches influence the perception of this technology in procedural laws. For instance, this form of *intelligence* should be rather conceptualized as a "thinking" tool that therefore can only support the "thinking" process of judges by providing arguments that judges and courts could add in their decision-making thoughts. This contribution argues that the integration of AI applications in courts must be subject to supervision and regulation. Both the EU and its Member States should keep an interest in the management, development, and implementation of this high-risk use of AI.

6 Within this large framework, Jura Golub evaluates the use of artificial intelligence in detecting deception or assessing the credibility in witnesses and experts' testimonies delivered through videoconferencing systems. While the taking of evidence is normally left to Member States procedural autonomy, Regulation 2020/1783 facilitates cooperation in the cross-border taking of evidence in civil matters, which favours videoconferencing for immediacy and simplicity. However, non-verbal cues are often more challenging to discern when a videoconferencing system stands between the witness and the judges or lawyers. The author conducts an assessment of the compliance of AI use in this context with fair trial principles, protection of personal data, as well as with the current proposal for an Artificial Intelligence Act. The author emphasizes the need for transparency as well as the need for a harmonized and consistent approach to the use of AI in the cross-border evidence collection with videoconferencing.

7 Federica Casarosa examines the role of platforms in content moderation policies and identifies the risks to freedom of expression that arise from these practices. This researcher then proceeds to answer the question of how the DSA seeks to mitigate these risks by providing for new procedural remedies against account suspension and termination, content removal, as well as monetization restrictions. These external, out-of-court, remedies are accompanied by procedural guarantees. On the one hand, the DSA stipulates the independence, impartiality and expertise required in the decision-making process. On the other hand, it establishes standards for accessibility, transparency, and fairness of the procedure. The dispute settlement bodies should receive an accreditation that certifies that these guarantees are being complied with. The author also calls for clarification on this accreditation process.

8 Within the context of platforms' content moderation policies, Pieter Wolters and Raphaël Gellert examine the notice-and-action mechanisms on digital platforms. Under the e-commerce Directive, online hosting services providers (i.e. digital platforms) were indeed exempt of liability for information created by third parties if they met two conditions. First, these providers had to remain passive and neutral, i.e. not modify or optimize the content. Second, they had to remove illegal content when they become aware of its illegal nature. However, the e-commerce Directive failed to provide procedures for notifying this illegal character, meaning that an individual had potentially to notify the infringement by post, and led to an underenforcement of compelling legislations on digital platforms. To address this issue, the DSA proposes a new notification procedure that compels digital platforms to act. The authors explore how this new procedure can protect victims while safeguarding fundamental rights (including the issues identified by Casarosa) as well as the economic interests of digital platforms. The authors conclude that there are still some gaps in the system, but that the DSA significantly improves the practices of moderation of online content.

9 Gregory Chan and Tan Yan Shen explore the outcome of online dispute resolution procedures and identify the gap that emerges between the policies and the resolution of cases across platforms. In terms of procedure, the authors examine the inequality of arms between the litigants, as buyers are often given more power than sellers on marketplaces. They also address the issue of disproportionate penalties resulting from ODR procedures. Indeed, platforms take sometimes a black-and-white approach to sanctioning users of the platforms, rather than adopting a more nuanced stance. The contributors also regret the lack of a uniform interpretation principle of the contracts between the users and the platforms, as well as between users themselves. They argue that ODR, therefore, hinders the consistent and systematic implementation of the law. Consequently, the authors recommend classifying the types of disputes that may arise and applying uniform guiding principles for assessing the merits of claims across platforms for the same types of disputes.

Part 2. Other articles on digital issues

10 The second part of this issue presents four articles not related to the call for papers on the administration of justice in the digital era. Nonetheless, they are also connected in some respects to the issues of delegating decision-making powers, implementing self-enforcing rules that remove the judge from the process, and to argumentative and interpretative techniques that judges need to use in the new digitalized context.

11 Considering the Copyright in the Digital Single Market Directive[6] ('CDSMD'), Martin Senftleben shows how the EU lawmaker has outsourced the protection of users' fundamental rights to private parties. The author addresses potential corrective measures that might mitigate this delegation of protection to the industry, including user complaint mechanisms, safeguards implemented in the CDSMD, Member States transposition measures seeking to address this issue, as well as the audit reports that the very large online platforms need to go through and submit to the European Commission under the Digital Services Act.

12 Dário Moura Vicente explores the issue of disinformation on the Internet. After examining the European Action Plan against Disinformation, the author analyses a specific instrument that was adopted in Portugal at the time of its presidency of the European Union: the *Portuguese Charter of Human Rights in the Digital Era*, which includes a right to protection against misinformation. The article providing for such a right raised serious constitutional concerns due to its potentially disproportionate effects on freedom of expression and information, and was eventually amended, leaving just a general duty of the State to protect society against disinformation. The author goes on by examining the role of self-regulation in this area and the protections from liability for online intermediaries, initially set forth in the e-Commerce Directive and now in the Digital Services Act, as well as the duties of care the DSA provides for.

13 Matteo Frigeri undertakes to assess the evolution of Design law regarding *digital files* that support 3D-printing processes, and particularly whether the online sharing of said files can be considered as *use of a design* under the Design Regulation[7] and thus a potential act of infringement. The author explores the relevant literature, case law and legislative history in this regard, suggests possible solutions, and examines how this issue is addressed by the current proposal put forward by the European Commission to update the existing legal framework.

14 Last but not least, Sergey Kasatkin researches the issue of automated execution of contracts, as exemplified by smart contracts. In such scenarios, there is in principle no need to call for the intervention of a judge in cases of non-compliance with contract terms. Rather, an automation code executes the terms of the signed contract. However, as the author notes, the code behind the automation is not always accessible to lay people, and the terms of the contract are not always provided in writing. The article underscores the importance of the White Paper that commonly accompanies a smart contract (especially in the case of Initial Coin Offerings), which plays a key role in the implementation of the contract.

6 European Parliament and Council Directive 2019/790 on copyright and related rights in the Digital Single Market, *OJ L 130,* 17 May 2019, p. 92–125.

7 Council Regulation 6/2002 on Community designs, *OJ L 3*, 5 January 2002, p. 1–24.

The Evolution of the Perception of Artificial Intelligence in the EU: The Case of Judicial Administration

by **Kalliopi Terzidou***

Abstract: Efficiency of judicial administration is one of the priorities of justice systems, it acts as a means to achieve effective administration of justice and wider access to courts through minimum spending of resources. One element associated with a satisfactory level of court efficiency is the integration and use of digital technologies by judicial staff. Artificial Intelligence (AI) stands out as a superior alternative to traditional digital technologies due to its use of Machine Learning (ML), to achieve designated goals. This article will trace the evolution EU policymakers' understanding of AI in the context of EU Member States' courts integrating AI systems to efficiently automate their judicial administration. By comparing AI definitions provided by EU bodies, specifically referencing the proposed AI Act, this article highlights the commonly accepted characteristics of AI. Additionally, it examines arguments put forth by leading computer scientists regarding the interpretation of "intelligence" in artificial artifacts. We will find that AI systems are perceived as systems employing ML and logic and knowledge-based approaches that are capable of mimicking basic human cognitive functions to autonomously automate manual tasks. These findings will be followed by remarks on the necessary steps for the integration of AI-based applications in EU justice systems.

Keywords: Artificial Intelligence; Judicial Administration; Justice; Efficiency; EU (European Union)

Recommended citation: Kalliopi Terzidou, The Evolution of the Perception of Artificial Intelligence in the EU: The Case of Judicial Administration, 14 (2023) JIPITEC 365 para 1.

A. Introduction

1 The advent of the COVID-19 pandemic functioned as a magnifying glass into the internal operation of courts and their inefficiencies in handling incoming applications and ongoing proceedings. Questions of prioritization of cases, selection of judges, and realization of (online) hearings had to be considered by national authorities competent for the organization of courts. Important factors for consideration included the protection of the rights of individuals, the resources available to courts for technical equipment, and the training of judicial staff to learn how to use digital systems.[1] Due to the suspension of physical presence in courthouses, the

* LL.M.; Doctoral Researcher; Faculty of Law, Economics, and Finance; University of Luxembourg. The present paper has been written in the context of the author's doctoral research, funded under the PRIDE funding program (DILLAN) of the Fonds National de la Recherche Luxembourg.

1 Council of Europe, 'The Functioning of Courts in the Aftermath of the Covid-19 Pandemic' (2020) <https://rm.coe.int/the-functioning-of-courts-in-the-aftermath-of-the-covid-19-pandemic/16809e55ed> accessed 15 August 2022.

use of digital technologies was important in ensuring that the judicial branch would remain accessible to citizens applying for court proceedings.

2 This response to the health crisis highlighted not only the contribution of digital technologies in the effective administration of justice but also the lack of their systematic integration and use by judicial staff. Firstly, digital systems were not tailored to the remote conduct of judicial administration and hearings. Courts preferred online videoconferencing platforms, such as Zoom or Skype, over their own systems to conduct virtual hearings due to the former's user friendliness, despite the risks of data protection breaches.[2] Secondly, judicial staff do not always possess the necessary digital skills to operate the systems due to their lack of training, therefore resorting to paper-based processes that might have been inadequate in dealing with remote proceedings during the health crisis. Thirdly, digital systems currently in use by courts are not interoperable to enable the exchange of information among national or even international judicial authorities. However, there are efforts to enhance interoperability among European states' justice systems: the e-CODEX project (e-Justice Communication via Online Data Exchange), was launched to facilitate the secure cross-border exchange of judicial information. This is achieved through the communication of encrypted data between connected gateways installed in the legal authorities of Member States, including a validation tool for electronic signatures.[3] Currently, though, these projects may not be as widely employed as necessary to achieve a satisfying level of interoperability throughout the EU.

3 Artificial Intelligence (AI) is a digital technology that is considered superior to traditional alternatives in automating manual tasks. Artificial agents have been characterized as autonomous in optimizing their performance, interactive with their environment by receiving input data and producing output values, and adaptive by altering their parameters to adjust to their current environment.[4] These characteristics can compensate for disadvantages of traditional digital systems by offering customized digital solutions for judicial staff and interoperability with external systems, further enhancing the efficiency of courts.

4 "Efficiency" is an economic concept that can be applied to courts to indicate the successful accomplishment of their objectives, particularly the administration of justice within a specific society, while utilizing minimal financial resources, time, and effort. Automation of tasks through technological means theoretically allows for minimum processing time of cases and administrative tasks, leading to less efforts by judicial staffin the execution of manual tasks. But this might not necessarily be the case, especially when considering the significant funds required for the procurement, purchase, installment, monitoring, and maintenance of the system, along with the training sessions necessary for the staff to familiarize themselves with its operation. Pending empirical studies, this article considers automation of judicial administration through the integration of AI systems as something that improves courts' efficiency.

5 The article will trace changes in the perception of AI technology by EU bodies overtime, in particular regarding attempts to increase the efficiency of judicial administration through the introduction of AI applications. This is achieved by collecting and comparing selected definitions of AI produced by EU bodies to determine the common understanding of the technology's characteristics, as well as some of its applications in the judicial administrative field. In this context, the proposed AI Act will be reviewed with a focus on the regulatory provisions on high-risk AI systems for the safety and fundamental rights of EU citizens. To further delineate the characteristics that render AI technology a factor towards a more efficient judicial administration, the meaning of "intelligence" is explored through a review of arguments made by leading authorities in the computer science field. The article concludes with thoughts on the successful integration of AI systems in EU Member-States' courts.

B. Defining AI in the Justice Field

6 There is no single definition of AI. Many actors, including international bodies, private corporations, and civil society organizations, have attempted to provide a definition to inform their policies, develop their products, or pursue their mandate respectively. However, no matter the type of actor, a working definition is important to ensure a common perception of AI systems by all members of the given organization. Especially on an international level, policies to regulate the development and use of AI must define early on what this technology entails, so Member States entering in relevant agreements

2 Anne Sanders, 'Video-Hearings in Europe Before, During and After the COVID-19 Pandemic' (2021) *International Journal for Court Administration* <https://iacajournal.org/articles/10.36745/ijca.379>, 12-14.

3 E-CODEX Website, 'Technical Solutions' <https://www.e-codex.eu/technical-solutions> accessed 16 August 2022.

4 Luciano Floridi and J.W. Sanders, 'On the Morality of Artificial Agents' (2004) Minds and Machines 14, no. 3 <https://doi.org/10.1023/B:MIND.0000035461.63578.9d> 357–362.

are aware of the scope of the regulations and align their interests accordingly. This section begins with an overview of AI definitions given by EU bodies to determine the general understanding of its features, moving to an overview of AI-based applications for the automation of judicial administration.

I. Understanding of AI by EU Policymakers

7 EU bodies are becoming gradually more interested in regulating aspects of AI use in the public and private sectors, considering not only the growing use of its applications but also its reported risks. AI systems have been accused, most notably, of the "black box" effect due to the opaqueness of their internal processes and/or the inability to explain these processes in an intelligible manner. Another observable risk is the production of biased outputs that lead to discrimination of certain protected groups in society, either due to the use of bias-charged data for the training of the system or the correlation of data that can indirectly reveal information on protected grounds, such as race or religion. During policymaking processes, EU bodies define the subject-matter of the legal act, resulting in diverse definitions of AI.

8 The *High-Level Expert Group on Artificial Intelligence* ("The Group") of the European Commission published a definition of AI in 2018, with the aim of establishing a common understanding of the term that can serve as a starting point for future AI policies on an EU level. The Group states that:

> "Artificial intelligence (AI) refers to systems designed by humans that, given a complex goal, act in the physical or digital world by perceiving their environment, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data and deciding the best action(s) to take (according to pre-defined parameters) to achieve the given goal. AI systems can also be designed to learn to adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)."[5]

9 This definition is an oversimplification of the technical nature of AI, but it still offers an insight into the characteristics of the technology. The Group places an emphasis on the process that AI systems follow to achieve the goal set by the developer. The algorithmic system is designed to perform a specific task, constituting its goal, and the developer must then train the algorithmic system with input data so it can provide an output. This process can be achieved through different techniques of AI. The definition refers to a non-exhaustive list, including "machine learning," "machine reasoning," and "robotics" techniques. An important technique that is not mentioned, but might be implied, is Natural Language Processing (NLP), which concerns the analysis of text or speech (Automatic Speech Recognition – ASR) training data, so tasks such as the filing of court documents or the transcription of a trial can be performed. NLP techniques fall under the wider spectrum of AI technology, while they can employ ML techniques for advanced statistical analysis, for example, to perform pattern recognition for the searchability of court documents.[6] They can also use Deep Learning (DL) approaches which are even less dependent on human intervention and can allow for the processing of larger sets of unstructured data to determine the distinctive features among different categories of data.[7] Another issue is that robotics is a branch of engineering that does not necessarily involve the use of AI for the execution of commands. Hence, it may not be considered as a distinct category of techniques that specifically involves AI.

10 In 2021, the European Commission published the Proposal for an AI Act to regulate its distribution on the market, application, and the use of AI systems in the EU, including rules on transparency, monitoring, and surveillance (Article 1).[8] Article 3 (1) of the Proposal defines AI systems as:

5 High-Level Expert Group on Artificial Intelligence, 'A Definition of AI' (2018) <https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf9> 7.

6 Gokul Prasath, 'Difference between Machine Learning, Artificial Intelligence and NLP' (2019) *Medium* (blog) <https://medium.com/@cs.gokulprasath98/difference-between-machine-learning-artificial-intelligence-and-nlp-d82ba64a7f32>.

7 IBM, 'What Is Machine Learning?' (2021) <https://www.ibm.com/cloud/learn/machine-learning> accessed 27 April 2022.

8 European Commission, 'Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts' (2021) <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52021PC0206> Recital 40.

> "...software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."

11 Annex I of the Proposal further specifies the techniques used for the development of AI software, being (i) ML approaches, including DL; (ii) logic and knowledge-based approaches, including knowledge representation and reasoning and expert systems, and (iii) statistical approaches, Bayesian estimation, and search and optimization methods. This definition is differentiated from The Group's attempt in that it does not provide a high-level explanation on how AI systems function to achieve a certain goal, making it difficult for a person without a basic computer engineering background to familiarize themselves with the subject matter of the Proposal. In addition, the Proposal's definition provides more concrete examples of AI techniques, excluding "robotics" and distinguishing between logic and knowledge-based approaches on the one hand, and search and optimization methods on the other. In the Group's definition, these two approaches coexisted under the category "machine reasoning." Their separation might be attributed to the fact that search and optimization methods might rely more on machine learning than machine reasoning, according to The Group's distinction. Logic and knowledge-based approaches seek to represent information (i.e. processed data) in a machine-readable manner, so the system can complete complex tasks, possibly using reasoning techniques that resemble human logic. However, machine reasoning approaches, such as ontologies, can be employed in search-related tasks, most notably to offer a repository of legal terms that are represented not only under their syntactic but also their semantic meaning, acting as available key words in search queries.[9]

12 Pending the joint adoption of the Proposal by the EU Parliament and the Council of the EU, the latter body has released several political agreements ("General Approaches"), establishing certain amendments to the text of the Proposal. In December 2022, the Council recommended an alternative definition for AI systems.[10] Article 3 (1) defines an AI system as:

> "...a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts."

13 The most notable difference from the Proposal's definition is the exclusion of the category of statistical approaches, placing the Council's definition in line with definitions provided by other international organizations.[11] These AI techniques might be considered as more traditional in comparison with ML and logic or knowledge-based approaches, thus not yielding the same challenges that require the regulatory interventions established in the Proposal, including risks to the safety and fundamental rights of EU citizens. Another reason might be the intention to establish a sufficiently wide regulatory sandbox for the promotion of innovation and for the creation of an attractive environment for business and investment within the EU. This is important since the Union should become competitive in relation to the U.S. and Chinese jurisdictions regarding the development and dissemination of AI systems in the market.

14 An interesting feature of the Council's definition is the mention of "generative AI systems," in relation to content production. Generative AI systems are generally regarded as general-purpose AI systems. According to Article 3 (1b) of the General Approach, a General Purpose AI System (GPAIS) "...is intended by the provider to perform generally applicable functions such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others; a general purpose AI system may be used in a plurality of contexts and be integrated in a plurality of other AI systems." The main difference between AI systems and GPAIS seems to be that while GPAIS are intended to be part of multiple AI systems and apply to multiple domains, traditional AI systems are stand-alone and designed for a specific goal ("...for a given set of human-defined objectives..."). However,

9 Joost Breuker, Andre Valente, and Radboud Winkels, "Legal Ontologies in Knowledge Engineering and Information Management," *Artificial Intelligence and Law* 12 (December 1, 2004): 241–77, <https://doi.org/10.1007/s10506-006-0002-1>, at 269-273.

10 General Secretariat of the Council, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - General approach, 6 December 2022, <https://data.consilium.europa.eu/doc/document/ST-15698-2022-INIT/en/pdf>.

11 *See*, for example, UNESCO, 'Recommendation on the Ethics of Artificial Intelligence' (2021) <https://unesdoc.unesco.org/ark:/48223/pf0000381137> 10, and; OECD, 'Scoping the OECD AI Principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)' (2019) <https://doi.org/10.1787/d62f618a-en> 7.

the indication in the definition that GPAIS "may" be used in multiple contexts and as a part of multiple AI systems implies that they might also be designed for a specific context and to fit a specific AI system, putting into question the generality of their nature.[12]

15 This distinction is important since Article 4b of the General Approach states that GPAIS may be used as "high-risk" AI systems or as their components. High-risk AI systems are regulated under Title III of the Proposal and denote systems that pose a high risk to the health and safety or fundamental rights of natural persons, depending on the performed function, purpose, and intended modalities of the system. These systems must be developed according to a set of requirements prescribed in Articles 8-15 of the Proposal. These requirements concern accountability, transparency, and technical safety goals, ranging from record-keeping (Article 12) to the provision of information to users (Article 13) and human oversight (Article 14). Apart from high-risk AI systems, the Proposal establishes different levels of risk, namely unacceptable (prohibited practices that contravene Union values and are likely to manipulate users' subconscious or take advantage of vulnerable groups), limited (slight risk of manipulation of users in not realizing that they do not interact with a machine, necessitating transparency obligations), and minimal (not considerable).[13]

16 The common elements of the EU bodies' definitions of AI are that the systems pursue specific goals through certain techniques, namely through ML and logic or knowledge-based approaches. It is evident that EU representatives started with a wider approach and gradually narrowed down the definition of AI systems, to the point of excluding statistical and related approaches. Despite the restriction of the scope of AI systems in ML and logic or knowledge-based techniques, the Council's definition might still be considered as technologically neutral to the extent that these techniques encompass a broad field of AI sub-techniques, functionalities, and applications, thus rendering the Proposal applicable to a variety of AI systems developed in the EU and/or addressed to EU citizens and guaranteeing the safety and rights of users throughout the entire lifecycle of the AI system.

12 Philipp Hacker, Andreas Engel and Theresa List, 'Understanding and Regulating ChatGPT, and Other Large Generative AI Models: With input from ChatGPT' (*Verfassungsblog*, 20 January 2023) <https://verfassungsblog.de/chatgpt/> accessed 7 March 2023.

13 European Commission, 'Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts,' Preamble 5.2.2. – 5.2.4.

II. The Use of AI Systems in Judicial Administration

17 "Judicial administration" or "administration of courts" represents the sum of tasks necessary for the internal organization of courts. These tasks can be purely managerial in nature, encompassing back-office duties for the operation of the courthouse and the management of personnel. At the same time, they can be ancillary in the adjudicatory work of judges, in other words assisting them with the systematization of case management and decision-making. Judicial administration is carried out by judicial staff, including judges, prosecutors, judicial assistants, and administrative personnel or clerks. AI systems designed to automate judicial administrative tasks have been classified in various ways throughout recent academic literature.

18 *Sourdin* makes a distinction among supportive, replacement, and disruptive technologies, under which AI technology may be used to support online information services on justice processes, replace physical court proceedings with online proceedings using videoconferencing tools, and informing judges' decisions applying prediction models, respectively.[14] *Reiling* distinguishes between three main categories of AI uses, being the organization of information through the recognition of patterns in documents and files to discover information, the provision of advice to individuals on possible solutions to their problem, and the "prediction" of the outcome of court proceedings.[15] *Terzidou* reviews AI uses according to the stage of proceedings they are contributing to, namely in pre-trial, hearing, and post-sentencing proceedings.[16] Examples include the provision of information on court proceedings using chatbots, the transcription of the courtroom procedure, and the anonymization of court decisions, respectively. A major part of the reviewed technologies has a managerial character in automating tasks that concern back-office duties, with the exceptions of document discovery

14 Tania Sourdin, 'Judge v Robot? Artificial Intelligence and Judicial Decision-Making' (2018) UNSW Law Journal 41, no. 4 <https://www.unswlawjournal.unsw.edu.au/article/judge-v-robot-artificial-intelligence-and-judicial-decision-making/> 1117-1119.

15 A. D. (Dory) Reiling, 'Courts and Artificial Intelligence' (2020) 11(2) International Journal for Court Administration 8 <https://papers.ssrn.com/abstract=3736411> 3-6. accessed 7 March 2023

16 Kalliopi Terzidou, 'The Use of Artificial Intelligence in the Judiciary and Its Compliance with the Right to a Fair Trial' (2022) 31 Journal of Judicial Administration <https://orbilu.uni.lu/handle/10993/51591> 157-158.

and predictive models representing the advisory potential of AI applications to judges' decision-making process.

19 To better illustrate the use of AI applications with an advisory role, predictive analytics are engineered into the systems to predict defendants' future behavior or the court's most probable decision outcome based on previous patterns. In the former scenario, algorithmic systems are reportedly used to measure the risk of convicted people reoffending, in order to decide whether they are eligible for parole. The COMPAS system determines the risk of defendants reoffending in the future based on a risk score that is determined through their responses to a 137-questions survey, complemented by information from their criminal record.[17] In the latter case, AI systems predict the whole or part of the hearing proceedings' outcome. *Aletras et al.* used ML and NLP techniques to predict the European Court of Human Rights decisions in cases concerning Articles 3, 6, and 8 of the European Convention of Human Rights, mainly relying on the facts of the case to reveal patterns in the case law document.[18] Additionally, the DataJust project, led by the French Ministry of Justice, aims at offering to the public indicative benchmarks for compensation in cases of physical harm, by processing court decisions to extract and exploit data concerning "the amounts requested and offered by the parties to the proceedings, the assessments proposed within the framework of procedures for the amicable settlement of disputes and the amounts allocated to victims by the courts."[19]

20 It is important to note that the above systems merely inform judges' decision-making by providing further grounds in their reasoning or assist individuals in deciding whether to resort to courts for the resolution of their case. In Europe, there is no application that replaces the role of judges in awarding binding and enforceable judgments. In 2019, a magazine article was released concerning the design of a robot judge for the adjudication of small claims disputes based on the analysis of information uploaded by the parties, a project allegedly coordinated by the Estonian Ministry of Justice.[20] This report, however, was subsequently characterized as "misleading" by the Ministry, stating that it does not pursue such a project.[21] The replacement of judges by AI systems automating the decision-making process would likely undermine the legitimacy of the trial and the acceptance of the final judgment, given that the systems cannot currently replicate the reasoning of judges, characterised by well-structured arguments on how legislative provisions and/or case law apply to the facts of the case.[22] The machine's logic in adhering to its pre-programmed rules cannot be compared with such reasoning, because it can be expressed only in technical terms that are not humanly intelligible and need to be treated by developers in order to circumvent the "black box" effect and derive some kind of explainability. Nevertheless, there are techniques that attempt to enhance algorithmic transparency and mimic human reasoning. These approaches are explored in the next section.

III. The Interest of the EU in AI-Assisted Judicial Administration

21 In the EU, Member States' courts express a preference in the development of AI-based applications with a managerial role, automating administrative tasks for the efficiency of the courthouse. National competent authorities are prioritizing the development of AI systems automating, in full or partially, the anonymization or pseudonymization of judgments, the searchability of court documents for legal research, the analysis of evidence, the filing of court documents, the transcription of the trial, the translation of court documents, and internal and external communications.[23]

17 Julia Angwin Mattu Jeff Larson,Lauren Kirchner,Surya, 'Machine Bias' *ProPublica* <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=l24Nh-wDyBgy53bhcy5jGvQh1IDRcxzE> accessed 24 January 2022.

18 Nikolaos Aletras et al., 'Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective' (2016) *PeerJ Computer Science* 2 <https://doi.org/10.7717/peerj-cs.93> 6-15.e

19 Justice.Fr, 'DataJust' <https://www.justice.fr/donnees-personnelles/datajust> accessed 25 January 2022.

20 Eric Niiler, 'Can AI Be a Fair Judge in Court? Estonia Thinks So' *Wired* <https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/> accessed 18 August 2022.

21 Ministry of Justice of Estonia, 'Estonia Does Not Develop AI Judge | Justiitsministeerium' <https://www.just.ee/en/news/estonia-does-not-develop-ai-judge> accessed 20 June 2022.

22 Jasper Ulenaers, 'The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge? (2020) *Asian Journal of Law and Economics* 11, no. 2 <https://doi.org/10.1515/ajle-2020-0008> 27-28.

23 Directorate-General for Justice and Consumers (European Commission) and Trasys International, 'Study on the Use of Innovative Technologies in the Justice Field: Final Report' (2020) LU: Publications Office of the European Union, <https://data.europa.eu/doi/10.2838/585101> 111-142.

22 The interest of Member States in integrating AI-based systems in their judiciaries is further reflected in Preamble 40 of the Proposal for an AI Act, stating that AI systems "...intended to assist judicial authorities in researching and interpreting facts and the law and in applying the law to a concrete set of facts..." should be qualified as high-risk, not including AI systems "...intended for purely ancillary administrative activities that do not affect the actual administration of justice in individual cases..." The Preamble provides examples of AI systems for "purely ancillary administrative activities, namely the anonymization of court documents, the communication between personnel, and the allocation of resources." This differentiation of administrative tasks validates the distinction marked above between AI applications for the automation of tasks related to back-office duties and tasks concerning the decision-making process, while highlighting the importance that the European Commission places on the high level of risk that AI systems have for the research, interpretation, and application of what the law might entail.

23 An illustration of a high-risk AI system used by judges for the purpose of retrieving legislative and case law resources in preparation for the hearing would be Open AI's chatbot, also known as ChatGPT. ChatGPT is, in fact, a language model trained with Reinforcement Learning techniques, upon which Open AI developed its chatbot, which reacts to users' prompts in a conversational manner and generates suitable responses.[24] There have already been reports on uses of the chatbot by judges, admittedly outside the EU, posing questions regarding the applicable rules to a given legal issue to facilitate their decision-making process, albeit also taking into consideration past case law to arrive to their final decision.[25] Even if the output of the chatbot is not the sole or main basis of the judge's final decision, these generative AI systems can be characterized as high-risk due to the challenges they pose to case management prior to and during the trial. It is possible that chatbots are not trained with sufficient or domain specific input data, or are trained with data collected through sources of misinformation, thus providing judges with insufficient and/or inaccurate legal information that might lead them to misapplications of the legislation and jurisprudence in a given case. Therefore, a careful design and development of generative AI systems must be conducted by developers and providers alike, in accordance with the Proposal's requirements on high-risk AI systems.

24 The review of the general understanding of AI through the EU bodies' definitions of the AI applications in the justice field revealed that AI systems are primarily considered to be based on ML and logic or knowledge-based approaches, applied in judicial administration to automate back-office tasks and assist judges with their decision-making process. The following section expands upon the concept of "intelligence" in relation to artificial artefacts as a further step in determining the components of AI systems that are most conducive to raising the efficiency of judicial administration in EU Member States' courts.

C. The Intelligence of AI Systems in Judicial Administration

25 "Intelligence" is an abstract concept that is normally associated with human beings. Yet, it is the second component of the term "Artificial Intelligence," hinting the ability of machines to mimic the cognitive functions of human beings. This section attempts to understand what "intelligence" means in relation to artificial artefacts through the review of arguments by leading computer scientists and of the operation of selected AI applications.

I. Perspectives on the Intelligence of Artificial Artefacts

26 The *Cambridge Dictionary* defines "intelligence" as "the ability to learn, understand, and make judgments or have opinions that are based on reason,"[26] competences generally associated with human beings. In the computer science field, *John McCarthy* claimed that intelligence is "the computational part of the ability to achieve goals in the world," specifying that AI does not have to restrict itself to biologically observable methods but can also involve computational methods that are not found in human beings.[27] He then explains that these computational methods cannot generally be characterized as intelligent because humans themselves cannot yet understand all the mechanisms of intelligence.

24 OpenAI, 'Introducing ChatGPT' <https://openai.com/blog/chatgpt> accessed 7 March 2023."plainCitation":"OpenAI, 'Introducing ChatGPT' <https://openai.com/blog/chatgpt> accessed 7 March 2023.

25 Luke Taylor, 'Colombian Judge Says He Used ChatGPT in Ruling' *The Guardian* (3 February 2023) <https://www.theguardian.com/technology/2023/feb/03/colombia-judge-chatgpt-ruling> accessed 8 March 2023.

26 Cambridge Dictionary, 'Definition of "Intelligence' <https://dictionary.cambridge.org/dictionary/english/intelligence> accessed 19 August 2022.

27 John McCarthy, 'What Is Artificial Intelligence?' < http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html> accessed 19 August 2022 2.

27 Earlier work has attempted to establish the machines' potential to display intelligence by mimicking human reasoning. *Turing* established a test, called the "Imitation Game," to conclude if machines, that is digital computers, can think or operate as a human would. The test required three participants: a human interrogator, a human respondent, and a machine respondent; if the interrogator cannot tell the difference between her interaction with the human and the machine, then the machine passes the test.[28] In the same paper, Turing mentioned two contrary opinions to his theory: *Lady Lovelace's* argument that a machine does not originate an act but can only perform based on pre-programmed orders, and *Professor Jefferson's* view that a machine is not driven by thoughts and emotions to perform a task nor can it be emotionally affected by its accomplishments or failures.

28 Under the above statements, AI applications for judicial administration could be viewed as "thinking" agents in terms of carrying out previously manual tasks in a way that humans would, but only because they are originally programmed to do so by human developers. Accordingly, AI systems cannot be considered fully autonomous since there is always a human in the loop operating the system, even if they alleviate much of the effort spent in the performance of a judicial task. For instance, speech-to-text systems are used to transcribe the trial by transforming recorded speech files uploaded to the server into text.[29] The clerk, however, has to upload these files to the system and remains in control of the application by verifying the accuracy of the transcribed text with her signature, while technical issues can be communicated to the IT expert that can make any necessary adjustments to the system.

29 The autonomy of an AI system is better perceived in its ability to interact and adapt to its environment through the improvement of its performance overtime, being constantly trained with new data inputs to build on its past performances. AI systems for information retrieval, that assist judges in finding legislation and jurisprudence by searching structured documents and files, can always improve their accuracy by being trained with larger datasets. The challenge of the optimization of AI systems trained with legal data is that legal documents are long, they display a complex structure and legal terminology, and datasets with domain-specific documents are rare.[30]

30 AI systems could also demonstrate their "thinking" ability by mimicking more complex cognitive tasks. Research projects are focusing on the reproduction of legal reasoning by artificial agents, a process that otherwise requires a considerable time and effort by legal professionals to perform. It has to be noted, however, that AI systems perform legal reasoning in a computational or mathematical manner; the concepts argued are closed-ended rather than open-ended, the context of argumentation is similarly well-defined rather than consisting of incomplete information, and the conclusions are objective and definite rather than subjective and open to further discussion and amendments.[31] As a result, the mechanical analysis of legal texts is distinct from the reasoning of legal professionals on abstract legal concepts and might render relevant AI systems unsuited for case management in the criminal branch, where judges must often deal with legal terms and concepts that are open to interpretation and difficult to computerize.

II. Intelligent AI Applications for the Automation of Judicial Administration

31 The "intelligence" of AI systems in (semi-) autonomously completing previously manual tasks through the imitation of basic cognitive features can be demonstrated in several judicial applications. Taking the example of AI systems for the anonymization or pseudonymization of judgments in compliance with personal data protection rules, NLP techniques might be employed for the annotation of entities and their replacement with labels in a consistent manner, so the same entity is assigned the same label throughout the text.[32] There is some mimicking of human intelligence in the processing of textual data to find personal information and replace it with the designated labels. However, human input is still needed to verify and, if needed, correct the output of the algorithm, especially in cases where

28 A. M. Turing, 'I.—Computing Machinery and Intelligence' (1950) Mind LIX, no. 236 <https://doi.org/10.1093/mind/LIX.236.433> 433-451.

29 Tanel Alumäe, 'Transcription System for Semi-Spontaneous Estonian Speech' (2012) *Human Language Technologies – The Baltic Perspective* <https://doi.org/10.3233/978-1-61499-133-5-10> 10-11.

30 Diego Collarana et al., 'A Question Answering System on Regulatory Documents' (2018) Legal Knowledge and Information Systems <https://doi.org/10.3233/978-1-61499-935-5-41> 42.

31 T. J. M. Bench-Capon and Paul E. Dunne, 'Argumentation in Artificial Intelligence' (2007) *Artificial Intelligence*, 171, no. 10 <https://doi.org/10.1016/j.artint.2007.05.001> 619-621.

32 Diego Garat and Dina Wonsever, 'Automatic Curation of Court Documents: Anonymizing Personal Data'(2022) *Information* 13, no. 1 <https://doi.org/10.3390/info13010027> 5-6.

there is a lack of consistency in the anonymization of the same entity throughout the text.[33]

32 Regarding examples on computational legal reasoning, compliance checking applications automate the assessment of a real-world incident in terms of its compliance with a norm, which in this context means the way a provision is applied. This can be achieved through ontologies, such as the OWL language for knowledge modeling in the Semantic Web, where real world incidents are represented as ontologies and norms are represented as restrictions to ontological properties, reflecting the legal restraints that individuals must comply with.[34] Therefore, legal reasoning is automated through ontologies, which further enables the explainability of AI systems, that is "... the ability to explain both the technical processes of an AI system and the related human decisions ..." in a humanly understandable way,[35] without resorting to ML methods that can only be viewed in numerical terms. Explainable processes can lead to accountability for the algorithmic outcomes and redesigning in cases of malfunctions or necessary updates.

33 In continuation of the discussion on the COMPAS system, an ontology could be created to represent the concept of "recidivism," which is then accompanied by different properties representing the indicators mentioned by the provider Northpointe, such as criminal history, criminal associates, and drug involvement.[36] The conceptualization of "recidivism" into an ontology and the tagging of its distinguished properties would allow users, in this case judges, to infer logical similarities among these properties in an explainable manner. In this way, they could understand how each indicator contributed to the predicted risk score, so as to detect instances of adverse bias when indicators based on protected grounds, such as race or religion, have contributed to the algorithmic output more than permitted by the threshold established by competent authorities.

34 The "thinking" process of AI systems is still of a mathematical nature and realized within the strict limits of the goals set by developers, confirming Lady Lovelace's argument on the inability of machines to originate an action. Machines are also not conscious in recognizing the reasons behind their actions and taking pride in their accomplishments according to Professor Jefferson, instead acting upon the programmed rules. Nevertheless, machines can still perform an action that could be realized by a human, mimicking minimum cognitive capabilities. Placing such a system under Turing's test, the human interrogator might not be able to distinguish between the machine and the human participants completing a manual task, thus proving that AI systems are intelligent in this restricted fashion. Combined with their autonomous character, though not autonomous enough to replace their users, AI systems could theoretically yield efficiencies in judicial administration by automating a considerable number of judicial tasks and thus minimizing time and effort spent in back-office duties and, ultimately, disposition time. In addition, AI predictive systems can improve the quality of the adjudication process by providing judges with additional grounds for their decisions, consisting in the system's outputs that can be assessed for possible adverse biases or other defects through techniques, such as ontologies, that render AI systems explainable.

D. Final Remarks

35 This paper highlighted the evolution of the understanding of AI by EU policymakers and its perceived efficiencies for the judicial administration of EU Member States' courts. In the first section, it was shown that the definition of AI systems by EU bodies has been gradually narrowed to refer to ML and logic or knowledge-based techniques. The literature review revealed that AI applications in judicial administration can be categorized in AI systems automating managerial, back-office tasks and in those that assist judges in legal research or in predicting post-sentencing parameters, including the amount of compensation to be attributed to the injured party. AI systems assisting judges during the decision-making process are considered as high-risk systems by the Proposal for an AI Act and must be developed in compliance with certain requirements of a technical and governance nature. In the second section, AI systems were claimed to be "intelligent" in terms of their computational ability to arrive to the goal set by human developers, mimicking basic cognitive functions, and of their autonomy in improving their performance overtime by being

33 *See*, for example, Alan Akbik, 'The Flair NLP Framework' Institut für Informatik <https://www.informatik.hu-berlin.de/en/forschung-en/gebiete/ml-en/Flair> accessed 11 July 2022.

34 Enrico Francesconi and Guido Governatori, 'Patterns for Legal Compliance Checking in a Decidable Framework of Linked Open Data' (2022) *Artificial Intelligence and Law* <https://doi.org/10.1007/s10506-022-09317-8> 6-7.

35 High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI' (2019) Publications Office of the EU <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1> 18.

36 Northpointe, 'Practitioner's Guide to COMPAS Core,' (2015) Northpointe, <https://s3.documentcloud.org/documents/2840784/Practitioner-s-Guide-to-COMPAS-Core.pdf> 27.

trained on new data inputs and "learning" from past performances.

36 Certain steps must be taken to ensure the successful integration of AI systems in judicial administration and, consequently, the realization of the potential efficiencies for time and effort management. More specifically, AI applications must adhere to relevant legal requirements, be securely developed, and follow specific rules for their sound integration and systematic use in courts. The use of AI systems in the justice field must primarily adhere to the right to a fair trial, meaning that they must support access to courts and safeguard the independence and impartiality of the judiciary, along with the fairness of the court proceedings.[37] Further legal requirements include the protection of personal information during the training and performance of the algorithm, so their processing is done in a lawful and transparent manner, for clearly stated purposes and to the extent necessary, retaining the data in an updated form and for the necessary amount of time.[38]

37 Moreover, AI systems must be technically secure and robust throughout their design, development, use, and possible redesign. The *High-Level Expert Group on AI* states that AI systems must adhere to several standards, including human oversight (continuous human control), technical robustness and safety (accuracy, reliability, and safety from cyberattacks), transparency (documentation and communication of the technical processes in a humanly understandable manner for accountability purposes), and non-discrimination (no reproduction of discrimination based on protected grounds, such as gender).[39] The Proposal for an AI Act further develops these standards according to the level of risk that the AI system presents, ranging from data management and documentation for high-risk systems to transparency measures for limited-risk systems.

38 Finally, the process of the integration of AI applications in courts must be regulated so AI systems can produce legal effects and accountability can be attributed when checking the outputs of AI systems against the existing legal certification. On a national level, few policy or legal documents exist for the regulation of the use of AI in the judiciary; however, national courts in Europe have ongoing AI projects for the automation of their judicial administration that, once concluded, will need to be officialized by a state act or equivalent to be integrated in national justice systems. On a regional level, the Proposal for an AI Act proves that EU bodies and Member States are interested in the uniform regulation of AI systems in the public sector, including the judicial branch, even in the case of high-risk AI applications that must conform with harmonized standards to be introduced to national courts.

37 Terzidou, 158-163.

38 *See,* European Commission, 'Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)' (2016) <http://data.europa.eu/eli/reg/2016/679/oj> Article 5.

39 European Commission, 'Ethics guidelines for trustworthy AI' 15-20.

Application of artificial intelligence (AI) in the assessment of the credibility of statements in the cross-border taking of evidence in civil and commercial matters

by **Jura Golub***

Abstract: Regulation (EU) 2020/1783 on 'co-operation between the courts of the Member States in the taking of evidence in civil and commercial matters' introduces taking evidence by videoconference or other distance communications technology as the "gold standard" in the process of direct cross-border taking of evidence by examining a person who is present in another Member State. This represents a step forward compared to the previous Regulation 1206/2001, as the provision for direct evidence taking through videoconferencing was rarely applied in practice. The direct taking of evidence through videoconference contributes significantly to the realisation of the principle of orality and immediacy in civil proceedings, as opposed to indirect methods of cross-border taking of collection. On the other hand, a question arises whether the principle of immediacy is weakened by using videoconferencing, given that there is a "digital barrier" between a witness and the court. When assessing the credibility of the statements made by parties, witnesses, and experts, psychological criteria in addition to logical criteria plays an important role in shaping the court's opinion on the truth of the assertion regarding the existence of certain facts. As a solution for consideration, there is a possibility of using an artificial intelligence system to detect deception during the direct taking of evidence by examining parties, witnesses, or experts. However, the admissibility of the above solution should be considered as a multi-faceted issue, particularly regarding aspects of the right to a fair trial, personal data protection rules, and the proposed provisions of the Artificial Intelligence Act.

Keywords: cross-border taking of evidence, judicial cooperation in civil and commercial matters, artificial intelligence, civil procedure, statement credibility, deception detection

Recommended citation: Jura Golub, Application of artificial intelligence (AI) in the assessment of the credibility of statements in the cross-border taking of evidence in civil and commercial matters, 14 (2023) JIPITEC 376 para 1

A. Introductory considerations

1 This forward-looking paper addresses the potential use of artificial intelligence as an auxiliary tool for the court to assess the credibility of statements in the cross-border taking of evidence in civil and commercial matters. In general, the assessment of the credibility of statements by using various technology tools occupies the attention of the scientific public in the field of criminal procedural law. EU procedural law is generally opposed to the use of tools such as polygraphs for assessing the credibility of statements in court proceedings.[1] However, the normative activity of the EU in the field of cross-border taking of evidence in civil and commercial matters coupled with the development

* Jura Golub, LLM, PhD Student and Research Assistant, University of Osijek – Faculty of Law Osijek, S. Radića 13, HR - 31000 Osijek, Croatia; E-mail: jgolub@pravos.hr, ORCID: https://orcid.org/0000-0002-2440-8081.

1 Robert Bradshaw, "Deception and detection: the use of technology in assessing witness credibility" [2021] 37 *Arbitration International* 711.

of systems for assessing the credibility of statements based on artificial intelligence make it necessary to consider the potential of using such systems. Taking evidence in any judicial proceeding is a prerequisite for establishing the facts of the case and thus for the correct application of substantive law. To achieve this, it is crucial to ensure access to evidence, which contributes to the actualization of the right of access to justice.[2] In civil disputes with a cross-border element, access to evidence is even more challenging, especially in the context of taking evidence by way of examination parties, witnesses, or experts. Long distances and considerable travel costs mean that a balance must be struck between the principles of economy and efficiency and the principle of immediacy when choosing the method of the cross-border taking of evidence.[3] In this balancing act, the courts of the Member States applying Council Regulation (EC) No 1206/2001 of 28 May 2001 on 'cooperation between the courts of the Member States in the taking of evidence in civil or commercial matters[4]' (hereinafter: Regulation 1206/2001) have usually opted for the method of indirect taking of evidence. According to data provided by the European Commission, during the mentioned period, in an average of 87.5% of cases, the court of one Member State requested the taking of evidence by the court of another Member State (the indirect taking of evidence), while in an average of 12.5% of cases, the direct taking of evidence was applied.[5] It is obvious that the direct method of taking evidence has failed with the application of Regulation 1206/2001, and thus the principle of immediacy as one of the fundamental principles of civil procedure.

2 In light of the identified shortcomings, Regulation (EU) 2020/1783 of the European Parliament and of the Council of 25 November 2020 on 'cooperation between the courts of the Member States in the taking of evidence in civil or commercial matters[6]' (hereinafter: the Revised Evidence Regulation or RER) was adopted and entered into force on 1 July 2022. The RER, *inter alia*, introduces the use of videoconferencing or other distance communications technology as the "gold standard" in the direct taking of evidence by examining persons from other Member States, with the aim of strengthening access to justice,[7] and of facilitating and speeding up the taking of evidence.[8]

3 The introduction of videoconferencing as the primary method of direct taking of evidence by examining a person has undoubtedly strengthened the principle of immediacy. However, it is necessary to consider whether this represents significant progress in strengthening immediacy as a principle of civil procedure and whether there is room for further improvement, especially when considering the development of modern technology. In particular artificial intelligence systems developed for the purpose of deception detection. Indeed, the available research shows that humans are able to detect deception, i.e., untrue statements, in only 57% of cases.[9] Given that there is a kind of "digital barrier" between the court that takes evidence directly by way of videoconferencing and the person being heard, it can be assumed that the judge's perception is further weakened when assessing the credibility of statements, even though the examination takes place in real time with audio and visual production. To date, several applicable AI-based deception detection solutions have been developed. Typically, the systems analyse facial micro expressions and eye tracking, and perform verbal and linguistic

2 European Law Institute and UNIDROIT (eds), *ELI – Unidroit Model European Rules of Civil Procedure* (OUP 2021) 136.

3 Commission, "COMMISSION STAFF WORKING DOCUMENT IMPACT ASSESSMENT Accompanying the document Proposal for a Regulation of the European Parliament and of the Council amending Council Regulation (EC) No 1206/2001 of 28 May 2001 on cooperation between the courts of the Member States in the taking of evidence in civil or commercial matters" SWD (2018) 285 final, 29.

4 Council Regulation (EC) No 1206/2001 of 28 May 2001 on cooperation between the courts of the Member States in the taking of evidence in civil or commercial matters [2001] OJ L 174/1 (hereinafter: Regulation 1206/2001)

5 Commission, "COMMISSION STAFF WORKING DOCUMENT EVALUATION Accompanying the Document Proposal for a Regulation of the European Parliament and of the Council Amending Council Regulation (EC) No 1206/2001 of 28 May 2001 on cooperation between the courts of the Member States in the taking of evidence in civil or commercial matters" SWD [2018], 11.

6 Regulation (EU) 2020/1783 of the European Parliament and of the Council of 25 November 2020 on cooperation between the courts of the Member States in the taking of evidence in civil or commercial matters (taking of evidence) (recast) [2020] OJ L 405/1 (hereinafter: Revised Evidence Regulation or RER)

7 Commission, "Proposal for a Regulation of the European Parliament and of the Council amending Council Regulation (EC) No 1206/2001 of 28 May 2001 on cooperation between the courts of the Member States in the taking of evidence in civil or commercial matters" COM (2018) N 388 final, 6.

8 Revised Evidence Regulation, recital 21.

9 Bradshaw (n 1) 714. According to Amit Katwala, "The Race to Create a Perfect Lie Detector – and the Dangers of Succeeding" *The Guardian* (London, 5 September 2019)

analysis of respondents.[10] Based on this, AI and machine learning are used in an automated process to evaluate the credibility of a single statement, thus eliminating any subjective human influence.[11] Individual AI-based deception detection systems are explained in detail in the next sections of this paper.

4 Considering the above, the main research question is whether artificial intelligence can contribute to strengthening the principle of immediacy in the cross-border taking of evidence through videoconferencing. In this context, the paper aims to determine the admissibility of the application of AI in assessing statement credibility in the cross-border taking of evidence, and this must be viewed as a multi-faceted issue. First, it is necessary to legally qualify the position of the system for assessing statement credibility in court proceedings. Can a deception detection system be considered a *sui generis* witness or expert, or something else? Furthermore, the admissibility of the application of AI in the assessment of statement credibility via videoconferencing must be examined from the perspective of the right to a fair trial guaranteed by the European Convention for the Protection of Human Rights and Fundamental Freedoms[12] (ECHR), the Charter of Fundamental Rights of the European Union[13] (CFR), the General Data Protection Regulation[14] (GDPR), and the conformity of the application with the draft Artificial Intelligence Act[15] (AI Act).

B. EU normative framework for the cross-border taking of evidence in civil and commercial matters using videoconferencing.

5 The legal basis for the regulation of the cross-border taking of evidence in the EU is Article 81(2)(d) of the Treaty on the Functioning of the European Union,[16] and thus the RER is an integral part of the normative framework of the European Union in the area of judicial cooperation in civil and commercial matters.[17] The aim of the European approach to the regulation of the cross-border taking of evidence is to create an appropriate legal and procedural framework that complements the effective resolution of cases with cross-border implications, i.e., the successful application of European private international law.[18] In addition to this purpose, a uniform legal and procedural framework for cross-border taking of evidence is important for the functioning of the internal market of the European Union.[19]

6 Prior to the implementation of the RER, Regulation 1206/2001 was applied in cross-border taking of evidence in civil and commercial matters.[20] In the context of this issue, it should be noted that Regulation 1206/2001, "coyly" and as an incentive, provided the possibility for direct taking of evidence by videoconferencing by the requesting court.[21] However, this possibility was rarely used in

10 Bradshaw (n 1) 709.

11 ibid

12 Consolidated Version of the Europan Convention on Human Rights [2021] <https://www.echr.coe.int/documents/convention_eng.pdf> accessed 24 July 2022 (hereinafter: ECHR)

13 Charter of Fundamental Rights of the European Union [2016] OJ C202/389 (hereinafter: CFR)

14 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) [2016] OJ L119/1 (hereinafter: GDPR)

15 Commission, „Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts" COM (2021) 206 final (hereinafter: AI Act)

16 Consolidated version of the Treaty on the Functioning of the European Union [2007] OJ C 202/1

17 Mirela Župan, "50 godina europske pravosudne suradnje u građanskim stvarima – 5 godina hrvatske primjene" [2019] 10(1) Godišnjak Akademije pravnih znanosti Hrvatske 475-476 <https://doi.org/10.32984/gapzh.10.1.20> accessed 25 July 2022

18 ibid 475-76.

19 Paula Poretti, "Council Regulation (EC) No 1206/2001 of 28 May 2001 on cooperation between the courts of the Member States in the taking of evidence in civil or commercial matters as a guarantee of the right to effective judicial protection" in Željka Primorac et al. (eds), *Economic and Social Development - 16th International Scientific Conference on Economic and Social Development – "The Legal Challenges of Modern World"* (Varazdin Development and Entrepreneurship Agency, Faculty of Law – University of Split and University North 2016) 219 <https://bib.irb.hr/datoteka/833529.esd_Book_of_Proceedings_Split_2016_Online.pdf> accessed 28 July 2022

20 Revised Evidence Regulation, art 34(1)

21 Jiri Valdhans and David Sehnalek, "The 1970 Hague Evidence

practice. In the previous section, it was statistically established that the direct taking of evidence was used in only 12.5% of cases that required cross-border taking of evidence.[22] However, the European Commission estimates that videoconferencing was specifically used in only 10-25% of cases where the direct method of taking evidence was applied.[23]

7 In light of this, the RER replaces Regulation 1206/2001 and takes a digital step forward, by introducing a number of solutions related to electronic communication between the Member State authorities, evidence transfer and the legal effect of electronic documents, and the use of videoconferencing in the context of the direct taking of evidence.[24] The RER *ratione materiae* applies in civil and commercial matters when the court of one Member State requests the competent court of another Member State to take evidence or when it requests the direct taking of evidence in another Member State.[25] The Revised Evidence Regulation applies *ratione teritorii* in all EU Member States, with the exception of Denmark.[26]

8 Since this paper focuses thematically on the possible application of AI in the direct cross-border taking of evidence by videoconference, it is necessary to consider the relevant provisions of the RER on this method of taking evidence. In relation to Regulation 1206/2001, the RER more imperatively mandates the use of videoconferencing or other distance communications technology when a court in one Member State requires the direct taking of evidence by examining a person located in another Member State.[27] It further requires that such taking of evidence shall be conducted on the condition that this technology is available to the court and if the court considers it to be appropriate in light of the particular circumstances of the case.[28] A court of a Member State that wishes to hear a person located in another Member State using videoconferencing submits a request to the central body or the competent authority of another Member State using an appropriate form.[29] The RER does not provide details of the procedure of examination held through videoconferencing, but refers the courts or the authorities of the Member States to mutual agreements regarding practical arrangements for the examination[30] Therefore, in any other situation, the general provisions on direct taking of evidence in Article 19 of the RER should apply to the procedure of direct taking of evidence by examining persons through videoconference. The direct taking of evidence is always carried out on a voluntary basis without the use of coercive measures, and the person being heard must be informed of this.[31] A decision on the request for the direct taking of evidence is made by the central body or the competent authority of the requested Member State, and the RER prescribes the time limits for the decision on the request.[32] In the event that the request for the direct taking of evidence is not decided within the prescribed time limit, the RER also provides for a positive presumption that the request shall be deemed to have been accepted. The central body orcompetent authority of the requested Member State has the power to refuse a request only in certain cases, i.e., if a request for the direct taking of evidence does not fall within the scope of the RER, if a request does not contain all the information required by the RER (Article 5), or if the direct taking of evidence is requested in a manner that is contrary to the fundamental principles of law of the requested Member State.[33]

9 It should be noted that the RER proposal explicitly mentions that the examination of a person conducted by videoconference must take place on court premises.[34] However, in the adopted version of the RER, such a provision was not explicitly included, which opens the possibility of further broad interpretations regarding where a person is heard. Nevertheless, it can be interpreted from

Convention, the European Union and the 2001 EU Evidence Regulation – Interfaces" in C.H. van Rhee and Alan Uzelac (eds), *Evidence in Cross Border Civil Litigation* (Intersentia 2015) 359.

22 See n 5.

23 Commission (n 5) 45.

24 Elena Alina Ontanu, "Normalising the use of electronic evidence: Bringing technology use into a familiar normative path in civil procedure" (2022) 12(3) Oñati Socio-Legal Series 594 <https://opo.iisj.net/index.php/osls/article/view/1370> accessed 28 July 2022. See also Revised Evidence Regulation, arts 7, 8, 20. Revised Evidence Regulation, arts 7, 8, and 20.

25 Revised Evidence Regulation, art 1.

26 ibid recital 38.

27 ibid art 20(1).

28 ibid

29 ibid arts 19(1) and 20(2).

30 ibid art 20(2).

31 ibid art 19(2).

32 ibid art 19(4)(5).

33 ibid art 19(7).

34 Commission, (n 7) 12.

the provision in Article 20 of the RER on 'mutual agreements between courts and competent authorities regarding practical solutions for the examination that the examination of a person should take place in the premises of the court.[35]

10 To further examine the admissibility of the application of AI, it should be noted that although the RER is part of European international procedural law, the procedural elements outlined above are mainly standardised in legal and technical terms to facilitate judicial cooperation and the taking of evidence. Similarly to Regulation 1206/2001, the RER does not regulate fundamental procedural issues, such as the admissibility and the probative value of evidence and other rules on the taking of evidence, but leaves these to national procedural autonomy.[36] Indeed, the RER provides that the requesting court shall conduct the direct taking of evidence in accordance with its national law.[37] Thus, the RER follows the generally accepted rule that the rules of evidence are to be assessed according to *lex fori*, i.e., according to the procedural law of the court taking a particular procedural action.[38] The direct cross-border taking of evidence under *lex fori* contributes to the uniform treatment of evidence throughout the entire procedure conducted in one Member State, irrespective of the fact that certain evidence is taken abroad.[39] The opposite is the case with the indirect taking of evidence, when the requested court executes the request according to *lex fori*, i.e., according to its evidence-taking rules, because in this case that court undertakes a specific procedural action. But, subsidiarily it may also execute the request in accordance with the national law of the requesting court, if the latter has requested it and if such taking of evidence is neither contrary to the national law of the requested court nor entails major practical difficulties.[40]

11 Finally, it should be noted that in view of the challenges of justice in the period of the COVID-19 pandemic and in order to achieve the digital objectives in the field of justice, the European Commission adopted the Proposal for a Regulation of the European Parliament and of the Council on the digitalisation of judicial cooperation and access to justice in cross-border civil, commercial and criminal matters, and amending certain acts in the field of judicial cooperation (hereinafter: Proposal on the digitalisation of judicial cooperation).[41] The aforementioned Proposal on the digitalisation of judicial cooperation proposes, *inter alia*, the use of videoconferencing for holding oral hearings in cross-border disputes in order to facilitate access to justice.[42] However, it is important to distinguish the scope of the Proposal on the digitalisation of judicial cooperation from the scope of the RER. The Proposal on the digitalisation of judicial cooperation provides for the introduction of videoconferencing for holding oral hearings when one of the parties to the proceedings is located in a Member State different from the one before whose court the proceeding is conducted.[43] Thus, it is only a question of facilitating the participation of the parties in cross-border proceedings via videoconference, but not about the taking of evidence, to which the provisions of the RER would continue to apply.[44]

C. The principle of immediacy vs videoconferencing systems in the cross-border taking of evidence.

12 In this section, we will consider the compatibility of the principle of immediacy with the use of videoconferencing systems in the direct cross-border taking of evidence by examining persons. It is a principle that has a long standing tradition across all European civil procedural law.[45] Indeed, as a civil procedure principle dealing with evidence-taking, the principle of immediacy imposes several requirements on the court. Among other things,

35 Ontanu, (n 24) 595.

36 Poretti, (n 19) 224.

37 Revised Evidence Regulation, art 19(8).

38 Franceso Parisi, Daniel Pi and Alice Guerra, "Access to Evidence in Private International Law" (2022) 23(1) Theoretical Inquiries in Law 14 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3964387> accessed 1 August 2022.

39 Đuro Vuković and Eduard Kunštek, *Međunarodno građansko postupovno pravo* (2nd edn, Zgombić&Partneri 2005) 188.

40 Revised Evidence Regulation, art 12(2)(3).

41 Commission, "Proposal for a Regulation of the European Parliament and of the Council on the digitalisation of judicial cooperation and access to justice in cross-border civil, commercial and criminal matters, and amending certain acts in the field of judicial cooperation" COM/2021/759 final (hereinafter: Proposal on the digitalisation of judicial cooperation).

42 Xandra Kramer, "Digitising access to justice: the next steps in the digitalisation of judicial cooperation in Europe" [2022] 56 *Revista General de Derecho Europeo* 5.

43 Commission (n 41) 27.

44 Kramer (n 42) 5.

45 European Law Institute and UNIDROIT (n 2) 114.

this principle requires the court understands the nature and content of evidence and to decide on its probative value.[46] There is no doubt that by introducing videoconferencing as a default form of direct evidence taking, the RER contributes significantly to strengthening the principle of orality and immediacy.[47] The court conducting proceedings in one Member State should no longer need to obtain information about a particular statement through the requested court of another Member State, considering that in practice, according to the previously mentioned statistics, indirect taking of evidence was the most widespread.[48] However, although it has been strengthened, the question arises of whether the principle of immediacy has been fully realised through videoconferencing. This question arises because videoconferencing as a technical solution still limits the direct observation of the court in terms of the immediate perception of the person testifying.[49]

13 In addition to the substantive and logical assessment of coherence, the realisation of the principle of immediacy allows the court to apply psychological criteria in assessing the probative value of statements. The court can pay attention to a respondent's gestures, the volume and tone of voice, as well as their relative persuasiveness in giving a statement, and thus it gets the opportunity to exercise the principle of free evaluation of evidence in its entirety.[50] The conducted research shows that the use of videoconferencing has an impact on the assessment of the credibility of statements.[51] Namely, the statements given by persons physically present in the courtroom are usually assessed by the court as more reliable and convincing than those made by videoconference.[52] A "digital barrier" in the form of a videoconference can lead to a wrong perception of the respondent's emotions, which can consequently have an impact on the establishment of facts by the court.[53]

14 The ELI/UNIDROIT Model European Rules of Civil Procedure suggest the use of videoconferencing in the cross-border taking of evidence in the EU as one of the possible options.[54] However, under the European Rules of Civil Procedure, the general position on the examination of witnesses or experts is that their oral statements are considered more reliable if those witnesses or experts are physically present in the courtroom when giving their statements.[55] ELI/UNIDROIT acknowledges that the use of videoconferencing contributes to the effectiveness and efficiency of the procedure. However, it is pointed out that the principle of immediacy is not fully achieved by the use of videoconferencing, as it is not equivalent to the physical presence of the respondent.[56] Therefore, it can be concluded that due to its shortcomings, the use of videoconferencing is a kind of substitute for physical contact between the court and the evidence.[57] In order to bridge the gap between physical and virtual presence in cross-border taking of evidence, and to fully realise the principle of immediacy, the next section considers some AI-based solutions that could possibly help to achieve this.

D. AI-based systems for assessing the credibility of statements.

15 In assessing the credibility of certain statements, AI is based on the application of machine learning such that the behaviour of respondents during their statements is compared to previously stored features of true or false statements collected

46 Siniša Triva and Mihajlo Dika, *Građansko parnično procesno pravo* (7th edn, Official Gazette 2004) 185.

47 Viktória Harsági, „Digital Technology and the Character of Civil Procedure" in Miklós Kengyel and Zoltán Nemessányi (eds), *Electronic Technology and Civil Procedure* (Springer 2012) 131.

48 See n 5.

49 Harsági (n 47) 131.

50 Triva and Dika (n 46) 186.

51 Alicia Bannon and Janna Adelstein, "The Impact of Video Proceedings on Fairness and Access to Justice in Court" (Brennan Center for Justice at New York University School of Law 2020) 6-7. <https://www.brennancenter.org/our-work/research-reports/impact-video-proceedings-fairness-and-access-justice-court> accessed 4 August 2022.

52 ibid

53 Amy-May Leach et al., "COVID-19 and the courtroom: how social and cognitive psychological processes might affect trials during a pandemic" (2021) 28(8) Psychology, Crime & Law 738.

54 European Law Institute and UNIDROIT (n 2) 171.

55 ibid 148.

56 ibid 161.

57 Georg E. Kodek, "Modern Communications and Information Technology and the Taking of Evidence" in Miklós Kengyel and Zoltán Nemessányi (eds), *Electronic Technology and Civil Procedure New Paths to Justice from Around the World* (Springer 2012) 274.

from respondents under controlled conditions.[58] According to previous research and the level of development, three basic techniques for assessing the credibility of respondents' statements can be distinguished: a) analysis of non-verbal behaviour, b) analysis of verbal behaviour, and c) analyses based on the brain imaging method (functional magnetic resonance imaging – fMRI).[59]

16 Analysis of non-verbal behaviour is usually based on the detection of false statements based on facial or eye movements.[60] Research shows that there is an interdependence between the expression of emotions and facial expressions, since facial expressions are neurologically controlled by two brain centres whose task is to control spontaneous and non-spontaneous facial movements.[61] In the case of true statements, emotions are spontaneous and consequently, facial expressions of the respondents are produced equally spontaneously.[62] However, if the respondent makes a false statement, both brain centres are activated and a neurological conflict occurs between spontaneous and non-spontaneous facial reactions, which are manifested in the form of micro expressions.[63] Furthermore, according to research, the eyes can also be a source for assessing the credibility of statements. Indeed, software has been developed that monitors eye tracking and blinking, as well as pupil dilation, and it uses these signs to assess the credibility of statements. According to some research results, it is reliable up to 90%.[64]

17 Analysis of verbal behaviour assesses the credibility of statements by measuring the respondent's voice stress level, which is higher in the case of deliberate deception, or even by linguistic analysis, which analyses the words spoken by the respondent and their frequency, which may imply a non-credible statement.[65] According to some research, linguistic analysis is reliable in deception detection in approximately 75% of cases.[66]

18 The European Union has also shown interest in non-verbal behaviourbased systems for assessing the credibility of statements. Namely, a virtual avatar was developed within the framework of the EU-funded iBorderCtrl project, which is based on the Automatic Deception Detection System (ADDS) whose purpose is to analyse the non-verbal behaviour of travellers.[67] The ADDS was tested in such a way that third-country nationals were questioned by an avatar via a web camera before arriving at the border crossing as part of the pre-registration process, in order to assess the credibility of their statements regarding the reasons for travelling.[68] The ADDS assessed the credibility of statements based on facial recognition technology and measurement of facial micro-expressions.[69] The accuracy of ADDS in detecting true statements was about 76%, while the reliability in detecting false statements was about 74%.[70]

19 Brain imaging-based analysis (fMRI) originated in the field of neuroscience. It was developed to detect misleading or deceptive statements based on blood flow in the brain, because it is believed that when a statement is false, parts of the brain are activated that are not normally active when the statement is true.[71]

20 In the context of this paper, systems that analyse both

58 M. U. Şen, V. Pérez-Rosas, B. Yanikoglu, M. Abouelenien, M. Burzo and R. Mihalcea, "Multimodal Deception Detection Using Real-Life Trial Data" (2020) 13(1) IEEE Transactions on Affective Computing 306 <https://ieeexplore.ieee.org/document/9165161/> accessed 4 August 2022

59 Tommaso Fornaciari and Massimo Poesio, "Automatic deception detection in Italian court cases" (2013) Artificial Intelligence and Law 306 <https://link.springer.com/article/10.1007/s10506-013-9140-4#citeas> accessed 5 August 2022

60 Bradshaw (n 1) 709.

61 Joan Pico, "The new challenges of evidence law in the fourth industrial revolution" in Koichi Miki (ed), *Technology, the global economy and other new challenges for civil justice* (Intersentia 2021) 486

62 ibid

63 ibid

64 Bradshaw (n 1) 709.

65 Fornaciari and Poesio (n 59) 307-308.

66 Fornaciari and Poesio (n 59) 308.

67 T. Krïgel, R. B. Schïtze and J. Stoklas, "Legal, ethical and social impact on the use of computational intelligence based systems for land border crossings" (2018) International Joint Conference on Neural Networks (IJCNN) 1 <https://ieeexplore.ieee.org/document/8489349> accessed 7 August 2022.

68 ibid 1-2.

69 Javier Sánchez-Monedero and Lina Dencik, "The politics of deceptive borders: 'biomarkers of deceit' and the case of iBorderCtrl" (2022) 25 (3) *Information, Communication & Society* 414 <https://www.tandfonline.com/doi/full/10.1080/1369118X.2020.1792530> accessed 7 August 2022

70 ibid 419.

71 Bradshaw (n 1) 709.

verbal and nonverbal behaviour are hypothetically considered for assessing the credibility of statements in the cross-border taking of evidence by videoconference. Indeed, for most systems, all that is needed is a web camera, a computer, a microphone, and an Internet connection,[72] i.e., essentially everything that is needed for taking evidence by videoconference, with the addition of automatic deception detection software.

E. The admissibility of using artificial intelligence in the assessment of the credibility of statements in the cross-border taking of evidence

21 The RER does not prescribe fundamental procedural elements for taking evidence, such as the admissibility and probative value of the evidence but leaves this to national procedural law.[73] However, it would be wrong to conclude that Member States are completely free with regard to the possible use of deception detection systems when taking evidence. Issues of admissibility of evidence are important to protect the fundamental rights of participants in the proceedings. Consequently, the admissibility of evidence can affect the effectiveness of cross-border judicial cooperation, which is closely related to the principle of mutual trust.[74]

22 Some Member States expressly regulate the inadmissibility of evidence by procedural law, alternatively this assessment of inadmissibility is developed through case law. Thus, for example, French law qualifies all evidence as inadmissible if obtained in an unfair manner.[75] French judges connect the unfairness of the evidence with the relevant provisions of the national Code of Civil Procedure, but also with the right to a fair trial guaranteed by the ECHR.[76] Moreover, if a decision involves an assessment of the behaviour of a particular person, French law does not allow judicial decisions to be based on the application of algorithms and automated processing of personal data.[77] According to Slovenian case law, the results of a polygraph as evidence in civil proceedings are considered inadmissible because the polygraph has elements of coercion, and it is up to the court to assess the reliability of an individual statement by applying the principle of free evaluation of evidence.[78] The situation is similar in Germany, where the results of the polygraph test are considered inadmissible evidence, even if the test was performed on a voluntary basis.[79]

23 Namely, it was previously said that the requesting court conducts the direct taking of evidence in accordance with the law of its Member State.[80] However, the central body or competent authority of the requested Member State is authorized to reject this request if, *inter alia*, it would be contrary to the fundamental principles of the law of requested Member State.[81] Therefore, the question arises whether the requested Member State would be authorized to reject the request for the taking of evidence with the application of a deception detection system, due to the contradiction with the fundamental principles of law of the requested Member State?

24 The answer to the question is not simple. However, in respecting national procedural peculiarities, the answer to the question could go in the negative direction. Namely, for direct taking of evidence by examining a person via videoconference, the court of the requested Member State can only provide technical support to the requesting court.[82] Furthermore, the direct taking of evidence by examining a person is always carried out on a voluntary basis, and the person testifying must be aware of the voluntariness of the testimony.[83] Therefore, it could be concluded that in the case of the application of the deception detection system by the requesting court, there would be no basis for the authority of the requested Member State to reject the request for the direct taking of evidence. Namely, the entire procedure is carried out before the court of the requesting Member State, which uses the

72 ibid

73 See n 33.

74 Župan (n 17) 473.

75 Vesna Rijavec and Tomaž Keresteš, „Restrictions on the Admissibility of Evidence" in C.H. van Rhee and Alan Uzelac (eds), *Evidence in Contemporary Civil Procedure* (Intersentia 2017) 98.

76 ibid

77 Florence G'sell, „AI Judges" in Larry A. DiMatteo, Cristina Poncibò and Michel Cannarsa (eds), *The Cambridge Handbook of Artificial Intelligence* (CUP 2022) 353.

78 Rijavec and Keresteš (n 75) 91-2.

79 ibid 92.

80 See n 37.

81 See n 33.

82 Revised Evidence Regulation, arts 19(6) and 20(2)

83 ibid, art 19(2)

deception detection system, while the requesting Member State only provides technical support in terms of computers, cameras, and microphones. Moreover, the examination of a person is carried out exclusively on a voluntary basis, and the person who needs to be heard is authorized to refuse to participate in the testimony. However, from the above it is still not possible to conclude that such a way of taking evidence would really be in accordance with the fundamental rights of the requesting Member State.

25 On the other hand, the situation regarding the issue of admissibility of evidence could be more challenging in the case of recognition and enforcement of judgements of one Member State in another Member State, if such a judgement originates from a procedure in which the deception detection system was applied. In that case, the Member States could, at the request of the interested party, refuse recognition and enforcement of the judgement due to conflict with public policy in the requested Member State, based on the Brussels Ibis Regulation[84]. Namely, the disparity of national procedural rules in the taking and evaluation of evidence *per se* is not a sufficient reason for establishing a violation of public policy.[85] However, the requested Member State may apply the public policy clause if recognition or enforcement of the judgement would violate rules considered essential in the legal order of the requested State or would constitute a violation of fundamental rights.[86] Fundamental rights are part of the general principles of law arising from the constitutions of the Member States and international treaties on the protection of human rights to which the Member States are parties.[87] Moreover, the Court of the European Union specifically indicates the importance of the ECHR and the right to a fair trial as a general principle of Community law.[88]

26 Considering the above and that the application of AI is considered for deception detection in the context of cross-border evidence collection based on the RER, the admissibility of the application of AI should be assessed against its compatibility with the right to a fair trial, which is guaranteed by the CFR and which Member States are obliged to respect when implementing EU law,[89] i.e., by the ECHR, to which all Member States are contracting parties.[90] Furthermore, with respect to personal data, it is established that any processing of personal data carried out in compliance with the RER must be compatible with the GDPR.[91] Therefore, in the next sections of the paper, the admissibility of the application of the deception detection system will be considered through the prism of the right to a fair trial and the GDPR provisions, and the compatibility of the system for assessing the credibility of statements with the draft AI Act will be considered as an additional contribution to this topic.

I. Right to a fair trial

27 It is known that the right to a fair trial (Article 6 of the ECHR, Article 47 of the CFR) consists of several elements. By analogy, the selected elements will be analysed in terms of the compatibility of the deception detection system with the right to a fair trial. Given that this paper analyses the possible application of a deception detection system in the cross-border taking of evidence, it is necessary to consider the views of the European Court of Human Rights (hereinafter: ECtHR) regarding the evidence itself. According to the case law of the ECtHR, the admissibility of evidence and the method of its assessment and probative value fall within the jurisdiction of national law and national courts.[92] However, this does not mean that national courts completely disregard the right to a fair trial with respect to the evidence-taking procedure. Indeed, the ECtHR assesses the fairness of the procedure as a whole, i.e., it assesses all aspects of the procedure, including the manner in which the evidence was

84 Regulation (EU) No 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (recast) OJ L351/1 (hereinafter: Bruxelles I bis Regulation), arts 45(1)(a) and 46.

85 Stefano Dominelli, „Unjustified Interruption of the Taking Evidence by the Court of Origin as a Ground to Refuse CrossBorder Enforcement Under the Brussels I Rules" (2022) 1(2) The Italian Review of International and Comparative Law 403 <https://doi.org/10.1163/27725650-01020009> accessed 11 August 2022

86 ibid 403-4.

87 ibid 404.

88 ibid

89 Charter of Fundamental Rights of the European Union [2016] OJ C 202, art 51(1)

90 Council of Europe, "Chart of signatures and ratifications of Treaty 005" (2022) <https://www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=005> accessed 11 August 2022

91 Revised Evidence Regulation, art 30(1).

92 Päivi Hirvelä and Satu Heikkilä, *Right to a fair trial* (Intersentia 2021) 104. See also *García Ruiz v Spain* App no 30544/96 (ECtHR, 21 January 1999); *Tiemann v France and Germany* App no 47457/99 47458/99 (ECtHR, 27 April 2000)

taken.[93] It follows from the above that there is nothing to prevent the introduction of a deception detection system into national procedural systems. However, this still does not mean that the right to a fair trial would not be violated in terms of assessing the fairness of the entire procedure and compatibility with other elements of the right to a fair trial.

28 The first controversial reason in the application of the deception detection system is the possible violation of the right to access the court. Namely, the right of access to a court guarantees that everyone has the right to have their civil rights and obligations decided by an independent and impartial court.[94] Given that the deception detection system would also have a certain influence in the procedure, the right could be violated. The right to access the court guarantees a decision by the court in a certain dispute.[95] It should be considered that the decision-making process includes a whole series of procedural actions that precede the rendering of a judgment. Evaluation of evidence is also one of such actions, the purpose of which is to determine the facts to which the law is applied. Therefore, if the deception detection system were to be applied in the assessment of the credibility of statements, it would be uncertain whether the court truly independently decided on disputed facts and the criteria that it utilized.

29 The next question that arises is in which cases is it necessary to foresee the use of a deception detection system. If only persons heard by videoconferencing were subjected to a deception detection system, then the party whose proposed witness is heard by videoconferencing would be put in an unequal position. The testimony of that witness would be subjected to a stricter assessment regime compared to other witnesses who testify in court in person. Such treatment could be in conflict with the right to the procedural equality of arms, which requires that each party be given a reasonable opportunity to present its case under conditions that do not place the party at a substantial disadvantage compared to the opponent.[96] According to the case law of the ECtHR, a different approach in dealing with the examination of witnesses from the opposing parties may call into question the principle of equality of the parties and constitute a violation of the right to a fair trial.[97] The right to adversarial proceedings before the court is also connected with the principle of equality of arms. The adversarial principle guarantees the parties to discuss all relevant procedural material before the court.[98] Namely, the parties have the right to be informed and to state their opinion on all evidence or statements presented in order to influence the court's decision.[99] Although optional for the court, the results of the deception detection system would certainly constitute a body of procedural material, and the parties should be able to discuss the content of these results. However, the expert knowledge of the parties, as well as judges and lawyers, about the technology of deception detection systems appears as a potential difficult problem to overcome. As a complex technology that is difficult to understand for most citizens, it could represent an obstacle in the discussion of the obtained results, and thus in the actual exercise of the right to adversarial proceedings.

30 There is also a danger that the application of the deception detection system will become a routine for the judge who does not independently assess the results obtained in relation to statement credibility, but automatically accepts them. This would constitute a possible violation of the right to an independent and impartial court and the right to a reasoned court decision. Namely, the court is obliged to properly consider the submissions, arguments and evidence presented by the parties, without prejudice to their assessment of whether they are material to its decision.[100] Therefore, it is the duty of the judge to examine each piece of evidence and reach a conclusion on its credibility and relevance. Furthermore, the court is obliged to justify its actions through the explanation of its decision.[101] Therefore, the court would have additional obligations to explain how it evaluated the obtained results of the deception detection system and why it accepted or did not accept the results of the deception detection system.

31 In the first part of the paper, it was mentioned that pursuant to the RER, the direct cross-border production of evidence is carried out on a voluntary

93 *Elsholz v Germany* App no 25735/94 (ECtHR, 13 July 2000)

94 Alan Uzelac, „Pravo na pravično suđenje u građanskim predmetima: Nova praksa Europskoga suda za ljudska prava i njen utjecaj na hrvatsko pravo i praksu" (2010) 60(1) Zbornik Pravnog fakulteta u Zagrebu 107.

95 *Lupeni Greek Catholic Parish and others v Romania* App no 76943/11 (ECtHR, 29 November 2016)

96 *Užukauskas v Lithuania* App no 16965/04 (ECtHR, 6 July 2010)

97 *Ankerl v Switzerland* App no 17748/91 (ECtHR, 23 October 1996)

98 Uzelac (n 94) 109.

99 *Ruiz-Mateos v Spain* App no 12952/87 (ECtHR, 23 June 1993)

100 *Carmel Saliba v Malta* App no 24221/13 (ECtHR, 29 November 2016)

101 *Suominen v Finland* App no 37801/97 (ECtHR, 1 July 2003)

basis. Accordingly, the application of deception detection systems should also rest on the voluntary consent of the person testifying. However, the question arises as to what happens if the witness is willing to testify, but without applying a deception detection system. Indeed, the court could prejudge the unreliability of an individual witness if he or she refuses to use the deception detection system, which could lead to subjective bias on the part of the court. According to the case law of the ECtHR, impartiality is determined using a subjective test that takes into account personal beliefs and behaviour of an individual judge, i.e., whether the judge had personal prejudice or bias in a particular case.[102] Therefore, if a witness refuses to be assessed by the deception detection system, this could establish unfounded subjective bias on the part of the judge, which could ultimately impact the dispute resolution process between the parties and violate the right to a fair trial. Also, the introduction of the deception detection system into the procedural law could raise doubts about the objective impartiality of the court as a judicial body, as it would call into question the public's trust in the courts, whose existence is necessary in a democratic society.[103] The aforementioned could contribute to the collapse of the mutual trust between citizens and the state, because the AI, through its application in court proceedings, would encroach on the very essence of the relationship between citizens and the state.[104]

32 The next question that arises is how to qualify the legal status of the deception detection system in procedural law. Is this system an expert *sui generis,* or an auxiliary tool of the court? It should be obvious that the deception detection system cannot be an expert, because experts are natural persons.[105] However, if it is taken into account that the deception detection system, based on its specific technical characteristics, makes an assessment of the credibility of statements, then it can be concluded that the deception detection system would be an auxiliary tool of the court for risk assessment, which, like an expert, observes and renders an opinion on the facts that are essential for assessing the veracity of allegations that are the subject of evidentiary proceedings.[106] Moreover, the parties always have the right to rely on the results and opinions of experts and even to raise the objection that the expert is biased. But how can you object to the deception detection system, or will the parties be informed of the individual statement credibility assessment results? With regard to the findings and opinions of experts that relate to a technical field that is not within the scope of knowledge of judges, the ECtHR took the position that such opinions are likely to have a dominant influence on the judge's evaluation of the facts.[107] Therefore, it would be necessary to give the parties an opportunity to look back at the results of the deception detection system, which would serve as an auxiliary tool of the court for risk assessment. Moreover, the influence on the judge would certainly be significant, and it would be necessary for the judge to discuss with the parties the relevance of the results of the deception detection system. The same is stated in the ELI/UNIDROIT Rules of European Civil Procedure. Namely, the ELI/UNIDROIT Rules allow the use of AI to the extent that it is compatible with the right to be heard. However, the Rules require that the use of AI be transparent, in such a way that the parties know that AI is being used and that they can discuss the nature, quality and conclusions that can be drawn from the application of AI.[108]

33 Furthermore, there is a real danger that the system could be biased against certain social groups based on gender, ethnicity or cultural affiliation. Therefore, the representativeness and quality of the stored data is critical to ensure that it faithfully represents all social groups.[109] For example, people of different genders may have different facial expressions, gestures or verbal expressions, while patterns of one gender predominate in the stored templates of the deception detection system.[110] All of this affects the reliability of the results obtained and the overall assessment of whether the procedure was fair, respecting other rights guaranteed by the ECHR. In addition, low-quality IT equipment or Internet connection may negatively affect the image or sound received during the cross-border taking of evidence, which may lead to unreliable results of the deception detection system. Through the ECtHR´s case law, bad acoustics, and even the image, can be reasons that may lead to a violation of the right to

102 *Micallef v Malta* App no 17056/06 (ECtHR, 15 October 2009)

103 *Wettstein v Switzerland* App no 33958/96 (ECtHR, 21 December 2000)

104 European Law Institute and UNIDROIT (n 2) 23.

105 Triva and Dika (n 46) 527.

106 ibid 526.

107 *Mantovanelli v France* App no 21497/93 (ECtHR, 18 March 1997)

108 European Law Institute and UNIDROIT (n 2) 23.

109 Jo Ann Oravec, „The emergence of "truth machines"?: Artifcial intelligence approaches to lie detection" (2022) 24(6) *Ethics and Information Technology* 6. <https://doi.org/10.1007/s10676-022-09621-6> accessed 11 August 2022

110 Bradshaw (n 1) 717.

a fair trial.[111]

II. General Data Protection Regulation (GDPR)

34 As mentioned above, deception detection systems can analyse a number of behavioural factors, including facial micro-expressions, eye tracking and voice cues. Given that these factors are evaluated using a range of technical tools, including video cameras, microphones and AI-based software, this could initially lead to a wrong impression that this is about biometric data processing. However, in the context of the GDPR, the application of the deception detection system would not fall under a stricter regime of processing special categories of data provided for in Article 9 of the GDPR.[112] Namely, the GDPR defines biometric data as "personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data".[113] As the purpose and function of a deception detection system is to determine the credibility of a particular statement, and not the identity of the respondent, it is obvious that the application of such systems does not fall within the scope of the stricter regime for the processing of special categories of data. Indeed, deception detection systems do not perform biometric comparisons, but compare individual factors, such as facial microexpressions, with factors of the same type that are crucial for determining the credibility of statements.[114]

35 However, the processing of personal data,[115] which is a broader term than biometric data, could fall within the context of deception detection systems under the provisions on automated individual decision-making, including profiling.[116] According to Article 4(4) of the GDPR, profiling means:

> "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements".

36 In accordance with the provisions of the GDPR, the data subject has the right not to be subjected to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.[117] Put simply, the aim of the said provision is to prevent decisions related to individuals from being made by machines whose content is not subject to human judgement.[118] The provision prohibiting automated data processing, including profiling, applies regardless of whether the final decision produces positive or negative effects, until its content is decided by a human being.[119] In relation to the prohibition of automated data processing and profiling, the GDPR prescribes certain exceptions. The aforementioned will still be permitted, *inter alia*, if it is authorised by European Union or Member State law to which the controller is subject and which also lays down appropriate measures to protect the rights and freedoms as well as legitimate interests of the data subject; or if the decision is based on the express consent of the respondent.[120]

37 There is no doubt that deception detection systems use personal data from which they extract certain

111 *Stanford v United Kingdom* App no 16757/90 (ECtHR, 23 February 1994)

112 Art 9(1) of the GDPR expressly prohibits, *inter alia*, the processing of biometric data, unless there exists one of the legal bases listed in art 9(2) of the GDPR.

113 GDPR, art 4(14)

114 Els J. Kindt, "Biometric data processing: Is the legislator keeping up or just keeping up appearances?" in Gloria González, Rosamunde Van Brakel, and Paul De Hert (eds), *Research Handbook on Privacy and Data Protection Law* (Edward Elgar Publishing 2022) 385.

115 Pursuant to art 4(1) of the GDPR, personal data means "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".

116 Keeley Crockett, Sean Goltz and Matt Garratt, "GDPR Impact on Computational Intelligence Research" (2018) International Joint Conference on Neural Networks (IJCNN) 4.

117 GDPR, art 22(1)

118 Paul Voigt and Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 180-181.

119 ibid 181-182.

120 GDPR, art 22(2)

factors based on which they analyse the behaviour of the person giving the statement, but also predict his or her reliability. Therefore, it is clear that, as prescribed by the GDPR, the operations of the deception detection system can theoretically be subsumed under profiling. In the context of this paper, this would refer to the case when the judge taking evidence in cross-border matters by examining a person would independently evaluate neither the results of the deception detection system, nor the entire statement of the respondent, but would base his or her decision solely on the obtained results without the possibility of influence.[121] Of course, with the fulfilment of one of the previously mentioned conditions that exceptionally allow the application of automated data processing systems. This consideration is only theoretical since in practical application it is incompatible with the right to a fair trial. Namely, if the judge really had to accept the results of the deception detection system without his or her decisive influence, then the system would become the judge. Therefore, it should be concluded that if the final decision on the credibility of the statement is made independently by the judge, regardless of the results of the deception detection system, then we would not deal with automated individual decision-making, i.e., profiling, after all.[122]

III. Artificial Intelligence Act Proposal

38 Considerations on the compliance of the deception detection system with the AI Act proposal follow from previous GDPR-related considerations. Namely, among other proposed solutions, the AI Act sets out the transparency requirements regarding emotion recognition systems[123] that use biometric data and imposes an obligation that the respondent shall be informed of his or her interaction with such a system.[124] Under the AI Act Proposal, deception detection systems could fall under the definition for emotion recognition systems as long as biometric data is not specified as a basis for emotion recognition[125]. Given that the AI Act takes over the definition of biometric data from the GDPR, it follows that deception detection systems would still not be considered emotion recognition systems.[126]

39 However, since the use of AI-based deception detection systems is considered from the aspect of potential use in cross-border evidence taking, the AI Act still requires scrupulous handling. Namely, all AI systems intended to assist judicial bodies in researching and interpreting facts and law and in applying law to a specific set of facts are considered high-risk systems.[127] Given that deception detection systems, as an auxiliary tool of the court, participate in research and interpretation of facts, because through the analysis of respondent behaviour they assess statement credibility, it is obvious that such systems would be considered high-risk in the context of the AI Act. However, the classification of an AI system in the judiciary as high-risk does not necessarily mean permission to use such systems.[128] Namely, according to clarifications of the AI Act, the use of high-risk systems should only be possible if it complies with the CFR and secondary law of the European Union and national laws of the Member States.[129] From the above, it should be clear that a possible application of a deception detection system in evidence taking, in addition to compliance with the CFR, would also require standardisation in national procedural law. The use of this system is directly related to the basic procedural elements related to the assessment of the probative value. As already mentioned above, the RER is restrained in this direction and leaves itsubject to national law.[130]

40 Given that AI systems intended for use in the judiciary are classified as high-risk, the AI Act sets more rigorous requirements for such systems. Acknowledging that deception detection systems, using machine learning, compare signs that point to a non-credible statement with signs from a stored data set, the data quality requirement is important

121 Keeley Crockett, Sean Goltz and Matt Garratt (n 116) 4.

122 *Arg a contrario.*

123 Pursuant to art 3(34) of the AI Act Proposal, emotion recognition system means "an AI system for the purpose of identifying or inferring emotions or intentions of natural persons on the basis of their biometric data".

124 Jan Czarnocki, "Will new definitions of emotion recognition and biometric data hamper the objectives of the proposed AI Act?" in: Brömme, A., Busch, C., Damer, N., Dantcheva, A., Gomez-Barrero, M., Raja, K., Rathgeb, C., Sequeira, A. & Uhl, A. (eds), *BIOSIG 2021 - Proceedings of the 20th International Conference of the Biometrics Special Interest Group* (Gesellschaft für Informatik e.V. 2021) 182.

125 See n 123.

126 Czarnocki (n 124) 182.

127 Michael Veale and Frederik Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act" (2021) 22(4) Computer Law Review International 102 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3896852> accessed 12 August 2022. See also art 6(3) of the AI Act and Anex III (8)(a) of the AI Act.

128 AI Act, recital 41.

129 ibid

130 See n 36.

in this context. Namely, the AI Act requires that the data sets meet the data quality criterion in such a way that they are relevant, representative, error-free and complete.[131] It is therefore necessary to ensure that the data stored in the deception detection system are regularly refreshed, with a complete and representative sample in terms of age, gender, race and other factors, in order to produce accurate results. Furthermore, *inter alia*, the AI Act requires human oversight of high-traffic AI systems.[132] The purpose of human oversight is to ensure that AI systems are subject to human control in order to reduce risks related to fundamental rights, health and safety.[133] Pursuant to the AI Act, human oversight requires a series of measures that enable users to understand the capabilities of the AI system, to interpret the results correctly, to stop, ignore or change the results at any time, and to intervene in the operation of the system.[134] Although the possibility of altering the results is reasonable and justified with the aim of protecting fundamental rights, attention should be paid to an interesting research study. Namely, the research was conducted under controlled conditions in order to determine whether the automatic deception detection system achieves greater accuracy in hybrid form, i.e., by the assessment of a judge who can reject the obtained results or adjust them within certain limits.[135] The research showed that human influence on the obtained results impairs their reliability, i.e., that judges are more inclined to classify answers as true even though they are not.[136] Therefore, based on the conducted research, the deception detection system proved to be more reliable than humans.[137]

41 In terms of human oversight, the concept of "automation bias" is interesting, which could also be a significant risk in the application of deception detection systems in cross-border evidence taking.[138] Namely, "automation bias" towards the AI Act is defined as "the tendency of involuntarily relying or over-relying on the output produced by a high-risk AI system (...), in particular for high-risk AI systems used to provide information or recommendations for decisions to be taken by natural persons."[139] Accordingly, the risk for judges in the application of the AI-based deception detection system consists in the risk of routine application of such systems. Over-reliance and the uncritical acceptance of the results would not be compatible with the right to a fair trial. In that case, AI would become the indirect judge, while the judge would be only a formal decision maker, whose decisions would be made on the basis of mechanical downloads of the deception detection system results.

F. Concluding remarks

42 The introduction of the direct cross-border taking of evidence by examining persons via videoconference will, as the primary method, undoubtedly contribute to the realisation of the right to access to justice. However, although the principles of orality and immediacy have been significantly strengthened, there is still room for strengthening the latter. The presented research shows that AI-based deception detection systems are nevertheless more accurate in terms of assessing the credibility of statements than the average person. Therefore, the research question from the introductory part of the paper should be answered in such a way that there is the potential for the application of a deception detection system in the cross-border taking of evidence. On the other hand, there is no shortage of counterarguments regarding the admissibility. Namely, open research questions within the framework of respect for the right to a fair trial still point to reticence about the application of the deception detection systems in civil proceedings. Although, according to research, the accuracy of deception detection systems is higher than human, the risks associated with the application of such systems are far greater. Unconditional commitment to fundamental rights, an essential component of which is the right to a fair trial, contributes to citizens' trust in the courts and strengthening their legitimacy. Also, if in the future the introduction of these systems into the judiciary were to be considered more seriously, then coherence in the legislative approach of the Member States is necessary. A unique approach in standardizing the use of AI in cross-border taking of evidence contributes to the preservation and strengthening of cross-border judicial cooperation and prevents the violation of mutual trust between Member States. It is also necessary to strengthen the digital competences of legal experts so that they can

131 Veale and Zuiderveen Borgesius (n 127) 103.

132 ibid

133 ibid

134 AI Act, art 14(4)

135 Bennett Kleinberg and Bruno Verschuere, "How humans impair automated deception detection performance" (2021) 213 *Acta Psychologica* 1-8 <https://www.sciencedirect.com/science/article/pii/S0001691820305746> accessed 28 August 2022.

136 ibid

137 ibid

138 AI Act, art 14(4)(b).

139 ibid

adequately understand and explain the work and effects of AI-based systems to parties. This mainly applies to the work of judges, for whom it is of crucial importance to understand the work of the AI and to resist the automation.

Note: The research reflected in this article was financed by the Young Researchers' Career Development Project – Training New Doctoral Students, funded by the Croatian Science Foundation.

Out-of-court dispute settlement mechanisms for failures in content moderation

by **Federica Casarosa**

Abstract: Content moderation is at the core of online platform activities. Many platforms allow users to post content that may or may not comply with the terms of service or that may violate national laws. In order to avoid these violations, online platforms have started to monitor content both ex post and ex ante. However, mistakes may still (frequently) happen.

In order to allow users to effectively contest decisions and compel platforms to restore content or accounts after erroneous decisions, online platforms should provide adequate due process mechanisms to appeal and seek redress. The EU has addressed this point by including specific provisions in the recently adopted Digital Services Act (DSA). In particular, Article 21 provides that complaints against online platforms can also be resolved through out-of-court dispute settlement mechanisms provided by certified bodies.

After analysing the role of online platforms in content moderation, this essay focuses on the types of dispute resolution mechanisms envisaged in the DSA. Assessing, on the one hand, the proposed criteria for effective out-of-court dispute settlement bodies according to the principles of fairness, accountability, independence and transparency and, on the other hand, the shortcomings that emerge from the certification mechanism defined in the DSA.

Keywords: Online Dispute Resolution; Digital Services; Content Moderation; Certification; Harmonisation

Recommended citation: Federica Casarosa, Out-of-court dispute settlement mechanisms for failures in content moderation, 14 (2023) JIPITEC 391 para 1

A. Introduction

1 Content moderation is at the core of online platform activities.[1] Every platform that provides a hosting service online allows users to post and disseminate content that may or may not comply with the platform's terms of service or, even worse, may be in contrast with national laws applicable in the user's country or in the country of establishment of the hosting platform. In order to avoid violations, online platforms have started to control content both *ex post*: by reducing and minimising the dissemination of unlawful content, and *ex ante:* by employing a preventive mechanism able to screen and eventually hamper uploads of content even before they are published.[2]

* Part-time professor at the Centre for Judicial Cooperation, EUI. Visiting Fellow at Mazaryk University. This contribution is based on research activity carried out in the framework of the DG Justice-supported project the e-Justice ODR scheme (GA n. 101046468).

1 Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (Yale University Press, 2018).

2 See James Grimmelmann, 'The Virtues of Moderation' (2015) 17 *Yale Journal of Law & Technology* 42. Note that automated content filtering has been used since the first years of internet development, as many tools have been deployed to analyse and filter content and among them the most

2 Content moderation, which is more and more frequently carried out using technical tools that include artificial intelligence systems, was not originally part of the legal obligations involved in the services provided by online platforms. In fact, online platforms were (and still are) not required to verify the content available on their platforms, as, for instance, the editors of online newspapers are. Nonetheless, as we will see, it has become a norm as a result of both economic decisions and incentives provided by policymakers.

3 However, content moderation is definitely not free from flaws depending on the ability of technological tools to recognise the substance of the content analysed (e.g. whether or not it qualifies as hate speech or aggressive expression) and also the context in which the content is expressed (e.g. a quotation from another person, a joke or a verbal attack). If these factors are not correctly evaluated then mistakes may occur which have a subsequent effect on the choice of the online platform to remove or disable access to the content.

4 What happens if content is wrongly removed? Many examples can be recalled, including the decision by Facebook to remove the well-known photo of the so-called 'Napalm girl.'[3] When public outcry points out the mistake it is easy to restore the status quo. Eventually this may also help technology improve, as in the above example the algorithm used by the social platform learned that the specific photo was not to be deemed to be pornography.

5 However, several less dramatic cases may emerge, leaving users to decide whether it is worth starting a quarrel with an online platform over why content has been removed. It must be acknowledged that some platforms have already started to provide different forms of resolution. For instance, the Facebook Oversight board employs a procedure applicable to a selected number of complaints[4] and the YouTube Content ID claim mechanism is an automated tool that is triggered by an inclusion of copyrighted material on the platform.

6 Recently the European Commission has addressed the issue of a reliable complaint handling mechanism that should reduce the negative impact of erroneous removals of content. The recently adopted Digital Services Act (DSA)[5] addressed the point by including specific provisions. In particular, Art. 20 provides that for decisions on content removal, suspension of service and account termination the internet platform should make available a free internal electronic complaint handling mechanism. This should not only be automated but also have human oversight. Alternatively, Art. 21 provides that complaints against online platforms can also be resolved using out-of-court dispute settlement mechanisms provided by certified bodies. In order to verify if the solution proposed in Art. 21 DSA will be an efficient tool to resolve cases of erroneous decisions by online platforms we need to clarify which standards are adopted in the legislation.

7 This contribution will therefore first analyse the role of online platforms in content moderation (Section B). Subsequently, it will describe the type of dispute resolution mechanisms envisaged in the DSA (Section C), assessing on the one hand the proposed criteria for effective out-of-court dispute settlement bodies according to the principles of fairness, accountability, independence and transparency and, on the other hand, the shortcomings that emerge from the certification mechanism defined in the DSA. Conclusions follow.

B. The role of online platforms in content moderation

8 The starting point to understand the role and obligations of online platforms regarding content moderation is the Directive on electronic commerce 2000/31/EC,[6] which will be applied at least until 2024.[7] In its Art. 14, this directive classifies online

common and well known are those for spam detection or hash matching. For instance, spam detection tools identify content received at one's email address, distinguishing between clean emails and unwanted content on the basis of certain sharply defined criteria derived from previously observed keywords, patterns and metadata. See Thamarai Subramaniam, Hamid A. Jalab and Alaa Y. Taqa 'Overview of Textual Anti-spam Filtering Techniques' (2010) 5 *International Journal of Physical Science* 1869.

3 See Hortense Goulard, 'Facebook accused of censorship of 'Napalm girl' picture,'' 9 November 2016 <https://www.politico.eu/article/norwegian-prime-minister-facebook-wrong-to-censor-vietnam-war-picture/>.

4 Kate Klonick 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (2020), *The Yale Law Journal* 129, 2450.

5 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act).

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

7 As will be explained later, the new legal framework provided

platforms as Information Society Service Providers (ISSP), and in particular as hosting providers.[8] Hosting providers are only exempted from liability for the content they store if they have neither actual knowledge of illegal activity or information nor awareness of facts and circumstances from which illegal activity and information are apparent. Only if they obtain such knowledge or awareness are hosting providers obliged to act expeditiously to remove or disable access to the information through a notice and take-down procedure. As a result, hosting providers are treated as pure passive and neutral actors that should not interfere in the storage and transmission of online content.[9] The Directive on electronic commerce goes even further and in Art. 15 it excludes an obligation to *ex ante* monitor content. This article has been further clarified in recent CJEU case law.[10] In its *Glawischnig-Piesczek v. Facebook* decision, the court affirmed that there is no violation of the prohibition of a monitoring obligation in Art. 15(1) of the Directive on electronic commerce even if a national court orders a platform to prevent the publication of "information with an equivalent meaning."[11]

9 However, seeing hosting providers as pure passive intermediaries is now an outdated vision of their role. Hosting providers still distribute user content and facilitate user interactions, although they are now more and more able to intervene in the experience that users have of their online activities.[12] Online platforms now provide a wide-ranging set of services including online advertising platforms, marketplaces, search engines, social media, creative content outlets, application distribution platforms, communication services, payment systems and platforms for the collaborative economy.[13] Although from a technical perspective, each of the above-mentioned cases has specific characteristics, from a substantial perspective delivery of these services allows online platforms to steer and control what users may disseminate.

10 How is this control exercised? The immediate answer is content moderation. As mentioned above, content moderation aims to verify if content hosted and stored on a platform is in line with its internal rules and conditions and with the applicable laws and regulation. This monitoring, which is exercised both *ex ante* and *ex post*, is not without consequences in terms of the choices available to users and also the ability of users to express themselves on online platforms.[14] The literature has highlighted that pre- and post-publishing moderation activities have strong impacts on the exercise of users' freedom of expression rights.[15] This has also been confirmed by

in the proposed Digital Services Act will depend on the date of its adoption by EU bodies. After it is published, its rules will apply 15 months after its entry into force or from 1 January 2024, whichever is later. See European Commission, 'The Digital Services Act: ensuring a safe and accountable online environment,' available at <https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en>.

8 According to Art 14, hosting "consists of the storage of information provided by a recipient of the service".

9 Directive 2000/31/EC, Recital 42.

10 CJEU, Case C-18/18, *Eva Glawischnig-Piesczek v. Facebook Ireland Limited* [2019] ECLI:EU:C:2019:821.

11 Case 18/18 (n 11) para 46. See Federica Casarosa 'When the algorithm is not fully reliable: the collaboration between technology and humans in the fight against hate speech,' in Oreste Pollicino and Andrea Simoncini (eds.) *Constitutional Challenges in the Algorithmic Society* (2021, Cambridge University Press) 298.

12 Note that the 'active hosting providers' qualification has also been developed in national jurisprudence. See the analysis of Italian jurisprudence on this point in Federica Casarosa 'Copyright Infringing Content Available Online – National Jurisprudential Trends' in Agusti Cerrillo i Martínez, Miquel Peguera, Ismael Peña-López, *et al.* (eds.) *Challenges and Opportunities of Online Entertainment. Proceedings of the 8th International Conference on Internet, Law & Politics. Universitat Oberta de Catalunya, Barcelona 9-10 July, 2012* (UOC-Huygens Editorial, 2012) 61.

13 For a taxonomy of the activities provided by online platforms, see European Parliament *Liability for online platforms* (2021, European Union publications : Brussels) IV.

14 Content moderation, although mostly interpreted as a form of monitoring of comments, posts and speech in general, can cover all the types of content that are shared on an online platform, such as, for instance, copyrighted material in the form of text, audio, video or also goods, as is clarified in CJEU, Joined cases C-236/08 to C-238/08, *Google France SARL and Google Inc. v Louis Vuitton Malletier SA* (C-236/08), *Google France SARL v Viaticum SA and Luteciel SARL* (C-237/08) and *Google France SARL v Centre national de recherche en relations humaines (CNRRH) SARL and Others* (C-238/08), 23 March 2010, ECLI:EU:C:2010:159.

15 Enguerrand Marique and Yseult Marique, 'Sanctions on digital platforms: Balancing proportionality in a modern public square', *Computer law & security review* 36 (2020), 1; Luc von Danwitz, 'The Contribution of EU Law to the Regulation of Online Speech', 27 *MICH. TECH. L. REV.* 167 (2020), 185; Kate Klonick, The New Governors: 'The People, Rules, and Processes Governing Online Speech', 131 *Harv. L. Rev.* 1598 (2018); Orla Lynskey, 'Regulation by Platforms: the Impact on Fundamental Rights', in Luca Belli and Nicolo Zingales (eds), *Platform regulations: how platforms are*

cases brought before national and European courts. The European Court of Human Rights (ECtHR) in the *Cengiz and Others v Turkey* case emphasised that online platforms such as Facebook, Twitter and YouTube provide an "unprecedented" means of exercising freedom of expression online.[16] This has also been confirmed more recently in *Delfi v Estonia*, in which the Strasbourg court affirmed that the internet is an "unprecedented platform for the exercise of freedom of expression."[17] Similarly, national courts have acknowledged that social networks can be equivalent to public spaces,[18] although the internet may lead to "inexpensive, easy, and instantaneous means whereby unscrupulous persons or ill-motivated malcontents may give vent to their anger and their perceived grievances against any person."[19] In order to cope with these risks, it is possible to affirm that when online platforms design the moderation rules they are contextually providing their own balancing of the rights and freedoms of users on the platform itself.[20]

11 From the point of view of online platforms, content moderation rules must strike a balance between the protection of free speech online and business interests. Clearly, platforms are eager to attract and retain users, not only in terms of the numbers of individuals registered but also in terms of content that circulates on the platform. Only if the users feel – relatively – free to express their opinions on platforms will they participate and indirectly contribute to its growth. However, in order to enhance users' perceptions that they are part of a network of like-minded people, the online platform may promote the visibility of selected content, leading to a proliferation of so-called 'filter bubbles.'[21] The ability to decide what users may or may not get in contact with has been acknowledged as a concentration of power in the hands of online platforms, which has triggered a wide academic debate regarding the legitimacy and effectiveness of pre- and post-moderation activities, not only considering the standards applied but also considering the technical tools applicable to such activities.[22]

12 The development of technology has also impacted the ability of online platforms to scan and identify suspicious content. Several studies have highlighted the increased adoption of artificial intelligence tools for content moderation.[23] The advantages of these technologies are of course lower costs, rapidity of analysis and, presumably, a high rate of correct evaluation of content. However, the effectiveness of the technology is limited by its ability to accurately analyse and classify content in its own context. The ability to parse the meaning of a text is highly relevant when making important distinctions in ambiguous cases such as, for instance, when differentiating between contemptuous speech and irony. For this task, the industry has now increasingly turned to machine learning to train its programs to become more context sensitive.[24]

regulated and how they regulate us (2017, FGV Direito Rio), 83.

16 ECtHR, *Cengiz and Ors v Turkey* Apps. nos. 48226/10 and 14027/11 (ECtHR, 1 December 2015), para 49.

17 ECtHR, *Delfi AS v Estonia* App. no. 64569/09 (ECtHR, Grand Chamber, 16 June 2015), para 110. For a more detailed analysis of the jurisprudence of the ECtHR on freedom of expression with specific application to social media, see Lorna Woods 'Social Media Jurisprudence: The European Court of Human Rights' in Federica Casarosa and Evangelia Psychogiopoulou (eds.) *Social Media and National Courts In Europe: A Fundamental Rights Perspective* (Routledge, forthcoming 2023), 48.

18 See Italian Court of Cassation decision no. 37596/2014, in which the Court affirmed that Facebook is to be considered a place open to the public as it constitutes a 'virtual' place open to access by anyone using the network. For more, see Federica Casarosa and Concetta Causarano, 'Social Media Before Higher Courts In Italy: A Thorough Adaptation of Existing Rules and Protection of Constitutional Rights Online' in Casarosa and Psychogiopoulou (n 18), 170.

19 See Mr Justice Peart's opinion in *Tansey v Gill* [2012] IEHC 42, as quoted by Elisabeth Farries in 'Social Media, Fundamental Rights and Courts: An Irish Perspective' in Casarosa and Psychogiopoulou (n 18), 152.

20 Giovanni De Gregorio and Oreste Pollicino, 'The European Constitutional Road to Address Platform Power' VerfBlog, 31 August 2021 <https://verfassungsblog.de/power-dsa-dma-03/>; Klonick (n 5).

21 The concepts of 'echo chambers' and 'filter bubbles' were identified as risks in internet communication since early 2000 by Cass Sunstein and Eli Pariser respectively. See CR Sunstein, *Republic.com* (Princeton University Press, 2001); and Eli Pariser *The Filter Bubble: What the Internet Is Hiding from You* (Penguin, 2011).

22 De Gregorio and Pollicino (n 20); David Kaye, "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression" Human Rights Council, A/HRC/38/35, 6 April 2018 <https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/35>; Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Public Affairs, 2019).

23 Emma Llanso (2019), Platforms want centralized censorship. That should scare you. Wired, 18 April. Available at: <https://www.wired.com/story/tumblr-porn-ai-adult-content/>; Tarleton Gillespie 'Content moderation, AI, and the question of scale,' *Big data and society* (2021).

24 Christoph Krönke, 'Artificial Intelligence and Social Media,'

13 It should not be a surprise that the development of technology has not always delivered the expected results. For instance, Appelman, Quintais and Fahy highlight that content moderation systems fail to safeguard freedom of expression in particular in cases of speech by minority and marginalised groups, black activist groups, environmental activist groups and other activists.[25]

14 These cases cannot be qualified as mere mistakes as the level of automation adopted by content moderation mechanisms allows online platforms to assess millions of posts (be it in textual or graphical representation) every week and even very low error rates can equate hundreds of thousands of mistakes every week.[26] Moreover, the biases that may – consciously or not – be embedded in the automated content moderation mechanism may lead to a risk of over-broad censorship.[27]

15 In order to allow users to effectively contest decisions and compel platforms to restore content or accounts after erroneous decisions (so called 'put-back'), online platforms should provide adequate due process mechanisms to appeal and seek redress, either by an internal complaint handling mechanism or by an out-of-court dispute settlement mechanism.[28] As was mentioned above, this last option is one of the most interesting innovations in the recent proposal for a Digital Services Act. Unfortunately, as we will describe in the next sections, the out-of-court dispute resolution mechanism envisaged in the proposed legislation does not meet expectations as it not only refrains from providing more detailed standards related to the due process guarantees but it also leaves issues with addressing the certification mechanisms that should apply to out-of-court dispute resolution providers.

C. The out-of-court dispute resolution mechanism envisaged in the Digital Services Act

16 On 16 December 2020 the European Commission published two linked proposals addressing the governance of digital data, namely the Digital Services Act (DSA) and the Digital Markets Act (DMA). Both proposals were already envisaged in the European Digital Strategy "Shaping Europe's Digital Future"[29] and were aimed at promoting fundamental rights in digital services and promoting technological innovation through the establishment of common rules for digital service providers in the European single market and beyond.[30] The final text of the DSA was adopted on 19 October 2022.[31]

17 The DSA aims to provide a dedicated horizontal regulatory framework for online platforms with rules on digital services in order to prevent unfair

in Thomas Wischmeyer and Timo Rademacher (eds.) *Regulating Artificial Intelligence* (Springer, 2019).

25 Naomi Appelman, João Pedro Quintais and Ronan Fahy, 'Using Terms and Conditions to apply Fundamental Rights to Content Moderation: Is Article 12DSA a Paper Tiger?' *VerfBlog* 1 September 2021 <https://verfassungsblog.de/power-dsa-dma-06/>. See also the case of Google's AI tool aimed at detecting toxic comments, which according to studies often classifies comments in African-American English as toxic. See Jonathan Vanian, "Google's Hate Speech Detection A.I. Has a Racial Bias Problem", *Fortune*, 16 August 2019 <https://fortune.com/2019/08/16/google-jigsaw-perspective-racial-bias/>.

26 Nicolas Suzor, *Lawless: The Secret Rules That Govern Our Digital Lives* (2019, Cambridge University Press); Daniel Holznagel, 'The Digital Services Act wants you to "sue" Facebook over content decisions in private de facto courts' *VerfBlog* 24 June 2021 <https://verfassungsblog.de/dsa-art-18/>.

27 The UN Special Rapporteur on freedom of expression has criticised these content moderation systems for their overly vague operating rules, inconsistent enforcement and over-dependence on automation, which can lead to over-blocking and pre-publication censorship. See also Kaye (n 23).

28 Marta Cantero Gamito, 'Regulation of online platforms,' (2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3971076>; Giancarlo Frosio, 'Why Keep a Dog and Bark Yourself? From Intermediary Liability to Responsibility' (2018) 26 *Oxford Int'l J. of Law and Information Technology* 1; Marten Schultz 'Six Problems with Facebook's Oversight Board. Not Enough Contract Law, Too Much Human Rights,' in Judith Bayer, Bernd Holznagel, Paivi Korpisaari and Lorna Woods (eds.) *Perspectives on Platform Regulation* (2021, Nomos Verlagsgesellschaft mbH & Co. KG) 145; Amy Schmitz 'Expanding Access to Remedies through E-Court Initiatives' (2019), 67 *Buffalo Law Review* 89.

29 European Commission, "Shaping Europe's Digital Future" (European Commission, February 2020) <https://ec.europa.eu/info/sites/info/files/communication-shaping-europes-digital-future-feb2020_en_4.pdf>.

30 European Commission, "Proposal for a Regulation on a Single Market For Digital Services (Digital Services Act)" COM(2020) 825 final (European Commission, December 2020), p 2, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>.

31 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act).

practices by online intermediaries and to reduce the power of gatekeepers.[32] Although the European Commission presented the DSA as an act reshaping the European rules on platform governance, its provisions addressing internet service provider liability cannot be qualified as innovative. Instead, the act makes an effort to integrate the Court of Justice's interpretation of the rules on liability.[33]

18 The DSA follows the same distinction provided in the Directive on electronic commerce between mere conduit, caching and hosting services. In the last category the DSA includes a subcategory of online platforms that are defined as operators bringing together sellers and consumers such as online marketplaces, app stores, collaborative economy platforms and social media platforms. Online platforms can only benefit from the liability exemption contained in Art. 6(1) DSA if the following conditions are met:

(a) the online platform does not have actual knowledge of the illegal activity or illegal content and, regarding claims for damages, is not aware of facts or circumstances from which the illegal activity or illegal content is apparent; and

(b) on obtaining such knowledge or awareness, it acts expeditiously to remove or to disable access to the illegal content.[34]

19 Moreover, online platforms will not lose the benefits of the liability exemption even when carrying out "voluntary own initiative investigations or other activities aimed at detecting, identifying and removing, or disabling access to illegal content," as is affirmed in Art. 7. Accordingly, the DSA confirms that online platforms may autonomously perform content moderation activities regarding information stored and transmitted through their platforms without a need to receive prior permission from judicial or other competent authorities.

20 However, the DSA introduces an additional step that addresses the procedure that online platforms should follow when content is removed or disabled. Art. 17 DSA provides that users should be informed of the removal of their content or disablement of access to it at the latest by the time of the decision, providing not only a statement of the fact, but also a clear and specific statement of the reasons that led to the platform's decision. Users should also receive information on the redress possibilities available to the recipient of the service in respect of the decision, particularly through internal complaint-handling mechanisms, out-of-court dispute settlement and judicial redress.

21 The inclusion of a specific rule addressing the availability of out-of-court dispute settlement mechanisms is not new in EU legislation, as several other recent European interventions have included a set of similar provisions. For instance, Art. 13 of Regulation (EU) 2019/1150 on promoting fairness and transparency for business users of online intermediation services[35] requires "providers of online intermediation services and organisations and associations representing them to, individually or jointly, set up one or more organisations providing mediation services [...] for the specific purpose of facilitating the out-of-court settlement of disputes with business users arising in relation to the provision of those services." Similarly, Art. 17(9) of Directive 2019/790 on copyright and related rights in the Digital Single Market[36] specifies that online content-sharing service providers shall provide an effective and expeditious complaint and redress mechanism, which is qualified as an out-of-court redress mechanism in cases of disputes between rightsholders asking for content removal and platforms. Another example comes from Directive 2018/1808 amending the Audio-visual Media Services Directive,[37] Art. 28b provides for out-of-court redress for the settlement of disputes between users and video-sharing platform providers.[38]

22 The DSA identifies a more detailed architecture

32 Miriam Buiten 'The Digital Services Act: From Intermediary Liability to Platform Regulation' (2021) <https://www.ibm.com/cloud/learn/machine-learning>.

33 Caroline Cauffman and Catalina Goanta 'A New Order: The Digital Services Act and Consumer Protection' *European Journal of Risk Regulation* (2021) 12(4), 758.

34 See CJEU, C-324/09, *L'Oréal SA and Others v eBay International AG and Others*, [2011], ECLI:EU:C:2011:474, part. Para 113.

35 Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services.

36 Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC.

37 Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities.

38 Jörge Wimmers 'The Out-of-court dispute settlement mechanism in the Digital Services Act – A disservice to its own goals' 12 (2021) JIPITEC 421

addressing the selection of potential conflicts that may emerge on the basis of decisions by online platforms. The cases are listed in Art. 20 DSA and include

> "(a) decisions whether or not to remove or disable access to or restrict visibility of the information;
>
> (b) decisions whether or not to suspend or terminate the provision of the service, in whole or in part, to the recipients;
>
> (c) decisions whether or not to suspend or terminate the recipients' account;
>
> (d) decisions whether or not to suspend, terminate or otherwise restrict the ability to monetise information provided by the recipients."[39]

23 When affected with these types of decisions by the platform, Art. 21 DSA requires online platforms to indicate a certified out-of-court dispute resolution provider that can solve the dispute. The article goes further and relies on the lawful behaviour of online platforms in the procedure ("engage in good faith") and recognises the decisions of the out-of-court dispute resolution providers as binding.[40]

24 Art. 21(3) identifies a set of due process guarantees that the out-of-court dispute resolution provider should ensure in order to be certified. The provision lists the following elements:

> "(a) it is impartial and independent, including financially independent, of providers of online platforms and of recipients of the service provided by providers of online platforms, including of individuals or entities that have submitted notices;
>
> (b) it has the necessary expertise in relation to the issues arising in one or more particular areas of illegal content, or in relation to the application and enforcement of terms and conditions of one or more types of online platform, allowing the body to contribute effectively to the settlement of a dispute;
>
> (c) its members are remunerated in a way that is not linked to the outcome of the procedure;
>
> (d) the out-of-court dispute settlement that it offers is easily accessible, through electronic communications technology and provides for the possibility to initiate the dispute settlement and to submit the requisite supporting documents online;
>
> (e) it is capable of settling disputes in a swift, efficient and cost-effective manner and in at least one of the official languages of the institutions of the Union;
>
> (f) the out-of-court dispute settlement that it offers takes place in accordance with clear and fair rules of procedure that are easily and publicly accessible, and that comply with applicable law, including this Article."

25 Moreover, in order to increase the incentives for users to submit their complaints, Art. 21 (3) specifies that if the final decision of the out-of-court dispute resolution provider results in favour of the user, the latter will receive a refund from the online platform covering the fees and expenses incurred.

I. The due process guarantees

26 Although several criticisms have already emerged in the literature addressing the doubts and ambiguities regarding the subjective and material scope of out-of-court dispute resolution mechanisms,[41] this contribution focuses on the due process guarantees that allow the out-of-court dispute provider to be certified.

27 Although Art. 21(3) DSA does not elaborate in detail on the elements listed, we can interpret these elements on the basis of applicable criteria in the existing literature.[42]

39 These cases can also be solved using internal complaint-handling mechanisms, as Art. 20 DSA requires online platforms to provide users with an internal complaint-handling system "for a period of at least six months following the decision." Where a complaint contains sufficient evidence that the information is not illegal and not incompatible with the terms and conditions of the provider, the provider shall reverse the decision.

40 Note that Article 21 only refers to online platforms.

41 Wimmers (n 38); Holznagel (n 27).

42 See also the Manila Principles on Intermediary Liability, a joint declaration by a group of civil society organisations. These provide some minimal guidelines on what a legitimate decision-making process should include. Most relevantly, the Manila Principles require that users be given an opportunity to appeal decisions to restrict content, and these processes should be as transparent as possible without harming the privacy rights of individuals. These procedural safeguards are the hallmark of legitimate decision-making. Under the standards of the rule of law, rules must be clear, well known and fairly applied, and they must represent some defensible vision of the common good. See Suzor (n 27); Pablo Cortes, *The Law of Consumer Redress in an Evolving Digital Market - Upgrading from Alternative to Online Dispute Resolution* (2017, Cambridge University Press); Jie Zheng, *Online Resolution of E-commerce Disputes - Perspectives from the European Union, the UK, and China* (2020, Springer)

28 The first element defined in Art. 21(3)(a) DSA is the impartiality and independence of the body vis-à-vis online platforms and users. In this case, independence can be evaluated by means of the membership rules that are applied by the out-of-court dispute settlement body. On the one hand, members of the body should have terms of office long enough to ensure the independence of their actions; on the other hand, members should disclose any circumstances that may, or may appear to, affect their independence or create a conflict of interest. An additional element related to independence is the availability of adequate financial and human resources to carry out their functions effectively.[43] This is an important issue as the financial resources of an out-of-court dispute settlement body may come from the fees allocated to the parties for the settlement procedure as mentioned also by letter (c). Therefore, it may be possible that in order to attract as many cases as possible there is a risk of preferring the positions of claimants in order to incentivise their participation.

29 The second element defined in Art. 21 (3)(b) is the necessary expertise, which requires knowledge and competence not only concerning the legal rules applicable to the case at stake but also concerning the terms and conditions that may have triggered the decision of the online platform. This very general requirement should be framed according to the object of the dispute settlement procedure. Therefore, the members of the out-of-court dispute settlement body will have a keen understanding of the law and its application when balancing conflicting fundamental rights. Accordingly, out-of-court dispute settlement bodies should select their members from trained lawyers who are familiar not only with the applicable Union and national laws but also with the relevant case law.[44]

30 The elements defined in Art. 21(3) (d) and (e) are connected to the provision of a settlement procedure that is easily accessible by users and that does not require them to make a high investment of time and resources. Out-of-court dispute settlement bodies may adopt several features in order to ease their accessibility by users. These clearly include the fee adopted in order to cover the cost of the procedure. However, this point is further specified in Art. 18(3), in which a clear preferential treatment for users is defined. If an out-of-court dispute settlement body decides the dispute in favour of the user then it is the online platform that bears all the costs (including any fees and procedural expenses) suffered by the user. However, if the decision is in favour of the online platform then the fees and procedural costs are allocated to each party.

31 Other practical elements can include the availability of sample documents able to clarify the type of information required, the availability of a (free or paid) online expert advisor and the possibility to select the type of case documentation to provide if, for instance users are only asked to fill in a template online or if the provider allows paper filings to be automatically converted into online forms. Additionally, the technological tools used for resolution of the dispute can be adapted to the preferences of the users, including mediation, blind bidding, videoconferencing, chat rooms etc. The selection of such tools may impact the ability of users to access the settlement body.[45]

32 The last element introduced in Art. 21(3)(f) is transparency and fairness of procedure. Out-of-court dispute settlement bodies should ensure that all the steps that lead to the decision are transparent and fair.[46] For instance, there must be clear rules on the procedure to select the person in charge of deciding the dispute, the factual circumstances that the deciding body will take into account and how the documentation will be handled and stored. Moreover, attention should be paid to the power of investigation that may be allocated to the out-of-court dispute settlement body and the power of the parties to contest the results of the investigation. Finally, out-of-court dispute settlement bodies should indicate the standards that would apply regarding evaluation of content as unlawful, in particular if they rely on the national or international provisions addressing the exercise of freedom of expression online.

33 This element is also relevant to avoid the risk of out-of-court settlement body shopping leading to a race to the bottom. If procedures are uniformly assessed according to the criteria of fairness and transparency there will be fewer opportunities for different

236 ff; Loïc Cadiet, Burkhard Hess, Marta Requejo Isidro (eds.) *Privatizing Dispute Resolution – Trends and limits* (2019, Nomos);

43 Kristina Irion, Wolfgang Schulz and Peggy Valcke, *The Independence of the Media and Its Regulatory Agencies: Shedding New Light on Formal and Actual Independence Against the National Context* (2014, Intellect, Bristol UK / Chicago USA).

44 Compare with the requirements provided by the Directive 2013/11/EU of the European Parliament and of the Council of 21 May 2013 on alternative dispute resolution for consumer disputes, in particular article 6. See the analysis in Cortes (n 42) 107.

45 Cortes (n 42) 254.

46 Compare also with the analysis of Caroline Daniels, 'Alternative Dispute Resolution for European Consumers: A Question of Access to and Standards of Justice', in Cadiet, Hess, Requejo Isidro (n 42) 257, part. 287.

quality standards across Europe. Then, users may select *any* certified out-of-court settlement body in *any* Member State.

34 As was clarified above, Art. 21 (3) DSA only provides a short list of elements that should be evaluated. However, more detailed standards can be identified. The following may act as recommendations for a checklist that can guide evaluation of the due process guarantees.

Impartiality and independence	- **Disclosure of financial resources** - **Duration of membership** - **Absence of conflict of interests**
Expertise	- **Selection of members (interview evaluation, prior experience, etc.)**
Accessibility, efficiency and cost-effectiveness	- **Availability of an advice portal (free or paid) providing sample legal documents and online expert advice** - **Case documentation (only online filings or both online and paper filings, according to the preferences of the users; paper filings are manually converted into online forms or automatically converted)** - **Availability of technological tools (blind bidding; videoconferencing; chat room)** - **Cost of procedure (initial fee)**
Transparency and fairness	- **Structure of the deciding panel (single member or more than one panel)** - **Selection of the members of the deciding panel (by parties; by internal allocation based on alphabetical order, least number of pending cases, language of both parties, area of expertise)** - **Case management (information included in each case; people able to access case files; order of appearance of cases in the list of cases)** - **Classification of cases (differentiation based on type of defendant; differentiation based on type of claim)** - **Possibility to merge cases (based on predetermined conditions; based on decisions by the parties)** - **Access to case files by the deciding panel and parties (all available documents; only to documents not marked as internal notes by a party)** - **Availability of online hearing (open to the public, on restricted access)** - **Possibility to opt out from the procedure**

35 These features can provide a starting point for the evaluation by a certification body identified in the DSA, but again the rules defined in the proposed legislation are not well detailed and disregard other important issues, including security of the dispute settlement platform,[47] the guarantees in case of use of artificial intelligence tools,[48] etc. Moreover, the listed criteria leave too much room for national adaptation which may run contrary to the objective of a fair and harmonised level of protection of users' rights in out-of-court dispute resolution mechanisms.

II. Certification of out-of-court dispute settlement bodies

36 According to Art. 21(3) DSA, in order to receive certification the existence of the list of elements that address the due process guarantees should be evaluated by the Digital Services Coordinator of the Member State where the out-of-court dispute settlement body is established.[49] The DSA sets up a certification mechanism that is fully allocated to the national authority designated by the Member States for consistent application of the DSA. However, the provisions in the DSA only indicate that Digital Service Coordinators should appraise each and every (general) requirement provided in Art. 21(3) and then the Commission of the list of certified bodies.[50] Regardless of the specificities that emerge from the fact that the out-of-court dispute settlement bodies address conflicts that may concern the freedom of expression of users, the certification process should be well-defined in order to ensure that users of the certified out-of-court dispute settlement can rely on evaluation given by the certification body.

37 It is evident that the certification mechanism described in the DSA lacks any additional specification in terms of the definition of applicable standards, the type of evaluation, the geographical scope of the certification scheme and the duration of the certification appraisal. This is a lost opportunity which cannot be justified by lack of knowledge or expertise, as in many other legislative interventions the Commission has engaged in a more structured description of the certification mechanism.

38 In EU law, there are several areas where certification mechanisms were adopted and have flourished. The Commission has fostered the use of certification in

47 Fahimeh Abedi, John Zeleznikow, Chris Brien, 'Developing regulatory standards for the concept of security in online dispute resolution systems', *Computer law & security review* 35 (2019) 1.

48 Hibah Alessa, 'The role of Artificial Intelligence in Online Dispute Resolution: A brief and critical overview', *Information & Communications Technology Law*, 31:3, (2022) 319.

49 The Digital Service Coordinator is defined in Article 49 DSA as the national competent authority in charge of verifying the application and enforcement of the DSA in each Member State.

50 Art. 21(8) DSA provides that "Digital Services Coordinators shall notify to the Commission the out-of-court dispute settlement bodies that they have certified in accordance with paragraph 3, including where applicable the specifications referred to in the second subparagraph of that paragraph, as well as the out-of-court dispute settlement bodies the certification of which they have revoked. The Commission shall publish a list of those bodies, including those specifications, on a dedicated website that is easily accessible, and keep it up to date."

the digital market by adopting the Cybersecurity Act (CSA)[51] and has also included certification schemes in the General Data Protection Regulation.[52] More recently, the Proposal for an Artificial Intelligence Act[53] describes the structure for certification of artificial intelligence systems. In all these cases the Commission has defined in a more or less elaborate manner[54] a certification procedure involving several actors and stakeholders that contribute to the definition of the standards adopted, and a detailed structure of actors in charge of accreditation and certification. The most developed is the certification mechanism defined in the CSA, which provides a clear preparatory phase for the definition of the standard and an equally detailed guidelines for the evaluation of the compliance with the standards. Accordingly, it would be the most suitable point of comparison to achieve not only consistency, but also reliability of the certification mechanism itself. In the following, the DSA procedure will be compared with the more detailed procedure described for cybersecurity certification.

39 In general terms, a certification scheme should involve at least two phases, a conformity assessment and an attestation of conformity, the latter being a statement that the underlying process, product or person complies with a set of pre-defined requirements that are identified on the basis of the objectives and reach of each certification scheme.

40 The certification scheme of the CSA is defined in a centralised process started by the European Commission involving both the ENISA and relevant stakeholders in the field which aims to achieve the most updated level of information security.[55] The result of this process is the adoption by the Commission of the certification scheme, which may include different assurance levels (basic, substantial or high)[56] that take into account the resilience of the ICT product, process or service in the face of potential security threats based on either past experience or potential vulnerabilities.

41 This level of detail regarding the procedure is absent in the DSA. Although the criteria for the certification scheme are already listed in the DSA, as explained in the previous section, there are several sub-criteria that the Digital Service Coordinators may identify in order to operationalise each item in the list provided. The different approaches that may emerge at the national level may run the risk of different safeguards being provided to the users of the out-of-court dispute settlement bodies.

42 Another step in the procedure is the conformity assessment. Art. 58 CSA requires each Member State to designate one (or more) financially and institutionally independent authorities to oversee the enforcing of rules included in European cybersecurity certification schemes and monitoring the compliance of ICT products, services and processes with the requirements of the European cybersecurity certificates. Accordingly, the certification authorities enjoy both investigative and enforcement powers allowing them to carry out investigations (i.e. audits) of conformity assessment bodies, European cybersecurity certificate holders and issuers of EU statements of conformity to verify their compliance,[57] and in cases of infringement to impose penalties in accordance with national law.[58]

51 Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act), OJ L 151, 7.6.2019.

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

53 Proposal for a Regulation of The European Parliament and of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM/2021/206 final.

54 Federica Casarosa 'Cybersecurity certification of Artificial Intelligence: a missed opportunity to coordinate between the Artificial Intelligence Act and the Cybersecurity Act,' *Int. Cybersecur. Law Rev.* (2022).

55 The process is qualified as a centralised one as on the basis of the Union rolling work programme the European Commission defines the strategic priorities for cybersecurity certification schemes. Accordingly, the Commission requests ENISA to draft a candidate scheme, setting very specific goals and requirements such as the subject matter and scope of the scheme, the types or categories of ICT products, systems and services covered, the purpose of the scheme and references to technical standards and specifications etc. These requirements are to be strictly followed by ENISA in order to ensure coherence and uniformity of the certification scheme structure, taking into account differences that may clearly emerge depending on the scope, sector and context of each scheme. See Article 47-49 CSA. For a more detailed analysis of the CSA certification scheme see Casarosa (n 54).

56 See Arts. 52 (6) and (7) CSA.

57 See Art. 58 (8) (b) Cybersecurity Act.

58 See Art. 58 (8) (f) Cybersecurity Act.

43 No similar provision is included in the DSA. On the one hand, the legislation relies on the resources and expertise of the Digital Services Coordinator at the national level: in this sense Art. 39 DSA provides a safety net. The article acknowledges that the Digital Services Coordinators should be independent in carrying out their tasks, and that the Member States should ensure that they have adequate technical, financial and human resources. On the other hand, nothing is stated about the powers of the body regarding evaluation of the certification schemes. There is no help in Art. 41 DSA on the powers of the Digital Services Coordinators, as it lists the powers of supervision and enforcement of the rules applicable to online platforms. No mention is made regarding supervision, investigation and sanctioning powers vis-à-vis out-of-court dispute settlement bodies.

44 Another interesting element is the validity of CSA-based certifications, which may vary but should never exceed the maximum of four years and requires a periodic review of the certification schemes adopted. Moreover, the certification scheme has a geographical scope that covers all EU Member States. This coverage is important not only as certification is recognised in any EU country market where the producer, manufacturer or service provider sells its product, process or service, but it also implies that certification can be obtained in any EU country regardless of the physical location of the requesting company.

45 The same cannot be said for the certification mechanism envisaged in the DSA. The out-of-court dispute settlement body can only be certified in the country where it is established. Although Art. 18(2) DSA acknowledges that the out-of-court dispute settlement body can provide its services in other EU languages, it is not expressly mentioned if the certification is to be recognised in other countries too. This is an evident lack of foresight as the attestation of conformity provided by the certifying body should allow services to be provided across Europe. It will be difficult for an out-of-court dispute settlement body to only provide its services on a country basis. Instead, it will aim to specialise in disputes emerging in some type of platforms (e.g. social networks or C2C marketplaces) in order to provide the service to any user regardless of nationality.

46 The certification mechanism applied to out-of-court dispute settlement bodies would create the conditions for offering transparency and increasing trust in the certified organisation, thus reducing the risks of fragmentation and differentiation in the standards applicable. This would be beneficial both for online platforms and for users. Given the complete absence of guidance regarding this certification process in the DSA, it will be up to each national legislator to fill the gaps in the EU legislation so as to create *ad hoc* procedures and more detailed standards.

D. Conclusion

47 The increasing relevance of online platform activities in users' lives has relevant consequences for the ability of online platforms to gather information about preferences, opinions, and, more generally, knowledge about us. In fact, every platform can screen and potentially filter what is disseminated online by users both *ex ante* and *ex post*. This content moderation activity is increasingly reliant on algorithms and artificial intelligence systems. However, these tools are not fool proof. There are many studies that analyse if, when and to what extent these tools make mistakes, as the subsequent effect is removal or disabling of online content.[59]

48 Of course, mistakes can occur. However, procedures that allow users to contest decisions of the online platform should be available. Internal complaint handling mechanisms are slowly emerging, but another promising alternative is out-of-court dispute settlement mechanisms that can be in charge of resolving disputes on content removal, suspension of service and account termination. The proposed legislation in the Digital Services Act shares this position and provides in its Art. 21 that users of online platforms shall be entitled to resolve the abovementioned types of disputes also through certified bodies providing their services in the EU.

49 The DSA provision not only pushes towards the creation of such out-of-court dispute settlement bodies but it also requires them to ensure due process guarantees, which are listed as the main criteria for certification of them. Although this is a commendable effort by the EU bodies to safeguard the position of users vis-à-vis the increasing power of online platforms, the provisions in the DSA run short of useful guidelines, which may hamper achievement of the objectives sought.

50 On the one hand, the list of criteria in Art. 21 is far from being immediately applicable and will require an effort by Digital Services Coordinators at the national level to operationalise the general elements into more practical features. Can each Digital Services Coordinator define its own criteria at the national level? It is more than probable that this issue will require coordination at the European

59 Article 19, 'The Social Media Councils: Consultation Paper' (2019) <https://www.article19.org/wp-content/uploads/2019/06/A19-SMC-Consultation-paper-2019-v05.pdf>; Stuart Benjamin 'Algorithms and Speech' 161 *University of Pennsylvania Law Review*, 1445-1493 (2013)

level. Otherwise, the harmonisation objective would be jeopardised.

51 Moreover, the certification mechanism provided in Art. 21 may also be qualified as a raw structure as the guidelines addressing the definition of applicable standards, the type of evaluation, the geographical scope of the certification scheme and the duration of the certification appraisal are very limited.

52 This seems to be a lost opportunity as out-of-court dispute settlement mechanisms will probably flourish as they are not only present in many recent legislative acts but they will also probably emerge more and more as an alternative way to resolve cases of user dissatisfaction.[60] In this context, certification may provide a very useful signal to users regarding due process guarantees and safeguard their position vis-à-vis platforms. Moreover, it is possible that users engaging in a copyright dispute may recognise among the out-of-court dispute settlement bodies one or more that have been certified according to the DSA procedure. In this case, certification may become an added value and steer the choice of users towards this provider.

53 What if the DSA certification mechanism (if improved and structured in a clearer way) also becomes the standard for bodies providing their services in other legal areas? Of course, this is clearly a step that will require further legislative interventions, but it may be possible that the path set by the DSA will lead in this direction.

60 Civil Justice Council's Online Dispute Resolution Advisory Group 'Online Dispute Resolution for Low Value Civil Claims' (2015) <https://www.judiciary.uk/wp-content/uploads/2015/02/Online-Dispute-Resolution-Final-Web-Version.pdf>.

Towards a better notice and action mechanism in the DSA

by **Pieter Wolters and Raphaël Gellert** *

Abstract: The adoption of the DSA will bring important changes in the content moderation landscape in the EU. By harmonising, codifying, and further developing a notice and action mechanism, the DSA addresses many content moderation-related challenges, and in so doing also affects the balance that existed thus far between the protection of victims of illegal content, the safeguarding of fundamental rights and the economic interests of hosting service providers. This contribution answers the following question: Does the notice and action mechanism of the DSA create an adequate balance between the various involved interests?

As far as the economic interests of hosting service providers are concerned, the harmonisation of the mechanism should certainly be a welcome change for economic operators. Further, even though the DSA contains many new procedural obligations, they entail reasonable efforts.

The requirement of a harmonised, efficient, effective and user-friendly notification procedures should fix the existing limitations and can be seen as an important step for the protection of the victims' interests. However, the lack of an obligation to provide a statement of reasons to notifiers is a missed opportunity.

Finally, the safeguards of content providers' fundamental rights are also enhanced. Not only through the creation of new redress mechanisms, but also through the hosting provider services' obligation to provide decisions that are objective, non-arbitrary, diligent and timely, and to justify them through a statement of reasons. Although the applicability of the safeguards is still too narrow in some respects, the new safeguards and their requirements should improve the current situation in which hardly any binding legal provisions exist.

All in all, even though it contains various shortcomings that prevent it from truly striking an adequate balance, the DSA's notice and action mechanism does represent a significant step forward for all the parties that have a stake in the moderation of online content.

Keywords: Digital Services Act; Notice-and-Action; Freedom of Expression; Illegal Content; Online platforms Notice-and-action

Recommended citation: Pieter Wolters and Raphaël Gellert, Towards a better notice and action mechanism in the DSA, 14 (2023) JIPITEC 403 para 1.

A. Introduction

1 On 15 December 2020, the European Commission published its long-awaited proposal for the Digital Services Act ('DSA proposal').[1] More than twenty years after the adoption of the e-Commerce Directive,[2] the DSA revised the European framework

* Pieter Wolters is an associate professor at the Radboud University and the Radboud Business Law Institute. Raphaël Gellert is an assistant professor at the Radboud University and the Radboud Business Law Institute. Both are affiliated to Radboud University's interdisciplinary hub on digitalization and society (iHub).

1 Commission, 'Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC' COM (2020) 828 final.

2 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information

for the liability and responsibilities of 'intermediary services'.[3] Following the 'General approach'[4] of the Council and 'Amendments'[5] by the European Parliament, the DSA was adopted on 19 October 2022.[6]

2 The core of the framework remains the same. Like the e-Commerce Directive, the DSA holds that these providers cannot be held liable for transmitting or storing information that is provided by the 'recipients of the service' (the 'content providers').[7] However, 'hosting' service providers can be held liable if they know about the illegal content and do not act expeditiously to remove or disable access to the information.[8]

3 At the same time, the DSA also introduces new obligations for the providers of intermediary services.[9] Notably, 'online platforms' and other providers of 'hosting' services have a duty to take reactive steps against 'illegal content'.[10] Article 16 of the DSA obligates providers to put a notice and action mechanism in place, allowing anyone to notify them of hosted illegal content. The hosting service providers must subsequently remove or disable access to the illegal content or face liability.[11] Although such a mechanism is already used by many online platforms and imposed by various specific European rules, national laws and codes of conduct (Section 3), the DSA harmonises, codifies and develops the existing practices and rules. It imposes a notice and action mechanism that applies to all hosting services[12] and for all types of illegal content. Furthermore, the DSA develops the mechanism by providing detailed rules and safeguards, including a statement of reasons (Article 17) and redress mechanisms (an internal complaint-handling system and a system for out-of-court dispute settlements).[13] For the purpose of this article, these safeguards are considered an integral part of the notice and action mechanism.

4 In accordance with the aims of the DSA, the notice and action mechanism is designed to strike a proper balance between the various competing interests.[14] In this article, we analyse how the mechanism has considered the various interests and whether it has succeeded in creating a proper balance. We answer the following question: *Does the notice and action mechanism of the DSA create an adequate balance between the various involved interests?*

5 We consider the balance as 'adequate' if the DSA addresses the limitations of the current legal framework (see Section 4) and creates a framework that leads to a proper balance of the various involved interests. The notice and action mechanism should strengthen *both* the protection of society and individual victims against illegal content *and* the involved fundamental rights without disproportionally affecting the economic interests of hosting service providers. As a minimum requirement, Article 52 of the Charter of Fundamental Rights of the European Union should be respected. Any limitation of a fundamental right should be proportional and respect the essence of this right. Within the bandwidth that the Charter provides, the heterogeneity of the various involved interests makes the determination of the 'best' way

society services, in particular electronic commerce, in the Internal Market [2000] OJ L178/1.

3 DSA, arts 1(2)(a), 3(g).

4 Council of the European Union, 'Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC - General approach' 13203/21.

5 European Parliament, 'Digital Services Act: Amendments adopted by the European Parliament on 20 January 2022 on the proposal for a regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC' P9_TA(2022)0014.

6 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ L277/1.

7 DSA, arts 3(b), 4-6; e-Commerce Directive, arts 12-14.

8 DSA, arts 3(g), 6(1)(a), (b); e-Commerce Directive, art 14(1)(a), (b). See DSA, arts 3(g), 4-5; e-Commerce Directive, arts 12-13 about the conditions under which providers of 'mere conduit' and 'caching' services can be held liable.

9 See also DSA, art 1(2)(b).

10 DSA, art 3(h), (i); n 15.

11 DSA, arts 6(1)(b), 16(3).

12 However, some exclusions exist for micro and small enterprises. DSA, art 19.

13 DSA, arts 20, 21.

14 DSA, recital 52. About the aims of the DSA in general, see also DSA, recitals 3, 4, 40, art 1(1). The e-Commerce Directive has the same goals. See eg Case C-360/10 *SABAM* [2012] ECLI:EU:C:2012:85, para 51; e-Commerce Directive, recital 41; Giancarlo Frosio and Sunimal Mendis, 'Monitoring and Filtering: European Reform or Global Trend' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (OUP 2020), 563.

to balance them subjective. However, it is possible to formulate general requirements that should be fulfilled. First, the protection of one interest (such as the protection of victims through the removal of illegal content) should not disproportionally affect other involved interests (such as the freedom of information). Furthermore, the balance is not adequate if the protection of one of the interests can be improved with either no or only minimal adverse effects to the other involved interests.

6 The article is structured as follows. Section B provides a short description of the most important involved interests. Next, we give a birds-eye overview of the current practices and rules (Section C) and their limitations (Section D). Section E discusses the notice and action mechanism in the DSA. It analyses how the mechanism has considered the various interests and whether it leads to an adequate balance. Section 6 provides a conclusion: Although the notice and action mechanism in the DSA is a significant step forward, it contains various shortcomings that prevent it from truly striking an adequate balance.

7 Importantly, this article is focussed on the notice and action mechanism of the DSA. For this reason, it is necessary to at least tentatively accept some of the propositions underlying the adoption of such a mechanism. Most importantly, it is necessary to tentatively accept that it has the potential to limit the dissemination of illegal content without unduly affecting other concerned interests. The conditions that are necessary for this result are discussed in Section B. Furthermore, the article does not discuss the role of other forms of 'content moderation' such as (voluntary) proactive monitoring and the role of 'trusted flaggers'.[15] Furthermore, we do not discuss provisions that are relevant but not directly part of the notice and action mechanism such as reporting and transparency obligations.[16]

B. The involved interests

8 A proper balance between the various interests can only be achieved through an adequate understanding of what these interests involve. For this reason, this Section gives an overview of the most important interests in relation to notice and action mechanisms. It subsequently discusses the protection of the victims of illegal content (Section B. I.), the fundamental rights of the recipients of the hosting services (Section B. II.) and the economic interests of hosting services (Section B. III.).

I. Protecting victims of illegal content

9 'Illegal content' has a broad definition. Pursuant to Article 2(h) of the DSA, it includes "any information, which, in itself or in relation to an activity, including the sale of products or the provision of services, is not in compliance with Union law or the law of any Member State, irrespective of the precise subject matter or nature of that law". It includes information that is illegal in itself, such as terrorist content, illegal hate speech or child pornography, but also information that relates to illegal activities such as online stalking, the sale of counterfeit goods, non-authorised use of copyright protected material or infringements of consumer law.[17]

10 Due to this broad definition, the 'victims' of illegal content also come in all shapes and forms. A victim can be a defrauded consumer, a child who is depicted in pornography, a rights holder whose content is disseminated without permission or a recipient that is exposed to shocking or otherwise inappropriate content. The victim can be a recipient of the hosting service, but this is not necessary. For example, a victim of hate speech may be targeted by reactions on his or her pictures on social media, but also by posts on a forum of which they are not a user. Finally, illegal content such as terrorist content is not always aimed at individual victims. It (also) threatens society as a whole. Because of the differences between these victims, their interests and needs may be different. However, there are also strong similarities.

11 First, the notice and action mechanism should be effective. It should lead to the speedy removal of the illegal content.[18] The longer the illegal content stays up, the more harm it can cause.[19] Furthermore, a successful notification should also provide some future protection. This can be directly achieved by preventing the illegal content from being

15 e-Commerce Directive, art 15; DSA, arts 3(t), 7-8, 22.

16 Eg DSA, arts 14, 15, 24, 42.

17 DSA, recital 12.

18 Some victims may not always be interested in removal. For example, rights holders may prefer monetization of the infringing content. Cf Annemarie Bridy, 'The Price of Closing the Value Gap: How the Music Industry Hacked EU Copyright Reform' (2020) 22 *Vand J of Ent & Tech L* 323, 330-331; Henning Grosse ruse-Khan, 'Transition through automation' in Niklas Bruun and others (ed), *Transition and coherence in Intellectual Property Law* (Cambridge University Press 2021) 160.

19 Eg, Joris van Hoboken and others, *WODC-onderzoek: Voorziening voor verzoeken tot snelle verwijdering van onrechtmatige online content* (IViR 2020) 73.

reuploaded,[20] but also indirectly by suspending or otherwise punishing the content providers. The notifier should be informed about the decision concerning the notified content.[21]

12 The victim may also benefit from further redress such as a right to damages from the content provider. A notice and action mechanism can be used to facilitate this right. Although a successful notification typically leads to the removal of the illegal content, a notifier could also require other actions from the hosting service provider such as the provision of information about the identity of the content provider. An order to share information about content providers of illegal content can already be obtained in some jurisdictions.[22] However, such an order is only useful if the hosting service providers know the identity of the content providers with some degree of certainty. This is not the case with most online platforms[23] and other hosting service providers.[24] Furthermore, the DSA does not impose a general[25] 'know-your-customer-obligation' on hosting service providers. For these reasons, a duty to share information about the content providers is not part of the notice and action mechanism in the DSA. It is also not discussed further in this article.

13 Second, the notice and action mechanism should be efficient. Submitting a notification should be free, accessible, fast and user-friendly.[26] Submitting a notification should not have any negative consequences. For example, a notifier should not have to fear retaliation from the user that uploaded the content. This can be achieved by allowing anonymous notifications.[27]

14 In the end, the attractiveness of the notice and action mechanism depends on its costs and benefits. Victims are less likely to submit notifications if it takes a long time and seldomly leads to speedy removal. In contrast, they will submit more notifications if it can be done with a few clicks and the illegal content is actually removed within a short timeframe.

II. Safeguarding fundamental rights

15 The protection of victims by the removal of online content comes at the expense of the freedom of expression and freedom of information of the recipients of the services. However, the limitation of these freedoms is not necessarily undesirable. Generally speaking, the fact that the content is illegal can justify a limitation of these rights. There is no fundamental reason to protect the online dissemination of such content through online intermediaries.[28] Furthermore, the illegal content may also affect fundamental rights. For example, child sexual abuse material affects the fundamental rights of children protected in Article 24 of the

20 A notice and *stay* down mechanism. See Section 3.

21 Commission Recommendation (EU) 2018/334 of 1 March 2018 on measures to effectively tackle illegal content online [2018] OJ L 63/50, point 8.

22 Cf HR 25 November 2005, ECLI:NL:HR:2005:AU4019 (*Lycos/Pessers*) (for providers that host websites, in the Netherlands); Case C-275/06 *Promusicae* [2008] ECLI:EU:C:2008:54 (for internet service providers, in Spain); Rb. Amsterdam (vzr.) 25 June 2015, ECLI:NL:RBAMS:2015:3984 (for Facebook, in the Netherlands); n 24.

23 Although Facebook has a 'real name' requirement, it is possible to use a pseudonym. About this requirement on Facebook and other online platforms, see eg Shun-Ling Chen, 'What's in a name? – Facebook's real name policy and user privacy' (2018) 28 *Kan J L & Pub Pol'y* 146.

24 For examples, see <https://www.kybc.eu/case-studies-research/>.

25 DSA, art 30 only contains a know-your-customer obligation in relation to (professional) traders for platforms that allow consumers to conclude distance contracts with traders.

26 Eg Recommendation on measures to effectively tackle illegal content online (n 21), point 5; Alexandre de Streel and others, *Online Platforms' Moderation of Illegal Content Online. Law, Practices and Options for Reform* (Study for the European Parliament PE 652.718, 2020) 40, 49, 69, 79.

27 Recommendation on measures to effectively tackle illegal content online (n 21), point 7; De Streel and others (n 26) 51.

28 Commission, 'Tackling Illegal Content Online. Towards an enhanced responsibility of online platforms' (Communication) COM (2017) 555 final, 2; European Parliament, 'Digital Services Act: Improving the functioning of the Single Market' (Resolution) P9_TA(2020)0272, point 6; Niva Elkin-Koren and Maayan Perel, 'Guarding the Guardians: Content Moderation by Online Intermediaries and the Rule of Law' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (OUP 2020) 671. Cf the concept of 'internet exceptionalism': some authors put more emphasis on the 'free' character of 'cyberspace', even at the expense of other legally protected interests. For example, see John P. Barlow, 'A Declaration of the Independence of Cyberspace' (Electronic Frontier Foundation, 8 February 1996) <https://www.eff.org/cyberspace-independence> accessed 9 September 2021; Dan Jerker B. Svantesson, 'Internet Jurisdiction and Intermediary Liability' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (OUP 2020) 692-693. See also P.T.J. Wolters, 'Search Engines, Digitalization and National Private Law' [2020] *ERPL* 795, 799 for more examples of internet exceptionalism in law.

Charter. In contrast, this section primarily focusses on the fundamental rights of *other* users.

16 Content moderation is often imprecise and it can lead to over-removal. This risk is especially relevant if the permissibility of certain content is unclear. For example, the line between an infringement of copyright and a permissible parody or between unfounded slander and legitimate critical journalism will not always be clear.[29] Furthermore, for many types of illegal content, this line may be drawn differently in each member state.[30] In these circumstances, a hosting service provider may be induced to err on the side of caution. For them, the direct legal risk of liability for permitting content that is ultimately deemed illegal outweighs the indirect[31] adverse effects of removing lawful content.[32] This leads to a limitation of the freedom of expression and freedom of information.

17 The removal of online content can also affect other fundamental rights and lead to discrimination. Content moderation may disproportionally affect certain groups. For example, a conservative country's hostile stance against LGBTQ-content may cause it to be removed due to incorrect or abusive notices, even when it is not illegal.[33] Furthermore, certain types of over-removal may be more damaging to the society as a whole. For example, the removal of news also affects the freedom of the press.[34]

18 The fundamental rights can be protected by only removing online content that is undoubtedly or 'manifestly' (see Section 3.1) illegal.[35] Furthermore, a notice and action mechanism (and content moderation in general) should include safeguards to prevent the removal of permissible content.[36] This does not mean that content moderation should never go beyond the removal of manifestly illegal online content. Different platforms with different content moderation practices can cater to different people and different types of content. For example, removing legally permissible insults may stimulate other recipients to express themselves more freely and thus facilitate freedom of information and freedom of expression.[37] At the same time, these moderation practices should be non-discriminatory, transparent, well-balanced, and consistently applied.[38]

19 By submitting a notification, the notifier forces the hosting service provider to judge whether the

29 Thibault Verbiest and others, *Study on the liability of internet intermediaries* (2007) 14-15; Georgios N. Yannopoulos, 'The Immunity of Internet Intermediaries Reconsidered?' in Mariarosaria Taddeo and Luciano Floridi, *The Responsibilities of Online Service Providers* (Springer 2017) 50; Christophe Geiger, Giancarlo Frosio and Elena Izyumenko, 'Intermediary Liability and Fundamental Rights' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (OUP 2020) 146; Tarlach McGonagle, 'Free Expression and Internet Intermediaries: The Changing Geometry of European Regulation' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (OUP 2020) 483; Maria Lillà Montagnani, 'A New Liability Regime for Illegal Content in the Digital Single Market Strategy' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (OUP 2020) 304; De Streel and others (n 26) 40, 43, 52.

30 Verbiest and others (n 29) 14-15; Montagnani (n 29) 304; De Streel and others (n 26) 40, 51, 56-57, 61.

31 Strict content moderation may affect the popularity of a service. Cf De Streel and others (n 26) 44. The terms and conditions of the hosting services are generally stricter than the law and also prohibit certain kinds of undesirable content that is not (always) illegal. For this reason, the removal of such content does not constitute a breach of contract towards the recipients. Eg Verbiest and others (n 29) 16; Yannopoulos (n 29) 50; Giancarlo Frosio, 'Mapping Online Intermediary Liability' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (OUP 2020) 26; De Streel and others (n 26) 10, 14, 40, 43, 61.

32 Eg Joined Cases C-682/18 and C-683/18 *Youtube* [2020] ECLI:EU:C:2020:586, Opinion of AG Saugmandsgaard Øe, para 189; European Commission, 'Impact assessment Accompanying the Proposal for a Regulation of the European Parliament and off the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC PART 1/2' SWD (2020) 348 final, box 1; Frosio (n 31) 26; Aleksandra Kuczerawy, 'From 'Notice and Takedown' to 'Notice and Stay Down': Risks and Safeguards for Freedom of Expression' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (OUP 2020) 527; McGonagle (n 29) 483; De Streel and others (n 26) 23.

33 DSA, recital 81; Alex Hern, 'TikTok's local moderation guidelines ban pro-LGBT content' (*The Guardian* 26 September 2019) <www.theguardian.com/technology/2019/sep/26/tiktoks-local-moderation-guidelines-ban-pro-lgbt-content> accessed 13 September 2021.

34 Cf Geiger, Frosio and Izyumenko (n 29) 143; De Streel and others (n 26) 83.

35 Geiger, Frosio and Izyumenko (n 29) 140-147; De Streel and others (n 26) 77.

36 Commission, 'Tackling Illegal Content Online' (n 28) 3, 20; Commission, 'Impact Assessment Digital Services Act' (n 32), points 51-52, 54.

37 Cf Commission, 'Impact Assessment Digital Services Act' (n 32), point 62.

38 Commission, 'Impact Assessment Digital Services Act' (n 32), points 54, 57.

content is permissible. Although it is broadly argued that hosting service providers should not be the ones to make these complex decisions, or to determine the balance between the protection of victims and fundamental rights and become the judges of online legality,[39] the fact that they carry the *responsibility* to separate illegal and permissible content after receiving a notification is not necessarily undesirable. After all, their services also facilitate the dissemination of illegal content.

20 At the same time, the ultimate *power* to make the distinction should not lie with the hosting service providers: it should lie with judges. In theory, both victims and content providers can go to a court when they disagree with a decision to (not) remove certain content.[40] In practice, this opportunity is used infrequently. The costs and efforts generally outweigh the benefits.[41]

21 This issue is exacerbated by the influence of hosting services, and online platforms in particular. Depending on the message or type of online content, platforms can become so ubiquitous that their services are the only way to effectively disseminate information.[42] In these situations, the platforms become the *de facto* judges about the permissibility of online content.[43]

22 Because of the *de facto* influence and responsibility of the hosting services, the dispute resolution in relation to content moderation affects the fundamental right to a fair trial of both the victim and the content provider. Although not every form of content moderation can, or should, be the same as a court proceeding, accessible forms of alternative dispute resolution should exist, be fair and have adequate safeguards. Finally, judicial oversight can be reinforced through transparency. When a notice and action mechanism is used, the hosting service provider should provide clear reasons for its decisions to both the notifier and (when an action is taken) the content providers. This allows both the involved parties and the courts to understand and critically assess the decisions. Furthermore, the costs and efforts to go to court should not be too high.

III. The economic interests of hosting service providers

23 The liability and responsibilities of hosting service providers come at the expense of their economic viability and their fundamental right to freedom of business.[44] Hosting service providers play an important role in the development of our information society. They facilitate freedom of expression and information, effective communication and the development of all kinds of economic activities.[45] The costs of liability and responsibilities can negatively impact their development and availability and (consequently) the development of the internet and the information society. They could cause the providers to abandon or limit their services or start charging a (higher) price.

24 The economic interests of the hosting service providers can be protected by only imposing a notice and action mechanism and certain requirements or safeguards when they are proportional and can be fulfilled at a reasonable cost.[46] Furthermore, the responsibilities should be clear, harmonised, consistently applied and technology-neutral.[47]

39 Eg Joined Cases C-682/18 and C-683/18 *Youtube* [2020] ECLI:EU:C:2020:586, Opinion of AG Saugmandsgaard Øe, para 187; Sophie Stalla-Bourdillon, 'Internet Intermediaries as Responsible Actors? Why It Is Time to Rethink the E-Commerce Directive as Well' in Mariarosaria Taddeo and Luciano Floridi (eds), *the Responsibilities of Online Service Providers* (Springer 2017), 290; Frosio (n 31) 26; Kuczerawy (n 32) 527; De Streel and others (n 26) 45; Svantesson (n 28) 693.

40 Cf e-Commerce Directive, art 18.

41 Eg Tim F. Walree and Pieter T.J. Wolters, 'The right to compensation of a competitor for a violation of the GDPR' (2020) 10 IDPL 346, 351, with references to further literature.

42 About this issue, see eg Commission, 'Impact Assessment Digital Services Act' (n 32), points 85-86; Yannopoulos (n 29) 46, 53-56; Geiger, Frosio and Izyumenko (n 29) 138-139; Kuczerawy (n 32) 527; McGonagle (n 29) 479-480; De Streel and others (n 26) 80-81; Mariarosaria Taddeo, 'The Civic Role of OSPs in Mature Information Societies' in Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (OUP 2020), 134-136.

43 N 39, 42.

44 About this right, see DSA, recital 52; Geiger, Frosio and Izyumenko (n 29) 148-149.

45 DSA, recital 1; Commission, 'Online Platforms and the Digital Single Market. Opportunities and Challenges for Europe' (Communication) COM (2016) 288 final, 2-3; Commission, 'Tackling Illegal Content Online' (n 28) 2.

46 DSA proposal 8, 13; DSA, recital 4; European Parliament, 'Improving the single market' (n 28), point 10.

47 DSA, recital 4; Commission, 'Online Platforms and the Digital Single Market' (n 45) 4; Commission, 'Impact Assessment Digital Services Act' (n 32), points 70-71, 75-76; European Parliament, 'Improving the single market' (n 28), points 10, 12, 14.

C. Notice and action in current law and practice

I. Notice and action in the e-Commerce Directive and national law

25 Pursuant to Article 14 of the e-Commerce Directive, a hosting service provider can be held liable if it has actual knowledge of the illegal content and does not act expeditiously to remove or disable access to the information. However, the e-Commerce Directive does not clarify how the actual knowledge is supposed to be obtained. Although a notice and action mechanism is an important tool for gaining this knowledge, the e-Commerce Directive does not impose an obligation to facilitate or respond to notifications.[48]

26 This obligation is imposed for specific situations by other European rules (Section C. II.), but also generally by various (but not all) national laws and codes of conduct. The details of these national obligations vary from member state to member state.[49] For example, some member states place formal requirements on the notifications, only obligating hosting service providers to remove content when the notification contains certain information and/or is made by a competent authority.[50]

27 A notification can only lead actual knowledge if it sufficiently specific. Unless the hosting service is specifically designed to facilitate the dissemination of illegal content,[51] a provider cannot be held liable for abstract knowledge that its service may be used for this purpose.[52] For this reason, a notification should contain a link to the illegal content.[53] In practice, this is usually facilitated by the notice and action mechanism (Section 3.3). If a single URL refers to a plurality of content, it might be necessary to provide more information. For example, a notifier should include a timestamp if the illegal content is included in a long (and otherwise permissible) video.[54]

28 Furthermore, the notification should trigger knowledge about *the illegal nature* of the content. In most member states, a hosting service provider can only be held liable if the illegal nature is sufficiently clear or 'manifest'.[55] This approach prevents over-removal (Section B. II.), but also causes more illegal content to stay available. Furthermore, it allows hosting service providers to escape or delay their responsibilities by claiming that the illegality of certain content is unclear. In contrast, more recent provisions such as § 3(2) of the German NetzDG impose an obligation to remove any illegal content.[56] In any case, the notification should be adequately substantiated and provide information about why the content is illegal. It should allow a diligent

48 Commission, 'Impact Assessment Digital Services Act' (n 32), point 91.

49 Commission, 'Impact Assessment Digital Services Act' (n 32), points 93-99; Verbiest and others (n 29) 41-47; Kuczerawy (n 32) 530.

50 Verbiest and others (n 29) 14-15, 36, 42-46. See also Stalla-Bourdillon (n 39) 291. This requirement can also depend on the type of liability. In the Netherlands, criminal liability is only possible when a hosting service provider ignores an order from a public prosecutor, while private law liability may also be imposed when the actual knowledge or awareness is acquired through another channel. Dutch Criminal code, art 54a; Dutch Civil code, art 6:196c.

51 *Piratebay* B 13301-06 (Stockholms tingsrätt 2009); Joined Cases C-682/18 and C-683/18 *Youtube* [2020] ECLI:EU:C:2020:586, Opinion of AG Saugmandsgaard Øe, para 191; Joris van Hoboken and others, *Hosting intermediary services and illegal content online. An analysis of the scope of article 14 ECD in light of developments in the online service landscape* (Study for the European Commission, 2018), 38-39; Frosio and Mendis (n 14) 552.

52 Case C-324/09, *eBay* [2011] ECLI:EU:C:2011:474, para 122; Joined Cases C-682/18 and C-683/18 *Youtube* [2021] ECLI:EU:C:2021:503, para 111; Verbiest and others (n 29) 37; Van Hoboken and others (n 51) 38.

53 For example, see the French 'Avia law', Avia Law Loi no 2020-766 du 24 juin 2020 visant à lutter contre les contenus haineux sur internet, art 2(I).

54 Cf European Parliament, 'Digital Services Act: adapting commercial and civil law rules for commercial entities operating online' (Resolution) P9_TA(2020)0273, Annex B, art 9(1)(a); Folkert Wilman, *The Responsibility of Online Intermediaries for Illegal User Content in the EU and the US* (Edward Elgar 2020) 301.

55 Joined Cases C-682/18 and C-683/18 *Youtube* [2020] ECLI:EU:C:2020:586, Opinion of AG Saugmandsgaard Øe, para 187, 190; Verbiest and others (n 29) 38-41; Stalla-Bourdillon (n 39) 290.

56 Netzwerkdurchsetzungsgesetz vom 1. September 2017 (BGBl. I S. 3352), das durch Artikel 274 der Verordnung vom 19. Juni 2020 (BGBl. I S. 1328) geändert worden ist <https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html> accessed 15 September 2021. Note that uncertainty about the illegal nature does affect the period for the analysis of the permissibility. A provider has 24 hours for obviously illegal content, '*offensichtlich rechtswidrigen Inhalt*', and seven days in other situations.

hosting service provider to realise that the content is illegal.[57] Although a notifier cannot be expected to provide a detailed legal clarification of the illegality, the notification should at least provide the necessary facts. For example, a notification about a copyright violation should clarify who owns the copyright and that no permission has been given.[58]

II. Notice and action in other European rules

29 The e-Commerce Directive does not provide for a notice and action mechanism. However, due to its increasing popularity, this mechanism has been formally adopted in a number of relevant European instruments. Whereas some of these instrument are binding (Section C. II. 1.), others are not (Section C. II. 2.). This section provides a rapid overview of the most relevant instruments.

1. EU binding instruments

30 Since its 2018 revision, the EU's Audiovisual media services Directive (AVMSD),[59] contains specific obligations for so-called Video Sharing Platform services (VSPs).[60] VSPs must offer transparent and user-friendly mechanisms to allow users to report and flag content,[61] which includes a follow-up explanation on the manner in which the flagging has been internally handled.[62]

31 Next, Article 25 of Directive 2011/93/EU on combatting children sexual abuse and exploitation (CSAED) obligates member states to take measures to remove or block access to websites that contain child sexual abuse material.[63] In order to comply with this obligation, various member states have implemented a notice and take down mechanism. This approach is based upon a network of organisations that serve as hotlines; the best known probably being INHOPE.[64]

32 Further, the recently adopted Copyright in the Digital Single Market Directive (CDSMD) provides for an advanced notice and take down mechanism known as notice and stay down. That is, so-called content-sharing service providers must not only take down illegal content, they must also make sure that such content cannot be re-uploaded after having been removed.[65]

2. EU non-binding instruments

33 In 2018 the European Commission adopted a Recommendation on Measures to Effectively Tackle Illegal Content Online ("Recommendation"). The latter builds upon the European Commission's 2017 Communication on Tackling Illegal Content Online, and is arguably a forerunner of the DSA. It contains a general notice and action mechanism that applies to all types illegal content and all hosting service providers.[66]

34 Several other soft-law instruments contain obligations that are more narrow in (either personal or material) scope. A so-called Code of Conduct on Countering Illegal Hate Speech online in the EU was adopted by some of the main platforms (e.g., Youtube, Microsoft, Facebook, Twitter) in 2016.[67]

57 Case C-324/09, *eBay* [2011] ECLI:EU:C:2011:474, para 122; Joined Cases C-682/18 and C-683/18 *Youtube* [2021] ECLI:EU:C:2021:503, para 115. See also Verbiest and others (n 29) 16; Stalla-Bourdillon (n 39) 291.

58 Cf Joined Cases C-682/18 and C-683/18 *Youtube* [2020] ECLI:EU:C:2020:586, Opinion of AG Saugmandsgaard Øe, para 188-189.

59 Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) (codified version) [2010] OJ L95/1 (herein after AVMSD).

60 AVMSD, art 1(1aa).

61 AVMSD, art 28b(3)(d).

62 AVMSD, art 28b(3)(e).

63 Directive 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA [2011] OJ L335/1 (hereinafter CSAED), art 25(1),(2).

64 European Commission, 'Report from the Commission to the European Parliament and the Council assessing the implementation of the measures referred to in Article 25 of Directive 2011/93/EU of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography' (report) COM (2016) 872 final 7.

65 Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC [2019] OJ L130/92. (CDSMD), art 17(4).

66 See, Recommendation on measures to effectively tackle illegal content online (n 16), point 1 and Chapter II.

67 Hate speech is defined with reference to the EU's 2008

According to this Code, notified content is first assessed on the basis of the applicable terms and conditions and only then on basis of the relevant legal framework.[68]

35 Another soft-law instrument providing for a notice and action mechanism is the so-called Memorandum of Understanding on the sale counterfeit goods, meant to prevent the violation of intellectual property rights in the context of counterfeit goods. It was adopted in 2011 and revised in 2016.[69]

36 In 2018 and with the support of the European Commission, a number of stakeholders adopted the Product Safety Pledge. The goal of this initiative is to be able to better detect products sold online unto the European Market and which do not comply with product safety requirements.[70] A notice and take down mechanism to allow users to flag unsafe products is one of the 12 commitments contained in the Pledge.[71]

III. Notice and action in practice

37 Despite the legal fragmentation and the lack of a general European obligation to put a notice and action mechanism in place, (almost) all major online *platforms*[72] implemented a procedure to facilitate notifications.[73] Typically, these mechanisms allow a user to 'flag' illegal content by clicking a dedicated button and subsequently selecting the reason for the perceived illegality. The effectiveness of the mechanism depends on its accessibility (Section 2.1), which varies from platform to platform.[74] The platforms typically have specific channels or procedures for law enforcement agencies and other privileged or 'trusted' flaggers.

38 Although differences exist between the various platforms and types of illegal content, most platforms react relatively fast. They claim to usually remove terrorist-related content and child pornography with-in one hour and other illegal content within 24 hours of the notification.[75] Although most platforms do allow them to appeal against the removal of their online content through a 'counter-notice' procedure, the content providers are not always notified or given a clear expla-nation about the reasons for the removal.[76]

D. Limitations

39 The current system of notice and action mechanisms is subject to various limitations and flaws, which affect each of the identified interests.

I. Limitation 1: lack of harmonization and preservation of the economic interests

40 A first limitation has been evidenced in Section C. The current framework is fragmented. There is no general European obligation to put a notice and action mechanism in place. The fragmentation is visible at various levels. The scope of the various mechanisms varies according to the member state, reason for the illegality and type of service. Furthermore, the mechanism's content and requirements themselves vary greatly. There is a lack of consistency and uniformity between the various mechanisms.

41 Some have minimum quality requirements concerning the notices, this includes reasons to

Counter-Racism Framework Decision, which refers to: "all conduct publicly inciting to violence or hatred directed against a group of persons or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin", Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law [2008] OJ L328/55, art 1(a).

68 Code of Conduct on Countering Illegal Hate Speech Online (2016) 2.

69 Memorandum of understanding on the sale of counterfeit goods via the internet (2016), para 11-19.

70 Product Safety Pledge: Voluntary Commitment of Online Marketplaces with Respect to the Safety of Non- Food Consumer Products Sold Online by Third Party Sellers (2018) 1.

71 Product Safety Pledge (n 68) 2.

72 Notice and action mechanisms are not common with other types of hosting services.

73 For a general description of these mechanisms, see eg Frosio and Mendis (n 14) 556; De Streel and others (n 26) 40, 46-51; Raphaël Gellert and Pieter Wolters, *The revision of the European framework for the liability and responsibilities of hosting service providers* (Report for the Dutch Ministry of Economic Affairs and Climate Policy, 2021) 27-28.

74 De Streel and others (n 26) 40, 48-49, 51.

75 See also De Streel and others (n 26) 44, 47, 49.

76 De Streel and others (n 26) 50; Gellert and Wolters (n 73) 28.

believe why the content is illegal or where it can be found,[77] others may make more general references to notices submitted in good faith.[78] Some mechanisms allow anonymous notices while others do not.[79] A few insist on the user-friendly nature of the mechanism,[80] while others also contain broader language on the efficient nature of the mechanism.[81] The time in which the service providers have to respond also differs from mechanism to mechanism. Some require action within 24 hours[82] or 5 working days,[83] while others simply refer to the lack of undue delay.[84] Whereas most notice and action mechanisms are strictly speaking notice and take down, some go a step further and are known as notice and stay down and require that the provider prevents the content from being uploaded again.[85]

42 Finally, one can also mention the specific systems of the Dutch Notice and Take Down Code of Conduct and of the NetzDG, which make specific distinctions not found in the other instruments discussed. The NetzDG distinguishes between manifestly and non-manifestly illegal content. The former must be removed within 24 hours and the latter within seven days.[86] The Dutch Code of Conduct differentiates between unequivocally lawful and not unequivocally unlawful content.[87]

77 Recommendation on measures to effectively tackle illegal content online (n 21) points 5-8; AVIA law, art 2; Gedragscode Notice-and-Take-Down 2018 inclusief addendum 1 <https://noticeandtakedowncode.nl/ntd-code/> last accessed 14 July 2022, art 4.

78 Memorandum of understanding on the sale of counterfeit goods via the Internet (n 67) point 15.

79 AVIA law, art 2; Recommendation on measures to effectively tackle illegal content online (n 21) point 5; NetzDG, § 3(1).

80 AVMSD, art 28b(3)(i).

81 Memorandum of understanding on the sale of counterfeit goods via the Internet (n 67) point 13.

82 NetzDG, § 3(2); Code of Conduct on Countering Illegal Hate Speech Online (n 65) 2.

83 Product Safety Pledge (n 68) 2.

84 Memorandum of understanding on the sale of counterfeit goods via the Internet (n 67) point 18; Gedragscode Notice-and-Take-Down 2018 (n 82) explanatory memorandum.

85 CDSMD, art 17(4)(b)-(c); Product Safety Pledge (n 68) 2.

86 NetzDG, § 3(2).

87 Gedragscode Notice-and-Take-Down 2018 (n 82), art 5-6.

II. Limitation 2: Lack of quality and protection of the victims

43 The observed discrepancies also point to a lack of consensus as to what constitutes a notice and action mechanism of sufficient quality as far as the protection of victims is concerned. The latter stems from the lack of clear requirements in the legal provisions, which are general at best. For instance the CDMSD refers to "sufficiently substantiated notice[s] from the rights holders",[88] whereas the AVMSD requires that VSPs put in place a "transparent and user-friendly mechanism" for flagging content.[89] These general provisions do not guarantee that the mechanism is effective and efficient (Section B. I).

44 In practice a lot will thus depend upon the hosting service providers' willingness and resources. The implemented mechanisms are not always sufficiently user-friendly, and suffer in particular from a lack of sufficient information about the processing of the notices.[90] At the other end of the spectrum, one can point to the Memorandum of Understanding for counterfeit goods, which has been interpreted as allowing for bulk notifications.[91] This might be user-friendly, but can also foster the submission of notifications that are not detailed enough to justify the removal of all notified content.[92] The lack of quality also affects the use of the existing notice and action mechanisms. A 2020 survey in the Netherlands has shown that one third of the people that have been affected by illegal content have used the notice and action mechanism. The survey has also shown that many people are unfamiliar with the mechanism and that potential users value (among other things) the accessibility, user-friendliness, speed and effectiveness of such a mechanism.[93] An increase of quality of and familiarity with the

88 CDSMD, art 17(4)(c).

89 AVMSD, art 28b(3)(d).

90 See, Gellert and Wolters (n 73) 63.

91 See, European Commission, 'Report on the functioning of the Memorandum of Understanding on the sale of counterfeit goods on the internet' (Staff Working Document) SWD (2020) 166 final/2 24.

92 Similar criticisms apply to the Code of Conduct on Countering Illegal Hate Speech, see De Streel and others (n 21) 49.

93 Van Hoboken and others (n 19) 55-57. About the use of such mechanisms in the United States, cf Jennifer M. Urban, Joe Karaganis and Brianna L. Schofield, *Notice and takedown in everyday practice. Version 2* (UC Berkely Public Law Research Paper No. 2755628) 2017.

mechanism could thus increase its use.

III. Limitation 3: lack of adequate safeguards for fundamental rights

45 Beyond the fragmentation of the framework and the lack of agreement on what constitutes an adequate notice and action mechanism from the victims' perspective, additional questions pertain to the safeguards that should accompany such mechanisms. The issue of safeguards exemplifies the way in which the three interests at stake are interwoven: adequate safeguards often entail more resources and have thus a bearing on the economic interests. Also, content providers are not the only ones who should benefit from such safeguards: victims are also entitled to a fair decision-making process. With that being said, the focus of this section is on the content providers.

46 Content providers frequently lack any guarantee that removal decisions are well-balanced, non-discriminatory, consistent, and more generally, fair (Section B. II.). Furthermore, in case they think a decision does not live up to these standards, they lack effective possibilities to challenge a notification or to contest decisions (typically, the removal of content) based on a notification.[94] Such possibilities ensure the protection of the third interest at play, namely that of the other users (and in particular content providers) by safeguarding their fundamental rights such as the right to a fair hearing, the right to equality of arms, or the right to adversarial proceedings.[95] Despite the importance of these possibilities, the AVMSD and the CDSMD are the two only binding European instruments that provide for them. However, both instruments limit themselves to general language without entering into the specifics of what such a mechanism should look like.[96] They therefore do not guarantee an effective protection.

47 Crucial to challenging decisions is the possibility to receive a motivation of the decision upon which a contestation of the decision can build (Section B. II.). However, as seen in section C. III. , the content providers are not always notified or given a clear explanation about the reasons for the removal.

94 Commission, 'Impact Assessment Digital Services Act' (n 27) 26; Kuczerawy (n 27) 535.

95 Kuczerawy (n 27) 535.

96 See, CDSMD, art 17(9), AVMSD art 28b(3)(i).

E. The notice and action mechanism in the DSA

48 The DSA harmonises, codifies and develops the notice and action mechanism. It provides that all hosting service providers should put a notice and action mechanism in place (Article 16). The goal of this section is to see whether the notice and action mechanism in the DSA sufficiently addresses the needs of the three discussed interests, and in so doing addresses the identified limitations.

I. The protection of the victims of illegal content

49 Articles 16 DSA contains harmonized requirements on the effectiveness and efficiency of the notice and action mechanism (see Section B. I.). The mechanisms must be user-friendly and easy to access, which entails among others that they should allow for exclusively electronic notices.[97] Notices submitted in accordance with the prescriptions of the DSA will lead to a presumption of knowledge of the illegality of the content. Although this does not directly obligate the hosting service providers to remove the content, they may face liability if they don't.[98] Hosting service providers must facilitate the submission of valid notices,[99] and must process the notices they receive in a timely, diligent, non-arbitrary, and objective manner.[100] They should also notify the notifier without undue delay of the receipt of and their decision on the notice.[101] These requirements go a long way in resolving limitation 2 (Section D. II.). By providing clear and detailed requirements, hosting service providers are obligated to ensure that their notice and action mechanism is adequate.

50 If the provider of an *online platform* decides not to

97 DSA, art 16(1).

98 DSA, art 16(3). An obligation to remove illegal content is imposed indirectly and implicitly through the obligation to apply and enforce their terms and conditions in a diligent, objective and proportionate manner and the obligation of the providers of very large online platforms to take measures to mitigate the risks of the dissemination of illegal content. DSA, arts 14(4), 34(1)(a), 28; P.T.J. Wolters, 'Privaatrechtelijke en consumentrechtelijke bescherming in het DSA-voorstel' [2022] TvC 18, 22.

99 DSA, art 16(2).

100 DSA, art 16(6).

101 DSA, art 16(4), (5).

remove the notified content, article 20 DSA grants the victim the right to lodge a complaint in the internal complaint-handling system of this platform. If this complaint is dismissed, it can take the dispute to a certified out-of-court dispute settlement body pursuant to article 21(1). The DSA thus follows the General approach of the Council. The notifier did not have these redress possibilities under the original DSA proposal.[102] We support this extended scope. After all, it would be an unfair limitation on the protection of the victims if they aren't able to contest a negative decision on their notice, especially in cases where they might not have enough resources to pursue the only other possible option, namely court proceedings. These redress possibilities and their limitation to online platforms are further discussed in Section E. II.

51 The DSA has also sought to ensure that potential victims would not abuse the notice and action mechanism. Article 16(3) states that 'notices referred to in this article' give rise to actual knowledge. Article 16 (2) requires that the mechanism facilitates the submission of notices that contain '*all* of the following elements', including the name and email address of the notifier. Furthermore, recital 53 states that the notice and action mechanism *should* ask the notifier to disclose its identity in order to avoid misuse. In contrast, recital 50 states that the mechanism should allow, but not require, the identification of the notifier. The DSA is thus unclear about the existence of a requirement that valid notices not be anonymous (except for cases of children sexual abuse material).[103] In this respect, the DSA may not adequately protect the interests of the victims. Non-anonymous notices can jeopardise online anonymity and may prevent victims from submitting notices out of fear of retaliation.[104] For this reason, requiring non-anonymous notices should be avoided as much as possible, except where unfeasible (e.g., alleging copyrights violations might require identification).[105]

52 Article 23(2) DSA contains additional measures against misuse of the notice and action system. However, rather than making the notice submission more cumbersome, these additional measures are of an *ex post* nature as they entail an obligation to suspend for a reasonable period of time the processing of notices and complaints from notifiers who frequently submit notices and complaints that are manifestly unfounded. This *ex post* nature is to be favoured. It provides safeguards for fundamental rights without limiting the effectiveness and efficiency of the notice and action mechanism and thus the protection of victims of illegal content.

53 However, it is important to make sure that these *ex post* measures can be applied relatively fast. In this light, the conditions of Article 23(2) DSA may be too strict. The processing of the notices can only be suspended if the notifier 'frequently' submits notices that are 'manifestly' unfounded and only after a prior warning. This suggests a high threshold. A prior warning should not be necessary if the notifier crosses this threshold and clearly acts in bad faith, especially because the suspension can only be 'for a reasonable period of time'. In this regard one should note that the DSA does not explicitly allow online platforms to determine a lower threshold compared to Article 23(2) via their terms and conditions.[106] Article 16(6) obligates the platforms to process 'any notices that they receive'. This implies that restrictions that go beyond Article 23(2) are not allowed.

54 Anonymous notices can further complicate the application of Article 23(2). However, even if a service provider does not know the (real) name and email address of the notifier, it may still be able to distinguish various notifiers through pseudonyms, IP-addresses or cookies. For this reason, we believe that the additional protection of victims of illegal content outweighs the additional risks of abuse, especially because the DSA also contains other safeguards for the fundamental rights of the (other) users.

II. Safeguarding fundamental rights

55 Various provisions of the DSA make sure that the notice and action mechanism does not disproportionally encroach on fundamental rights. First, a notice only leads to actual knowledge, and thus potentially to liability, if it is sufficiently substantiated, precise and allows a diligent provider to identify the illegality without a detailed legal examination. Under this rule, a provider cannot be held liable if the illegality is uncertain,[107] the

102 Cf DSA proposal, art 17, 18; General approach, art 17, 18.

103 Note that the DSA proposal was more explicit about this requirement by specifically referring to the elements of Article 14(2) in Article 14(3).

104 Section 2.1. On the value of anonymity online, see, e.g., A Michael Froomkin, 'From Anonymity to Identification' (2015) 01 *Journal of Self-Regulation and Regulation* 120.

105 On this point, see European Parliament, 'Adopting commercial and civil law rules', Annex B, art 9(1)(e).

106 See, DSA, art 14(1). Cf DSA, art 23(1), recital 64, which allows online platforms to establish stricter measures in relation to the removal of illegal content.

107 Either because the facts or the law is unclear, cf Gellert and

hope being that hosting service providers will have less incentives to precautionarily remove content (see Section B. II.). As far as the quality of the decisions themselves are concerned we can also point to Article 16(6) DSA, which requires that decisions be taken in a diligent, non-arbitrary, and objective manner (Section E. I.). These safeguards also apply to the automated moderation of content. However, the exact meaning of these safeguards for automated moderation remain unclear. It may be necessary to formulate more specific requirements. For example, the General Data Protection Regulation and the proposed AI Act require specific safeguards in relation to possible errors and bias of automated means.[108] In contrast, Article 16(6) DSA only provides that a hosting service provider should inform the notifier of the use of automated means.

56 Next, Article 17 of the DSA provides for transparency with regards to decisions to remove or disable content. Hosting service providers shall inform content providers of a decision to demonetise or restrict the visibility of their content or to suspend or terminate the provision of the service or account. They must also provide the content providers with a clear and specific statement of reasons.[109] This requires to indicate the type of decision, the legal ground relied upon (or the Terms and Conditions provision), as well as the facts and circumstances supporting it (and the redress possibilities).[110]

57 It should not be construed as exceedingly affecting the economic interests of the hosting services providers since Article 17(4) DSA clarifies that it should be "as precise and specific as reasonably possible under the given circumstances". We believe that general statement about the reason for the removal would comply with such an obligation. A hosting service provider should make clear which rule is violated by the content, but is not obligated to provide a detailed analysis.

58 In contrast, the DSA does not explicitly state that the hosting service provider should provide a statement of reasons to the notifier. It is only obligated to inform the notifier of the decision and redress possibilities (Section E. I.). This distinction is not justified. Like content providers, notifiers require a statement of reasons to understand the decision and to effectively exercise their redress possibilities. In this light, the lack of an explicit obligation in relation to the notifiers is a missed opportunity to strengthen judicial oversight and protect the victims of illegal content (Section B. II.) and fully address this limitation in current practice (Section D. III.).[111]

59 The DSA also devotes considerable attention to the redress possibilities of both the content provider and the notifier. It provides both for an internal complaint-handling mechanism and an out-of-court dispute settlement mechanism.[112]

60 Article 20 of the DSA is dedicated to the internal complaint-handling mechanism. As a fundamental rights safeguard it should provide content providers with adequate guarantees,[113] even though one cannot hold it to the same standards (e.g., independence, impartiality) as a regular court.[114] In this regard, the DSA refers to a free complaint-handling mechanism, which should be effective, 'easy to access, user-friendly', and should 'enable' and 'facilitate' the submission of 'sufficiently precise and adequately substantiated' complaints.[115] The complaints should also be processed in a 'diligent, non-discriminatory, and non-arbitrary' manner.[116] Finally, the decision must be taken under the control of appropriately qualified staff pursuant to Article 20(6) DSA. Unlike the initial decision on the notice (Article 16(6) DSA), it cannot be made solely on the basis of automated means.

61 The internal-complaint handling mechanism is limited to online platforms. This can be seen as an undue limitation on safeguarding the fundamental rights. After all, there is a case to be made that the internal complaint-handling mechanism

Wolters (n 73) 28-30.

108 Regulation 2016/679/EU of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data [2016] OJ L119/1, art 22(2)(b), (3); Article 29 Data Protection Working Party, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679* (WP251rev.01, 2018) 27; Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts' COM(2021) 206 final, 2, 11, recitals 33, 40, 44, 50, arts 10(2)(f), 15(3).

109 DSA, art 17(1).

110 DSA compromise, art 17(3)(a), (b), (d), (e), (f).

111 On this topic, see also Naomi Appelman and others, 'Access to Digital Justice: In Search of an Effective Remedy for Removing Unlawful Online Content' (2021) *Amsterdam Law School Legal Studies Research Paper No. 2021-35, Institute for Information Law Research Paper No. 2021-06*.

112 DSA, arts 20, 21.

113 See, Wilman (n 54) 373-374.

114 Wilman (n 54) 371.

115 DSA, art 20(1), (3).

116 DSA, art 17(4).

should be available to all users, as many relevant situations take place outside of online platforms. For example, a website offering legal content may be a victim of overzealous intellectual property-based notifications.[117]

62 Such an extension of the internal complaint-handling mechanism's scope might however be overly burdensome for the hosting service providers. However, this depends upon the way in which such mechanism is conceived. On the basis of the DSA provisions, this mechanism can take many shapes, but nothing says that all internal complaint-handling mechanisms should look like a court proceeding.[118] An alternative solution would be one that is closer in spirit to *ex post* counter notices,[119] in which the content providers have an opportunity to explain why the content is not illegal and have their content restored. Modelling the internal complaint mechanism on counter notices and essentially making the submission of complaints a similar procedure as the submission of notices could go a long way in addressing some of the concerns relating to the economic interests of hosting service providers. Furthermore, it could also facilitate rapid response times in order to avoid situations where content is rapidly deleted but only slowly reinstated. Article 17(3) of the DSA refers to 'timely' responses. Stronger language such as 'without undue delay' might be more useful, without going so far as giving strict deadlines such as the European Parliament's position which allocates 10 working days to reply to such a complaint.[120]

63 The out-of-court dispute settlement provided in the DSA strives to provide adequate safeguards for the fundamental rights of the content providers by allowing notifiers and content providers to take a dispute to the certified out-of-court dispute settlement body of their choice.[121] The DSA further provides safeguards by imposing requirements in terms of impartiality and independence,[122] expertise,[123] online accessibility,[124] and procedural fairness.[125]

64 A couple of potential limitations on the safeguarding of fundamental rights and caveats can be highlighted here. Similar to the internal complaint-handling mechanism, out-of-court dispute settlement only applies to online platforms. However, here too there are many relevant situations outside of online platforms: many victims or interested parties in the context of 'regular' hosting service providers may also not have the sufficient resources concerning court proceedings (e.g., victim of online children sexual abuse and exploitation material, or a website offering legal content).

65 Contrary to the internal complaint-handling mechanism, the DSA does not require that the out-of-court dispute settlement be fully paid by the service provider. Here, one must observe that the system in the adopted DSA is much more friendly to the user (the notifier or content provider) than previous iterations. The DSA proposal was not entirely clear on the requirements of the fees. It merely stated that such fees could not exceed the costs.[126] The DSA distinguishes between the fees charged to online platform providers (which should be reasonable and not exceed the costs) and the fees charged to users (which should be either inexistent or nominal and should be refunded by the online platform provider if the dispute is decided in their favour).[127] This system seems to strike an adequate balance between the various interests at stake. A limited fee can prevent frivolous use of the mechanism and thus protect the economic interests of the service providers without unduly limiting independent oversight and thus the protection of the victims and fundamental rights.

117 On this topic, see for instance, Jason Mazzone, *Copyfraud and Other Abuses of Intellectual Property* (Stanford University Press 2011).

118 Cf Meta's Oversight Board, <https://www.oversightboard.com/>, last accessed 14 July 2022.

119 About counter notices and their limitations, see eg Commission, 'Impact Assessment Digital Services Act' (n 32) 26; João Pedro Quintais and others, 'Safeguarding User Freedoms in Implementing Article 17 of the Copyright in the Digital Single Market Directive: Recommendations from European Academics' (2019) 10 *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 280; Wilman (n 54) 370-374; Kuczerawy (n 32) 531-532, 535, De Streel and others (n 26) 49.

120 Amendment 227.

121 DSA, art 21(1).

122 DSA, art 21(3)(a), (c). But see also The Greens/EFA, 'Regulation on procedures for notifying and acting on illegal content and for content moderation under terms and conditions by information society services', (2020), art 22(2), for additional requirements.

123 DSA, art 21(3)(b).

124 DSA compromise, art 21(3)(d). The information about the mechanism should also be accessible online, see DSA compromise, art 21(1).

125 DSA compromise, art 21(3)(f).

126 DSA proposal, art 18(3).

127 DSA, art 18(5).

66 Finally, contrary to the original DSA proposal which only referred to a 'swift, efficient, and cost-effective' procedure,[128] the adopted DSA has supplemented this general formulation with more concrete timelines.[129] Namely, a decision should be taken within a reasonable period of time that does not exceed 90 days (or 180 days for the more complex cases). This text builds upon the European Parliament's position.[130]

III. The economic interests of hosting service providers

67 Article 16 DSA imposes a harmonised notice and action mechanism on all hosting service providers that applies to all types of content. Although some obligations to implement notice and action mechanisms already exist (Sections C. I. and C. II.), the DSA certainly increases the obligations of the hosting service providers. Furthermore, the requirements in terms of effectiveness and efficiency (Section E. I.) and the redress possibilities and safeguards for fundamental rights (Section E. II.) can impose significant costs. In this light, the DSA represents a new balance between the various interests. The economic interests of the providers of hosting services have been limited in favour of the protection of victims and fundamental rights.

68 It is noteworthy that the DSA does not exempt micro and small enterprises from the obligation to implement a notice and action mechanism.[131] This is justified, since the dissemination of illegal content through small hosting services can also have a significant adverse effect on victims.[132] Micro and small enterprises are exempted from the additional obligations for providers of online platforms pursuant to Article 19 DSA. The DSA does distinguish between online platforms and other hosting services. Articles 20, 21 and 23 DSA only apply to online platforms. In contrast, the notice and action mechanism applies to all hosting service providers. The reason for this inconsistency is not necessarily justified (Section E. II.). Recital 41 DSA merely states in general terms that the obligations should be adapted to the type and nature of the intermediary service. In any case, the additional obligations for online platforms lead to a better protection of both the victims of illegal content and fundamental rights at the expense of the economic interests of the providers of online platforms.

69 At the same time, the new provisions do not completely ignore the economic interests of the service providers. First, the harmonised nature allows them to implement one mechanism throughout Europe. However, the lack of harmonisation (Limitation 1, Section D. I.) is only partially resolved by the DSA since the more specific rules in vertical instruments still apply (and can force hosting service providers to implement special procedures).[133]

70 Second, the rules in relation to the use of automated means for the processing of notices also balance the economic interests of the hosting service providers with the other involved interests. Hosting service providers are allowed to process notices automatically pursuant to Article 16(6) DSA. Similarly, the limitation of the statement of reasons to what is reasonable also facilitates automated and standardised motivations. In contrast, decisions in respect to the internal complaint-handling system cannot be made solely on the basis of automated means (Section E. II.).

71 Third, the DSA contains a number of provisions against abuses, specifically directed at online platforms. Article 23(2) allows them to refuse to process notices or complaints insofar as they are manifestly unfounded and originate from the same user who submits them frequently. This can reduce the economic burden of processing these unfounded notices and complaints. However, the high threshold of this provision (Section E. I.) means that online platforms can only benefit from it in limited cases. Furthermore, Article 23(2) does not apply to micro and small online platforms or other hosting services, which would force them to keep processing these notices pursuant to Article 16(6). Obviously, this goes against the intention of the DSA, which is to limit the obligations of micro and small online platforms and other hosting services. We therefore argue that these service providers can also suspend the processing of manifestly unfounded notices if the conditions of Article 23(2) DSA apply.

72 Further, and in relation to out-of-court dispute settlement, the adopted DSA added a new provision against abusive procedures which allows online platforms providers to refuse to engage in proceedings if the issue has already been previously resolved.[134] Also, the online platforms are entitled to a reimbursement of their fees if the notifier or content provider acted manifestly in bad faith

128 DSA proposal, art 18(2)(d).

129 The general formulation is still kept, see DSA, art 18(3)(e).

130 Amendment 240.

131 Cf DSA, arts 12(4), 19.

132 Gellert and Wolters (n 73) 95-96.

133 DSA proposal 4-5, 9; DSA, art 2(4).

134 DSA, art 21(2).

pursuant to Article 21(5) DSA.

73 Next, a notice only leads to actual knowledge, and thus potentially to liability, if it is sufficiently substantiated and precise, and allows a diligent provider to identify the illegality without a detailed legal examination (Section E. II.). Although the exact requirements of a 'diligent' provider will still need to be ironed out, it is clear that this provision limits the liability risks of the providers.

74 Finally, the provisions on the notice and action mechanism contain open-ended norms. They refer to a 'diligent' provider,[135] demand 'timely' action[136] as well as 'good faith' engagement[137] and limit the statement of reasons to what is 'reasonably' possible.[138] These open-ended norms make sure that the economic interests of the hosting service providers can also be taken into account when interpreting the various obligations. It allows for differentiation based on the size of the service provider. In addition to the explicit exemptions and additional obligations, the open-ended norms allow for stricter demands on large hosting service providers and less stringent requirements on smaller providers.[139] For example, it could affect the requirements in relation to the diligence, non-arbitrariness and objectiveness of the automated means that are used to process notices.[140]

F. Conclusions

75 The adoption of the DSA will bring important changes to the content moderation landscape in the EU. By harmonising, codifying, and further developing a notice and action mechanism, the DSA addresses many content moderation-related challenges, and in so doing also affects the balance that existed thus far between the protection of victims of illegal content, the safeguarding of fundamental rights and the economic interests of hosting service providers.

76 As far as the economic interests of hosting service providers are concerned (Section B. III.), the harmonisation of the mechanism should certainly be a welcome change for economic operators (Section D. I.). Further, even though the DSA contains many new procedural obligations, they entail reasonable efforts (Section E. III.). One can point to the limitation of the statement of reasons to what is reasonable or the possibility to use automated tools. By stating that a notice only gives rise to 'actual knowledge' if it allows a diligent operator to identify the illegality, the DSA limits the operators' liability and also contributes to the safeguarding of fundamental rights since it limits overcautious removals of content. Finally, the DSA also contains provisions against abusive notices and complaints. However, the high threshold of Article 23(2) limits its effectiveness. Interestingly, this provision does not help small and micro online platforms and other hosting service providers (Section E. III.). This is paradoxical since the economic interests also explain why these actors fall outside of the scope of the redress mechanisms. Whereas limiting the costs that these new fundamental rights safeguards can entail is a legitimate goal, we have also shown that there is a real fundamental rights interest to extend these mechanisms to all hosting service providers. We have also shown that this is possible and feasible depending upon the manner in which these mechanisms are conceived (Section E. II.).

77 Beyond excluding certain economic operators from redress mechanisms, the DSA also refuses notifiers the possibility to receive a statement of reasons for a (negative) decision taken pursuant to their notice. Whereas the adopted DSA can be seen as an improvement concerning notifiers' right compared to the DSA proposal (they can now also benefit from the redress mechanisms), we consider the lack of an obligation to provide a statement of reasons as a missed opportunity. Furthermore, one can still lament the unclarity about the effect of anonymous notifications. Beyond that however, the requirement of a harmonised, efficient, effective, and user-friendly notification procedures should fix the existing limitations (Section D. II.) and can be seen as an important step for the protection of the victims' interest (Sections B. I. and E. I.).

78 Finally, the safeguards of content providers' fundamental rights are also enhanced (Sections B. II. and E. II.). Not only through the creation of new redress mechanisms, but also through the hosting provider services' obligation to provide decisions that are objective, non-arbitrary, diligent and timely, and to justify them through a statement of reasons. Although the applicability of the safeguards is still too narrow in some respects, the new safeguards and their requirements should improve the current situation in which hardly any binding legal provisions exist (Section D. III.).

79 All in all, even though it contains various shortcomings that prevent it from truly striking

135 DSA, arts 16(3), (6), 17(4), 23(3)

136 DSA, arts 16(6), 20(4), 23(3).

137 DSA, art 21(2).

138 DSA, art 17(4); Section 5.2.

139 See also Gellert and Wolters (n 73) 97.

140 DSA, art 16(6).

an adequate balance, the DSA's notice and action mechanism does represent a significant step forward for all the parties that have a stake in the moderation of online content.

Online-Dispute Resolution - Paving the way towards harmonising the Birksian archipelago[1] of obligations?

by **Gregory Chan and Tan Yan Shen** *

1 The term "Birksian" is a reference to the works of Professor Peter Birks and his theories which are regarded as the baseline for modern private law theory, originating from: PBH Birks Unjust Enrichment (Oxford: Clarendon Press, 2nd edn, 2005). Helpfully summarised by Professor Duncan Sheehan and Professor TT Arvind, Birksian thinking favours "timeless principles" to generate lower-order rules used by legal decision-making, and "a suspicion of policy as a means of avoiding proper analysis of the principles and rationale of the law": Duncan Sheehan, TT Arvind. "Private Law Theory and Taxonomy: reframing the debate". (2015) 35 Legal Studies 3, 480-501.

Abstract: It is only natural that the rise of e-commerce is coupled with an increasing number of disputes; eBay alone has seen a record 60 million cases opened under its online dispute-resolution ('ODR') scheme. While this can be regarded as the first step towards the creation of an online rule-of-law, such ODR mechanisms are often shrouded in uncertainty.

In that regard, this paper explores ODR mechanisms in both established, and in, what we describe as 'informal' marketplaces, such as commerce on Reddit and Discord. This paper first asks whether these ODR mechanisms give rise to its own jurisprudence possibly inconsistent with "offline" rules of law, and whether such a bifurcation of "online" and "offline" rules of law is normatively desired. Next, it then queries the limitations of various policies and regulations which attempt to strengthen ODR mechanisms. It contends that various policies are disconnected from their practical implementation and constraints which ODR platforms face.

Ultimately, it concludes that a more nuanced approach is required if such frameworks were to be harmonised across Courts through the proposed taxonomy. Current international recommendations, while a good starting point, should be condensed to certain principles which may be adopted across platforms, while preserving site-autonomy across different types of platforms.

Keywords: ODR, Online Dispute Resolution, Legal Technology, Dispute Resolution, Access to Justice, Forum Marketplace, E-Commerce, Access to Justice

Recommended citation: Gregory Chan and Tan Yan Shen, Online-Dispute Resolution - Paving the way towards harmonising the Birksian archipelago of obligations?, 14 (2023) JIPITEC 420 para 1

A. Introduction

1 With the shifting tide of commerce towards the online realm, there has been an increased conversation about the role that e-commerce places in our lives. Online shopping seems to have overtaken traditional brick and mortar stores, revolutionising the ways that companies have conducted their businesses.[1] According to the United Nations Commission on International Trade Law, 2020 saw a 20% increase in the trade volume of e-commerce compared to the

* All information contained in this paper represents the views and opinions of the authors, and does not necessarily represent the views of the publishers or affiliated organisations. The content in this paper is not to be taken as formal legal advice, and is written for academic purposes. Any errors present are solely the fault of the authors.

1 Anjali Gupta, 'E-Commerce: Role of E-Commerce in Today's Business', (2014) 4 International Journal of Computing 1. <http://www.ijccr.com/January2014/10.pdf> accessed 31 July 2022.

previous year.[2] Unfortunately, as popularity in online trade rises, it seems inevitable that there will be a growing number of conflicts. Hence, e-commerce platforms have worked to develop their own unique forms of dispute-resolution through their platforms. Colloquially, these mechanisms are 'Online Dispute-Resolution' ('ODR'). However, due to the diversity of e-commerce sites, ODR has become site specific, operating very differently across the multitude of e-commerce platforms. Such creates inconsistencies across decisions taken, which would invariably lead to frustrated users and a lack of certainty across ODR platforms.

2 As such, this essay seeks to explore the growing trend of ODR mechanisms across various e-commerce platforms and identify core trends across various e-commerce sites. Ultimately it highlights that there seems to be a disconnect between users, regulators, and platform administrators in the administration of ODR. This, in turn, leads to inconsistency across various platforms, which frustrates the implementation and development of an online code-of-conduct and an established Rule of Law. To that end, it posits that a more generalised approach is perhaps preferable in ODR sites - allowing platforms to maintain their autonomy while ensuring a degree of legal certainty and procedural safeguards.

3 Following, this paper first provides an overview of ODR mechanisms across various e-commerce sites, and attempts a brief taxonomy of e-commerce platforms for the purposes of this paper in Section B. Section C considers both procedural and substantive issues in the implementation of ODR platforms across formal and informal e-commerce sites. Section D goes on to identify potential solutions which could be implemented, highlighting the constraints of current regulatory proposals while making its own. Section E concludes.

B. An overview of Online-Dispute Resolution

4 At the outset, it must be recognised that the ODR can take place across a multitude of platforms, and is not strictly limited to e-commerce. For the purposes of this paper, it is thus important to clarify certain definitions and distinctions that will be used in later sections.

I. Online Dispute Resolution

5 Generally speaking, ODR mechanisms are not limited to the e-commerce industry. It is a broad term used to reflect a novel form of dispute-resolution available on the internet, not requiring parties' physical presence for adjudication. Proponents of ODR have cited the process as a means of achieving access to justice for civil suits - avoiding the costly legal fees, and achieving efficient dispute-resolution.[3] To that end, various jurisdictions such as Singapore have been taking advantage of this, launching a successful ODR platform for employment related claims under the Tripartite Alliance for Dispute Resolution program.[4] This trend of the growing use of ODR indeed suggests a promising future for this form of alternative dispute-resolution. Undoubtedly, Singapore's application of ODR is one for public purposes, established under a statutory framework for employment laws in Singapore. A different situation would inevitably arise if private companies were to utilise such platforms on a different scale, and without statutory safeguards.

6 This is where the story begins. Across e-commerce platforms, ODR mechanisms are commonplace to resolve disputes between users, as well as between third parties. For instance, eBay's ODR mechanism operates under their Resolution Centre, and was designed with high-volume claims in mind.[5] Indeed, eBay currently averages at approximately 60 million disputes a year.[6] In a similar vein, e-commerce rival Amazon has a similar ODR mechanism operating on the Amazon Pay platform, for sales made on its website.[7] Other e-commerce giants such as

2 UNCTAD, *Global E-commerce Jumps to $26.7 trillion, COVID-19 boosts Online Sales*, (UNCTAD.org, 3 May 2021). <https://unctad.org/news/global-e-commerce-jumps-267-trillion-covid-19-boosts-online-sales> accessed 28 July 2022.

3 Robert J Condlin, 'Online Dispute Resolution: Stinky, Repugnant, or Drab?', (2017) Faculty Scholarship 1576, 717-758, <https://digitalcommons.law.umaryland.edu/fac_pubs/1576> accessed 1 August 2022.

4 Ministry of Manpower, *Employment Standards Improve in 2021 Through Proactive Tripartite Efforts*, (2022, Employment Practices), <https://www.mom.gov.sg/newsroom/press-releases/2022/0718-employment-standards-report-2021> accessed 30 July 2022.

5 Louis F. Del Duca Colin Rule Kathryn Rimpfel, 'eBay's De Facto Low Value High Volume Resolution Process: Lessons and Best Practices for ODR Systems Designers' (2014) 6 Y.B Arb & Mediation, 204-219. <https://elibrary.law.psu.edu/cgi/viewcontent.cgi?article=1060&context=arbitrationlawreview> accessed 28 July 2022.

6 Mizzou Law. *Library Guides: Online Dispute Resolution: Companies Implementing ODR*. (2018, Missouri School of Law) <https://libraryguides.missouri.edu/c.php?g=557240&p=3832247> accessed 28 July 2022.

7 See Amazon's ODR platform under Amazon Pay, available at: <https://pay.amazon.com/help/201751580>.

Etsy,[8] Alibaba,[9] and RedBubble[10] have similar high-volume mechanisms in place. These high-volume, high-efficiency models are often regarded as a fundamental characteristic of these ODR mechanisms. Online disputants are known to be highly focused on efficiency; empirical studies have indicated users would prefer to lose a case over a few days, than win a case over a few weeks.[11] However, the trade-off from efficiency is the quality of ODR on these platforms, both procedurally as well as substantively of each individual case. This will be further discussed later in the paper.

II. E-commerce platforms

7 While it is impossible to provide an overview of the profiles of every site due to space limitations, this paper highlights 2 core distinctions that the authors have identified - formal marketplaces, as well as informal marketplaces.[12] At its core, we propose this distinction between these e-commerce platforms lies in the purpose for which the platform was set up for. Formal marketplaces were set up for the purposes of e-commerce, whereas informal marketplaces were established for other purposes, but evolved to include e-commerce on their platforms as an extension of its purpose.

8 Looking through the former, core examples of forum marketplaces include eBay, and Etsy. These marketplaces can be characterised through their use of End-User Licensing Agreement ('EULAs') to delineate the rights of users when operating on their sites, particularly in the areas of commercial arrangements. For instance, eBay's EULA incorporates terms for fees and taxes of users posting listings, conditions for international trade, as well as policies for the trading of goods.[13] Similarly, Etsy's EULA warrants terms for the use of Etsy as a platform for sale, including provisions for their 'House Rules for Sellers' and 'House Rules for Buyers'.[14] These EULAs form the primary characterisation for what has been identified as formal marketplaces, perhaps best described as 'top-down governance'.

9 On the other hand, informal marketplaces operate through a 'bottom-up governance'; albeit cliche, they can be described as "by users, for users". These marketplaces often operate as forums, before transitioning towards operating as a marketplace through what can be identified as the 'natural expansion'.[15] As a result, e-commerce on these platforms is largely user-driven; platform owners and administrators themselves often do not have a stake in commercial activity here; there are no associated listing fees for users, or any governing EULAs which accommodate for trade. One such informal marketplace operates on the site Reddit. While the site describes itself as a 'online discussion site',[16] sub-communities around various hobbies have themselves created marketplaces as a consequence of growing popularity, and an alternative for users to subvert the strict requirements of formal marketplaces. These include r/mechmarket, a marketplace for mechanical keyboards, r/BoardGamesExchange for the sale of board games, as well as the various trading card marketplace subreddits for popular card games including Yu-Gi-Oh![17] and Magic the Gathering[18]. However, Reddit's EULA does not make any provision for the sale of goods on their sites.[19] Instead, governance of these

8 See here, Etsy's ODR platform: <https://help.etsy.com/hc/en-us/articles/360016126873?segment=selling>.

9 See here, Alibaba's ODR platform: <https://service.alibaba.com/page/knowledge?pageId=128&category=9207656&knowledge=20154304&language=en>

10 See here, RedBubble's ODR platform: <https://help.redbubble.com/hc/en-us/articles/202982715-Resolving-Conflict-with-another-Member>

11 Arno R. Lodder, John Zeleznikow, '*Enhanced Dispute Resolution Through the use of Information Technology*' (2010, Cambridge University Press).

12 For a further elaboration on this distinction, see Gregory Chan, 'Online Dispute Resolution: Beginnings of an Online Rule of Law' (2022) *Rule of Law* 3, 2-9. <https://ruleoflaw.lse.ac.uk/articles/abstract/35/> accessed 20th July 2022.

13 See here, eBay's EULA that can be found at: <https://www.ebay.com/help/policies/member-behaviour-policies/user-agreement?id=4259>. At 5, provisions on listing fees and taxes. At 6, clauses on listing conditions for sellers, at 8, on policies of buying and selling goods.

14 Here, see Etsy's EULA at: <https://www.etsy.com/legal/terms-of-use/#services>. At 2, see provisions for buyers and sellers according to their EULAs.

15 (n 14), at 7.

16 Katie Elson Anderson, 'Ask me anything: what is Reddit?' (2015) 32 *Library Hi Tech News* 5. <https://www.emerald.com/insight/content/doi/10.1108/LHTN-03-2015-0018/full/html?casa_token=zo_SCVCYIIYAAAAA:-cBugD1x1XvWIzFnVy9a7URLnGtC0QPEu2fjzAlcevU6a9wJ0f-9JsESK-bLBmQpuj8qYTAnUr8Ck89DLpfw8NTXdFsa_bLTjtgDAElcxuQSmsAXSVKq> accessed 31 July 2022.

17 See here at <http://old.reddit.com/r/YGOMarketplace>.

18 See here at <http://old.reddit.com/r/MTGSales>.

19 See here at <https://www.redditinc.com/policies/user-

marketplaces turn to user-created conventions, rather than binding policies.

III. Categories of disputes on these platforms

10 It must lastly be noted that trade disputes are not the only claims which operate and are resolved by ODR claims. While such user-user disputes are the crux of what occurs on e-commerce sites, 2 further types of disputes are similarly relevant in the field of e-commerce, namely user-user reputation-based disputes and user-third party intellectual property disputes.[20]

11 Reputation-based disputes can be summarised as disputes over the reviews that traders leave for each other on these platforms. On both formal and informal marketplaces, administrators and moderators have developed a unique 'reputation-based' system, where users are able, and often required to, leave feedback for each other based on their sales experience with other parties.[21] However, disputes arise when one party leaves misleading, or false feedback on these platforms that were intended vexatiously. Consequently, these innocent users are portrayed as distrustful, harming their standing and potentially resulting in false sales. These ODR platforms thus have been used by platforms to require users to modify their feedback (if claimants are successful), or moderators use their platform privileges to outrightly remove these misleading statements.

12 The next type of ODR claim is of a different nature, and primarily involves a third party, as opposed to direct buyers and sellers - intellectual property disputes. Predominantly, these claims involve third parties alleging that the seller is selling counterfeit products, or those of stolen designs, and, in light of the anonymity which the internet gives them, has no recourse under traditional means of dispute-resolution. For such claims, third parties are required to file complaints through the platform's ODR mechanism to enforce their intellectual property rights against these sellers.[22] However, it must be noted that such claims offer limited recourse, on both established, and informal marketplaces; the most that moderators or administrators are able to do remain to be the taking down of such posts made by users. Of course, there are rare situations where companies have chosen to enforce their intellectual property rights against the platform as a whole, seeking specific reliefs against the sellers. One such instance was in *Tiffany v eBay*[23] on the sale of counterfeit products on eBay's platform. However, this challenge was denied by the New York Court of Appeal, citing the difficulties of the platform in policing future sales of such products. Hence, it would follow that, while recourse is available on such platforms between users and third parties, they remain rather limited in nature.

13 While these sectors are worth mentioning for completeness, this paper will primarily focus on the traditional user-user dispute for the sale of goods. This follows the traditional fact-pattern of e-commerce scams, through misrepresentation of the conditions of goods, failure to ship the goods, defective products, and other sale-related disputes.[24] However, even on this perhaps clearer front, there exists complex nuances which will be explored in the subsequent section on both formal and informal marketplaces.

agreement-september-12-2021>.

20 Collin Rule, 'Designing a Global Online Dispute Resolution System: Lessons Learned from eBay. (2017) 13 University of St. Thomas Law Journal, 354-370. <https://heinonline.org/HOL/LandingPage?handle=hein.journals/usthomlj13&div=21&id=&page=> accessed 23rd July 2022.

21 For an analysis of eBay's reputation system, see: Kat Busch and others, 'Psychology of Trust on the Internet', (2010-2011, Stanford University). <https://cs.stanford.edu/people/eroberts/cs201/projects/2010-11/PsychologyOfTrust/rep2.html> accessed 31 July 2022. See also here, for Etsy's reputation system, available at: <https://cs.stanford.edu/people/eroberts/cs201/projects/2010-11/PsychologyOfTrust/rep2.html> But see also for example here, for informal marketplaces on r/MechMarket on Reddit, on <https://old.reddit.com/r/mechmarket/comments/wd06su/august_confirmed_trade_thread/>. Other instances of similar mechanisms include <https://old.reddit.com/r/YGOMarketplace/> using a flair-based system (accessed 1 August 2022).

22 See, for example, the eBay IP mechanism known as VeRo available at: <https://www.ebay.com/sellercenter/ebay-for-business/verified-rights-owner-program>. On Etsy, the IP disputes mechanism is available at: <https://www.etsy.com/legal/ip/>. For informal platforms, on r/mechmarket, it is written on the guidelines of use of the platform, that "It is up to the discretion of r/MechMarket mods on whether the claims are relevant and valid regarding any action taken for infringing posts." Taken from: <https://www.reddit.com/r/mechmarket/wiki/rules/rules?v=307a046c-234f-11e9-8765-0e7e4515df94>. Accessed 1 August 2022.

23 600 F.3d 93 (2nd Cir 2010).

24 M Niranjan Murthy and others, 'Analysis of E-Commerce and M-Commerce: Advantages, Limitations and Security issues', (2013) 2 *International Journal of Advanced Research in Computer and Communication Engineering* 13.

C. ODR applied: Challenges and Difficulties on Various Platforms

14 Having laid out the foundation, challenges within these marketplaces in the field of trade disputes on e-commerce sites can be explored. Particularly, this paper sheds light on the ways that the two different categories of platforms identified tackle various procedural and substantive issues in the application of ODR.

I. Procedural Matters

15 The crux of procedural matters in ODR lies in the mechanisms by which parties are heard, and bring their disputes to the relevant adjudicators available on various platforms. While the procedural aspects of such claims differ from platform to platform, one key trend can be noted across the board - that sellers are systematically disadvantaged. This occurs either through a lack of equality of arms or being subject to disproportionate penalties.

1. (In)equality of arms

16 Beginning first with issues around equality or arms. The crux of such concerns lie in the lack of procedural due process. While alluded to earlier when comparing private ODR platforms with those established under statutory provisions, the lack of procedural safeguards across these platforms give cause for concern.

17 To delve further in, one should first note the procedure of an ODR claim on these platforms. Traditionally, claims against sellers are started by buyers for defective goods or products that do not match the listed description, often after a mandatory period of mediation between the two parties.[25] However, as opposed to traditional service of court documents, these claims are submitted to the platform that would inform the seller of the existence of such a claim. While this seems necessary in light of the anonymity which these platforms offer through the internet, this first step already presents issues. Firstly, ODR platforms often do not have a mechanism for buyers to challenge the appropriate forum for disputes. Their reasoning for this is sound - the emphasis on efficiency, coupled with provisions stipulated in EULAs that are buyer-focused.[26] However, if buyers stray away from the stipulated ODR mechanism, and engage a third party service provider involved in the transaction, this would inevitably create issues. For instance, buyers may call their credit card companies alleging their card's misuse, thus, having their credit card company give chargebacks and effectively refunding the purchase. This leaves the buyer with the goods purchased, and his money back, while leaving the sellers with no recourse.[27] While safeguards can be put in place, the fundamental problem turns to the anonymity of these e-commerce sites; it becomes impossible for sellers to be represented in such ODR claims. This issue is similarly more prevalent on informal marketplaces, where these sites often do not store a site-specific payment mechanism, and opt for third party financial services, such as PayPal.[28] By bringing a claim under PayPal (or other third party financial service provider) as opposed to the platform-specific e-commerce site, buyers are able to circumvent both the sellers and administrators who are often able to accrue evidence on both sides, and create conditions favourable to their case with no alternative recourse for sellers.

18 However, even if an appropriate forum is chosen in accordance with stipulated EULAs or through parties consent on informal platforms, ODR platforms themselves do not afford equality of arms to both parties. For instance, on Facebook Marketplace (a formal marketplace by characterisation of the implementation of their EULA and top-down governance[29]), only buyers are able to file ODR claims through the 'Commerce Manager' system. In that vein, after the mandatory mediation period has elapsed, the buyer may start a claim against the

25 See here, Facebook's ODR mechanism requiring mandatory mediation period: <https://www.facebook.com/business/help/1167434420087941?id=353836851981351>. See also, on AirBNB's platform for <https://www.airbnb.com.sg/help/article/767/how-the-resolution-centre-helps-you>, requiring a period of mandatory negotiation between consumers and service providers before stepping in to arbitrate a dispute between the parties.

26 Mohammed A. Aslam, "B-2-C Pre-dispute Arbitration Clauses, E-commerce Trust Construction and Jenga: Keeping Every Cog and Wheel" (2013) 7 *Masaryk University Journal of Law and Technology* 1, 1-18.

27 Yue Guo and others, 'To sell or not to sell: Exploring sellers' trust and risk of chargeback fraud in cross-border electronic commerce', (2017) 28 *Information Systems Journal* 2, 359-383.

28 See, for instance, r/mechmarket, that encourages users to use third party financial services such as PayPal for their transactions. Available at: <https://www.reddit.com/r/mechmarket/wiki/payment>. See also here on <http://www.reddit.com/r/YGOMarketplace> on the sidebar which lists the Subreddit's rule. At 5

29 See here, Facebook's EULA and more specific rules governing ODR mechanisms: <https://www.facebook.com/policies/purchase_protection>.

seller. The buyer is given the opportunity to state their case and provide the details of the claim in the claim form. After the buyer has submitted his/her claim, the seller is not given an opportunity to submit a defence or adduce evidence to support his/her defence.[30] Rather, the platform will review the claim and the messages passed between the parties on the platform, and make a decision after only hearing from one party, and considering the messages sent in attempts of settlement. While Facebook provides a mechanism for sellers to appeal any decision,[31] and thus, perhaps akin to adducing a defence, this is undoubtedly too little too late for sellers; appropriate procedural safeguards should be guaranteed at the start of the process, rather than at the tail end of it. The importance of procedural law remains to ensure due process and fairness; that each individual receives the same treatment across the adjudication process. However, giving parties different rights at different stages of the proceedings would only serve to create tension between e-commerce business owners, and various customers on the market.

19 Further constraints arise in situations where procedural aspects are governed under EULAs. For example, on the Amazon Pay platform, when a claim is submitted by the buyer, the seller has to cooperate with that claim "in good faith".[32] It is unclear what such "good faith" refers to in this context and whether the duty of such an obligation would vary with the seriousness of the claim filed against the seller. This obligation of good faith is independent of the substantive content of the claim itself – while a poor defence submitted in good faith would only result in the seller losing the dispute, a defence submitted in bad faith would not only mean that the seller would lose the dispute, but also face severe penalties such as a restriction or termination of their account.[33] However, there are no such obligations on the buyer. Indeed, there is often nothing prohibiting the buyer from submitting multiple frivolous (or even fraudulent) claims against a seller in the hope that the platform might view one or more of these claims to be strong enough to overcome the seller's defence. At the same time, given the obligation of good faith on the seller, it is unsure whether the seller can respond to these claims against buyers in a dismissive manner since that may flout the vaguely worded obligation of "good faith".

20 On informal marketplaces, such procedural safeguards are, to an extent, alleviated through the implicit trust that users have in moderators. As opposed to administrators, moderators are merely users on the platform, and rarely have a financial incentive to decide the disputes in one way.[34] Hence, it seems that, on most informal marketplaces, moderators do consider the evidence in a holistic manner before making a final decision on the matter.[35] However, the question fundamentally remains as to whether this element of trust is sufficient in these circumstances, particularly as these forums are largely amorphous, and have flexible procedures. In addition, such 'trust' may entail users' belief in their moderator's competence to grant them the public acceptance of their authority to handle such disputes, rather than a mechanism that ensures that due process will be guaranteed in all disputes. Hence, safeguards should be in place to ensure due process, rather than trusting that due process will be granted, in such informal marketplaces.

2. Disproportionate Penalties

21 The last point which brings about inequality in the procedural rights lies in the harshness of remedies available for a parties' potential breach of due process requirements. Namely, that the failure by the seller to respond to a claim in a manner deemed proper by the platform would lead to a penalty that is disproportionate compared to that faced by the buyer reticent in providing information to sustain his/her claim.

30 Ibid. Notably however, on Facebook's marketplace, the policy reads: "*When using onsite checkout, if a seller or individual seller has not responded or resolved your issue after 2 business days, you can submit a claim for our review on the third business day. When you file a claim, answer the questions presented, and include details regarding your issue within the form. We'll review your claim, including any messages that you and the seller sent to each other along with supporting documentation from the buyer and the seller. We'll typically respond within 48 hours.*"

31 See here, more details regarding Facebook's policy regarding disputes at: <https://www.facebook.com/business/help/1167434420087941?id=353836851981351>. Accessed 31 July 2022.

32 See here, Amazon Pay's dispute policies available at: <https://pay.amazon.co.uk/help/201751580> Accessed 1 August 2022.

33 ibid.

34 See here, an analysis on eBay's fees for sales, as well as use of ODR mechanism at (n 13), 6.

35 See, for example here, a publication by the moderators of r/MechMarket on the parent SubReddit r/MechanicalKeyboards about an investigation around the Group Buy about the Lyra <https://www.reddit.com/r/MechanicalKeyboards/comments/nfnbau/warning_about_santigo_customs_lyra_monoflex_gb/>. Accessed 31 July 2022.

22 This is particularly problematic on established e-commerce sites, particularly when users depend on them for their livelihoods. On the Etsy platform for example, it is mandatory for sellers to "participate in a case against [their shops]". Similarly, on the Amazon Pay platform, if the seller does not "respond timely to a dispute or does not honour a commitment made to resolve a dispute within a reasonable amount of time", Amazon Payments may "place a hold on funds in a seller's account".[36] As such, it is rather evident that the penalties levied on the sellers far outpace those which are levied on the buyers for similar breaches of obligations. Indeed, these penalties are often levied on areas beyond the dispute itself (e.g. by striking out the seller's defence or finding the case in favour of the buyer in default) and involve matters relating to the seller's ability to continue their operations on the platform (e.g. existence on the platform or access to their funds or account on the platform). Even if the penalties levied on the seller and buyer in such cases are the same, the *effect* of the penalties on the sellers would still be, in the usual case, far heavier since many sellers on the platform are often are there "for the long run" and have built up not only a system of operations, but also commercial reputation for themselves. A suspension of their accounts, even if temporary, might mean disruption in their business and would bear a detrimental impact on their reputation. A detrimental impact on their ability to continue operations on such platforms would thus have a more severe impact on them as compared to a buyer on such platforms, who may only occasionally visit such platforms to purchase goods or services and can create a new account with relative ease.

23 Of course, it is not necessarily the case that due process is infringed just because the penalties on the sellers and buyers are unequal in the case of breach. Such a disparity between the treatment of the parties may be justified if it is proportionate to any legitimate aim sought. In the instant case, heavy penalties on the sellers may have a role to play in deterring potential fraudulent sellers from entering into an agreement to sell the goods without ultimately delivering said goods to the buyer. Ostensibly, fraudulent sellers do not challenge the buyer's claims since where the goods were not delivered, did not match the description, or were defective due to fraud, there is unlikely to be any serious defence or evidence to support such defences. Thus, placing harsh penalties on sellers who do not cooperate in the dispute resolution process may weed out fraudulent sellers by removing their ability to conduct their business on the platform or collect the money the buyer has paid.

24 However, while it might be reasonable to weed out potential fraudsters, such measures are disproportionate. First, by suspending the accounts of those who are slow to reply, the platform risks pre-judging sellers who may legitimately be slow to reply. This is especially the case since usually, sellers are only given a few days to reply to a potential dispute and may not be able to craft a defence, gather evidence, or even take notice of the fact that a claim has been formally entered against them.

25 Second, there is no need to take such drastic measures to deter potential fraudsters. If it is indeed the case that fraudsters are less likely to challenge claims brought forth by the buyers, it would be enough, in the interests of justice pertaining to the case, that the buyers are able to win their claims by default if the seller does not respond to the claim within a set amount of time. If the measures bearing impact beyond the specific dispute such as the suspension of an account due to suspicions of fraud are to be taken, they can, and should be taken where there is evidence of such fraud arising from the adjudication of the case, or where there is an established pattern of suspicious activity such as where there are multiple successful claims against the seller or where the seller has had a history of not responding to the claims against him/her. This way, the platform can balance between upholding the rule of law through upholding the equality of arms and still maintaining a robust anti-fraud regime. Such a model of anti-fraud monitoring is in fact put in place for the buyers on the Amazon Pay platform. Where the buyer submits three or more complaints that are subsequently ruled invalid by Amazon Payments, their account may be terminated[37]. It is evident that these platforms are capable of using such a system to deter fraud instead of relying on draconic sanctions on its users to deter fraud.

26 The situation varies for forum-based marketplaces. For one, many of these forums are built around enthusiasts of different things, ranging from board games (r/boardgamesexchange) to keyboards (r/mechmarket). The specialised and community-based nature of such forums breed an "intrinsic degree of trust"[38] between users and moderators of these forums, and parties are more comfortable discussing the case with the moderators, and moderators feel an increased degree of accountability[39]. Thus,

36 See here, Amazon's policies available at <https://pay.amazon.co.uk/help/201751580>.

37 ibid.

38 Casey Fiesler and others, 'Reddit Rules! Characterising an Ecosystem of Governance' (2018) 12 Conference on Web and Social Media 1. <https://ojs.aaai.org/index.php/ICWSM/article/view/15033> accessed 1 August 2022.

39 Joseph Seering and others, 'Moderator engagement and community development in the age of algorithms', (2019)

there is less of a need to impose highly restrictive penalties on sellers should they not respond within an extremely short time frame for fear of fraud given the increased degree of accountability by the moderators and the trust that has been built up amongst the users in the forum. Therefore, while sellers may still be banned for failing to cooperate with a dispute, there are no strict rules on the timeline according to which they should respond to such a dispute. Further, buyers now also have the responsibility of providing evidence to support their claims and similar punishments are levied on them should they fail to provide evidence to substantiate their claims[40]. It hence appears that the rules on forum-based marketplaces appear fairer to both parties, taking into circumstances of their unique predicament.

27 However, this situation is not ubiquitous across all forum-based marketplaces. On r/hardwareswap for example, moderators take the approach of "ban first, and ask questions later" when dealing with suspected scammers.[41] This goes further than many of the informal marketplaces in that the penalty is applied immediately where there is a dispute, and the burden of proof is on the seller to show that he/she is not engaging in fraud. Further, this is to be done at the moderator's discretion, and there are very few rules on what would cause a moderator to ban a user. The lack of uniformity and certainty between different forums and within a forum itself thus leaves much to be desired.

II. Substantive Matters

28 The substantive rules applied to a dispute also arguably run contrary to the rule of law due to a lack of clarity over what the exact rules are and how they are to be interpreted. In that regard, three points are thus noted. First, EULAs and various subsidiary rules are not comprehensive enough to cover all situations where a dispute within the parameters of the EULA may arise, and there is little information on how the existing rules are to be applied. Second, given that each platform essentially has its own *sui generis* set of rules that depart from broader principles of contract law, it is difficult to reconcile broader principles of traditional contract law to pinpoint what rules may apply in the event that the EULAs or the subsidiary rules are silent on an issue.

29 EULAs, while commonly regarded as infamous lengthy documents that are often ignored by users,[42] are said to govern user-behaviour, providing the 'dos and don'ts' across online platforms. However, they can be said to have a special place on online marketplaces, acting as the equivalent of a "constitution" to serve as the basic contract law principles for parties looking to contract on these platforms.[43] However, as with constitutions of sovereign nations, the EULAs and subsidiary rules on both established platform-based marketplaces and forum-based marketplaces are insufficiently comprehensive enough to cover all the situations where a dispute may arise.

30 Indeed, EULAs are, by nature, limited documents, and one cannot expect drafters to cover all possible circumstances which may arise. That would undoubtedly be unfeasible, and impractical. It is thus best left to the dispute-resolution platform equipped to handle cases on its merits, as perhaps best reflected in Courts of law in sovereign nations. However, the same cannot be said for ODR platforms on e-commerce sites, particularly given the significant uncertainty and opaqueness of ODR mechanisms. Perhaps in that vein, EULAs can be said to have greater importance on ODR platforms. However, its incompleteness, as well as vagueness of the basis of its decisions present issues for both consumers and vendors. On the Amazon platform for example, buyers may obtain a refund or exchange of an item if it is "materially different" from what the buyer has described. While there are provisions that state situations where a goods may be "materially different" from what the buyer has described, Amazon Pay has recognised that this checklist is non-exhaustive and may not cover all scenarios.[44] While traditional jurisprudence in major jurisdictions would provide some guidance in normal courts,[45] such criteria remain fundamentally

21 New Media and Society 7, 1417-1443. <https://journals.sagepub.com/doi/abs/10.1177/1461444818821316> accessed 29 July 2022.

40 Failure to Provide this Evidence by EITHER PARTY can result in Permanent Ban from the trading platform: <https://www.reddit.com/r/mechmarket/wiki/rules/rules#wiki_disputes>

41 See here, for example, on the informal marketplace Hardware Swap on Reddit: <https://www.reddit.com/r/hardwareswap/wiki/rules/rules>

42 Yannis Bakos, Florencia Marotta-Wurgler and David R Trossen, "Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts", (2014) 43 *Journal of Legal Studies* 1, <https://doi.org/10.1086/674424> Accessed 26 July 2022.

43 (n 14).

44 (n 7).

45 For an analysis of comparative contract law, and interpretation of the term 'material difference', see: Arthur Taylor von Mehren, 'The "Battle of the Forms":

unclear across Amazon's ODR platform. One then inevitably wonders if it is the case that all defects would be considered material no matter how minute they may be or even if they do not pertain to the utility, value, or even aesthetic of the product for a return or an exchange to be triggered. The lack of clarity over the issue of "materiality" in the defect is common across other platforms such as eBay and Etsy. Similarly, this is also the case for forum-based, informal marketplaces. On r/mechmarket for example, the right of rejection is available for "defects" and "damage".[46] However, there is no clear indication of whether this needs to be a "material defect" or "material damage" for the right of rejection to be triggered, or whether any defect or damage would trigger the right to rejection. Such uncertainty in the substantive matters of ODR claims on various platforms would inevitably create a degree of confusion among users. In that regard, legal certainty seems to be undermined in these areas.

31 Further, it is unclear what the rules of interpretation are on these platforms. In major legal jurisdictions around the world, the parol evidence rule exists in different forms to bar the use of pre-contractual negotiations in the interpretation of the contract.[47] However, on platforms such as Facebook Marketplaces and r/BoardGameExchange, the conversation between the buyer and seller may be admitted as evidence in a dispute.[48] Similarly, it is unclear whether other e-commerce platforms have access to the messages between users in a similar light, and whether weight is afforded to these communications. Given that the rules on these platforms appear to potentially deviate significantly from the rules found in major jurisdictions worldwide and are silent on how exactly these rules are to apply, there is potential legal uncertainty in the rules, which militates against the rule of law.

32 The situation is far more pronounced in forum-based marketplaces. On such marketplaces, there is a distinction between different types of obligations that may arise. Popular obligations may include an ordinary sale of goods, "giveaways" (which creates an obligation on the giver to give away the product despite there being no consideration passing between the parties),[49] deals (an obligation to provide a discount on the goods offered) or sharing deals, which similarly include Group Buys (an obligation to purchase goods with another person to take advantage of a discount),[50] and fundraisers (an obligation to sell goods and donate the monies received to a charity).[51] These obligations are entered into under very formulaic conditions and have their own sui generis rules applying to them that parties cannot contract out of. As such, it appears that there is no singular rule of obligations in forum-based marketplaces, but rather a Birksian "archipelago" of different obligations with their own rules applying in such marketplaces. Such an archipelagic array of obligations do not mirror the various contractual obligations found in different jurisdictions around the world since they are formed in the unique circumstances of forum-based marketplaces, and are only applicable in those circumstances. At the same time, information on the application of these rules are scant and it is unsure what each of these different obligations entail when a user moves from one forum to another. This thus poses another set of challenges for legal certainty and the rule of law.

33 Of course, it is not the case that for the rule of law to be upheld all rules and laws must be laid down in stone before a contract is entered into. In some jurisdictions, the law is developed through a "gradual expansion" upon the adjudication of individual cases and such systems are nonetheless still regarded as certain enough to uphold the rule of law. These platforms, however, are not of the same ilk. Decisions made in individual cases are not published on these platforms such that it is not possible to infer from these cases what the rules applied are. The incompleteness of the rules on these platform is recognised by Amazon, which has stated that the platform will "ultimately determine material difference at [its] discretion"[52]. Similarly, on r/BoardGamesExchange, it is emphasised that "Should any ambiguous scenario arise, the Mods

A Comparative View', (1990) 38 *The American Journal of Comparative Law* 2, 265-298.

46 See here, at: <https://www.reddit.com/r/mechmarket/wiki/buying>.

47 Tony Cole, 'The Parol Evidence Rule: A Comparative Analysis and Proposal', (2003) 26 UNSW Law Journal 2, 680-703.

48 For Facebook Marketplace, see <https://www.facebook.com/business/help/1167434420087941?id=353836851981351>. Similarly, on r/BoardGamesExchange, see the rules available at: <https://www.reddit.com/r/BoardGameExchange/wiki/scam_awareness>.

49 See, for example, 'Giveaway' posts on r/MechMarket at: <https://old.reddit.com/r/mechmarket/comments/vcjz4s/giveaway_tofu60_gold_case_hotswap_pcb_switches/>

50 See here, an example of a Group Buy: <https://old.reddit.com/r/mechmarket/comments/wfinv0/gb_good_or_evil_rubberhose_by_deskpads_gallery/> .

51 See here, for instance, at <https://old.reddit.com/r/mechmarket/comments/t746t3/fundraiser_mechmarket_ukraine_crisis_relief/>.

52 (n. 8).

will deliberate and will have final say over the resolution." However, this seems to present itself as an excessive use of discretion as a 'gap-filling mechanism'. While the retroactive characteristics of the law would allow these decisions to build on one another and create firm rules for the future, such emphasis on discretion inadvertently only creates inconsistencies through adjudication. Hence, resulting in further frustrations among users and hampening the development of a possible online rule of law.

D. Problems

34 Having identified such issues across various ODR platforms is the first step towards resolving such matters. However, in proposing any solutions, it must be recognised that there are similarly certain limitations on the implementation of any feasible solution. To that end, this section first considers such limitations and concurrently addresses the current proposed regulatory framework across the world. Finally, we propose a set of solutions in light of these constraints.

I. ODR Constraints and current regulatory initiatives.

35 One core prominent feature of ODR lies in the ease of implementation. Yet, there are fears that any reform or regulatory initiative would overcomplicate ODR platforms. From the perspective of the layperson, the more complex an ODR system becomes, the less accessible and more time consuming ODR becomes. Online disputants are highly focused on efficiency; empirical studies have indicated users would prefer to lose a case over a few days, than win a case over a few weeks.[53] An overly complex system would thus require system administrators or 'tribunals' to be overburdened with formalities or in reviewing extensive evidence adduced by parties. This further impedes any appropriate dispensing of an effective remedy. Similarly, most mechanisms are platform specific and purport to operate as the only available recourse; an overcomplication may even result in potential disputants dropping cases in light of these complications. Such overcomplications create favourable conditions for respondents, which creates contradictions within the fundamental purpose of site-specific ODR.

36 This similarly follows the work of UNCITRAL Working Group 2's policy recommendations for ODR systems. While their work and initiatives of drafting a uniform code for ODR platforms is to be commended,[54] the implementation of international instruments across site-specific ODR platforms presents too high a hurdle. Indeed, to layperson users, these instruments present themselves as 'confusing legalese', which is rarely fully read and understood. However, a further issue can be identified where these instruments operate and run contrary to customs and traditions found on platforms. For instance, the UNCITRAL Working Group has written extensively about the incorporation of various international commercial codes such as the CISG or PICC, and methods to obtain user consent.[55] However, as has previously been pointed out, the specialist knowledge required to implement these doctrines remains too high a barrier for administrators to effectively dispense justice under such instruments.[56] It should further be noted that Article 2(1) of the CISG expressly indicates that the Convention would not apply to goods for personal use, reflecting the buyers' intention at the time of conclusion of the contract. From the *travaux preparatoire*, the International Commercial Court notes this provision was required for the CISG to be acceptable to many States.[57] Thus, it would be difficult for States to accept any potential amendment derogating from this provision, merely to extend the CISG to e-commerce.

37 However, apart from international instruments, there has been growing relevance of regional instruments which seek to regulate ODR mechanisms. Of note, Article 17 of the European Union's 'E-Commerce Directive' presents a unique take towards ODR - requiring member-states to adopt adequate procedural guarantees in ODR claims.[58] This ground-up approach however, has seen little success. Particularly, critics note the vagueness of what 'procedural guarantees' are defined as

53 Arno Lodder, John Zeleznikow, *Enhanced Dispute-Resolution Through the Use of Information Technology*, (2010, Cambridge University Press), 1-32.

54 United Nations Commission on International Trade Law Fifty-fourth session. "Legal issues related to the digital economy – dispute resolution in the digital economy" 2021. A/CN.9/1064/Add.4

55 Ibid, at 23.

56 (n 13).

57 United Nations Conference on Contracts for the International Sale of Goods. "Documents of the Conference and SummaryRecords of the Plenary Meetings and of the Meetings of the Main Committees." A.CONF.97/19. (Vienna, 10 March - 11 April 1980)

58 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce')

within ODR claims, and what these notions relate to.[59] In a similar vein, the E-commerce Directive Assessment Report does not query the applicability of these mechanisms. While it does cite the *Cornelius de Visser*[60] judgement noting that internal market clauses do not apply where the service provider is unknown,[61] such a position is rather unsatisfactory. Merely looking at the territorial applicability of such directives creates significant potential for abuse of non-compliance. Further, there is significant uncertainty over the classifications of 'e-commerce' to which the ODR directive seeks to govern. While these regulations would alleviate concerns for what perhaps is perceived as traditional B2C e-commerce by established business on their dedicated platform, what the Directive neglects to consider lies in both established marketplaces, and informal marketplaces. A further distinction should also be made between businesses which utilise these platforms as an extension of their services, and individuals who perhaps have one-off sales; requiring such individuals to comply with such formalities would indeed result in significant backlash.

38 In that light, the bloc's modernisation attempts through the recently proposed Digital Service Act 2020 showcases a more troubling interventionist approach taken towards online platforms. Particularly relevant within ODR, lies in Article 17, 22, and 23.[62] Article 17 and 23 requires online platforms to produce reports for ODR users about decisions taken. Article 22 limits Union-based users access to the platform only when personal details are provided, including their name, address, telephone numbers (Article 22(a)), and bank account details of the trader Article 22(c)). While these extensions do not alleviate the issues raised above, it seemingly makes things worse. In particular, the extensiveness of these new provisions would create a degree of concern among intermediary store fronts, and matters for compliance. Of note, the references to 'online platforms' create a degree of uncertainty. As discussed, the distinction between various forms of online and offline marketplaces would create fundamental issues in itself. While established marketplaces may be able to comply with such provisions, informal forum-based marketplaces likely lack the infrastructure to do so; the storing of such personal data, especially bank account numbers, would require significantly greater infrastructure development on those sites for systems more than merely storing log-in details of accounts. In a similar vein, informal marketplaces would in itself require a formalised system of ODR to comply with transparency and reporting mechanisms to govern the 'marketplaces' which have developed on those sites. The terminology of 'trader' also remains ambiguous within the directive. On both established and informal forums, there are 'business accounts' which run as an extension of established business - businesses which use these platforms as a secondary means to marketing their products.[63] While the provisions of the directive would make sense to govern the practices of such businesses, they present a significant hurdle for individual users. Yet, some individuals who operate on these platforms, but maintain high volumes of trade and use these platforms as a 'full-time job' must similarly be distinguished from the 'one-off' trader. This is largely a threshold issue, but requires further clarity within legislation. It would similarly make sense for such formalised rules to apply to such established traders, but not for the layperson. Similarly, one wonders if the broader term of 'traders' would similarly apply to buyers

39 Lastly, the implementation of such a directive across online platforms would create inconsistencies across the rules governing sale agreements. The underlying nature of e-commerce lies in global trade. Hence, the imposition of such requirements would create a fundamentally different atmosphere for Union-based traders, and global traders operating under a different set of legislation. While such likely makes matters complex for online platforms, the more prevailing issue lies in applicable law when a Union-based trader and a non-union-based trader contracts for goods. The current solution avoids this through the implementation of EULA's to avoid such discrepancies on some marketplaces, to others largely ignoring these claims. Yet, EULAs themselves often do not extend to the sales contract

59 Pablo Cortes, *Online Dispute Resolution for Consumers in the European Union*, (2010, Taylor & Francis Group London).

60 C-292/10, ECLI:EU:C:2012:142.

61 Alexandre de Streel, Martin Husovec, 'The e-commerce Directive as the cornerstone of the Internal Market Assessment and options for reform' (May 2020, Policy Department for Economic, Scientific and Quality of Life Policies Directorate-General for Internal Policies, PE 648.797). <https://www.europarl.europa.eu/RegData/etudes/STUD/2020/648797/IPOL_STU(2020)648797_EN.pdf>, 19. Accessed 29 July 2022.

62 European Commission, *Proposal for a Regulation of the European Parliament and the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC*. (2020/0361 (COD), Brussels). <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=EN> Accessed 31 July 2022.

63 Mersut Savul, Ahmet Incekara and Sefer Sener, 'The Potential of E-commerce for SMEs in a Globalizing Business Environment'. (2014) 150 *Procedia - Social and Behavioral Sciences*, 35-45. <https://doi.org/10.1016/j.sbspro.2014.09.005> Accessed 27 July 2022.

between parties to an ODR, and merely remain as a regulatory framework for compliance purposes. Nonetheless, requiring compliance of these strict Union-based rules would surely open the door for other nations to legislate. Of note, the recent United States Supreme Court decision in *AT&T Mobility LLC v Concepcion*[64], where small claims 'arbitration' were discussed. These small-claims mechanisms are often likened to ODR across academic literature,[65] and the contrasting position between the US and the EU's ODR rules have been lengthily discussed.[66] In light of such concerns, the constraints with procedural mechanisms and safeguard thus arises once more. Conflicting standards and applicable laws remain at the forefront of any regional solution which can be proposed.

40 The scope of such ODR reforms must also be defined. While various e-commerce sites and forums utilise third party payment-services, such as PayPal to enforce chargebacks,[67] use of third-party sites presents a different challenge altogether. Particularly, chargeback mechanisms which credit-card companies can adopt. While buyers in traditional e-commerce disputes can utilise such mechanisms to obtain a refund, the chargeback policies remain at the discretion of such companies.[68] Indeed, there are often significant limitations in obtaining a credit card chargeback in e-commerce, attributed to the ambiguities surrounding a dispute. In that regard, 3 practical hurdles in e-commerce chargeback claims - quality discrepancies of descriptions versus item received, responsibility of return shipping cost, and timely delivery.[69] Indeed, the latter 2 remain uniquely related to e-commerce. As such, focus of ODR reform has to remain fundamentally within the realms of site-specific remedies. While such could similarly extend to situations when third party payment-service platforms are used (distinguishing them from banks with chargeback policies), further consideration must be had between the interaction between the different sites involved in claims.

41 A further consideration ties into the enforceability of ODR mechanisms in domestic courts. Most experts agree that another reason which may hamper the development of ODR is the legal uncertainty regarding ODR enforcement. While significant conversation has been had on whether enforcement is required,[70] The consensus seems to follow that ODR mechanisms expect compliance, but do not ensure compliance.[71] Notably, the OECD Code of Conduct for ODR Tribunals is silent on this issue as a whole.[72] It is in this vein that Elizabeth Thornberg has argued that governments should enforce these decisions as these online tribunals perform public functions;[73] thus, national court intervention assists the enforcement of a contractual settlement to maintain the utility of ODR services. Certainly, institutions such as the ICANN and the UDRP have worked collaboratively to incorporate an enforcement mechanism in domain-name related disputes.[74] In that vein, arguments have been made to develop such mechanisms to further ODR in consumer-related disputes.[75]

42 Thus, bringing a decision by an ODR tribunal to domestic Courts for enforcement may be ideal; the strong institutional support provided by Courts would indeed be of aid, particularly where certain remedies awarded are discretionary on the parties. This is particularly problematic on forum marketplaces which operate with user-trust, where lack of enforcement means fraudulent users merely face a platform ban as opposed to any compensatory damages. Hence, an enforcement mechanism would allow these traders recourse. However, user-

64 563 U.S. 333 (2011)

65 Amy J Schmitz, 'Evolution and Emerging Issues in Consumer Online Dispute Resolution (ODR)' (June 27, 2022). Ohio State Legal Studies Research Paper No. 714, <https://ssrn.com/abstract=4147917 or http://dx.doi.org/10.2139/ssrn.4147917> Accessed 30 July 2022.

66 Amy Schmitz, 'ODR to Address Exceptionalism in Arbitration' (2013, University of Colorado Law). <http://conferences.law.stanford.edu/codr2013/wp-content/uploads/sites/9/2016/09/Schmitz-Stanford_SchmitzHO.pdf> Accessed 29th July 2022.

67 (n 29).

68 ibid.

69 Lucille M Ponte, 'Boosting Consumer Confidence in E-Business: Recommendations for Establishing Fair and Effective DIspute Resolution Programs for B2C Online Transactions', (2002) 12 Albany Law Journal of Science and Technology 2, 441-492. <https://www.mediate.com/Integrating/docs/Abernethy.pdf> accessed 30th July 2022.

70 Jie Zheng, 'Enforcement of ODR Outcomes', in: Jie Zheng (2020) *Online Resolution of E-Commerce Directives*, 291-344.

71 Elizabeth G Thornburg, 'Fast, Cheap & Out of Control: Lessons from the Icann Dispute-Resolution Process, (2001) 7 Journal of Small & Emerging Business Law, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=321500> Accessed 28th July 2022.

72 Ester van den Heuvel, 'Online Dispute Resolution as a Solution to Cross-border E-Disputes) (2000) <https://www.oecd.org/digital/consumer/1878940.pdf> pg 22

73 (n. 70), 54.

74 Uniform Domain-Name Dispute Resolution Policy, Para 4(k).

75 (n 59), 82-83.

anonymity presents a significant bar here, especially when users are identified by their online personas as opposed to those in person. Hence, there are practical limitations in serving a claim to individuals where it becomes almost impossible to identify them. Further, one must nonetheless consider the practical effects of bringing an e-commerce claim on an e-commerce transaction. The high cost of litigation and legal fees, alongside the lengthy duration of trial should not be understated. Similarly, where these disputes are often cross-border, issues surrounding the interaction between the various laws, and the platform's EULA would inevitably arise. However, in the off-chance that claimants wish to pursue domestic litigation, it seems evident that domestic courts are prepared to handle such claims. In England, the *JK v MK*[76] the decision on the enforceability of ODR has shown a rather pragmatic and prudent approach. Mostyn J notes that where ODR platforms can show a set of due process rules (in *JK*, a lack of a conflict of interest by the ODR 'tribunal'), the Courts are willing to enforce the decision. While the case concerned an ODR divorce platform and Mostyn J was careful to limit this to similar platforms, it remains likely that the Courts enforce similar mechanisms.[77] Nonetheless, this exemplifies that enforceability of ODR decisions remains a largely moot point. Any competing claims of 'setting aside' an ODR tribunal's decision should remain at the discretion of domestic Courts, where claims are pursued.

II. Proposed solutions

43 In light of such constraints, this essay thus makes a few suggestions to reform potential ODR mechanisms, while respecting the unique systems and cultures that are prevalent on the different marketplaces. Ultimately, any reform to ODR systems should take a user-centric approach, prioritising user-friendliness and user-experience, when maintaining a sense of procedural and substantive fairness.

1. Systems of Classification

44 Prior to enacting legislation governing online platforms, it remains key to distinguish what 'online platforms' would we be referring to. As discussed, the various online marketplaces operate distinctly from one another. Established marketplaces and informal marketplaces have different forms of ODR mechanisms, levels of enforceability of decisions, and sale mechanisms. Particularly important however, lies in the vastly diverse cultures of users within these platforms, creating different user-experience within these platforms. While this is trite on established marketplaces, such as the distinction between eBay and Etsy,[78] it becomes even more prevalent within informal marketplaces. For instance, Facebook Marketplace users operate on a peer-peer basis, but utilise the social-media aspect of the platform (common friends, location, ability to view sellers' personal profile, etc) to create a sense of 'trust' among users.[79] Conversely, user-expectation on Reddit's marketplace forums are based primarily on party autonomy, coupled with significant moderator intervention across the board.[80] Yet, even within Reddit's numerous hobbyist marketplaces, cultural differences among users are present on different 'Subreddits'.[81] As such, operating a blanket definition of 'online platform' with similar obligations would create a significant degree of backlash among users.

45 Apart from user-expectations, the different roles and responsibilities of administrators on such sites would benefit from a degree of classification. As identified previously, while administrators on established marketplaces are often employees of a particular team, the situation is vastly different for platform moderators on informal platforms. It would be simpler for established corporations to 'train' employees to comply with legislative mechanisms. However, on informal platforms, moderators are often trusted members of a community, appointed by other more 'senior' moderators. While they may seem akin to employees, in reality, they are often volunteers, with no relation to the platform which these informal marketplaces operate on. Particularly, Reddit's EULA expressly notes that

76 [2020] EWFC 2.

77 (n 14), 12-13.

78 See here, a comparative study on these digital business platforms: Arvind Rangaswamy and others, 'The Role of Marketing in Digital Business Platforms' (2020) 51 *Journal of Interactive Marketing*, 72-90. <https://doi.org/10.1016/j.intmar.2020.04.006> Accessed 27th July 2022.

79 Ahmad Anshorimuslim Syuhada, 'Online Marketplace for Indonesian Micro Small and Medium Enterprises based on Social Media, (2013) 11 *Procedia Technology*, 446-454. <https://doi.org/10.1016/j.protcy.2013.12.214> Accessed 29th July 2022.

80 Hanlin Li, Brent Hecht, Stevie Chancellor, 'All that's happening behind the scenes: Putting the Spotlight on Volunteer Moderator Labor in Reddit' (2022) 16 Proceedings of the Sixteenth International AAAI Conference on Web and Social Media. <https://ojs.aaai.org/index.php/ICWSM/article/view/19317> Accessed 31 July 2022.

81 See, for example, between the communities at (n 18), (n 19), (n 20).

> "Moderating a subreddit is an unofficial, voluntary position that may be available to users of the Services. We are not responsible for actions taken by the moderators."[82]

46 Hence, imposing legislative reforms which require these moderators to perform certain obligations, or be potentially privy to sensitive information (as required by Article 17 of the Digital Service Act) would likely be too onerous. Similarly, regular users would also be sceptical if they are required to entrust such legal obligations to mere volunteers, or to the platform to share with these volunteers. Such would undoubtedly both users and prospective moderators away from the platform. It must however, be noted that the situation is very different across the vast majority of informal marketplaces which each having their own system, this disparity would only give rise to greater considerations.

47 To that end, there first needs to be a distinction between established e-commerce sites, established marketplaces, and informal marketplaces. The nuances that arise between both users and administrators of these sites should be recognised to more effectively

2. A base set of 'governing principles'

48 The development of a model ODR code-of-conduct to be implemented has been discussed at length, and proposed at various stages.[83] However, what remains core here lies in the lack of appropriate stakeholder consultation; particularly - that of the nature of these ODR administrators. Indeed, the development of a set of governing principles should incorporate the considerations of users on both formal and informal marketplaces,[84] and bear in mind that such provisions should be developed from the perspective of laypersons rather than legal practitioners. It seems to follow then that any set of guiding principles should be highly intuitive, with a focus on access to justice, and simplicity of implementation.

49 In that vein, procedural safeguards seem to be the core consideration of policy-makers and users - to ensure that their case has been heard appropriately. Perhaps this could be attributed to the age-old maxim that "*justice must not only be done, but seem to be done*."[85] Nonetheless, truer words cannot be said about the ODR process. It is in this vein that Colin Rule, head developer of the ODR system at eBay argues that efficiency, consistency and certainty to create public confidence should be the priority of any ODR system.[86] Yet, one must nonetheless further consider that ODR remains as the only available recourse for users in e-commerce. This moves ODR into a necessity as opposed to a feature. Hence, if allegations of bias are thrown around by users, this compromises on the public confidence which ODR platforms have been created for.

50 Therefore, the substance of the dispute must similarly be considered. Basic contract law principles such as offer and acceptance, simple breaches of duties, and fraud should be adopted to form the backbone for such conducts. Similarly, wider evidential matrixes not limiting the evidence should be incorporated into these platforms. Lastly, reasons for decisions by tribunals should be given out to parties, whether extensive reasons or merely a few lines of text. Such allows users to better trust the ODR process, and feel as though their cases have been adequately heard amidst a backdrop of substantial principles.

E. Conclusion

51 The role of e-commerce technologies and the impact on the global economy which it has brought about cannot be understated. Indeed, both businesses and consumers have taken to the internet, moving away from traditional brick and mortar stores. However, as with any growing economic landscape, disputes between businesses and users would inevitably arise.

52 Perhaps rather novelly, this paper drew attention to the different forms of e-commerce platforms that are used by consumers. Namely, traditional and established forums, but also, informal, and forum-based marketplaces. Undoubtedly, significantly more literature has been written on the former rather than the latter. However, it is hoped that this paper would mark the beginnings of greater studies on that front. Nonetheless, in completing a comparative analysis of these marketplaces and how ODR is handled across these platforms, it can be said that both models of e-commerce sites have largely inadequate safeguards for how ODR is handled, both procedurally, as well as substantively.

82 (n 21), Para 8.

83 For a list of the various ODR standards proposed, see: <https://odr.info/standards/>.

84 (n 14), 13.

85 *Rex v Sussex Justices* [1924] 1 KB 256 perr Lord Hewart.

86 Amy Schmitz and Colin Rule, *The New Handshake: Online Dispute Resolution and the Future of Consumer Protection* (2017, American Bar Association Section on Dispute Resolution), 44. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3106913> Accessed 29th July 2022.

This is made worse with the current trajectory of such ODR platforms that seems to present itself as a consumer-bias system, neglecting the position of sellers and disadvantaging them both procedurally and substantially. In that regard, policymakers and regulators have attempted to take the stage and rescue sellers, but also maintain the rights of buyers through various instruments. Yet, it seems that these mechanisms, while sound in principle, are largely disconnected from the wider user-base on these platforms, which leads to potentially greater issues arising on these e-commerce platforms and creating a wider divide in user-bases.

53 Ultimately, to these ends, this paper sought to address this through 2 core mechanisms - the classification of marketplaces and e-commerce sites, and the implementation of guiding principles and a generalised approach. These solutions presented strive to preserve the autonomy and characteristics of the various e-commerce sites which attract their user-base, while maintaining public confidence as well as a degree of legal certainty through fundamental principles of commerce. These solutions, while not concrete in nature, were designed as the first steps towards what could potentially be regarded as a harmonised framework for ODR, while maintaining the nuances across these platforms and preserving the intuitive, accessible, but similarly effective and efficient nature of ODR. In that vein, further research in this field is similarly welcome, particularly from sociological and economics perspectives and especially in the field of informal forum-based marketplace, to explore greater community sentiments towards how ODR is conducted.

Guardians of the UGC Galaxy – Human Rights Obligations of Online Platforms, Copyright Holders, Member States and the European Commission Under the CDSM Directive and the Digital Services Act

by **Martin Senftleben** *

Abstract: With the shift from the traditional safe harbour for hosting to statutory content filtering and licensing obligations in Article 17 of the CDSM Directive, EU copyright law imperils the freedom of users to upload and share their content creations. Seeking to avoid overbroad inroads into freedom of expression, EU law obliges online platforms and the creative industry to take into account human rights when coordinating their content filtering actions. Platforms must also establish complaint and redress procedures for users. The European Commission will initiate stakeholder dialogues to identify best practices. These "safety valves" in the legislative package, however, may prove to be mere fig leaves. Instead of safeguarding human rights, the EU legislator outsources human rights obligations to the platform industry. At the same time, the burden of policing content moderation systems is imposed on users who are unlikely to bring complaints in each individual case. The new legislative design may thus conceal human rights violations instead of bringing them to light. The Digital Services Act rests on a similar – equally problematic – approach. Against this backdrop, the analysis addresses the risk of human rights interference, which is exacerbated by the fact that the Court of Justice, in its Poland decision, upheld the regulatory approach underlying Article 17, rather than exposing and discussing the corrosive effect of human rights outsourcing. Luckily, the new rules in the CDSM Directive and the Digital Services Act also contain several safeguards that allow EU Member States and the European Commission to actively take measures against the erosion of human rights.

Keywords: copyright, freedom of expression, transformative use, parody, pastiche, human rights outsourcing, content moderation, proportionality, complaint and redress, audit reports

Recommended citation: Martin Senftleben, Guardians of the UGC Galaxy – Human Rights Obligations of Online Platforms, Copyright Holders, Member States and the European Commission Under the CDSM Directive and the Digital Services Act, 14 (2023) JIPITEC 435 para 1

A. Introduction

1 User-generated content ("UGC")[1] is a core element of many internet platforms. With the opportunity to upload photos, films, music and texts, formerly passive users have become active contributors to (audio-)visual content portals, wikis, online marketplaces, discussion and news fora, social networking sites, virtual worlds and academic paper repositories. Today's internet users upload a myriad

* Ph.D.; Professor of Intellectual Property Law and Director, Institute for Information Law (IViR), University of Amsterdam; Of Counsel, Bird & Bird, The Hague, The Netherlands.

1 For a definition and description of central UGC features, see OECD, 12 April 2007, "Participative Web: User-Created Content", Doc. DSTI/ICCP/IE(2006)7/Final, available at https://www.oecd.org/sti/38393115.pdf (last visited on 12 August 2023), 8-12.

of literary and artistic works every day.[2] A delicate question arising from this user involvement concerns copyright infringement. UGC may consist of self-created works and public domain material. However, it may also include unauthorized takings of third-party material that enjoys copyright protection. As UGC has become a mass phenomenon and a key factor in the evolution of the modern, participative web,[3] this problem raises complex issues and requires the reconciliation of fundamental rights ranging from the right to (intellectual) property[4] to freedom of expression and information, and freedom to conduct a business.[5] Users, platform providers and copyright holders are central stakeholders.[6]

2 With the adoption of Article 17 of the Directive on Copyright in the Digital Single Market ("CDSMD" or "CDSM Directive"),[7] specific EU legislation seeking to regulate the UGC galaxy has become a reality. Article 17 puts an end to the traditional notice-and-takedown system and the corresponding liability privilege for providers of hosting services.[8] Under Article 17(1) CDSMD, online content-sharing service providers ("OCSSPs")[9] are directly liable for infringing user uploads. To avoid liability risks, they must enter into agreements with copyright owners. In practice, this regulatory approach leads to the application of an amalgam of licensing and filtering obligations.[10] If an OCSSP does not manage to conclude sufficiently broad licensing agreements with rightholders in line with Article 17(1) and (4)(a) CDSMD, Article 17(4)(b) and (c) CDSMD offers the prospect of a reduction of the liability risk in exchange for content filtering. The OCSSP can avoid liability for unauthorized acts of communication to the public or making available to the public when it manages to demonstrate that it:

> "made, in accordance with high industry standards of professional diligence, best efforts to ensure the unavailability of specific works and other subject

2 For example, statistics relating to the online platform YouTube report over one billion users uploading 500 hours of video content every minute. Cf. https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/#:~:text=This%20equates%20to%20approximately%2030%2C000,for%20online%20video%20has%20grown (last visited on 12 August 2023).

3 OECD, supra note 1, 8-22.

4 Article 17(2) CFR.

5 Articles 11 and 16 CFR. Cf. CJEU, 16 February 2012, case C-360/10, Sabam/Netlog, para. 51.

6 As to the debate on user-generated content and the need for the reconciliation of divergent interests in this area, see M.R.F. Senftleben, "Breathing Space for Cloud-Based Business Models – Exploring the Matrix of Copyright Limitations, Safe Harbours and Injunctions", *Journal of Intellectual Property, Information Technology and E-Commerce Law* 4 (2013), 87 (87-90); M.W.S. Wong, "Transformative User-Generated Content in Copyright Law: Infringing Derivative Works or Fair Use?", *Vanderbilt Journal of Entertainment and Technology Law* 11 (2009), 1075; E. Lee, "Warming Up to User-Generated Content", *University of Illinois Law Review* 2008, 1459; B. Buckley, "SueTube: Web 2.0 and Copyright Infringement", *Columbia Journal of Law and the Arts* 31 (2008), 235; T.W. Bell, "The Specter of Copyism v. Blockheaded Authors: How User-Generated Content Affects Copyright Policy", *Vanderbilt Journal of Entertainment and Technology Law* 10 (2008), 841.

7 Directive (EU) 2019/790 of the European Parliament and of the Council 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 of the European Communities* 2019 L 130, 92.

8 Article 17(3) CDSMD. For a discussion of this regulatory approach, see M.R.F. Senftleben, "Institutionalized Algorithmic Enforcement – The Pros and Cons of the EU Approach to Online Platform Liability", *Florida International University Law Review* 14 (2020), 299 (308-312); N. Elkin-Koren, "Fair Use by Design", *UCLA Law Review* 64 (2017), 1082 (1093); M. Husovec, "The Promises of Algorithmic Copyright Enforcement: Takedown or Staydown? Which Is Superior? And Why?", *Columbia Journal of Law and the Arts* 42 (2018), 53 (76-84).

9 Article 2(6) and Recitals 62, 63 CDSMD. Cf. A. Metzger/M.R.F. Senftleben, "Understanding Article 17 of the EU Directive on Copyright in the Digital Single Market – Central Features of the New Regulatory Approach to Online Content-Sharing Platforms", *Journal of the Copyright Society of the U.S.A.* 67 (2020), 279 (284-286).

10 M.R.F. Senftleben, "Bermuda Triangle: Licensing, Filtering and Privileging User-Generated Content Under the New Directive on Copyright in the Digital Single Market", *European Intellectual Property Review* 41 (2019), 480 (481-485); M. Husovec/J.P. Quintais, "How to License Article 17? Exploring the Implementation Options for the New EU Rules on Content-Sharing Platforms under the Copyright in the Digital Single Market Directive", *Gewerblicher Rechtsschutz und Urheberrecht – International* 70 (2021), 325; M. Leistner, "European Copyright Licensing and Infringement Liability Under Art. 17 DSM-Directive Compared to Secondary Liability of Content Platforms in the U.S. – Can We Make the New European System a Global Opportunity Instead of a Local Challenge?", *Zeitschrift für Geistiges Eigentum/Intellectual Property Journal* 12 (2020), 123 (123-214); C. Geiger/B.J. Jütte, "Towards a Virtuous Legal Framework for Content Moderation by Digital Platforms in the EU? The Commission's Guidance on Article 17 CDSM Directive in the light of the YouTube/Cyando judgement and the AG's Opinion in C-401/19", *European International Property Review* 43 (2021), 625 (625-635).

matter for which the rightholders have provided the service providers with the relevant and necessary information,..."[11]

3 Although the provision contains neutral terms to describe this scenario, there can be little doubt in which way the "unavailability of specific works and other subject matter" can be achieved: the use of algorithmic filtering tools seems inescapable.[12]

4 In the legislative process leading to this remarkable climate change in the EU, the human rights impact of the departure from the traditional notice-and-takedown model has not gone unnoticed. The wording of Article 17 CDSMD itself shows that the new legislative design gave rise to concerns about encroachments upon human rights and, in particular, freedom of expression and information. Article 17(10) CDSMD stipulates that, in stakeholder dialogues seeking to identify best practices for the application of content moderation measures, "special account shall be taken, among other things, of the need to balance fundamental rights and of the use of exceptions and limitations."[13] After the adoption of the CDSM Directive, the preparation of the Digital Services Act ("DSA")[14] offered further opportunities for the EU legislature to refine and stabilize its strategy for safeguarding human rights that may be affected by algorithmic content filtering tools. Article 14 DSA – regulating terms and conditions of intermediary services ranging from mere conduit and caching to hosting services[15] – reflects central features of the EU strategy. Article 14(1) DSA requires that providers of hosting services – the category covering UGC platforms – inform users about:

> "any policies, procedures, measures and tools used for the purpose of content moderation, including algorithmic decision-making and human review, as well as the rules of procedure of their internal complaint handling system."[16]

5 This information duty already indicates that users are expected to play an active role in the preservation of their freedom of expression and information. Article 14(4) DSA complements this transparency measure with a fundamental rule that goes far beyond sufficiently clear and accessible information in the terms and conditions. Providers of intermediary services, including platforms hosting UGC:

> "shall act in a diligent, objective and proportionate manner in applying and enforcing the restrictions [that they impose in relation to the use of their service in respect of information provided by the recipients of the service], 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."[17]

6 In other words: in the case of upload and content sharing restrictions following from the employment of content moderation tools, the UGC platform is bound to safeguard the fundamental rights of users, including freedom of expression and information. As a guiding principle, Article 14(4) DSA refers to the principle of proportionality ("proportionate manner")[18] that plays a central role in the reconciliation of competing fundamental rights under Article 52(1) of the EU Charter of Fundamental Rights ("Charter" or "CFR").[19]

7 At first glance, it makes sense to impose the obligation to safeguard fundamental rights of users on UGC platforms. In *UPC Telekabel Wien*, the CJEU already laid groundwork for this approach. Discussing website blocking orders, the Court stated that, when an internet service provider was subject to an injunction requiring the blocking of a website whose users notoriously infringed copyright, it had

11 Article 17(4)(b) CDSMD.

12 See CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 53, where this assumption has been confirmed.

13 Article 17(10) CDSMD.

14 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), *Official Journal of the European Union* 2022 L 277, 1.

15 See the definition of "intermediary services" in Article 3(g) DSA.

16 Article 14(1) DSA.

17 Article 14(4) DSA.

18 Article 14(4) DSA.an

19 Charter of Fundamental Rights of the European Union, *Official Journal of the European Communities* 2000 C 364, 1. Article 52(1) CFR reads as follows: "Any limitation on the exercise of the rights and freedoms recognised by this Charter must be provided for by law and respect the essence of those rights and freedoms. Subject to the principle of proportionality, limitations may be made only if they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others."

to ensure compliance with the fundamental right of internet users to freedom of information.[20] More specifically, the measures adopted by the internet service provider had to be strictly targeted, in the sense that they had to serve to bring an end to a third party's infringement of copyright "but without thereby affecting internet users who are using the provider's services in order to lawfully access information."[21] The Court added that, failing the implementation of a sufficiently targeted blocking mechanism, the provider's interference in the freedom of information would be unjustified in the light of the objective pursued.[22] Considering this earlier case law, the task of safeguarding fundamental rights of users is thus neither new nor surprising for internet service providers.

8 The crux of the approach chosen in Article 14(4) DSA, however, clearly comes to the fore when raising the question whether the possibility of imposing human rights survival obligations on internet service providers, such as UGC hosting platforms,[23] exempts the state power itself from the noble task of ensuring the observance of fundamental rights. Can the legislator legitimately outsource the obligation to safeguard fundamental rights, such as freedom of expression and information, to private parties? And can the legislator – when passing on that responsibility – confidently leave the task of defending the public interest in this sensitive area in the hands of companies belonging to the platform and creative industry? Arguably, an outsourcing strategy, such as the strategy reflected in Article 17(4)(b) and (c) CDSMD and Article 14(1) and (4) DSA, is highly problematic if it is not accompanied by robust and reliable control mechanisms that allow public authorities to verify the effectiveness of the measures taken by the private party concerned (content sharing platforms in the case of UGC) and the alignment of these measures with the broader public interest (following section II). Instead of focusing on control by public authorities, however, EU legislation leaves measures against excessive content blocking primarily to users (section III). The Member State obligation to safeguard quotations, parodies and pastiches etc. in Article 17(7) CDSMD and the audit system established in Article 37 DSA are welcome exceptions to this rule (section IV).

B. Outsourcing of Human Rights Obligations

9 As already indicated, legislation that applies outsourcing strategies refrains from providing concrete solutions for human rights tensions in the law itself. Instead, the legislator imposes the burden on private entities to safeguard human rights that may be affected by the legislative measure at issue, such as the statutory content filtering obligation in Article 17(4)(b) and (c) CDSMD. In the case of UGC, the addressees of this type of outsourcing legislation are online platforms – OCSSPs – that offer users a forum for uploading and sharing their creations. Discussing the increasing tendency to take refuge in human rights outsourcing, Tuomas Mylly has observed that "gradually, intermediaries and other key private entities become more independent regulators."[24] He describes central characteristics of this process as follows:

> "Courts are starting to rely increasingly on private entities to balance and adjust rights on technological domains but seek to secure formal appeal rights for users. Similarly, when legislatures shift decision-making power to intermediaries, they try to maintain some of the safeguards of traditional law and write wish-lists for private regulators. The executive pushes private regulation further to compensate for its policy failures and enters – at the request of the legislature – into regulatory conversations with private regulators to issue "guidance" in the spirit of co-regulation, thus establishing an enduring link to private regulators."[25]

10 Arguably, Article 17 CDSMD and Article 14 DSA offer prime examples of provisions that outsource human rights obligations to private entities – with the features Mylly describes. As explained above, Article 14(4) DSA places an obligation on intermediaries to apply content moderation systems in "a diligent, objective and proportionate manner."[26] In addition to this reference to the principle of proportionality, the provision emphasizes that online platforms are bound to carry out content filtering with due regard to the fundamental rights of users, such as freedom

20 CJEU, 27 March 2014, case C-314/12, UPC Telekabel Wien, para. 55.

21 CJEU, id., para. 56.

22 CJEU, id., para. 56.

23 See the concept of hosting services in Article 3(g)(iii) DSA.

24 T. Mylly, "The New Constitutional Architecture of Intellectual Property", in: J. Griffiths/T. Mylly (eds.), *Global Intellectual Property Protection and New Constitutionalism – Hedging Exclusive Rights*, Oxford: Oxford University Press 2021, 50 (71).

25 Mylly, supra note 24, 71.

26 Article 14(4) DSA. Article 14(1) DSA explicitly refers to content moderation measures.

of expression.[27] With regard to copyright limitations that support freedom of expression,[28] more specific rules follow from specific copyright legislation. According to Article 17(7) CDSMD, the cooperation between OCSSPs and the creative industry in the area of content moderation[29] must not result in the blocking of non-infringing UGC, including situations where UGC falls within the scope of a copyright limitation. Confirming Mylly's prediction that the executive power will enter into regulatory conversations with private entities to establish best practices and guiding principles, Article 17(10) CDSMD adds that the European Commission shall organize stakeholder dialogues to discuss best practices for the content filtering cooperation:

> "The Commission shall, in consultation with online content-sharing service providers, rightholders, users' organisations and other relevant stakeholders, and taking into account the results of the stakeholder dialogues, issue guidance on the application of this Article, in particular regarding the [content moderation] cooperation referred to in paragraph 4."[30]

11 In the quest for best practices, the stakeholder dialogues shall take "special account"[31] of the need to balance fundamental rights and the use of copyright limitations. As in Article 14(4) DSA, reference is thus made to human rights tensions. The private entities involved – copyright holders and OCSSPs – are expected to resolve these tensions in the light of the guidance evolving from the co-regulatory efforts of the European Commission.

12 Evidently, industry "cooperation" is the kingpin of this outsourcing scheme for human rights obligations. To fully understand risks that may arise from this regulatory approach, it is important to analyse Article 17 CDSMD in more detail. At the core of the obligation to filter UGC – and industry cooperation that is necessary to implement this obligation in practice – lies the grant of a specific exclusive right in Article 17(1) CDSMD that leads to strict, primary liability of OCSSPs for infringing content that is uploaded by users:

> "Member States shall provide that an online content sharing service provider performs an act of communication to the public or an act of making available to the public when it gives the public access to copyright protected works or other protected subject matter uploaded by its users."[32]

13 By clarifying that the activities of UGC platform providers amount to communication to the public or making available to the public, the new legislation collapses the traditional distinction between primary liability of users who upload infringing content, and secondary liability of online platforms that encourage or contribute to infringing activities. Under Article 17(1) CDSMD, it no longer matters whether the provider of a UGC platform had knowledge of infringement, encouraged infringing uploads or failed to promptly remove infringing content after receiving a notification. Instead, the platform provider is directly and primarily liable for infringing content that arrives at the platform.

14 In this way, EU legislation incentivizes rights clearance initiatives. To reduce the liability risk, the platform provider will have to obtain a license for UGC uploads. Evidently, this is an enormous task. Even though it is unforeseeable which content users will upload, the license should ideally encompass the whole spectrum of potential posts. While this dimension of the licensing obligation may be good news for users (whose activities would fall within the scope of the license and, therefore, no longer amount to infringement),[33] it creates a rights clearance task which platform providers can hardly ever accomplish in respect of all conceivable user contributions.[34]

15 Inevitably, the licensing imperative chosen in Article 17(1) CDSMD culminates in the introduction of filtering tools. As copyright holders and collecting societies are unlikely to offer all-embracing umbrella licenses,[35] OCSSPs must rely on algorithmic tools to ensure that content uploads do not overstep the limits of the use permissions they managed to obtain.[36]

27 Article 14(4) DSA.

28 CJEU, 1 December 2011, case C-145/10, Painer, para. 132; CJEU, 3 September 2014, case C-201/13, Deckmyn, para. 26. See also CJEU, 29 July 2019, case C-476/17, Pelham, para. 32, 37 and 59.

29 See the interplay of creative industry notifications and filtering measures applied by the platform industry that results from Article 17(4)(b) and (c) CDSMD.

30 Article 17(10) CDSMD.

31 Article 17(10) CDSMD.

32 Article 17(1) DSMD.

33 Article 17(2) CDSMD.

34 Cf. M.R.F. Senftleben, "Content Censorship and Council Carelessness – Why the Parliament Must Safeguard the Open, Participative Web 2.0", *Tijdschrift voor Auteurs-, Media- & Informatierecht* 2018, 139 (141-142).

35 Cf. Senftleben, supra note 8, 305-307.

36 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para 53.

From the perspective of freedom of expression and information, this amalgam of licensing and filtering is highly problematic.[37] Outside the licensing deals which UGC platforms have concluded, algorithmic enforcement measures will curtail the freedom of users to participate actively in the creation of online content.

16 The more specific regulation of content moderation in Article 17 CDSMD confirms that the EU legislator has willingly accepted inroads into freedom of expression and information to achieve the goal of subordinating UGC to the control of copyright holders. As explained, the law does not shy away from imposing institutionalized – statutory – content filtering obligations.[38] In the absence of licensing arrangements, Article 17(4)(b) and (c) CDSMD offers OCSSPs the prospect of a reduction of the liability risk in exchange for content filtering. The fundamental rights tension caused by this regulatory approach is evident. In decisions rendered prior to the adoption of Article 17 CDSMD, the CJEU has stated explicitly that in transposing EU directives and implementing transposing measures:

> "Member States must [...] take care to rely on an interpretation of the directives which allows a fair balance to be struck between the various fundamental rights protected by the Community legal order."[39]

17 Interestingly, the application of filtering technology to a social media platform hosting UGC already occupied centre stage in *Sabam/Netlog*. The case concerned Netlog's social networking platform, which offered every subscriber the opportunity to acquire a globally available "profile" space that could be filled with photos, texts, video clips etc.[40] Claiming that users make unauthorized use of music and films belonging to its repertoire, the collecting society Sabam sought to obtain an injunction obliging Netlog to install a system for filtering the information uploaded to Netlog's servers. As a preventive measure and at Netlog's expense, this system would apply indiscriminately to all users for an unlimited period and would have been capable of identifying electronic files containing music and films from the Sabam repertoire. In case of a match, the system would prevent relevant files from being made available to the public.[41] Given these underlying facts, the *Sabam/Netlog* case offered the CJEU the chance to provide guidance on a filtering system that has become a standard measure with the adoption of Article 17(4)(b) CDSMD.[42]

18 However, the CJEU did not arrive at the conclusion that such a filtering system could be deemed permissible. Instead, the Court saw a serious infringement of fundamental rights. It took as a starting point the explicit recognition of intellectual property as a fundamental right in Article 17(2) CFR. At the same time, the Court recognized that intellectual property must be balanced against the protection of other fundamental rights and freedoms.[43] Weighing the right to intellectual property asserted by Sabam against competing fundamental rights of Netlog's users, namely their right to the protection of their personal data and their freedom to receive or impart information,[44] The Court recalled that the use of protected material in online communications may be lawful under statutory limitations of copyright in the Member States, and that some works may have already entered the public domain, or been made available for free by the authors concerned.[45] Given this corrosive effect on fundamental rights, the Court concluded:

> "Consequently, it must be held that, in adopting the injunction requiring the hosting service provider to install the contested filtering system, the national court concerned would not be respecting the requirement that a fair balance be struck between the right to intellectual property, on the one hand, and the freedom to conduct business, the right to protection of personal data and the freedom to receive or impart information, on the other (see,

37 For a more candid statement, see M.R.F. Senftleben, "The Original Sin – Content 'Moderation' (Censorship) in the EU", *Gewerblicher Rechtsschutz und Urheberrecht – International* 69 (2020), 339-340.

38 For a more detailed discussion of this development, see Senftleben, supra note 8, 299-328; Elkin-Koren, supra note 8, 1093.

39 CJEU, case C-275/06, Productores de Música de España (Promusicae)/Telefónica de España SAU, para. 68.

40 CJEU, 16 February 2012, case C-360/10, Sabam/Netlog, para. 16-18.

41 CJEU, ibid., para. 26 and 36-37.

42 As to the different levels of content monitoring that can be derived from CJEU jurisprudence, see M.R.F. Senftleben/C. Angelopoulos, *The Odyssey of the Prohibition on General Monitoring Obligations on the Way to the Digital Services Act: Between Article 15 of the E-Commerce Directive and Article 17 of the Directive on Copyright in the Digital Single Market*, Amsterdam: Institute for Information Law/Cambridge: Centre for Intellectual Property and Information Law 2020, 7-16.

43 CJEU, ibid., para. 41-44.

44 Articles 8 and 11 CFR. See CJEU, ibid., para. 48-50.

45 CJEU, ibid., para. 50.

by analogy, Scarlet Extended, paragraph 53)."[46]

19 This case law confirms that the filtering obligation arising from Article 17(4)(b) CDSMD is highly problematic. As a way out of the dilemma, the EU legislature walks the fine line of distinguishing between monitoring all UGC in search of a whole repertoire of works,[47] and monitoring all UGC in search of specific, pre-identified works.[48] *Sabam/Netlog* concerned a filtering obligation targeting all types of UGC containing traces of works falling within the Sabam rights portfolio.[49] The drafters of Article 17(4)(b) CDSMD seem to make an attempt to avoid this prohibited general monitoring obligation (and escape the verdict of a violation of fundamental rights) by establishing the obligation to filter "specific works and other subject matter for which the rightholders have provided the service providers with the relevant and necessary information."[50]

20 At this point, the above-described element of industry cooperation enters the picture. The content filtering system established in Article 17 CDSMD relies on a joint effort of the creative industry and the online platform industry. To set the filtering machinery in motion, copyright holders in the creative industry must first notify "relevant and necessary information"[51] with regard to those works which they want to ban from user uploads. Once relevant and necessary information on protected works is received, the OCSSP is obliged to include that information in the content moderation process and ensure the filtering – "unavailability"[52] – of content uploads that contain traces of the protected works. It is this cooperation which, according to Article 17(7) CDSMD, must not result in the prevention of UGC that does not infringe copyright, including situations where UGC is covered by a copyright limitation. The same cooperation constitutes the central item on the agenda of stakeholder dialogues which the Commission must initiate under Article 17(10) CDSMD to identify best practices.

21 The problem of the whole cooperation concept, however, lies in the fact that, unlike public bodies and the judiciary, the central players in the cooperation scheme – the creative industry and the online platform industry – are private entities that are not intrinsically motivated to safeguard the public interest in the exercise and furtherance of fundamental rights and freedoms. Despite all invocations of diligence and proportionality – "high industry standards of professional diligence" in Article 17(4)(b) CDSMD; "diligent, objective and proportionate" application in Article 14(4) DSA – the decision-making in the context of content filtering is most probably much more down to earth: the moment the balancing of competing human rights positions is confidently left to industry cooperation, economic cost and efficiency considerations are likely to occupy centre stage. Arguably, they will often prevail over more abstract societal objectives, such as flourishing freedom of expression and information.

22 A closer look at the different stages of industry cooperation resulting from the regulatory model of Article 17 CDSMD confirms that concerns about human rights deficits are not unfounded. As explained, the first step in the content moderation process is the notification of relevant and necessary information relating to "specific works and other subject matter"[53] by copyright holders. In the light of case law precedents, in particular *Sabam/Netlog*,[54] use of the word "specific" can be understood to reflect the legislator's hope that copyright holders will only notify individually selected works. For instance, a copyright holder could limit use of the notification system to those works that constitute cornerstones of the current exploitation strategy. The principle of proportionality and high standards of professional diligence also point in the direction of a cautious approach that confines work notifications to those repertoire elements that are "specific" in the sense that they generate a copyright holder's lion's share of revenue.[55] In line with this approach, other elements of the work catalogue could be kept available for creative remix activities of users. This, in turn, would reduce the risk of overbroad inroads into freedom of expression and information.

46 CJEU, ibid., para. 51.

47 CJEU, ibid., para. 26 and 36-37.

48 Cf. Senftleben/Angelopoulos, supra note 42, 8-9.

49 CJEU, ibid., para. 26.

50 Article 17(4)(b) DSMD. The intention to obviate the impression of a prohibited general monitoring obligation also lies at the core of Article 17(8) DSMD. This provision declares that UGC licensing and filtering "shall not lead to any general monitoring obligation."

51 Article 17(4)(b) CDSMD.

52 Article 17(4)(b) CDSMD.

53 Article 17(4)(b) CDSMD.

54 CJEU, 16 February 2012, case C-360/10, Sabam/Netlog, para. 51.

55 For a corresponding concept of "normal exploitation" in the sense of the three-step test in copyright law, see M.R.F. Senftleben, *Copyright, Limitations and the Three-Step Test – An Analysis of the Three-Step Test in International and EC Copyright Law*, The Hague/London/New York: Kluwer Law International 2004, 189-194.

23 In practice, however, rightholders are highly unlikely to adopt this cautious approach. The legal basis for requiring a focus on individually selected works lies in the fact that the legislator has used the expression "*specific* works and other subject matter"[56] in Article 17(4)(b) CDSMD. Proportionality and diligence considerations only form the broader context in which the specificity requirement is embedded. Strictly speaking, the requirement of "high industry standards of professional diligence" in Article 17(4)(b) CDSMD concerns the subsequent filtering step taken by an OCSSP to ensure the unavailability of notified works – not the primary notification sent by copyright holders.

24 Like the requirement of "high industry standards of professional diligence"[57], the imperative of "diligent, objective and proportionate" application in Article 14(4) DSA relates to platform content moderation measures that restrict user freedoms – not the rightholder notification system that sets the filtering process in motion. The success of the risk reduction strategy surrounding the word "specific" in Article 17(4)(b) CDSMD is thus doubtful. In the cooperation with OCSSPs, nothing seems to prevent the creative industry from sending copyright notifications that cover each and every element of long and impressive work catalogues. UGC platforms may thus receive long lists of all works which copyright holders have in their repertoire. Adding up all "specific works and other subject matter" included in these notifications, the conclusion seems inescapable that Article 17(4)(b) DSMD may culminate in a filtering obligation that is very similar to the filtering measures which the CJEU prohibited in *Sabam/Netlog*. The risk of encroachments upon human rights is evident.

25 Turning to the second step in the content moderation process – the act of filtering carried out by OCSSPs to prevent the availability of notified works on UGC platforms – it is noteworthy that proportionality and diligence obligations are directly applicable. As explained, the requirements of "high industry standards of professional diligence"[58] and "diligent, objective and proportionate"[59] application only form the broader context surrounding the notification of specific works by rightholders. When it comes to the content moderation process as such, however, these diligence and proportionality rules impact the activities of OCSSPs directly: the UGC filtering process must be implemented in a way that complies with these diligence and proportionality requirements.

26 As to the practical outcome of UGC filtering in the light of diligence and proportionality requirements, however, it is to be recalled that OCSSPs will most probably align the concrete implementation of content moderation systems with cost and efficiency considerations. Abstract commandments, such as the instruction to act in accordance with "high standards of professional diligence"[60] and in a "proportionate manner in applying and enforcing [UGC upload] restrictions"[61] can hardly be deemed capable of superseding concrete commercial cost and efficiency necessities. Tuomas Mylly accurately characterizes litanies of diligence and proportionality requirements as "wish-lists for private regulators."[62] On its merits, the legislator whitewashes statutory content filtering obligations by adding a diligence and proportionality gloss to reassure itself that the drastic measure will be implemented with sufficient care and caution to avoid the erosion of human rights. The success of this ingredient of the outsourcing recipe is doubtful. In reality, the subordination of industry decisions to diligence and proportionality imperatives – the acceptance of more costs and less profits to reduce the corrosive effect on freedom of expression and information – would come as a surprise. Instead, OCSSPs can be expected to be rational in the sense that they seek to achieve content filtering at minimal costs.

27 Hence, there is no guarantee that industry cooperation in the field of UGC will lead to the adoption of the most sophisticated filtering systems with the highest potential to avoid unjustified removals of content mash-ups and remixes. A test of proportionality is unlikely to occupy centre stage unless the least intrusive measure also constitutes the least costly measure. A test of professional diligence is unlikely to lead to the adoption of a more costly and less intrusive content moderation system unless additional revenues accruing from enhanced popularity among users offsets the extra investment of money.

28 In addition, EU legislation itself sends mixed signals. Article 17(5) CDSMD provides guidelines for the assessment of the proportionality of filtering obligations. The relevant factors listed in the provision, however, focus on "the type, the audience and the size of the service," "the type of works or other subject matter" and "the availability of suitable and effective means and their cost for

56 Article 17(4)(b) CDSMD (emphasis added).

57 Article 17(4)(b) CDSMD.

58 Article 17(4)(b) CDSMD.

59 Article 14(4) DSA.

60 Article 17(4)(b) CDSMD.

61 Article 14(4) DSA.

62 Mylly, supra note 24, 71.

service providers."[63] Hence, cost and efficiency factors have made their way into the proportionality assessment scheme. Paradoxically, it is conceivable that these factors encourage the adoption of cheap and unsophisticated filtering tools that lead to excessive content blocking.

29 An assessment of liability questions also confirms that excessive filtering risks must be taken seriously. A UGC platform seeking to minimize the risk of liability is likely to succumb to the temptation of overblocking.[64] Filtering more than necessary is less risky than filtering only clear-cut cases of infringement. After all, the primary, direct liability for infringing user uploads following from Article 17(1) CDSMD is hanging above the head of OCSSPs like the sword of Damocles. The second step of the industry cooperation concept underlying Article 17 CDSMD is thus at least as problematic as comprehensive notifications of entire work catalogues. The OCSSP obligation to embark on content filtering to police the borders of use permissions and prevent content availability in the absence of licenses raises serious concerns about interferences with human rights, in particular freedom of expression and information.

30 Surveying the described human rights risks that arise from the industry cooperation scheme in Article 17 CDSMD, the conclusion is inescapable that, despite all invocations of diligence and proportionality as mitigating factors, the outsourcing strategy underlying the EU regulation of content moderation in the CDSM Directive and the DSA is highly problematic. Instead of safeguarding human rights, the regulatory approach is likely to culminate in human rights violations. Against this background, it is of particular importance to analyse mechanisms that could bring human rights deficits to light and remedy shortcomings.

C. Concealing Human Rights Deficits Caused by Reliance on Industry Cooperation

31 The question of mechanisms that allow the detection and correction of human rights deficits in content moderation leads back to the information duty laid down in Article 14(1) DSA.[65] Under this provision, UGC platforms are obliged to make information on content moderation "policies, procedures, measures and tools"[66] available to users. This must be done in "clear, plain, intelligible, user-friendly and unambiguous language."[67] Moreover, the information must be publicly available in an easily accessible and machine-readable format.[68] These information and transparency obligations can be regarded as exponents of a broader human rights preservation strategy.[69] The broader pattern comes to the fore when the information flow generated in Article 14(1) DSA is placed in the context of the complaint and redress mechanism for unjustified content filtering that forms a building block of Article 17 CDSMD. Article 17(9) CDSMD requires that OCSSPs put in place:

> "an effective and expeditious complaint and redress mechanism that is available to users of their services in the event of disputes over the disabling of access to, or the removal of, works or other subject matter uploaded by them."[70]

32 With regard to the role of users in the human rights arena, the complementary character[71] of Article 17(9) CDSMD and Article 14(1) DSA yields important insights: the legislator confidently leaves the identification and correction of excessive content blocking to users. A relatively low number of user complaints, however, may be misinterpreted as an indication that content filtering does hardly ever encroach upon freedom of expression and

63 Article 17(5) DSMD.

64 Cf. M. Perel/N. Elkin-Koren, "Accountability in Algorithmic Copyright Enforcement", *Stanford Technology Law Review* 19 (2016), 473 (490-491). For empirical studies pointing towards overblocking, see Sharon Bar-Ziv/Niva Elkin-Koren, "Behind the Scenes of Online Copyright Enforcement: Empirical Evidence on Notice & Takedown", *Connecticut Law Review* 50 (2017), 3 (37); Jennifer M. Urban/Joe Karaganis/Brianna L. Schofield, "Notice and Takedown in Everyday Practice", *UC Berkeley Public Law and Legal Theory Research Paper Series*, Version 2, March 2017, available at https://ssrn.com/abstract=2755628, 2.

65 Further information and transparency obligations are found in elsewhere in Article 14, namely in paras (2), (3), (5) and (6).

66 Article 14(1) DSA.

67 Article 14(1) DSA.

68 Article 14(1) DSA.

69 Examples can be found in the GDPR and Terrorist Content Regulation.

70 Article 17(9) CDSMD.

71 Article 2(4)(b) and Recital 11 DSA. For an extensive analysis of this topic, see J.P. Quintais/S.F. Schwemer, The Interplay between the Digital Services Act and Sector Regulation: How Special Is Copyright?, *European Journal of Risk Regulation* 13 (2022), 191; Alexander Peukert et al., European Copyright Society – Comment on Copyright and the Digital Services Act Proposal, *International Review of Intellectual Property and Competition Law* 53 (2022), 358.

information even though limited user activism may be due to overly slow and cumbersome procedures (following section 1). Instead of addressing this problematic concealment mechanism, the CJEU has confirmed the validity of the content moderation rules in Article 17 CDSMD. In this context, the Court has qualified elements of the problematic outsourcing and concealment strategy as valid safeguards against the erosion of freedom of expression and information. Instead of uncovering human rights risks, the Court, thus, preferred to condone and stabilize the system (section 2). Under these circumstances, only legislative countermeasures taken by EU Member States (section 3) and content moderation assessments in audit reports for the European Commission (section 4) give some hope that violations of human rights may finally be prevented despite the corrosive outsourcing and concealment scheme underlying the regulation of content moderation in the EU.

I. Reliance on User Complaints as Part of a Concealment Strategy

33 As explained, Article 17(9) CDSMD and Article 14(1) DSA both make users the primary addressees of information about content moderation systems and potential countermeasures. Article 17(9) CDSMD stipulates that OCSSPs shall inform their users "in their terms and conditions that they can use works and other subject matter under exceptions or limitations to copyright and related rights provided for in Union law."[72] In addition to this specific rule dealing with copyright limitations, Article 14(1) DSA applies: users shall receive information on upload and content sharing restrictions arising from the employment of content moderation tools.[73] If they want to take measures against content restrictions, Article 17(9) CDSMD ensures that complaint and redress mechanisms are available to users of OCSSP services "in the event of disputes over the disabling of access to, or the removal of, works or other subject matter uploaded by them."[74]

34 Again, this regulatory model is not new. In *UPC Telekabel Wien*, the CJEU sought to ensure that, in the case of website blocking measures, the national courts in EU Member States would be able to carry out a judicial review. This, however, was only conceivable if a challenge was brought against the blocking measure implemented by an internet service provider:

> Accordingly, in order to prevent the fundamental rights recognised by EU law from precluding the adoption of an injunction such as that at issue in the main proceedings, the national procedural rules must provide a possibility for internet users to assert their rights before the court once the implementing measures taken by the internet service provider are known.[75]

35 Therefore, the rights assertion option for users served the ultimate purpose of paving the way for judicial review. In Article 17(9) CDSMD, this pattern reappears. Users can avail themselves of the option to instigate complaint and redress procedures at platform level and, ultimately, go to court. The DSA also contains specific user complaint and redress rights. Complementing Article 17(9) CDSMD,[76] Article 20 DSA sets forth detailed rules for internal complaint handling on UGC platforms. Article 54 DSA confirms that users are entitled to compensation for any damage or loss they suffered due to an infringement of DSA obligations. As pointed out above, one of these obligations follows from Article 14(4) DSA. This provision obliges UGC platforms to apply content moderation measures in a proportionate manner – with due regard to freedom of expression and information. In addition, Article 86(1) DSA affords users the opportunity to mandate a non-profit body, organization or association to exercise their complaint, redress and compensation rights on their behalf. According to their statutes, these non-profit institutions must have a legitimate interest in safeguarding DSA rights and obligations.

36 However, the broad reliance placed on user activism – ranging from complaints to damage claims and work with non-profit bodies – is surprising. Evidence from the application of the DMCA counter-notice system in the U.S. shows quite clearly that users are unlikely to file complaints in the first place.[77]

72 Article 17(9) CDSMD.

73 For more general transparency obligations, see Article 15(1) DSA and the discussion of these more general obligations in section 4.

74 Article 17(9) CDSMD.

75 CJEU, 27 March 2014, case C-314/12, *UPC Telekabel Wien*, para. 57.

76 As to the complementary character of Article 20 DSA, see Article 2(4)(b) and Recital 11 DSA. Cf. Quintais/Schwemer, supra note 71, 358.

77 See the study conducted by J.M. Urban/L. Quilter, "Efficient Process or "Chilling Effects"? Takedown Notices Under Section 512 of the Digital Millennium Copyright Act", *Santa Clara Computer and High Technology Law Journal* 22 (2006), 621, showing, among other things, that 30% of DMCA takedown notices were legally dubious, and that 57% of DMCA notices were filed against competitors. While the DMCA offers the opportunity to file counter-notices and rebut unjustified takedown requests, Urban and Quilter find that instances in

Data from recent transparency reports covering the largest UGC platforms confirm the assumption of user inactivism.[78] If users have to wait relatively long for a final result, it is foreseeable that a complaint and redress mechanism that depends on user initiatives is incapable of safeguarding freedom of expression and information. Moreover, an overly cumbersome complaint and redress mechanism may thwart user initiatives from the outset. The hope that users will bring damage claims and collaborate with non-profit institutions to assert their rights, thus, finds little support in the real world. While it cannot be ruled out that some users will exhaust the full arsenal of complaint, redress and compensation options, it seems unrealistic to assume that user complaint mechanisms have the potential of revealing the full spectrum and impact of free expression restrictions that result from automated content moderation systems.

37 In the context of UGC, it must also be considered that it is often crucial to react quickly to current news and film, book and music releases. If the complaint and redress mechanism finally yields the insight that a lawful content remix or mash-up has been blocked, the decisive moment for the affected quotation or parody may already have passed.[79] From this perspective, the elastic timeframe for complaint handling – "shall be processed without undue delay"[80] – also gives rise to concerns. This standard differs markedly from an obligation to let blocked content reappear promptly. As Article 17(9) CDSMD also requires human review, it may take quite a while until a decision on the infringing nature of content is taken. Considering these features, the complaint and redress option may appear unattractive to users.[81]

38 Instead of dispelling concerns about human rights deficits, the reliance on user complaints, thus, constitutes a further risk factor. Apart from being ineffective as a remedy for human rights violations, the complaint and redress mechanism in Article 17(9) CDSMD may allow authorities to hide behind a lack of user activism. It may be that users refrain from complaining because they consider the mechanism too cumbersome and/or too slow. However, when taking the number of user complaints as a yardstick for assessing human rights risks, a relatively low number of user complaints may be misinterpreted as evidence that content moderation does not lead to excessive content blocking. As long as users refrain from taking action, human rights deficits stay under the radar. The oversimplified equation "no user complaint = no human rights problem" offers the opportunity of praising an overly restrictive content moderation system as a success. Instead of shedding light on human rights deficits, the complaint and redress mechanism can be used strategically to disguise encroachments upon freedom of expression and information.

39 The outsourcing problem described in the preceding section (inappropriate reliance on OCSSPs and copyright holders as human rights guardians) is thus aggravated by heavy reliance on complaint and redress mechanisms which users are unlikely to embrace. Leaving measures against the erosion of freedom of expression and information to users, the legislator cultivates a culture of concealing human rights deficits. Reliance on user complaints as indicators of human rights violations is simply inadequate. Even if users lodge a complaint, any redress, moreover, remains an *ex post* measure: a remedy that reinstates freedom of expression and information only after initial harm – in the form of unjustified UGC impoverishment – has occurred. The EU approach is thus wanting for at least two reasons: the outsourcing of human rights obligations to private entities and the expectation that users will take countermeasures against human rights violations.

II. Confirmation of the Outsourcing and Concealment Strategy in CJEU Jurisprudence

40 This outcome of the risk assessment raises the additional question whether other institutions in the platform governance arena could fulfil the role of

which this mechanism is used are relatively rare. However, cf. also the critical comments on the methodology used for the study and a potential self-selection bias arising from the way in which the analyzed notices have been collected by F.W. Mostert/M.B. Schwimmer, "Notice and Takedown for Trademarks", *Trademark Reporter* 101 (2011), 249 (259-260).

78 See the analysis conducted by M.R.F. Senftleben/J.P. Quintais/A. Meiring, "Outsourcing Human Rights Obligations and Concealing Human Rights Deficits: The Example of Monetization Under the CDSMD and the DSA", *Berkeley Technology Law Journal* 38 (2023), III.B.1 (forthcoming).

79 Apart from the time aspect, complaint systems may also be implemented in a way that discourages widespread use. Cf. Perel/Elkin-Koren, supra note 64, 507-508 and 514. In addition, the question arises whether users filing complaints are exposed to copyright infringement claims in case the user-generated quotation, parody or pastiche at issue (which the user believes to be legitimate) finally proves to amount to copyright infringement. Cf. N. Elkin-Koren, supra note 8, 1092.

80 Article 17(9) CDSMD.

81 Cf. Senftleben, supra note 10, 484.

human rights guardians more reliably. The judiciary seems a logical candidate. Interestingly, the CJEU already had the opportunity to discuss violations of freedom of expression and information that may arise from content moderation under Article 17(4)(b) and (c) CDSMD. In *Poland/Parliament and Council*, the Republic of Poland brought an annulment action against the content filtering branch of Article 17 CDSMD.[82] More specifically, Poland argued that OCSSPs were bound under Article 17(4)(b) and (c) CDSMD to carry out preventive – *ex ante* – monitoring of all user uploads. To fulfil this Herculean task, they had to employ automatic filtering tools. In Poland's view, EU legislation imposed this preventive monitoring obligation on OCSSPs "without providing safeguards to ensure that the right to freedom of expression and information is respected."[83] The contested provisions, thus, constituted a limitation on the exercise of the fundamental right to freedom of expression and information, which respected neither the essence of that right nor the principle of proportionality. Hence, the filtering obligations arising from Article 17(4)(b) and (c) CDSMD could not be regarded as justified under Article 52(1) CFR.[84]

41 Discussing these annulment arguments, the CJEU pointed out that prior review and filtering of user uploads, indeed, created the risk of limiting a central avenue for the online dissemination of UGC. The filtering regime in Article 17(4)(b) and (c) CDSMD imposed a restriction on the ability of users to exercise their right to freedom of expression and information which was guaranteed by Article 11 CFR and Article 10 of the European Convention on Human Rights ("ECHR").[85] However, the Court considered that such a limitation met the requirements set forth in Article 52(1) CFR – mandating that any limitation on the exercise of the right to freedom of expression and information had to be legally established and had to preserve the essence of those freedoms.[86] The Court was satisfied that the limitation arising from the filtering obligations in Article 17(4)(b) and (c) CDSMD could be deemed justified in the light of the legitimate objective to ensure a high level of copyright protection to safeguard the right to intellectual property enshrined in Article 17(2) CFR.[87]

42 More specifically, the Court identified no less than six freedom of expression safeguards in the regulatory design of Article 17 CDSMD – safeguards which, in the Court's view, gave sufficient reassurance that freedom of expression and information would not be unduly curtailed. A key aspect in this assessment of Article 17 CDSMD is the first point. The Court assumed that the introduction of automated content filtering tools would not prevent users from uploading lawful content, including UGC containing traces of protected third-party material that was permissible under statutory exceptions to copyright.[88] In this context, the Court recalled its earlier ruling in *Sabam/Netlog* from which it followed 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, guaranteed in Article 11 of the Charter, and would not respect the fair balance between that right and the right to intellectual property."[89]

43 Hence, the Court was confident that, in the light of its case law, OCSSPs would refrain from introducing content filtering measures unless these systems could reliably distinguish between lawful parody and infringing piracy – unless they were capable of leaving all kinds of lawful uploads unaffected.[90]

44 The second point made by the Court addresses statutory exceptions to copyright more directly. In line with earlier decisions, the CJEU confirmed that copyright limitations supporting freedom of expression, such as the right of quotation and the exemption of parody, constituted "user rights."[91]

82 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 24. For a more detailed discussion of the decision, see J.P. Quintais, *Between Filters and Fundamental Rights: How the Court of Justice saved Article 17 in C-401/19 - Poland v. Parliament and Council*, Verfassungsblog (2022), https://verfassungsblog.de/filters-poland/ (last visited 5 April 2023); M. Husovec, "Mandatory filtering does not always violate freedom of expression: Important lessons from Poland v. Council and European Parliament", *Common Market Law Review* 60 (2023), 173.

83 CJEU, id., para. 24.

84 CJEU, id., para. 24.

85 CJEU, id., para. 55, 58, 82.

86 CJEU, id., para. 63 et seq., referring to the principle of proportionality.

87 CJEU, id., para. 69.

88 CJEU, id., para. 86.

89 CJEU, id., para. 86. Cf. CJEU, 16 February 2012, case C-360/10, Sabam/Netlog, para. 50-51.

90 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 86.

91 CJEU, id., para. 87-88; CJEU, 29 July 2019, case C-516/17, Spiegel Online, para. 50-54; CJEU, 29 July 2019, case C469/17, Funke Medien NRW, para. 65-70. Cf. Tanya Aplin and Lionel

To avoid the dismantling of these free expression strongholds, EU Member States had to ensure that automated filtering measures did not deprive users of their freedom to upload content created for the purposes of quotation, criticism, review, caricature, parody, or pastiche.[92] On this point the judgment endorsed, by reference, the Advocate General Opinion stating that filters "must not have the objective or the effect of preventing such legitimate uses," and that providers must "consider the collateral effect of the filtering measures they implement" as well as "take into account, ex ante, respect for users' rights."[93]

45 As a third aspect that mitigated the corrosive effect of Article 17(4)(b) and (c) CDSMD on freedom of expression and information, the Court pointed out that the filtering machinery was only set in motion on condition that rightholders provided OCSSPs with the "relevant and necessary information"[94] concerning protected works that should not become available on the UGC platform. In the absence of such information, OCSSPs would not be led to make content unavailable.[95] The fourth point highlighted by the Court was the clarification in Article 17(8) CDSMD that no general monitoring obligation was intended.[96] The fifth point was the complaint and redress mechanism allowing users to bring unjustified content blocking to the attention of the platform provider.[97] Finally, the Court recalled that Article 17(10) CDSMD tasked the European Commission with organizing stakeholder dialogues to ensure a uniform mode of OCSSP/rightholder cooperation across Member States and establish best filtering practices in the light of industry standards of professional diligence.[98]

46 Qualifying all six aspects as valid safeguards against an erosion of freedom of expression and information, the Court concluded that the regulatory design of Article 17 CDSMD included appropriate countermeasures to survive Poland's annulment action.[99] Still, the Court cautioned EU Member States, as well as their authorities and courts, that transposing and applying Article 17 CDSMD, they had to follow a fundamental rights-compliant path.[100]

47 Undoubtedly, the *Poland* decision is a milestone that contains several important clarifications. With regard to the above-described human rights risks arising from the outsourcing and concealment strategy underlying Article 17 CDSMD, however, it is disappointing. A critical assessment of the regulatory scheme is missing. The Court did not seize the opportunity to unmask human rights risks that, as explained in the preceding sections, are inherent in the heavy reliance on industry cooperation. The Court also refrained from reflecting on human rights risks that could arise from the ineffectiveness of complaint and redress mechanisms for users. Instead of exposing the outsourcing and concealment strategy and addressing human rights deficits, the Court rubberstamped not only the broader regulatory design but also its individual elements. Singling out no less than six aspects of Article 17 CDSMD and declaring them valid safeguards against violations of freedom of expression and information, the Court readily accepted the very ingredients of

Bently, *Global Mandatory Fair Use: The Nature and Scope of the Right to Quote Copyright Works*, Cambridge: Cambridge University Press 2020, 75-84; C. Geiger/E. Izyumenko, "The Constitutionalization of Intellectual Property Law in the EU and the Funke Medien, Pelham and Spiegel Online Decisions of the CJEU: Progress, but Still Some Way to Go!", *International Review of Intellectual Property and Competition Law* 51 (2020), 282 (292-298).

92 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 87. With regard to the particular importance of the inclusion of the open-ended concept of "pastiche," see M.R.F. Senftleben, "User-Generated Content – Towards a New Use Privilege in EU Copyright Law", in: T. Aplin (ed.), *Research Handbook on IP and Digital Technologies*, Cheltenham: Edward Elgar 2020, 136 (145-162); Senftleben, supra note 8, 320-327; E. Hudson, "The pastiche exception in copyright law: a case of mashed-up drafting?", *Intellectual Property Quarterly* 2017, 346 (348-352 and 362-364); F. Pötzlberger, "Pastiche 2.0: Remixing im Lichte des Unionsrechts", *Gewerblicher Rechtsschutz und Urheberrecht* 2018, 675 (681); J.P. Quintais, *Copyright in the Age of Online Access – Alternative Compensation Systems in EU Law*, Alphen aan den Rijn: Kluwer Law International 2017, 235-237.

93 Opinion of Advocate General Saugmandsgaard Øe, 15 July 2021, case C-401/19, Poland/Parliament and Council, para. 193.

94 Article 17(4)(b) CDSMD.

95 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 89.

96 CJEU, id., para. 90. See Article 17(8) CDSMD; Article 8 and Recital 30 DSA. Cf. Senftleben/Angelopoulos, supra note 42, for a more detailed discussion on the prohibition of general monitoring obligations.

97 CJEU, id., para. 94. See Article 17(9) CDSMD.

98 CJEU, id., para. 96-97. As to existing best practices guidelines, see Communication from the Commission to the European Parliament and the Council, *Guidance on Article 17 of Directive 2019/790 on Copyright in the Digital Single Market*, COM/2021/288 final.

99 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 98.

100 CJEU, id., para. 99.

the Article 17 recipe that create the outsourcing and concealment risks discussed above. In the *Poland* ruling, the Court went far beyond condoning the approach chosen in Article 17 CDSMD. The CJEU expressly confirmed its validity – and the positive, mitigating effect of all its elements.

48 This central problem of uncritical rubberstamping in the *Poland* decision clearly comes to the fore when the six free expression safeguards are re-evaluated in the light of the above-described outsourcing and concealment risks. With regard to the necessity of distinguishing between lawful/unlawful content uploads (first point highlighted by the Court),[101] a reality check is sought in vain in the judgment. From a legal-theoretical perspective, the CJEU assumption – namely that filtering systems must not be applied as long as they cannot reliably distinguish permitted parody from infringing piracy – may be right and correct. The lack of incentives to refrain from the employment of such overblocking systems in practice, however, does not enter the picture. The Court does not even mention that, instead of discouraging the use of unsophisticated filtering machines, Article 17(1) CDSMD, quite clearly, gives a very strong impulse to implement automated filtering systems regardless of their capacity to distinguish between lawful and unlawful content. As pointed out above, the risk of direct liability for infringing UGC uploads is hanging above the head of OCSSPs like the sword of Damocles. Overblocking allows OCSSPs to avert this risk, escape direct liability under Article 17(1) CDSMD and avoid lengthy and costly lawsuits. Adopting an excessive filtering approach, they only have to deal with user complaints which are unlikely to come in large numbers. Practically speaking, the implementation of an underblocking approach to safeguard freedom of expression and information is thus unlikely. In its imaginary and pure universe of legal-theoretical assumptions, the Court may assume that content filtering will only occur when automated systems are capable of separating the wheat from the chaff. To whitewash the Article 17 approach on the basis of such unrealistic assumptions, however, creates a human rights risk of its own.

49 The same can be said about the inclusion of rightholder notifications in the list of effective free expression safeguards (third point made by the Court).[102] As pointed out above, nothing in Article 17 CDSMD prevents copyright owners from notifying long lists – entire catalogues – of protected works. Adding up all repertoire notifications arriving at OCSSPs, it seems naïve to assume that the notification mechanism laid down in Article 17(4)(b) CDSMD will never lead to a filtering volume that is comparable with the general filtering obligation which the Court prohibited in *Sabam/Netlog*.[103] From this perspective, the ban on general filtering obligations in Article 17(8) CDSMD (fourth safeguard identified by the Court)[104] can also be unmasked as mere cosmetics. The fifth safeguard which the Court accepted,[105] is the complaint and redress mechanism that causes the corrosive concealment risk described above. The sixth and final safeguard – stakeholder dialogues seeking to establish best practices[106] – has also been analysed above. It is a toothless tiger. Article 17(10) CDSMD is silent on measures which the Commission could take to enforce the best practices guidelines following from meetings with stakeholders. It remains unclear why the Court is willing to accept this type of fig-leaf measures as a valid free expression safeguard.

50 On balance, the Court has not only missed an important opportunity to reveal and address human rights risks that arise from the outsourcing and concealment strategy underlying Article 17 CDSMD. Choosing the most favourable interpretation of Article 17 features as a reference point for its assessment of human rights risks, and refusing to consider the practical reality of industry cooperation and the practical impact of the overblocking incentive resulting from the risk of direct liability for infringing UGC, the Court has made itself an accomplice in the outsourcing and concealment strategy that puts freedom of expression and information at risk.

III. Member State Legislation Seeking to Safeguard Transformative UGC

51 The foregoing critique of the six free expression safeguards which the CJEU identified in its *Poland* decision did not address the second point made by the Court: the obligation placed on EU Member States to ensure that transformative UGC – consisting of quotations, parodies, pastiches etc. – survives the implementation of automated content filtering systems.[107] The reason for this omission is simple:

101 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 86.

102 CJEU, id., para. 89.

103 CJEU, id., para. 86; CJEU, 16 February 2012, case C-360/10, Sabam/Netlog, para. 50-51.

104 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 90.

105 CJEU, id., para. 93.

106 CJEU, id., para. 96.

107 CJEU, id., para. 87-88.

in contrast to other Article 17 aspects, this element appears as a valid safety valve that could effectively safeguard freedom of expression and information in practice. This insight does not change the critical assessment of the *Poland* judgment. With regard to outsourcing and concealment risks, the decision remains a missed opportunity to address and minimize human rights risks.

52 As to the valid second point in the *Poland* phalanx of free expression safeguards – the need to preserve copyright limitations for creative remix activities, in particular use for the purposes of "quotation, criticism and review," and "caricature, parody and pastiche"[108] – Article 17(7) CDSMD plays a central role. The provision leaves no doubt that EU Member States are expected to ensure that automated content filtering does not submerge areas of freedom that support the creation and dissemination of transformative user productions that are uploaded to UGC platforms. The second paragraph of Article 17(7) reads as follows:

> "Member States shall ensure that users in each Member State are able to rely on any of the following existing exceptions or limitations when uploading and making available content generated by users on online content-sharing services:
>
> (a) quotation, criticism, review;
>
> (b) use for the purpose of caricature, parody or pastiche."[109]

53 Use of the formulation "shall not result in the prevention" and "shall ensure that users [...] are able" give copyright limitations for "quotation, criticism, review" and "caricature, parody or pastiche" an elevated status. In Article 5(3)(d) and (k) of the Information Society Directive 2001/29/EC ("ISD"),[110] these use privileges were only listed as limitation prototypes which EU Member States are free to introduce (or maintain) at the national level. The adoption of a quotation right[111] and an exemption of caricature, parody or pastiche[112] remained optional. Article 17(7) CDSMD, however, transforms these use privileges into mandatory breathing space for transformative UGC – at least in the specific context of OCSSP content moderation.[113] This metamorphosis makes copyright limitations in this category particularly robust: they "shall" survive the application of automated filtering tools.

54 Under Article 17(7) CDSMD, EU Member States are the guardians of these user rights.[114] This regulatory decision comes as a welcome surprise. In contrast to the prevailing preference for solutions based on outsourcing (passing on human rights responsibilities to private entities) and concealment (relying in user complaints to remedy human rights deficits), Article 17(7) CDSMD entrusts the Member States – the state power itself – with the important task of guaranteeing ("shall ensure") that, despite content filtering on OCSSP platforms, users can share creations made for the purposes of "quotation, criticism, review" and "caricature, parody or pastiche." In this regard, the *Poland* decision adds an important nuance. In its discussion of safeguards against an erosion of freedom of expression and information, the CJEU qualified the complaint and redress mechanisms mandated by Article 17(9) CDSMD as *additional* safeguards against content overblocking:

> "the first and second subparagraphs of Article 17(9) of Directive 2019/790 introduce several procedural safeguards, which are additional to those provided for in Article 17(7) and (8) of that directive, and which protect the right to freedom of expression and information of users of online content-sharing services in cases where, notwithstanding the safeguards laid down in those latter provisions, the providers of those services nonetheless erroneously or unjustifiably block lawful content."[115]

55 Hence, user complaint mechanisms evolving from Article 17(9) CDSMD only constitute additional *ex post* measures. As they allow corrections of wrong filtering decisions only after the harm has occurred,

108 Article 17(7) CDSMD. Cf. Senftleben, supra note 10, 485-490; P.B. Hugenholtz/M. Senftleben, *Fair Use in Europe. In Search of Flexibilities*, Amsterdam: Institute for Information Law/VU Centre for Law and Governance 2011, 29-30.

109 Article 17(7) CDSMD.

110 Article 5(3)(d) and (k) of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001, on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society, *Official Journal of the European Communities* 2001 L 167, 10).

111 Article 5(3)(d) ISD.

112 Article 5(3)(k) ISD.

113 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 87. Cf. J.P. Quintais/G. Frosio et al., "Safeguarding User Freedoms in Implementing Article 17 of the Copyright in the Digital Single Market Directive: Recommendations From European Academics", *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 10 (2020), 277 (278-279).

114 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 87-88.

115 CJEU, id., para. 93.

they can hardly be considered sufficient. First and foremost, it is necessary to have *ex ante* mechanisms in place that allow permissible content uploads – quotations, parodies, pastiches etc. – to survive automated content scrutiny. This is an important guideline for EU Member States. Implementing Article 17 CDSMD, they must ensure that UGC containing quotations, criticism, review, caricatures, parodies or pastiches[116] appear directly on the platform.

56 In practice, this goal can be achieved by introducing mandatory flagging options for users. To ensure *ex ante* content availability – without exposure to content filtering tools – domestic legislation in EU Member States can enable users to mark quotations, parodies, pastiches etc. as permissible content uploads and oblige OCSSPs to make these uploads directly available on the UGC platform. An example of national legislation following this approach can be found in Germany.[117] Alarmingly, however, the central importance of the state responsibility arising from Article 17(7) CDSMD seems to have escaped the attention of many other EU Member States. The German implementation model has not become widespread. Instead, the majority of Member States opted for a national transposition that does not offer users specific legal tools, such as statutory flagging options, to benefit from the exemption of quotations, parodies, pastiches etc.[118]

The Netherlands, for instance, gave preference to a literal implementation of Article 17 CDSMD. Effective *ex ante* mechanisms – capable of placing quotations, parodies, pastiches etc. beyond the reach of content filtering systems from the outset – are sought in vain. Instead, the Dutch legislator places reliance on complaint and redress mechanisms even though this legal instrument only allows users to take measures *ex post*: after quotations, parodies, pastiches etc. have been filtered out and the UGC spectrum has been impoverished.[119] In the light of the *Poland* decision, it is doubtful that this implementation is adequate. As explained, the CJEU characterized *ex post* complaint and redress mechanisms as additional safeguards that supplement – but cannot replace – *ex ante* safeguards, such as the statutory flagging options in Germany.[120]

IV. European Commission Taking Action on the Basis of Audit Reports

57 As many EU Member States seem reluctant to translate their human rights responsibility under Article 17(7) CDSMD into statutory *ex ante* mechanisms that immunize quotations, parodies, pastiches etc. against content filtering measures, it is important to look beyond the regulatory framework in the CDSM Directive. An analysis of Article 17 CDSMD does not exhaust the full spectrum of legal tools that could contribute to the preservation of freedom of expression and information in content moderation contexts. In line with the interplay between the CDSM Directive and the DSA configurated in Article 2(4)(b) and Recital 11 DSA, it is possible to factor DSA provisions into the equation when the CDSM Directive does not contain more specific rules.

58 A legal tool that does not appear in the CDSM Directive is the possibility for the executive to exercise control over content moderation systems on the basis of audit reports. In the DSA, this avenue for public authorities seeking to fulfil a watchdog function *ex officio* has been developed in Article 37. With respect to very large online platforms ("VLOPs")[121] and very large online search engines

116 Article 17(7) CDSMD.

117 See Sections 11(1), no. 1 and 3, 9(1) and (2), and 5(1) of the German Act on the Copyright Liability of Online Content Sharing Service Providers, available in official English translation at: https://www.gesetze-im-internet.de/englisch_urhdag/index.html.

118 For studies of national implementations of Article 17, see J.P. Quintais/P. Mezei et al., *Copyright Content Moderation in the EU: An Interdisciplinary Mapping Analysis*, reCreating Europe Report 2022, https://papers.ssrn.com/abstract=4210278 (last visited 12 August 2023); C. Angelopoulos, *Articles 15 & 17 of the Directive on Copyright in the Digital Single Market Comparative National Implementation Report*, Cambridge: Centre for Intellectual Property and Information Law 2022, Chapter 2 develops a conceptual framework and interdisciplinary methodological approach to examine copyright content moderation on online platforms and its potential impact on access to culture. The analysis clarifies our terminology, distinguishes between platform "governance" and "regulation", elucidates the concept of "online platform", and positions our research in the context of regulation "of", "by" and "on" platforms.Chapter 3 carries out a legal mapping of the topic of this report at EU level. Our focus here is the legal regime of art. 17 of the Copyright in the Digital Single Market Directive (CDSMDavailable at: https://informationlabs.org/copyright/ (last visited on 12 August 2023).

119 Article 29c(7) of the Dutch Copyright Act (*Auteurswet*).

120 CJEU, 26 April 2022, case C-401/19, Poland/Parliament and Council, para. 93. As to the German legislation, see the description above and German Act on the Copyright Liability of Online Content Sharing Service Providers, id., Sections 11(1), no. 1 and 3, 9(1) and (2), 5(1).

121 In accordance with Article 33(1) DSA, an online platform is qualified as a VLOP when it has a number of average

("VLOSEs"),[122] Article 37(1) DSA orders annual audits to assess compliance, among other things, with the obligations set forth in Chapter III of the DSA. Interestingly, one of the obligations laid down in Chapter III concerns the "diligent, objective and proportionate"[123] application of content moderation systems in line with Article 14(4) DSA.

59 Supplementing the complaint and redress system of Article 17(9) CDSMD that depends on user initiatives, Article 37 DSA may thus offer an important alternative basis that allows the executive power to prevent human rights violations. Article 37(3) DSA ensures that auditors establishing the report are independent from the VLOPs and VLOSEs under examination. In particular, it prevents organizations from performing an audit when they have a conflict of interest with the VLOP or VLOSE concerned, or with a legal person connected to that service provider. The audit report must contain an opinion – in the categories "positive," "positive with comments," and "negative" – on whether the VLOP or VLOSE has complied with the obligations and commitments under Chapter III DSA, including the above-described human rights and proportionality obligations laid down in Article 14(1) and (4) DSA.[124] If the audit opinion is not "positive," auditors are bound to include operational recommendations and specify the measures necessary to achieve compliance. They must also recommend a timeframe for achieving compliance.[125] In such a case, the VLOP or VLOSE concerned must adopt, within one month from receiving the recommendations, an audit implementation report. If the VLOP or VLOSE does not intend to implement the operational recommendations, it must give reasons for not doing so and set out alternative measures that it has taken to address the instances of non-compliance identified in the audit report.[126]

60 As to the role of the European Commission, Article 42(4) DSA is of particular importance. This provision obliges VLOPs and VLOSEs to transmit audit reports and audit implementation reports to the Commission without undue delay. If, based on this information, the Commission suspects a VLOP or VLOSE of infringing Article 14 DSA, it can initiate proceedings pursuant to Article 66(1) DSA. It may request further information, conduct interviews and inspect premises to learn more about the suspected infringement.[127] In case of a "risk of serious damage for the recipients of the service," Article 70(1) DSA entitles the Commission to order interim measures on the basis of a *prima facie* finding of infringement. If the Commission finally establishes non-compliance with "the relevant provisions of this Regulation" – including the human rights safeguards in Article 14(4) DSA – in a decision pursuant to Article 73(1) DSA, it may impose fines of up to six percent of the VLOP's or VLOSE's total worldwide annual turnover in the preceding financial year.[128] For the imposition of fines, Article 74(1) DSA requires a finding that the service provider under examination has infringed Article 14(4) DSA intentionally or negligently.

61 Considering this cascade of possible Commission actions, the potential of the audit mechanism in Article 37 DSA must not be underestimated. The audit system may be an important addition to the canon of norms in the CDSM Directive and, in particular, a promising counterbalance to outsourcing/concealment risks arising from the regulatory design of Article 17 CDSMD. Like the Member State legislation discussed in the preceding section, Commission interventions evolving from the problem analysis in an audit report are welcome departures from the strategy to pass on human rights responsibilities to private companies or users: the state power itself – in this case the Commission as the executive body of an international intergovernmental organization – remains directly responsible for detecting and remedying human rights deficits.

62 A potential blind spot of the described audit cascade leading to investigations, however, is this: in order to offer sufficient starting points for Commission action, audit reports addressing content moderation systems must go beyond a general problem analysis. The audit opinion must convincingly discuss a platform's failure to satisfy human rights obligations evolving from Article 14(4) DSA. It must contain a concrete assessment of the risk of human rights violations and a sufficient substantiation of that risk. Hence, auditors should be bound to devote sufficient attention to human rights implications of content moderation. They must insist on detailed

monthly active service recipients in the EU that is equal to, or higher than, 45 million, and has been designated as a VLOP by the European Commission pursuant to Article 33(4) DSA.

122 In accordance with Article 33(1) DSA, a search engine is qualified as a VLOSE when it has a number of average monthly active service recipients in the EU that is equal to, or higher than, 45 million, and has been designated as a VLOSE by the European Commission pursuant to Article 33(4) DSA.

123 Article 14(4) DSA.

124 Article 37(4)(g) DSA.

125 Article 37(4)(h) DSA.

126 Article 37(6) DSA.

127 Articles 67 to 69 DSA.

128 Article 74(1) DSA.

information on the practical implementation of content filtering tools that allows a proper assessment of the actual impact on users. An audit opinion merely scratching the surface – remaining at the superficial level of general platform policies and procedures to somehow tick off the point of freedom of expression risks – is not enough.

63 Luckily, the DSA itself points in this direction anyway. The general transparency obligation set forth in Article 15(1) DSA already obliges UGC platforms to publish annually clear and easily comprehensible content moderation reports. These reports must include information on the number of illegal content notices that have been submitted,[129] categorized by the type of alleged illegal content concerned and the number of notices submitted by trusted flaggers,[130] and information on any action taken pursuant to the notices, differentiating whether the action was taken on the basis of the law or the provider's terms and conditions. The reports must also specify the number of notices processed by using automated means and the median time needed for taking the action.[131] If automated content moderation tools have been deployed, the reports must include a qualitative description, a specification of the precise purposes, indicators of the accuracy and the possible rate of error of the automated means used in fulfilling those purposes, and any safeguards applied.[132]

64 Arguably, the source material for audit reports in the sense of Article 37(1) DSA must be richer than this standard information which UGC platforms must make available under Article 15(1) DSA anyway. Article 37(2) DSA points out that VLOPs and VLOSEs must afford auditors the cooperation and assistance necessary for conducting the audit in an effective, efficient and timely manner. This includes the obligation to provide access to all relevant data and give answers to oral or written questions. It would thus come as a surprise if audit opinions only reflected the generally available information flowing from Article 15(1) DSA. If this becomes necessary, the Commission can also ensure sufficient focus on the examination of human rights deficits by adopting a delegated act on the basis of Article 37(7) DSA that creates clarity about the necessity to devote particular attention to human rights questions in audit reports and seek all information necessary for this purpose.

D. Conclusion

65 On balance, the closer inspection of content moderation rules in the CDSM Directive and the DSA confirms a worrying tendency of reliance on industry cooperation and user activism to safeguard human rights. Instead of putting responsibility for detecting and remedying human rights deficits in the hands of the state, the EU legislature prefers to outsource this responsibility to private entities, such as OCSSPs, and conceal potential violations by leaving countermeasures to users. Considering the pattern of regulatory outsourcing and concealment decisions in the CDSM Directive and the DSA, it is justified to speak of a broader outsourcing and concealment strategy that endangers the fundamental rights of users. The risk of human rights encroachments is compounded by the fact that, instead of exposing and discussing the corrosive effect of human rights outsourcing, the CJEU has rubberstamped the regulatory approach in Article 17 CDSMD. In its *Poland* decision, the Court has even qualified problematic features of the outsourcing and concealment strategy as valid safeguards against the erosion of freedom of expression and information.

66 As a welcome departure from the Court-approved outsourcing and concealment scheme, Article 17(7) CDSMD obliges Member States to ensure that transformative UGC, containing quotations, parodies, pastiches etc., survives content filtering and appears on online platforms. In addition, audit reports evolving from Article 37 DSA can offer important information for the European Commission to identify and eliminate human rights violations. Both exceptions to the rule of outsourcing to private entities, however, are currently underdeveloped. Many EU Member States refrained from taking specific legislative action to protect transformative UGC from content filtering measures. The success of the DSA cascade of European Commission interventions – from audit reports to non-compliance decisions and fines that ensure human rights compliance[133] – is unclear. Therefore, it would be premature to sound the all-clear. To safeguard human rights in the UGC galaxy, the state power itself must become much more active. Litanies of due diligence and proportionality obligations for private entities and reliance on user activism are not enough.

129 Article 16 DSA.

130 Article 15(1)(b) DSA.

131 Article 15(1)(b) DSA.

132 Article 15(1)(e) DSA.

133 Articles 66 to 74 DSA.

Protection against Disinformation on the Internet: A Portuguese Perspective

by **Dário Moura Vicente***

Abstract: Disinformation, largely enhanced by the advent of the Internet and social networks, is one of the most serious challenges to the proper functioning of democratic systems today. In societies based on freedom of expression, protection against disinformation raises complex problems of reconciling of values. This has been highlighted in particular by recent developments in Portuguese law. This article presents an overview of these issues and of the extent to which tort liability may be utilized under Portuguese law as a potential means of protection against disinformation.

Keywords: Digital services, disinformation, exemptions from liability, freedom of expression, tort liability.

Recommended citation: Dário Moura Vicente, Protection against Disinformation on the Internet: A Portuguese Perspective, 14 JIPITEC 453 para 1.

A. Outline of the problem

1 In an essay published at the dawn of the new century, Oliveira Ascensão noted the following about the then emerging information society:

> "We are at a time when amazing possibilities open up – which man can use or not, or even misuse."[1]

2 Among the most vivid expressions of the latter alternative is the phenomenon, now more topical than ever, of disinformation. This will be taken here to mean the creation, presentation, and dissemination of demonstrably false or misleading information, for profit or with the intention of deceiving its addressees, and which may cause damage to public interests.

3 Disinformation is not to be confused with illegal content disseminated by the media, which includes, among many others, hate speech, incitement to commit crimes, child pornography or the unauthorised reproduction of works and performances protected by copyright and related rights. While disinformation may well comprise such content, it typically includes other material, which is not covered by any specific legal prohibitions because it is false or misleading.

4 Strictly speaking, disinformation is not a new phenomenon: it has probably existed since power and control over scarce resources were fought over among human communities. But the advent of the Internet and social media has exponentially increased the scale on which it is practised and

* Full Professor at the Faculty of Law of the University of Lisbon. President of the Portuguese Intellectual Property Law Association.

1 See "E agora? Pesquisa do futuro próximo", in *Estudos sobre Direito da Internet e da Sociedade da Informação*, Coimbra, 2001, pp. 45 ff. (p. 65).

the damage it can cause to society at large and to individuals within it.[2]

5 Several recent events reveal the harmful potential of misinformation disseminated through those media. Among many others are the 2016 Brexit referendum,[3] the US presidential elections of 2016[4] and 2020,[5] the opposition moved in several countries to the Covid 19 vaccination campaigns,[6] and the military invasion of Ukraine by Russia in 2022.[7]

B. The values at stake. In particular, freedom of expression

6 One may however ask how disinformation can be regulated. The question is not easy to answer, since essential values related to the rule of law are at stake here. These include freedom of expression, as enshrined in the European Convention on Human Rights,[8] in the Charter of Fundamental Rights of the European Union,[9] and in the constitutions of its Member States.[10]

7 No less relevant, though, are the integrity of the democratic process, which recent history proves to be severely undermined by the systematic and large-scale dissemination of disinformation; national security, potentially weakened by public decisions and policies based on false or distorted information; and the free and informed consent of citizens in the exercise of their civil rights and liberties, which can be vitiated by disinformation.[11]

8 The legal regulation of disinformation must therefore be obtained by weighing the relative import in individual cases of these values; none of them is an absolute, and democratic integrity is not an end in itself, but rather is instrumental in serving the effectiveness of popular sovereignty and the democratic principle.[12] This is why the Portuguese Constitutional Court has recognized that the protection of freedom of expression ceases "where it may jeopardize the essential content of another right or intolerably affect social morality or the fundamental values and principles of the constitutional order".[13]

9 It is accordingly crucial to ensure that the democratic principle is not undermined by the abusive exercise of freedom of expression through disinformation practices – an actual risk in liberal democracies, as shown by the examples mentioned above and as was already noted by Hannah Arendt more than 70 years ago:

> "Propaganda is one, and possibly the most important, instrument of totalitarianism for dealing with the nontotalitarian world."[14]

2 See Ariana Expósito Gázquez, "La (des)información en la red", *Revista digital de Derecho Administrativo,* 2022, pp. 259 ff.

3 Take, for example, the statement by Gisela Stuart, leader of *Vote Leave*, to *BBC Radio 4* in April 2016, accordingly to which "[e]very week we send £350m to Brussels...I would spend it on the NHS", which was defined as "potentially misleading" on 21 April 2016 by the Chair of the UK Statistics Authority.

4 In respect of which Colin Stretch, General Counsel of Facebook, admitted before the US Congressional Judiciary Committee on 30 October 2017 that "[p]osts from Russian-backed Facebook accounts from January 2015 to August 2017, by Facebook's estimation, reached potentially half of the 250 million Americans who are eligible to vote".

5 See, e.g., Donald Trump's statement on Twitter on 29 November 2020: "No way we lost this election!".

6 On which there was no shortage of claims such as "[t]he Covid-19 vaccines are designed to make us into genetically modified organisms" (quoted by Jack Goodman and Flora Carmichael in "Coronavirus: False and misleading claims about vaccines debunked", *BBC News*, 26 July 2020).

7 Mention should be made, among many others, of the statement made on 1 March 2022 by Sergey Lavrov, Russian Foreign Minister, in a speech before the United Nations Human Rights Council according to which: "The Russian special military operation in Ukraine seeks to save people, demilitarize and denazify this state in order to prevent such things from happening again".

8 Article 10.

9 Article 11.

10 As is the case of the Portuguese Constitution: see Article 37.

11 See in this regard, assimilating disinformation in contemporary societies to the shadows projected on the cave wall of Plato's well-known allegory, Iolanda Rodrigues de Brito, "The world of shadows of disinformation: the emerging technological caves", *Revista da Faculdade de Direito da Universidade de Lisboa,* 2022, pp. 365 ff.

12 See, on this point, Jónatas Machado, *Liberdade de expressão. Dimensões constitucionais da esfera pública no sistema social*, Coimbra, 2002, pp. 79.

13 Judgment No. 81/84, *Diário da República,* II series, No. 26, of 31 January 1985, pp. 1025 ff.

14 See *The Origins of Totalitarianism,* Orlando, etc., 1951, p. 344.

C. The European Disinformation Action Plan

10 It was in the context of some of the experiences alluded to above that the European Union adopted in 2018 the *European Action Plan Against Disinformation.*[15]

11 Recognising the threat that disinformation poses to democratic processes and other public goods, such as the environment or the health and safety of Union citizens, this document sets as its fundamental objective the formulation of a coordinated response to disinformation, articulated around four pillars: (a) improving the capabilities of Union institutions to detect, analyse and expose disinformation; (b) strengthening coordinated and joint responses to disinformation; (c) mobilising the private sector to tackle disinformation; and (d) raising awareness and improving societal resilience.

12 Each of these pillars is in turn broken down into separate projected actions by the EU institutions and the Member States aimed at mitigating the risks of misinformation.

13 The plan is, in any case, a programmatic text, from which no rules directly applicable to concrete situations capable of being characterised as acts of disinformation can be extracted. It is rather the expression of a policy, which needs to be translated into legal instruments of European or national scope. The latter will be addressed in the following.

D. The Portuguese Charter of Human Rights in the Digital Age

14 Among the statutes adopted by the Member States of the European Union with incidence on the matter under discussion is the *Portuguese Charter of Human Rights in the Digital Era*, approved during the Portuguese presidency of the Union by Law No. 27/2021, of 17 May 2021.

15 This law enshrines several individual rights related to access to and use of digital media. Among them, the following stand out: (a) The right of access to the digital environment, under which the State is responsible for promoting, among other things, the creation of a social tariff for access to Internet services applicable to economically vulnerable users; (b) The guarantee of access to and use of the Internet: the intentional interruption of Internet access, whether partial or total, or the limitation of the dissemination of information or other content are prohibited, except in circumstances provided for by law; (c) The right to privacy in a digital environment: everyone has the right to communicate electronically using encryption and other forms of identity protection or to avoid the collection of personal data, namely to exercise civil and political liberties without censorship or discrimination; (d) The use of artificial intelligence and robots is to be guided by the respect for fundamental rights; (e) The right to Internet neutrality: content transmitted and received in the digital environment should not be subject to discrimination, restriction or interference in relation to the sender, recipient, type or content of the information; (f) The right to be forgotten; (g) The right to cybersecurity; (h) The right to creative freedom and content protection in the digital environment; and (i) The right to protection against abusive geolocation.

16 Several provisions of the Charter, as has been noted,[16] seem redundant in the light of the Constitution and ordinary law, which regulates, for example, the protection of personal data, as well as copyright and related rights against their misuse in the digital environment. It is true that, according to Article 16(1) of the Portuguese Constitution, the fundamental rights enshrined therein do not exclude any others set out in applicable international laws and legal rules. However, most of the fundamental rights enshrined in the Constitution extend to the digital environment. A specific regulation of the exercise of those rights in this area would not therefore seem strictly necessary.

E. The right to protection against misinformation

17 Nevertheless, an exception to this redundancy is found in Article 6 of the Charter, which enshrines the right to protection against disinformation, which Paragraph 2 of that provision defines in the following terms:

> "any demonstrably false or misleading narrative created, presented and disseminated for obtaining an economic advantage or deliberately deceiving the public, and which is likely to cause public harm, namely a threat to democratic political processes, public policy-making processes and public assets."

18 Disinformation would include, according to Paragraph 3, "the use of manipulated or fabricated texts or videos, as well as practices for flooding electronic mailboxes and the use of networks of fictitious followers".

15 JOIN (2018) 36 final, of 5 December 2018.

16 See Domingos Soares Farinho, "The Portuguese Charter of Human Rights in the Digital Age: A Legal Appraisal", *Revista Española de la Transparencia*, 2021, pp. 85 ff.

19 Under Article 6(1), the State is to ensure compliance in Portugal with the European Disinformation Action Plan, in order to protect the society against *de jure* or *de facto* natural or legal persons who produce, reproduce or disseminate a narrative considered to be disinformation.

20 Furthermore, according to Paragraph 5, everyone is entitled to submit complaints against the entities that perform the acts referred to in Article 6 of the Charter, and have them examined by the Media Regulatory Authority. In those cases, also in accordance with Paragraph 5, the means of action referred to in Article 21 of the Charter (which regulates popular action for defence of the provisions of the Charter) and the provisions of Law No. 53/2005, of 8 November 2005, which created the said Authority, concerning complaints and sanctions, are applicable.

21 The State was also charged, under the terms of Paragraph 6, to support the creation of fact-checking structures by duly registered media organs and to encourage the attribution of quality seals by trustworthy entities endowed with the status of public utility.

F. Questions of constitutionality. Revision of the Charter

22 Were it accepted that, under Article 6 of the Charter, the Portuguese Media Regulator ("Entidade Reguladora da Comunicação Social") would be authorised to order the rectification or removal of information classified by it as disinformation, then that provision would permit a restriction, of indefinite scope, on freedom of expression.

23 However, under the terms of Article 18, Paragraphs 2 and 3, of the Constitution, such a restriction – like that of any other fundamental right –, since it is not expressly provided for in the Constitution, is only permissible under strict conditions.[17] These include the requirement that restrictions on fundamental rights be justified by the need to safeguard other constitutionally protected rights or interests; that they be proportionate; that they be defined by law; and that they do not reduce the extent and essential content of the rights, freedoms and guarantees.

24 It is therefore not surprising that, on 29 July 2021, the President of the Republic asked the Constitutional Court to review the constitutionality of Article 6 of the Charter.[18] Among the reasons invoked by the President of the Republic for the potential unconstitutionality of this provision was the deficient legal definition of the concept of disinformation. According to the President, the concepts used in the law for this purpose were too vague and indeterminate and could, as a result, restrict the content of freedom of expression disproportionately, this in violation of Article 18 of the Constitution and the parliamentary law reserve established therein. Article 6 would, on the other hand, involve the risk of censorship: the use of vague and indeterminate concepts to define disinformation could, in effect, have a censorial result, which would also be unconstitutional. Finally, it was noted in the request for a constitutionality review of the provision in question that it failed to indicate the scope of action and the attributions of the structures that would be responsible for supervising the verification of the veracity of facts reported in the media in order to confer "quality seals".

25 On 18 May 2022, the Portuguese Ombudsperson ("Provedora de Justiça") also requested the Constitutional Court to review the constitutionality of Article 6, Paragraphs 5 and 6, of the Charter.[19] As stated in the respective request:

> "[W]ithout legal criteria for its action, a specification of the concrete measures that, in this field, it may adopt, as well as a specifically designed architecture for the control of the exercise of these new powers that would minimally protect and safeguard the exercise of freedom of expression and information, the legal provision for the intervention of ERC [Entidade Reguladora da Comunicação Social] in the field of the fight against disinformation is intolerable in a democratic State based on the rule of law."[20]

26 On the other hand, the Ombudsperson stressed that the law did not ensure that fact-finding structures, which may benefit from support from the State, would be in a position to guarantee their independence from the Government, the Administration, and other public powers.[21]

27 On 17 June 2022, the Liberal Initiative Party proposed to Parliament that Article 6 be repealed, on the grounds of the risks of censorship that it allegedly

17 See, for an in-depth discussion of the subject, Jorge Reis Novais, *As restrições aos direitos fundamentais não expressamente autorizadas pela Constituição,* Coimbra, 2003, especially pp. 289 ss.

18 See the text of the request at https://www.presidencia.pt.

19 The text of the request is available at https://www.provedor-jus.pt.

20 Paragraph 57.

21 Paragraph 66.

entailed.[22] On the same date, the Socialist Party submitted a more limited proposal for amendment, which was eventually approved by Parliament.[23] Following these initiatives, Law No. 15/2022, of 11 August 2022, amending Law No. 27/2021, was adopted. As a result of that amendment, Article 6(1) of the Charter now reads as follows:

> "The State shall ensure compliance in Portugal with the European Disinformation Action Plan in order to protect society against natural or legal persons, de jure or de facto, who produce, reproduce or disseminate narratives considered to be disinformation."

28 All remaining paragraphs of Article 6 were repealed. The title of the provision retains the reference to a "right to protection against disinformation". However, the current Paragraph 1 enshrines, at most, a duty of the State to protect society against disinformation. Thus, no specific substance is given to that right by the provision as it stands. Moreover, no specific means are provided to enforce it. In particular, no public entity is entrusted with monitoring and curbing specific acts of disinformation. The position of Portuguese law in this respect is now therefore fundamentally the same as before the Charter was approved.

G. Self-regulation as an alternative?

29 Private enforcement of a right to protection against disinformation is, in principle, permitted. Information society service providers, in particular those providing virtual hosting services on online platforms, storing therein and disseminating to the public information produced by the recipients of those services, may therefore adopt their own policies in that regard, which are often set out in their terms and conditions. These are important forms of self-regulation of the activity developed by these economic agents.

30 However, significant doubts arise regarding the extent to which it will be possible to ensure effective protection against disinformation if this task of the State is delegated to the providers of information society services – not least because there would appear to be no consensus among the proprietors of the companies that operate those platforms as to the admissibility and scope of a control of the exercise of freedom of expression through the aforementioned policies.[24]

31 Even when such control is implemented by companies operating online platforms and social networks, general contractual terms enshrining such control may prove to be inconsistent with core rules of the legal system, as is illustrated by a recent judgment of the German Federal Court.[25] This ruled that the provider of a social network is, in principle, entitled to require that users of its network respect objective and verifiable communication standards which go beyond legal requirements, and also may reserve the right to take certain measures in case of violation of those standards, including deletion of individual contributions and blocking access to the network. However, the Court added in this respect:

> "[The] social network provider must undertake in its terms and conditions to inform the user of the removal of its contribution, at least immediately thereafter, and of a possible blocking of its user account in advance, to inform it of the reason for this and to give it the opportunity to make a counterclaim, which is followed by a new decision with the possibility of making the removed contribution accessible again. If there is no clause to this effect in the terms and conditions, these are ineffective pursuant to § 307 Paragraph 1, sentence 1, of the Civil Code."[26]

32 This was precisely the case under discussion in the judgment, which found that *Facebook*'s Terms and Conditions provided for no such a possibility.

H. Liability for misinformation?

33 One may moreover ask whether liability in tort can be invoked against perceived disinformation.

34 Article 37(4) of the Portuguese Constitution enshrines the principle according to which everyone is guaranteed the right to be compensated for damage suffered as a result of offences committed in the exercise of freedom of expression. But such a compensatory claim can only be granted if the general requirements of tort liability are met.

22 Bill No. 179/XV/1.ª, available at https://www.parlamento.pt.

23 Bill No. 180/XV/1.ª, available at https://www.parlamento.pt.

24 As evidenced by Elon Musk's statement after having acquired a well-known American social network: "The bird is freed" (Twitter, 28 October 2022).

25 Judgment of 29 July 2021, case No. III ZR 179/20, available at https://juris.bundesgerichtshof.de.

26 This provision reads as follows: "Provisions in standard business terms are ineffective if, contrary to the requirement of good faith, they unreasonably disadvantage the other party to the contract with the user".

35 In light of Article 483, Paragraph 1, of the Portuguese Civil Code, only the following variants of unlawful acts give rise to tort liability: (a) The infringement of a subjective right of another person (for example, a right of personality); (b) The infringement of a protective legal provision (which must specifically protect private interests, as is the case, for example, with rules on unfair competition, and not only public interests); and (c) The abuse of rights, whose unlawfulness is provided for in Article 334 of the Civil Code.[27]

36 Except to the extent that it may be construed as an abuse of freedom of expression,[28] disinformation, in the sense mentioned above, will hardly fall into any of these categories.

37 In particular, it would not seem possible to construe disinformation, for the purposes of the first variant of unlawfulness provided for in Article 483(1) of the Civil Code, as a violation of a subjective right to protection against disinformation – which, as seen above, the revision of the Portuguese Charter of Human Rights in the Digital Age carried out by Law No. 15/2022, of 11 August 2022, has eradicated from that normative text.

38 But even if this were not the case, it is important to bear in mind that the Portuguese Law on Electronic Commerce (Decree-Law No. 7/2004, of 7 January 2004), which implemented the European Union Directive on the same subject,[29] establishes important exemptions from liability for information society service providers that host harmful content in their infrastructures.

39 In fact, pursuant to Article 12 of that Law, online intermediary service providers are under no general obligation to monitor the information that they transmit or store, nor to investigate possible offences practised within their scope; and under the terms of Article 16, Paragraph 1, an intermediary provider of the server storage service will only be liable for the information stored, under the common rules, where it has actual knowledge of an obviously illegal activity or information and fails to act expeditiously to remove or to disable access to such information.

40 This reflects the fundamental rule that, for a quarter of a century, has governed this matter across the Atlantic and which has been said to be "one of the most valuable tools for protecting freedom of expression and innovation on the Internet",[30] even if, paradoxically, it is inserted in a statute originally intended to limit that freedom – section 230 of the US *Communications Decency Act*, of 1996, according to which:

> "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."

41 This provision resonated, in what concerns copyright infringements, in the US Digital Millennium Copyright Act, enacted in 1998, section 512 of which protects compliant Internet service providers against liability arising from the making available online of copyright-protected material, while adding to it a notice and take down mechanism that is regulated as follows:

> "A service provider shall not be liable for monetary relief, or, except as provided in subsection (j), for injunctive or other equitable relief, for infringement of copyright by reason of the storage at the direction of a user of material that resides on a system or network controlled or operated by or for the service provider, if the service provider—
>
> (A) (i) does not have actual knowledge that the material or an activity using the material on the system or network is infringing; (ii) in the absence of such actual knowledge, is not aware of facts or circumstances from which infringing activity is apparent; or (iii) upon obtaining such knowledge or awareness, acts expeditiously to remove, or disable access to, the material;
>
> (B) does not receive a financial benefit directly attributable to the infringing activity, in a case in

27 In this sense, Antunes Varela, *Das Obrigações em geral,* vol. I, 10th ed., Coimbra, 2003, pp. 533 ff. Reservations are however formulated regarding the third variant of unlawfulness mentioned in the text by António Menezes Cordeiro, *Tratado de Direito Civil,* vol. VIII, Coimbra, Almedina, 2014, pp. 454 ff. For a comparative outlook, see our *Comparative Law of Obligations,* Cheltenham/Northampton, 2021, pp. 275 ff.

28 As admitted by Mafalda Miranda Barbosa in situations where "the facts in question are manifestly and consciously false, and were disseminated to obtain an advantage at the expense of sacrificing the information of others": see "*Fake news* e *fact-checkers*: uma perspetiva jurídico-civilística", *Revista de Direito da Responsabilidade,* 2021, pp. 733 ff. In the sense that freedom of expression, although not requiring the truth of the facts expressed or the logical correctness of the reasoning, "does not allow, however, the conscious act of deceiving others", see Elsa Vaz de Sequeira, "Responsabilidade Civil e liberdade de expressão", *Revista de Direito da Responsabilidade,* 2021, pp. 63 ff.

29 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market.

30 Such is the view expressed, e.g., by the Electronic Frontier Foundation: cf. https://www.eff.org/issues/cda230.

which the service provider has the right and ability to control such activity; and

(C) upon notification of claimed infringement as described in paragraph (3), responds expeditiously to remove, or disable access to, the material that is claimed to be infringing or to be the subject of infringing activity."

42 This *knowledge-based approach* to liability of hosting service providers later found acceptance in the European Directive on Electronic Commerce, whose Article 14, Paragraph 1, states that:

"Where an information society service is provided that consists of the storage of information provided by a recipient of the service, Member States shall ensure that the service provider is not liable for the information stored at the request of a recipient of the service, on condition that: (a) the provider does not have actual knowledge of illegal activity or information and, as regards claims for damages, is not aware of facts or circumstances from which the illegal activity or information is apparent; or (b) the provider, upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the information."

43 Albeit not always peacefully accepted,[31] the rule in question is justified, as this author has noted elsewhere,[32] by the need to render the Internet and the development of e-commerce viable, which otherwise would be significantly impaired if the service providers in question were to be held liable without limitation for financial losses caused by the contents they host on their servers, but fail to control.[33]

44 Protection against misinformation through tort liability of the providers of such services is thus excluded in a wide range of situations, in Portugal and in several other countries.

I. The EU Digital Services Act and disinformation

45 It is of interest, in this regard, to inquire whether the recent European Union Digital Services Act,[34] which will apply from 17 February 2024, will alter the situation described above.

46 This European legal instrument, which took the form of a Regulation, also aims to combat disinformation.[35] However, it preserves, albeit amended, the approach underlying the liability exemptions provided for in the e-commerce Directive, Articles 12 to 15 of which are replaced by Articles 4, 5, 6 and 8 of the Regulation.[36]

47 Such is the purpose notably of Article 6(1), according to which the virtual hosting service provider will not be liable for the information it stores at the request of the recipients of its services, provided that: (a) it has no actual knowledge of illegal activity or illegal content and, as regards claims for damages, is not aware of facts or circumstances from which the illegal activity or illegal content is apparent; or (b) upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the illegal content.[37]

48 Article 8 of the Regulation states that there is to be no general obligation to monitor the information

31 See, for example, António Araújo's interrogation in *Diário de Notícias,* 6 November 2022: "The issue is not one of freedom of expression or censorship, as Elon Musk and other false 'libertarians' like him claim. The question is one of liability: is it acceptable that a commercial company disseminates lies on a planetary scale, spreads hatred among millions, makes billions in profits with that and is not held accountable?".

32 See *Problemática internacional da sociedade da informação,* Coimbra, 2005, p. 321.

33 See, on this, in more recent literature, Folkert Wilman, "The EU's system of knowledge-based liability for hosting service providers in respect of illegal user content – between the e-Commerce Directive and the Digital Services Act", *JIPITEC – Journal of Intellectual Property, Information Technology and E-Commerce Law*, 12 (2021) 3; and Pedro de Miguel Asensio, *Derecho Privado de Internet*, 6th ed., Madrid, 2022, pp. 294 ff.

34 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC.

35 See in particular recitals 2, 9, 69, 83, 84, 88, 95, 104, 106 and 108.

36 See Article 89 of Regulation (EU) 2022/2065. In light of Article 8(4) of the Portuguese Constitution, pursuant to which "[t]he provisions of the treaties that govern the European Union and the norms issued by its institutions in the exercise of their respective competences are applicable in Portuguese internal law in accordance with Union law and with respect for the fundamental principles of a democratic state based on the rule of law", the Regulation's provisions on the liability of Internet service providers will necessarily prevail over the abovementioned domestic provisions on the same subject that transposed the E-Commerce Directive.

37 See, on the Regulation's provisions on hosting service providers' liability, Pedro de Miguel Asensio, *Manual de derecho de las nuevas tecnologías: Derecho digital*, Cizur Menor, 2023, pp. 69 ff.

which providers of intermediary services transmit or store, nor will an obligation actively to seek facts or circumstances indicating illegal activity be imposed on those providers.

49 However, the Regulation enshrines a number of new ancillary duties for the service providers concerned, which seek to mitigate the risks of their infrastructures being used for disinformation purposes.[38]

50 These include the requirement for providers of very large online platforms and very large online search engines to diligently identify, analyse and assess all systemic risks arising, in the European Union, from the design or functioning of their service and its related systems, including algorithmic systems, or from the use of their services (Article 34(1)). This risk assessment includes, according to point (c) of the same provision, "any actual or foreseeable negative effects on civic discourse and electoral processes, and public security".

51 Under Article 35(1), the same service providers must put in place reasonable, proportionate, and effective mitigation measures, tailored to the specific systemic risks identified under Article 34, taking into account in particular the impact of such measures on fundamental rights. These measures may include, according to point (c), "adapting content moderation processes, [...] and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified". Service providers must nevertheless, pursuant to Article 14(1), include in their terms and conditions information on any restrictions that they impose in relation to the use of their services in respect of information provided by the respective recipients, notably any policies, procedures, measures or tools used for the purpose of content moderation.

52 In addition, under Article 45(1) of the Regulation, the Commission and the European Board for Digital Services are to encourage and facilitate the drawing up of voluntary codes of conduct at Union level to contribute to the proper application of this Regulation, considering in particular the specific challenges of tackling different types of illegal content and systemic risks. A *co-regulation model* on disinformation is thus enshrined in the Regulation. In order to implement it, a strengthened *Code of Practice on Disinformation* was adopted in 2022.[39]

53 The Regulation also contains significant measures to ensure compliance with the obligations it imposes on information society service providers. To this end, according to Article 49, Member States must designate one or more authorities responsible for the supervision of intermediary service providers and for the enforcement of the Regulation. Article 74(1) empowers the European Commission to impose fines on providers of very large online platforms or very large online search engines of up to 6% of their total annual worldwide turnover in the preceding financial year if it finds that those providers have intentionally or negligently infringed the relevant provisions of the Regulation, failed to comply with a decision ordering interim measures pursuant to Article 70, or failed to comply with a commitment made binding on them by a decision of the European Commission adopted pursuant to Article 71.

54 Notwithstanding the said exemption from liability in respect of information transmitted or stored at the request of a recipient of the service, intermediary service providers are liable, as pointed out in Recital 121 of the Regulation, for the losses suffered by recipients of the service that are caused by an infringement of the obligations set out therein. Compensation for such damage is established in accordance with the applicable national law, to be determined in accordance with common conflict-of-law rules. These include those contained in the Rome II Regulation, which in principle subjects such liability to the law of the country where the damage occurred.[40]

55 In Portugal, a breach of the provisions of the Digital Services Act – namely those relating to content moderation – may give rise to tort liability of service providers towards the recipients of their services to the extent that they qualify as rules imposing specific European duties of care for the protection of private interests.

38 See, on this point, Joris van Hoboken & Ronan Ó Fathaigh, "Regulating Disinformation in Europe: Implications for Speech and Privacy", *UC Irvine Journal of International, Transnational, and Comparative Law*, 2021, pp. 9 ff.

39 Available at https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation.

40 See Regulation (EC) No. 864/2007 of the European Parliament and of the Council of 11 July 2007 on the law applicable to non-contractual obligations, Article 4(1) of which states that: "Unless otherwise provided for in this Regulation, the law applicable to a non-contractual obligation arising out of a tort/delict shall be the law of the country in which the damage occurs irrespective of the country in which the event giving rise to the damage occurred and irrespective of the country or countries in which the indirect consequences of that event occur". See, on the subject, Dário Moura Vicente, "Responsabilidade civil por ilícitos comunicacionais transfronteiras: desenvolvimentos recentes", *Revista de Direito da Responsabilidade*, 2021, pp. 798 ff.

J. Conclusions

56 Open societies are particularly vulnerable to disinformation, which has an enormous potential to undermine democratic processes. However, in these same societies, the legal regulation of disinformation raises, as follows from the above, difficult problems regarding the reconciliation of freedom of expression with the protection of the public and private interests affected by it.

57 Recent developments in Portuguese legislation reflect these difficulties. In any case, primacy has been given in Portugal to freedom of expression; restrictions of it remain exceptional.

58 A subjective right to protection against misinformation is thus far from being effectively recognised in current law and is even less likely to be protected by tort liability.

59 Except when it comprises illegal information, disinformation hardly fits, in fact, into the categories of unlawfulness provided for in Portuguese law as possible grounds for tort liability; and in any event the exemptions from liability on which, for more than two decades, the legal framework for e-commerce has been based, apply to intermediary providers of information society services that convey it.

60 Only to the extent that the provisions of the Digital Services Regulation can be classified as rules for the protection of private interests will the breach of the duties of diligence enshrined therein give rise to liability of service providers, and only towards the recipients of the services.

61 Within the European Union, protection against disinformation thus appears today to be largely dependent on compliance by information society service providers with their duties of care in relation to the content they disseminate and, in particular, on their capacity to self-assess the "systemic risks" inherent in their activity, and to take preventive measures to mitigate them, in particular through content moderation procedures, as well as on the ability of the competent public bodies to monitor and sanction non-compliance with such duties.

62 Whether this will be a sufficiently robust response to the challenges currently posed by disinformation to democratic societies remains, for the time being, an open question.

Online sharing of Digital Design files as "use of a design"? A reassessment of the current regime of liabiliy

by **Matteo Frigeri** *

Abstract: EU Design law often appears as lacking the same strong identity that characterises trademark and copyright rights. Divergent conceptions over the scope of protection of these rights have persisted, disguised behind the pretence of a fully harmonised legal framework.

New developments in technology, social practices and business models now force us to question the extent to which design protection could apply to new forms of digital creation, distribution, and consumption of designs.

As the European Commission carries out a reappraisal of whether Design law is sufficiently flexible to remain relevant in the digital economy and what protection it can offer to rightsholders against acts of illegal online sharing of files, this article will attempt to critically assess the jurisprudence, literature, and legislative history of design legislation to determine whether immaterial forms of "use of a design" may constitute infringing acts – with a particular focus on the online sharing of Digital Design files.

This review demonstrates that the extension of protection to forms of immaterial exploitation of designs may have been an unintended result facilitated by the ambiguous wording of the legislation.

The last section of the article assesses the potential liability for the sharing of a DD file in a platform environment, a question also recently considered by the Commission's study. After recognising the crucial role of the "appearance" of a design as a condition of liability, the article discusses how this may cause Design law to be inconsistent or ineffective in tackling the online sharing of designs. In the conclusion of this article, a few possible solutions are canvassed. It is submitted that the current Commission Proposal does not satisfactorily address the conceptual issues outlined in the article, risking rather being a short-sighted and unprincipled response to a much broader necessity: a general reconceptualisation of what design should protect in the digital ecosystem

Keywords: EU Design law, Digital designs, Online sharing platforms, Peer-to-peer sharing

Recommended citation: Matteo Frigeri, Online sharing of Digital Design files as "use of a design"? A reassessment of the current regime of liabiliy, 14 (2023) JIPITEC 462 para 1

A. Introduction

1 While defining "design" is notoriously difficult, the Design Regulation ("Regulation")[1] provides a remarkably concise and clear explanation: "the appearance of the whole or a part of a product"[2]. In this simple definition, a tension can be observed between the immaterial appearance of a design and the material existence of a product; this opposition already anticipates the *leitmotif* of the discussion: how far does Design law venture into the digital domain? To what extent is the current regime of liability fit for purpose?

1 Council Regulation 6/2002/EC of 12 December 2001 on Community designs (2001) OJ L 003/1 (Regulation). Unless specified, this article will only look at the Regulation. The analysis may however may similarly be applied – *mutatis mutandis* – to the Design Directive.

2 ibid 3(a).

2 What is evident from this definition is the pivotal role played by the appearance of a design and the economic value that it attaches to products in the market[3]. This has prompted several scholars to claim that an infringement may arise from the mere use of the appearance of a design, without any physical interaction with the product (the "Abstract view" of protection)[4].

3 The standing of this theory seems to be already entrenched in the doctrinal architecture of Design law as a result of: 1) the inclusion in the Regulation of a limitation for the "acts of reproduction for the purpose of making citations or of teaching"[5]; 2) its consistency with several judicial decisions at both the national[6] and European level[7]; 3) the growing efforts by the industry to register and protect Digital Designs[8]; 4) its strong support in the academic literature[9] and, finally, 5) the increasing economic relevance of acts of immaterial exploitation of designs in the new ecosystem developing around 3D printing technology[10].

4 At the time of writing, this general evaluation of the doctrinal foundation of Design law is made even more pressing by the recent Commission Proposal for amending the Design Regulation ("Commission Proposal")[11]. While the industry's concerns regarding the growing threat of the use of 3D printing technology has been addressed in the newly introduced Article 19 (d), less clear is how this new provision will impact the protection of purely Digital Designs – namely, designs intended exclusively to be used in digital form or not intended to be printed. In the following discussion, possible futures of design protection will be canvassed.

5 Considering that the Commission Proposal aims to provide a clarification of the current scope of Design law[12], it is paramount that any amendment of the existing regime does not undermine the current level of legal certainty[13]. Looking at the present system, the study carried out by the Commission in 2016 ("Legal Review") highlighted the existing confusion over the definition of the subject matter of design protection – in particular, regarding the concept of product[14]. The available empirical evidence also suggests that the design community finds the law confusing, blaming courts for this state of affairs[15]. A historical perspective reveals that, while courts bear some responsibility[16], the uncertain scope of Design law seems to be a more endemic problem. Two factors help us to explain this situation.

6 The drafting of the Regulation took place in a state of diverging national practices, with such strong differences that any attempt at harmonisation

3 Commission, 'Green Paper on the Legal Protection of the Industrial Design (Green Paper)' (1991) III/F/5131/91-EN, para 2.1.2.

4 See Ana Nordberg and Jens Schovsbo, 'EU Design Law and 3D Printing: Finding the Right Balance in a New E-Ecosystem' in Ballardini et al. (eds), *3D Printing, Intellectual Property and Innovation: Insights from Law and Technology* (1st edn, Kluwer Law International 2017); Natalia Kapyrina, 'Limitations in the Field of Designs' (2018) 49 IIC – International Review of Intellectual Property and Competition Law 41; Mikko Antikainen, 'Differences in Immaterial Details: Dimensional Conversion and Its Implications for Protecting Digital Designs Under EU Design Law' (2021) 52 IIC – International Review of Intellectual Property and Competition Law 137. The Commission also endorses this theory in his review: Commission, 'The Intellectual Property Implications of the Development of Industrial 3D Printing (Commission study), (2020) doi/10.2873/85090.

5 Regulation (1) art 20(1)(c).

6 A notable case is BGH GRUR 2014, 175 Geburtstagszug (the Birthday Train case), a German case in which the registered design for the shape of a train was relied on to prevent reproduction of images of the train on the company's commercial brochure.

7 Joined cases C-24/16 and 25/16 *Nintendo v. BigBen* ECLI:EU:C:2017:724.

8 Rainer Filitz, Joachim Henkel and Jörg Ohnemus, 'Digital Design Protection in Europe: Law, Trends, and Emerging Issues' [2017] ZEW – Centre for European Economic Research Discussion Paper No. 17-007 para 3.1.

9 See generally footnote 4.

10 Nordberg and Schovsbo (n 4) para 13.02.

11 Commission, 'Proposal for a Regulation of the European Parliament and of the Council amending Council Regulation (EC) No 6/2002 on Community designs and repealing Commission Regulation (EC) No 2246/2002 (Commission Proposal)' COM (2022) 666 final.

12 ibid 2.

13 Commission, 'Staff Working Document Impact Assessment Report of the Commission Proposal' SWD (2022) 368 final, 108.

14 Commission, 'Legal Review on Industrial Design Protection in Europe (Legal review)' MARKT2014/083/D, 12, 57-60.

15 Alexander Carter-Silk and Michelle Lewiston, 'The Development of Design Law –- Past and Future: From History to Policy' (2012) SSRN Electronic Journal 118.

16 See section "*IV. Nintendo v. BigBen: towards a judicial recognition of the 'abstract' protection theory at the European level?*"

was deemed "hopeless"[17]. The difficulty in coming to a common agreement stemmed from opposing normative conceptions of what Design law should protect: a clash between the "copyright approach" to design (epitomized by French Design law) and the "patent approach" (characteristic of the legislation of the Nordic countries)[18]. The problem was exacerbated by the variety of industrial interests that Design law was meant to protect, a factor that played an evident role in shaping early proposals[19].

7 Despite these early obstacles, the adopted solution consisted in introducing a new design legislation with its own autonomous identity and rationale. The doctrinal foundations of this new legislative instrument were laid in the proposal for a "European Design law", devised by the Max-Planck-Institute working group ("MPI Proposal")[20]. Despite a promising consistency and clarity of purpose, the principles expressed in the MPI Proposal were arguably tainted during their transposition into the EU legislation. During this process – later analysed more in detail – several amendments were introduced that have allegedly altered or at least blurred the scope of protection afforded by the legislation, most notably by including an exception to the right to *reproduce* a design for the purpose of *citation*[21].

8 Questions on whether Design law could extend to "images appearing on a computer screen as a result of a program being loaded"[22] – in other words, purely Digital Designs – were surprisingly already being discussed shortly after the enactment of the Regulation; the technological advancements of the past 20 years have however opened up possible new forms of exploitation of designs – either by using them purely in a digital format (e.g., in the context of gaming) or with a view to print them as a new product – that were not fully anticipated at the time. New online platforms and business models have proliferated in response to the increase in accessibly priced 3D printing technology[23], the entrenchment of new social practices (e.g., the Maker Movement[24]) based on the online sharing of Digital Design files ("DD file"), and the distribution of new software for the creation and modelling of DD files[25].

9 It is therefore useful to look at how seamlessly the Regulation has evolved to reflect these developments. The Commission's regulatory response has largely been anticipatory rather than reactive. In fact, it mostly addresses what is the industry's fear of future mass-infringement of designs rather than a present and documented threat. These concerns should however not be dismissed as unrealistic. DD files are already being illegally downloaded via platforms such as Pirate Bay[26], and legal claims for design infringement have been brought against DD file-sharing platforms[27]. As a result of the mass adoption of 3D printing technology, the lowering of barriers to entry in terms of skills and tools required to create designs, as well as an increase in the economic value of designs destined for pure digital consumption (e.g., digital products available in the Metaverse[28]), it is likely that litigation will increase if these platforms succeed in reaching a

17 'Rosconi Designs Working Party Report' (1992) 2143/IV/62.

18 Annette Kur and Marianne Levin, 'The Design Approach Revisited: Background and Meaning' in Jens Schovsbo, Annette Kur and Marianne Levin (eds), *The EU Design Approach - A Global Appraisal* (Edward Elgar Publishing 2018) 4-6.

19 A notable example is the proposal of the "Treviso Group" in 1989, which was modelled on copyright law and had been favoured by the textile industry, a key market sector in northern Italy where the proposal originated. See Herman Cohen Jehoram, 'Cumulative Design Protection, a System for the EC?' (1989) 11 *European Intellectual Property Review* 83.

20 Reported in Michael Ritscher, Auf dem Wege zu einem europäischen Musterrecht, GRUR Int. 1990, 559–586.

21 Article 20 (1)(c) of the Regulation.

22 Anette Kur, 'Protection of Graphical User Interfaces Under European Design Legislation' (2003) 34/1 *International Review of Industrial Property and Copyright Law* 50, 58.

23 An important milestone in this regard was the expiry of the first patents in late 2000, which coincided with an increase in sales. See A Brief History of 3D Printing at https://www.3dhubs.com/guides/3d-printing/ and Mendis et al., 'Introduction – From the Maker Movement to the 3D printing era: opportunities and challenges' in Mendis et al., *3D Printing and Beyond* (Edward Elgar Publishing 2019).

24 It could be described as a series of activities characterised by the use of digital tools and desktop fabrication machines (e.g., 3D printers) to design and produce objects, combined with *an instinctive online sharing of such creations*. See Chris Anderson, *Makers: The New Industrial Revolution* (Random House 2012) 21–22.

25 Dinusha Mendis and Phil Reeves, The Current Status and Impact of 3D Printing Within the Industrial Sector: An Analysis of Six Case Studies (Intellectual Property Office, 2015).

26 Pedro Malaquias, 'Consumer 3D Printing: Is the UK Copyright and Design Law Framework Fit for Purpose?' (2016) 6 *Queen Mary Journal of Intellectual Property* 321, 324.

27 ibid 325.

28 'What is the metaverse?' <https://about.facebook.com/what-is-the-metaverse/>

broader audience.

10 Establishing more certainty over the liability of online users and platforms is necessary to safeguard the system of incentives for the creation and distribution of quality designs whilst promoting digital "creativity and innovation"[29]. The aim of this article is to evaluate to what extent the current design regime offers protection to rightsholders against the sharing of a DD file, reviewing the jurisprudence, the legislative history of the Regulation, and the academic literature. Some tentative recommendations on possible solutions to reduce the uncertainty over the scope of protection of Design law will also be outlined. Further, as the writing of this article coincides with the submission of the Commission Proposal to its first reading, an opinion will be expressed on whether legislation in its current form sufficiently addresses the concerns individuated.

B. New frontiers: 3D printing technology and online sharing of DD files

I. 3D Printing and the Maker Movement – the threat of the "zero marginal cost society"

11 While it is important to reiterate that Digital Designs intended for a purely digital consumption are likely to become an increasingly relevant category of designs[30], there is no denying that the threat – or opportunity – of 3D Printing[31] was a key motivation for the legislative reform[32]. At its most simple level, this technology consists in the reproduction of a digital model as a three-dimensional object by adding several layers of material[33].

12 Its origins can be traced back to the creation of objects with the use of a laser in the late 1960s[34]. Since its early days, the ability to create objects "impossible to mould" and unlock "effortless" creative ability were identified as the main advantages[35]. Beyond the steady improvement of the technology and its reduction in terms of costs, the appearance of online platforms where DD files are created, shared, and downloaded has profoundly altered the economic dimension of 3D printing, shifting it towards a model where production is decentralised from an industrial to a much more granular level: the individual.

13 These new business models were also the catalyst for the growth of new social practices, such as the "Maker Movement": a broad description of a series of activities characterised by the use of digital tools and desktop fabrication machines (e.g., 3D printers) to design and produce objects, combined with an *instinctive online sharing of such creations*[36]. This movement is connected to the development of Open Design – the open collaborative approach for design creation predicated on sharing information online[37] – and the FaBLabs network – a series of spaces enabling makers to have access to the necessary equipment to make (almost) everything[38].

14 The profound impact that these new developments may have in the future is well-captured by Neil Gershenfeld when he comments that the "personal fabrication [of objects] will bring the programming of the digital worlds we've invented to the physical world we inhabit"[39]. In other words, the merging of the digital and physical worlds opens new possibilities and reduces scarcity[40] by ushering us

29 Matthew Adam Susson, 'Watch the World "Burn": Copyright, Micropatent and the Emergence of 3D Printing' [2013] *Innovation Law & Policy eJournal*, 39.

30 See Antikainen (n 4) 140.

31 For the sake of simplicity, we will treat 3D Printing and additive manufacturing as interchangeable.

32 Commission Communication, 'Making the most of the EU's innovative potential. An intellectual property action plan to support the EU's recovery and resilience' (2020) COM(2020) 760 final, 6-7.

33 Tuomi et al., '3D Printing History, Principles and Technologies' in Ballardini et al. (eds), *3D Printing, Intellectual Property and Innovation: Insights from Law and Technology* (2017 Wolters Kluwer) 1-2.

34 Terry Wohlers, 'Early Research and Development' http://www.wohlersassociates.com/history.pdf

35 David Jones, 'Ariadne' Column (1974) *New Scientist* 80.

36 Chris Anderson, *Makers: The new industrial revolution* (New York: Crown Business 2012) 20-21.

37 Séverine Dusollier and Thomas Margoni, 'Open design' in Cornu-Volatron et al. (eds), *Dictionnaire des Biens Communs*, (2nd edn, Presses Universitaires de France 2021).

38 FabLabs originated from the mind of Neil Gershenfeld, himself inspired by the famous MIT course called How to Make (Almost) Anything at the MIT Center for Bits and Atoms.

39 Neil Gershenfeld et al., *Designing reality: How to survive and thrive in the third digital revolution* (Basic Books 2017) 17.

40 Mark A. Lemley, ,IP in a World without Scarcity' (2015) 90/2

into what has been called a "zero marginal cost society"[41].

15 From the perspective of rightsholders, this scenario poses a serious risk of losing the ability to control the distribution and manufacture of products incorporating their designs, thus undermining their economic incentives to invest in the production of quality designs. In addition, 3D printing is also likely to contribute to an increase in infringements by simplifying the production chain of counterfeiting products and shortening its distribution channels[42].

II. Online Sharing Platforms

16 There exists an increasing number of platforms catering to different needs and customers. Among the platforms currently registering the highest number of users we find Shapeways[43] and Thingiverse[44]. Both platforms allow a growing number of users to create, edit and share digital designs, mostly as 3D printable models. They also act as an online repository of designs, hosting a high number of files.[45]. More generally, both platforms have the effect of democratising the design creation process by empowering individuals to create their own designs and express their creativity[46].

17 Transactions between platform users are regulated by both legal and social norms. In a relatively recent report (2015), it was found that 65% of designers active on online platforms do not use any type of license to protect their rights when sharing their designs, notwithstanding the encouragement by these platforms to use licences such as Creative Commons, Commons Attribution and GNU Public Licences[47]. As pointed out by Mendis, it may sometimes be a deliberate choice by the designers to not claim any rights in their works[48]. Alternatively, it could be interpreted as indirect evidence of the designers' desire to self-regulate by adopting codes of conduct and internal rules.[49].

III. The elements of a Digital Design file

18 The sharing of a DD file is an integral part of the 3D Printing Process. A DD file contains the digital representation of a design, which is often created with the assistance of Computer-Aided Design (CAD) software, a common standard used in many different industries[50].

19 The information on the DD file created using the CAD software can then be saved in different file formats; the most common in 3D printing are the native DWG extension[51] and the neutral STL[52]. They both act as a blueprint for the design, allowing it to exist digitally without any physical embodiment. A difference is that the DWG extension is used whenever the design is created and modelled exclusively digitally, whereas the STL extension is the standard format used for files scanned from an existing physical object.

20 Although they both contain the description of the surface geometry of the design, only the DWG file contains metadata allowing us to review the creation process and subsequently edit the design. On the other hand, the STL file is more limited in its capacity to represent the design; for example, it lacks information on colour and texture[53]. It follows that the choice of the file format is likely to affect the overall impression of the design – a crucial test for determining the scope of protection.

New York University Law Review 460, 461-3.

41 Jeremy Rifkin, *The Zero Marginal Cost Society* (Griffin 2014).

42 Nordberg and Schovsbo (n 4) 275.

43 https://www.shapeways.com/. The scale of their operations is impressive: as of December 2020, the company manufactured more than 21 million parts , with more than 1 million customers worldwide. See Shapeways's Press Release of Report First QUarter 2022 Financial Results. Accessible at: https://investors.shapeways.com/news-events/press-releases/detail/51/shapeways-to-report-first-quarter-2022-financial-results.

44 https://www.thingiverse.com/

45 https://www.thingiverse.com/about.

46 See Thomas Margoni, 'Not for Designers: On the Inadequacies of EU Design Law and How to Fix It' 4 (2013) *JIPITEC* 3 225.

47 Dinusha Mendis and Davide Secchi, 'A Legal and Empirical Study of 3D Printing Online Platforms and an Analysis of User Behaviour (Intellectual Property Office 2015), 43.

48 ibid.

49 Nordberg and Schovsbo (n 4) 278.

50 Although throughout the article the more general term DD file is used, it often implies the use of a CAD file.

51 'The DWG File Specification' (Scan2CAD 2017) https://www.scan2cad.com/blog/dwg/file-spec/

52 Tuomi et al. (n 33) para 1.04.

53 'STL files'. <https://www.adobe.com/creativecloud/file-types/image/vector/stl-file.html#>

21 This is a powerful reminder of the current limitations of this technology. In fact, except in the case of very simply shaped objects, the output of the 3D printing process is rarely a finished product; the scanning and printing of the object also entail a significant loss of detail, often capturing only the general external shape of an object[54]. The 3D printing infrastructure is also complex and still relatively expensive, especially for specific materials such as metals[55]. For all these reasons, and despite the prevailing policy discourse, it is not difficult to imagine that the unauthorised use of purely Digital Designs – either as NFTs or in a gaming context – is likely to become a more significant issue for rightsholders than 3D printing in the near future. For this reason, it is even more important to establish whether the sharing of a DD file may amount to the "use of a design".

C. Design law and Digital Designs

22 In order to understand Design law, we must appreciate its justification and. These fundamental questions underpin the notion of what Kur and Levin have dubbed the "Design approach"[56], as expressed in the original MPI proposal. Facing a highly fragmented internal market, Design law promotes and protects the marketing of high-quality products: in saturated markets composed of highly substitutable products, the function of designs resides in its diversification effect – the "opportunity for differential advantage in the marketplace" that ultimately influences consumer choices[57]. However, and differently from trademarks, the market function of a design is not to convey a message (e.g., origin) but rather to appeal by virtue of its appearance.

23 The MPI proposal became the blueprint for the current EU design legislation[58]. The unique identity of this right has been recently confirmed by the European Commission Impact Assessment, where it was said that well-designed products "create a competitive advantage for the producers".[59]

54 Nordberg and Schovsbo (n 4) 278.

55 ibid.

56 Kur and Levin (n 18).

57 Mariëlle Creusen and Jan Schoormans, 'The different roles of product appearance in consumer choice' (2005) 22/1 *Journal of product innovation management* 63.

58 Kur and Levin (n 18) 7-8.

59 Commission, 'Inception impact assessment of the Review of the Design Directive and Community Design Regulation' (2020) Ares(2020)7065286, 1.

I. The legal definition of a Design – sufficiently flexible to encompass Digital Designs?

1. Design as the appearance (of the registration) of a product

24 At the heart of Design law lies the notion of the "appearance" of a product[60]. There is no requirement for designs to be either aesthetically pleasing nor should any consideration be paid to the cognitive effect of the design on consumers. The definition of 'designs' encompasses both 2D designs (e.g., an image or ornaments) and 3D designs (e.g., models)[61].

25 There is a general consensus in the literature that Design law only protects the visual features of a design to the exclusion of the other senses[62]; the argument rests on the limiting effect of the word "appearance", which implies that the design must be capable of being perceived visually, as well as on the modus of assessment of individual character as described in Recital 14, whereby the determination is to be made by reference to an "informed user *viewing* the design"[63]. It is also worth mentioning that considerable differences exist in the jurisprudence of EU domestic courts on this point[64].

26 Despite that a literal interpretation of the original Green Paper seems to suggest that all features perceivable by the human senses should be in principle treated as features protectable by design rights[65], there is strong support for requiring that such features result from the appearance of a design in order to be considered[66]. This confirms the overarching importance of the "appearance" of a design in delimitating the subject matter which can be protected by the Design law[67].

60 Regulation (n 1) art 3. See Charles-Henry Massa and Alain Strowel 'Community design: Cinderella revamped' (2003) 25/2 *European Intellectual Property Review* 68, 71.

61 Green Paper (n 3) 64.

62 Bently et al., *Intellectual Property Law*, (Fifth edn, Oxford, Oxford University Press 2018) 744; David Musker, *Community Design Law Principles and Practice*, (Sweet & Maxwell 2002) 12.

63 ibid.

64 Legal Review (n 14) 54-64.

65 Green Paper (n 3) para 5.6.1.1.

66 Nordberg and Jens Schovsbo (n 4) 281.

67 Legal Review (n 14) 157.

27 The appearance of a design is to be protected as represented in the application for registration, highlighting the crucial role of the registration in specifying the features of the design and laying claim to its protection[68]. While courts may consider actual examples of the registered design as embodied in products, the scope of protection is exclusively determined by the representation of the design as registered[69].

28 The choices made when registering a design can have important consequences, as the judgement in PMS International v Magmatic[70] demonstrates. In this judgement, the court describes how, for example, graphically representing the design "in monochrome, with grey-scale shading" will be interpreted by courts as a claim to the design in all possible colour variations[71]. The utmost importance attributed to these choices reflects the fact that the applicant can set "the level of generality at which the design is to be considered"[72]. In other words, "the selection of the means for representing a design is equivalent to the drafting of the claims in a patent: including features means claiming them"[73]. The technical means adopted to represent a design are also of consequence. For example, a CAD file is better capable to show "subtle shadings and colours as well as decoration"[74].

2. Assessing the Novelty of a Digital Design

29 At its core, the concept of novelty means that an identical design – or one that differs only in immaterial details – must not have been made available to the public before the date of filing[75]. Under the Regulation, "making available to the public" is treated as synonymous with "disclosure", a concept broadly defined[76] as generally covering all "acts which make the design public[77].

30 This broad interpretation is counterbalanced by the "safeguard clause", an inbuilt limitation that specifies that a disclosure should be disregarded if it could not have become known "in the normal course of business to the circles specialised in the sector concerned"[78]. Once again, the appearance of the design plays an essential role in determining what may amount to a disclosure: both the literature[79] and the jurisprudence[80] support the proposition that a written description cannot suffice to disclose a design.

31 Although it is currently rare for designs to fail due to lack of novelty, this proviso could gain importance if the uploading of a DD file on a website will be treated as tantamount to an absolute disclosure. Interestingly, the case law seems to be pointing in this direction. For example, in a decision of the EUIPO's Board of Appeal – *Crocs v Holey Soles Holdings* – the effect of uploading an image of a registered design on the company website was deemed to disclose the design to the audience targeted by the website[81].

32 In so far as it remains publicly accessible, information uploaded on public websites or online databases should therefore be considered a disclosure[82]. In addition, access restrictions are not sufficient to make the disclosure obscure as long as the requirements for access can be reasonably met by the professional's circle concerned[83]. For this reason, it is safe to assume that DD files uploaded to a website amount to a disclosure as long as it is capable to

68 Bently et al. (n 62) 758.

69 *Samsung Electronics (UK) Ltd v Apple Inc. (No 1)* [2012] EWHC 1882 (Pat) para 8.

70 PMS International Group Plc v Magmatic Ltd [2016] UKSC 12, 2016 RPC 11.

71 ibid para 18.

72 Lewison J, *Procter & Gamble v Reckitt Benckiser* (UK) Ltd [2007] FSR 13, para 48.

73 Martin Schlotelburg, 'The Community Design: First Experience with Registrations' (2003) 25/9 *European Intellectual Property Review* 383, 385.

74 Jacob LJ, *Procter & Gamble* (73) para 40.

75 Regulation art 5.

76 Regulation art 7. See EUIPO Third BoA *Watt Drive Antriebstechnik v. Nanotehnologija* (2013) Case R 1053/2012–3 para 13–18.

77 Arnold J, *Magmatic v PMS International Group* [2013] EWHC 1925, para 33.

78 Regulation art 7(1).

79 Bently et al. (n 62) art 765.

80 Joined Cases T-22/13 and T-23/13 *Senz Technologies v. OHIM* EU:T:2015:310, para 24.

81 EUIPO Third BoA *Holey Soles Holdings Ltd V Partenaire Hospitalier International (Phi)* (2010) R 9/2008-3.

82 Uma Suthersanen, *Design Law: European Union and United States of America* (2nd edn, Sweet & Maxwell 2010) 95.

83 EUIPO Invalidity Division *Napco Beds B.V.v Koninklijke Auping B.V.* (2015) 000009312.

reveal the outer appearance of the design[84].

3. The Product requirement – are Digital Design files products?

33 In the Regulation, a product is defined as "any industrial or handicraft item, including *inter alia* ... graphic symbols and typographical typefaces"[85].

34 There is little by way of clarification of what an industrial or handicraft item may be, with commentators struggling to determine how far the concept of product may stretch[86]. A tension is apparent: the intuitive association of products with material objects is contradicted by the addition of symbols and typefaces within the scope of the definition.

35 The EUIPO guidelines do not provide a conclusive view on how to solve this conundrum, although they note that "designs of screen displays and icons, graphic user interfaces and other kind of visible elements of a computer program"[87] are in principle eligible for registration under Class 14-04 of the Locarno Classification. This class has experienced a steady growth in applications, despite that a considerable share of them can be attributed to a limited number of enterprises (e.g., Microsoft)[88]. This growth highlights the increasing commercial value of digital designs. While Class 14-04 offers a modest degree of certainty to specific categories of digital products (e.g., GUIs), it remains unclear where the boundaries between products and non-products are to be drawn, and on which side DD files may fall. Three potential interpretations can be envisaged.

36 First, we could resolve the tension by treating all industrial or handicraft items as products, affording protection to articles that do not fall within this "narrow definition" only when a direct or indirect specific category is available – e.g., the inclusion of a graphical symbol as a basis for treating GUIs as a "product". This is an approximation of the approach adopted by Margoni[89].

37 On the other hand, we could try to infer a common interpretation of what a product is by identifying the common element – *ejusdem generis* – in the list of items included in the Regulation. While this approach has much to commend, it suffers a severe limitation: the lowest common denominator is difficult to find.

38 A third option, suggested by Antikainen, is to treat all digital designs as products, "as long as their appearance is visible"[90]. The advantage of this option is to avoid arbitrary distinctions and ensure that Design law finds wide application in the digital world. However, the price to pay for the adoption of this solution is that the "product requirement" becomes redundant, confined to a simple obligation to identify the most suitable Locarno class under which to register the design.

39 In light of this, it should be considered how DD files could be potentially registered. Even when adopting a conservative interpretation of the product requirement, there are several options to register a DD file. A first possibility would be to register a DD file under the "printed matters" classification (Class 19-08), drawing an analogy with the registration of blueprints for architectural structures – such as gardens and buildings[91].

40 Another option is to register a digital file – e.g., a CAD file – as a "blueprint" (Class 19-08). The EUIPO guidelines treat the blueprint and the physical object represented by the technical drawing as distinguishable. Since design only protects the appearance of the product as registered, the blueprint of, for example, a house would not disclose the appearance of an actual house, only of the blueprint for the house[92].

41 However, it must be noted that the Commission report (2020) casts doubts on both solutions. Relying on Article 3 of the Regulation, the report notices how a DD file does not possess the features described in Article 3(a) – inter alia, it has no "lines, contours, shape, texture"[93]. As such, it cannot be a product.

42 While the argument has some traction, it arises from an unduly formalistic analysis of the definition of

84 Viola Elam, 'CAD Files and European Design Law' 7 (2016) *JIPITEC* 146 para 73.

85 Regulation art 3(b).

86 Bently et al. (n 62) 745.

87 EUIPO, 'Guidelines for Examination of Registered Community Designs' (2022) para 4.1.3.

88 Henkel et al., 'Digital design protection in Europe: Law, trends, and emerging issues' (2017) *ZEW Discussion Papers* no 17–007, 9.

89 Margoni (n 46) 228.

90 Antikainen (n 4) 148.

91 Nordberg and Schovsbo (n 4) 282.

92 EUIPO Guidelines (n 87) para 4.1.1.

93 Commission study (4) 63.

a product, ignoring the inherently flexible nature of the product requirement (as discussed above). A better approach would be to more generally recognise that a DD file per se cannot be protected because they are not visible. What can be protected is only the digital representation – "the appearance" – caused by the execution of the software. This would shift the focus from the product – a highly uncoherent concept – to what is actually visible and worthy of protection.

43 At least in the context of sharing DD file, the third option proposed by Antikainen appears most suitable in so far as it guarantees that digital designs are treated coherently and in a technologically neutral way. In addition, this approach would force us to question what useful purpose the product requirement is serving. The marginal role of this requirement and its inability to block registrations suggest either that the purpose is unclear or that it is ineffectively pursued.

44 However, a possible role for the product requirement seems to remain. Not limiting protection by any specific product entails that the design corpus we consider when assessing the validity of a design is equally unrestrained, causing therefore more designs to be potentially declared invalid[94]. Reform in this area of the law should therefore not be undertaken lightly.

4. The exclusion of computer programs from the definition of design

45 Computer programs cannot constitute a product for the purposes of Design law, yet no definition delimiting the scope of this exclusion is provided[95]. A possible explanation for this omission is the desire to respect the principle of technological neutrality. It is clear that the notion of computer program should include – as a minimum – the object and the source code; Nordberg and Schovsbo maintain it should also include the preparatory works as well as the visual representation of the algorithms[96].

46 An official justification for the exclusion of computer programs from the definition of "product" can be found in the Explanatory Memorandum attached to the initial 1993 Regulation proposal[97]: the Commission wanted to ensure that the protection of computer programs was to be regulated exclusively by the Software Directive[98], avoiding any cumulation based on the "look and feel" of the computer program[99]. The non-protection of the overall visual appearance of a computer program does not however exclude the application of Design law to individual graphic elements[100]. This interpretation mirrors seamlessly the judgement of the CJEU in C-393/09 *BSA*[101].

47 It remains therefore possible that the "results of running a computer program" (e.g., the design of symbols displayed on the screen) could be protected, as well as any specific graphic designs for individual elements such as icons[102]. For this reason, the exclusion of computer programs should not be an obstacle to the protection of a DD file.

D. The scope of protection of Digital Designs

48 Upon registration, protection is extended to any design producing the same overall impression on the informed user[103]. This distinctive overall impression is also known as the individual character of a design[104]. Unlike in trademark law, there is no requirement for similarity of products: protection covers all categories of products[105]. However, the nature of the product to which the design is applied must be taken into consideration when assessing its overall impression, as well as the industrial sector to which it belongs[106].

94 Bernard Volken, 'Requirements for Design Protection: Global Commonalities' in Hartwig Henning (ed) *Research Handbook on Design Law* (Edward Elgar Publishing 2021) 12.

95 Regulation (n 1) art 3(b).

96 Nordberg and Schovsbo (n 4) 279.

97 EU Commission, 'Proposal for a European Parliament and Council Regulation on the Community Design (1993 Regulation Proposal)' COM (1993) 344.

98 Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (2009) OJ L 111 16–22 (Software Directive).

99 1993 Regulation Proposal (n 97).

100 Kur (n 22).

101 C-393/09 BSA v. Ministerstvo Kultury ECLI:EU:C:2010:81.

102 Commission study (4) 61.

103 Regulation (n 1) art 10.

104 ibid art 6.

105 C-361/15 P *Easy Sanitary Solutions v Group Nivelles and EUIPO* ECLI:EU:C:2017:720 para 96.

106 Regulation (n 1) recital 14.

49 The assessment consists of a four-step examination: 1) identify the sector to which the product(s) belong; 2) construct and delineate the profile of the informed user of those products[107]; 3) assess the designer's degree of freedom in the creation of the design; and 4) compare the designs at issue in terms of their overall impression[108].

50 It is submitted that the identification of the sector of the DD file (step 1) and the determination of the degree of freedom attributed to the designer (step 3) present the most interesting conceptual issues with regard to DD files in terms of implications for assessing the scope of protection. For this reason, after briefly discussing the characteristics of the informed user (step 2), the following sections will focus on how the uncertainty of the sector is likely to affect the identity of the informed user and what the constraints to the freedom in the creation of digital designs are. There is an underlying common to these questions: is the current conceptual architecture sufficiently flexible to adapt to digital designs?

I. Imagining the informed user – towards an informed user of Digital Design files?

51 The informed user determines the standard by which the design is to be judged. The attributes and knowledge imputed to this fictitious character affect the importance to be attributed to differences in the designs[109]. Positioned in between an expert in the sector and the "average consumer"[110], the informed user "has knowledge of the design corpus and the design features included in the designs existing in the *sector* concerned"[111], is interested in the products, and shows "a relatively high degree of attention when he uses them"[112]. In addition, whether the informed user would make a direct comparison between the designs depends on the practices and customs in the "sector concerned", as well as the "handling to which [the product in question] is normally subject".[113] Although it is unclear how the "sector concerned" is to be identified exactly, recital 14 suggests that, in assessing the individual character, attention must be paid to "the industrial sector to which [the products in which the design *is* applied/incorporated] belongs"[114].

52 The uncertainty in the identification of the sector affects the analysis of the identity of the informed user, as illustrated by the following example. Imagine that the registered design for a bottle opener is faithfully reproduced in a CAD file. The question would then be: who is the informed user? Should the sector be inferred from the product in which the design is applied (the infringing product) or the product represented by the design as per the registration, or again, the product in which the design was intended to be incorporated, as specified in Article 36(2) of the Regulation? In answering these questions, Elam submits that in the future the informed user could be identified in "a user of a 3D platform, who wants to manufacture the bottle opener"[115]. The consequence of such a finding would be to attribute to the informed user knowledge of the "specific methods and techniques" of the creation of Digital Designs[116]. In turn, this would likely alter the assessment of the overall impression produced on the informed users, especially when differences between designs can be attributable to the specific technique or nature of the program used.

II. The Freedom of the Digital Designer

53 Under Art 10(2), the margin of freedom enjoyed by the designer when developing the design – the design freedom – is a crucial element in the assessment of

107 It should be noted that informed user is a legal construct. This fictional character is constructed in accordance with the purpose of the products in which the design is intended to be incroproated; the informed user then is used to determine the the degree of awareness of the prior art and the level of attention in the comparison of the designs. See C-281/10 *PepsiCo v Grupo Promer Mon Graphic* (GC) ECLI:EU:C:2011:679 para 53, 55, 59; T-9/07 *Grupo Promer Mon Graphic v OHMI – PepsiCo (Représentation d'un support promotionnel circulaire)* (GC) ECLI:EU:T:2010:96 para 62.

108 T-526/13 *H&M Hennes & Mauritz v OHMI – Yves Saint Laurent (Sacs à main)* (GC) ECLI:EU:T:2015:614 para 32-34.

109 Bently et al. (n 62) 775.

110 *PepsiCo* (n 107) para 53.

111 *Samsung Electronics (UK) Ltd v Apple Inc. (No 1)* [2012] EWHC 1882 (Pat) para 34, referring to *PepsiCo* (n 107) para 54 and 59.

112 *PepsiCo* (n 107) 59.

113 ibid para 55 and C-102/11 P *Herbert Neuman v EUIPO/José Manuel Baena* ECLI:EU:C:2012:641 para 57.

114 Regulation (n 1) recital 14.

115 Elam (n 84) para 85.

116 ibid 93.

the scope of protection[117]. Several factors may curtail the freedom of the designer. These limitations are not confined to the technical function of the product but encompass all other constraints affecting the design[118] such as, for example, the customs, expectations, and regulations in the industrial sector of the product concerned[119], as well as the saturation of the market in terms of already existing designs for the particular product[120]. As a guideline, we can say that the more freedom attributed to the designer, the more differentiation will be required before a product can be considered to produce a different overall impression vis-à-vis other designs[121].

54 It is often maintained that purely Digital Designs generally enjoy a very high degree of freedom[122]; however, this oftenneglects important constraints and limitations under which the designers are operating. An illustrative example of this is the *TeamLava case*[123] where the court properly identified the multiple limitations that the designer had to respect when developing the design for computer icons, such as the size of the screen and other technical specifications.

55 The picture becomes more complex when we look at designs specifically developed to be suitable for 3D printing ("Hybrid Designs"). In such a case, the printer's specifications (e.g., height, size), and the physical limits of the material used (e.g., the 'minimum wall thickness')[124] may act as constraints. At the same time, these limitations are partially offset by the ability to create complex geometries which significantly enhances the designer's freedom[125].

56 A more serious challenge to the existing legislative framework is that, in some cases at least, it is not possible to distinguish between a purely Digital Design from a Hybrid Design without first inquiring into the actual intentions of the designer. It is therefore highly problematic that the design freedom – and consequently, the scope of protection – may depend on the subjective intentions of the designer.

57 However, a practical solution can be envisaged: As long as the appearance of the Digital Design is determined by the product it purports to represent, the degree of design freedom should reflect the technical or functional considerations normally attached to the designing of the product[126]. Although admittedly this approach raises several conceptual problems, these difficulties stem from the ambiguity of the product requirement and the unresolved conflict between immaterial and material forms of exploitation of designs.

III. The overall impression test in the context of dimensional conversions

58 This section considers the effects of the dimensional conversion (3D to 2D, or vice versa) on the overall impression produced by a design: would an informed user perceive a 2D design as producing a different overall impression than its counterpart in 3D form? In keeping with the example of the screwdriver, would the digital reproduction (e.g., reproduced by an eBook reader) of the appearance of its design infringe the registered design?[127]

59 It is possible to argue that a dimensional conversion necessarily entails a different overall impression as the informed user is unlikely to be confused[128]. An opposite argument would be that a mere digital conversion cannot produce a different overall effect as the purpose of such reproduction is to faithfully replicate the existing design in a 2D form[129]. Due to the paucity of rulings addressing this issue[130], it is not possible to conclusively settle which position should be preferred. However, replacing the overall impression test with a confusion test is a dangerous course to take as the latter may be considerably more

117 Regulation (n 1) art 10(2) and recital 14.

118 *Procter & Gamble* (n 72) para 29. See also Bently et al. (n 62) 779.

119 11/08 Kwang Yang Motor v OHIM (2011) (GC) ECR II-265 para 27 and 33; *Grupo Promer* (n 107) para 67 and 70.

120 Elam (n 84) para 95.

121 *Kwang Yang Motor* (n 119) para 33.

122 Antikainen (n 4) 155–56.

123 EUIPO Third BoA *TeamLava LLC v. King.com Limited* (2016) Case R 1951/2015-3 para 43.

124 Elam (n 84) para 96.

125 ibid 97.

126 Antikainen (n 4) 156.

127 ibid 45.

128 Margoni (n 46) para 45.

129 Malaquias (n 129); Antikainen (n 4).

130 Darren Smyth, 'How Is the Scope of Protection of a Registered Community Design to Be Determined?' (2013) 8 *Journal of Intellectual Property Law & Practice* 258.

stringent[131].

60 It is also important to note that the informed user, in assessing the overall impression, will automatically disregard elements "that are totally banal and common to all examples of the type of product in issue", concentrating instead on "features that are arbitrary or different from the norm"[132]. This could mean that the informed user may not notice differences attributable to a change of format, or other features which could be deemed trivial, common, or conventional.

61 Moreover, while dimensional conversion could be relevant for unregistered designs[133], this is less so for registered designs. After all, the scope of protection of the design is determined by the design as registered[134] while the existence of a physical product embodying that design is not necessary for protection to be granted[135]. In other words, most of the cases of design infringement involve some form of "dimensional conversion": namely, a comparison between the graphical representation of the design as registered[136] and the infringing 3D product[137,138].

62 Looking at the matter from a more technical perspective, the overall impression of a design may be substantially affected by the technique used to convert it – e.g., either by printing or digitalising it with the use of a 3D scanner[139]. For example, limitations in the technology itself may cause a loss of detail or intensify the presence of noise in the scan of the surface of the object.

63 Finally, the ability of the applicant to determine the technical means of representation, as well as the level of specificity and detail of the design represented[140] is likely to considerably affect the scope of protection. Whether dimensional conversions are covered by the registered designs is therefore not an issue that can be resolved in the abstract without reference to a specific design but rather depends on an evaluation on a case-by-case basis. There seems to be no reason why dimensional conversions should not fall within the scope of protection of design rights.

E. Drawing the boundaries of the right to "use a design" – a critical review of the "abstract protection theory"?

64 Article 19 states that a design registration confers on its holder the exclusive right to "use a design", a concept which includes at least the right to authorise the "making, offering, putting on the market, importing, exporting, or using of a product in which the design is incorporated or to which it is applied"[141].

65 Bently maintains that design rights should be limited to activities of the same nature as those listed in Article 19[142]; it follows from this reasoning that there is no infringement of a design without *the use of a product*, a conclusion further reinforced by a literal interpretation of recital 14 of the Regulation. Under this approach – the "concrete" view of protection ("Concrete view") – "use of a design" becomes synonymous with "use of a *product* in which the design is incorporated/applied".

66 An opposite position is taken by the proponents

131 Lack of confusion is not sufficient to exclude a finding of same overall impression, although confusion could be evidence of it.

132 *Grupo Promer* (n 107) para 74.

133 Under Article 11, it is *inter alia* the publication of the design which triggers its protection as an unregistered design (UCD).

134 The new proposal for a Design Regulation further reinforces this by specifying in Article 18a that only the 'features of the appearance ... of a design which are shown visibly in the application for registration' shall be protected. See Commission Proposal (n 11) art 18a.

135 Elam (n 84) para 52.

136 Council implementing Regulation No 6/2000/EC (2002) No 2245/2002 art 4.

137 Adopting a dicta by Kitchen LJ: "The scope of the protection must be discerned from the graphical representation and the information it conveys". Kitchen LJ, *Magmatic v PMS International Group* [2014] EWCA Civ 181 para 31.

138 The courts have not treated the informed user has having any problem dealing with such cases so we should not expect, following this logic, any more difficulty in perceiving the distinctive character of two designs when both are in 2D – e.g., the registered design compared with a digital 2D reproduction.

139 For example, 3D Laser Scanning allow to digitalise only object surfaces within "the line of sight" of the instrument, excluding therefore the internal – albeit visible – features. See '3D Laser Scanning Limitations' <https://www.engineersedge.com/inspection/3d_laser_scanning_limitations.html> accessed 14 May 2022.

140 Procter & Gamble (73) 48.

141 Regulation (n 1) art 19.

142 Bently et al. (n 64) 972.

of the so-called "abstract" view of protection ("Abstract view"), which argues that '*in addition* to the making, offering, ... of a design' the exclusivity *also* covers *immaterial forms of use* of a design[143]. Such an interpretation, the argument goes, is consistent with the intention of the drafters to not unduly limit the concept of "use of a design" in anticipation of future technological developments[144]. Under this theory, the scope of design protection extends to the "design as such", independently of the product in which it is incorporated.

I. Examining the doctrinal arguments in favour of the "abstract" protection theory

67 Kapyrina provides one of the most elaborated arguments in favour of extending the scope of protection to immaterial uses of the design[145]. The argument goes as follows: Recital 7 of the Regulation directs Member States to grant "enhanced protection" for the purpose of encouraging innovation and the development of new products; this "enhanced protection" extends beyond the design rights as construed in the pre-harmonisation era in the jurisprudence of the CJEU, which limited design protection to the right to "prevent third parties from manufacturing and selling or importing, without its consent, *products incorporating the design*"[146]. According to Kapyrina, the adoption of the Regulation marked a shift in the interpretation of the CJEU, as evidenced by the court's explicit recognition that design rights grant protection to '*the appearance of the product*'[147].

68 It must nonetheless be noted that this argument relies on a selective reading of the case law. In particular, the author relies on C-238/87 *AB Volvo* case[148] to demonstrate how – pre-harmonisation – the Concrete view was largely accepted as valid by the CJEU, a position from which it departed in post-harmonisation cases such as C-23/99 *Commission c/France*[149]. However, it should be noted how in C-238/87 AB Volvo the preliminary question referred to the Court concerned a UK Registered Design; in specifying that the product must be incorporated in the design, the CJEU merely took notice of the fact that, under the national law then in force, a design needed to be "applied to an article by any industrial process or means"[150]. Rather than a policy change, the different formulation used in the in C-23/99 *Commission c/France*[151] may be attributed instead to the differences in the definition of design in the Directive[152]. Whether this also imports a shift in the scope of protection is exactly the question in need of an answer. Finally, the case is an infringing proceeding on quantitative restrictions of goods and does not purport to give an interpretation on the scope of protection of design rights and, most importantly, does not concern a form of immaterial exploitation of a design – the cited portion of the judgement refers instead to "the manufacturing, sale and importation of products"[153].

69 Looking now to more recent developments in the jurisprudence, the German Case I ZR 56/09 *Deutsche Bahn v Fraunhofer-Gesellschaft*[154] is often cited as a judicial recognition of the Abstract view[155]. In this case, the German Federal Court found that the reproduction of the design of the train (ICE 3) in the trade fair catalogue infringed the rights conferred by the registered design under § 38 (1) Geschmacksmustergesetz[156, 157].

70 Considering that the wording of § 38 (1) is identical

143 Antikainen (n 4).

144 Mario Franzosi (ed), 'European Design Protection: Commentary to Directive and Regulation Proposals' (1996) 20 *European Intellectual Property Review* 131.

145 Kapyrina (n 4).

146 C-238/87 *AB Volvo & Erik Veng* ECLI:EU:C:1988:477.

147 C-23/99 *Commission c/France* ECLI:EU:C:2000:500 para 42.

148 *AB Volvo* (n 146).

149 *Commission c/France* (n 147).

150 Registered Design Act 1949, s 1(1).

151 *Commission c/France* (n 147) para 42.

152 Directive 98/71/CE of the European Parliament and of the Council of 13 October 1998 on the legal protection of designs (1998) L 289/28 art 1(a).

153 *Commission c/France* (n 147) para 42: "Use of the appearance of the original design" refers to the manufacturing of products made to that design.

154 BGH ZR 56/09 *Deutsche Bahn v Fraunhofer-Gesellschaft* (7 April 2011).

155 David Stone, *European Union Design Law: A Practitioners' Guide* (Oxford: Oxford University Press, Incorporated 2016) 470; Nordberg and Schovsbo (n 4) 284-5.

156 *Deutsche Bahn (n 154)* para 29-30.

157 Gesetz über den rechtlichen Schutz von Mustern und Modellen (Geschmacksmustergesetz) (2004) BGBl. I S. 390. See Nordberg and Schovsbo (n 4) 285.

to Article 19(1), this finding reinforces the idea that it is possible to interpret the Regulation as extending to immaterial uses of the design[158]. More precisely, the adoption of this interpretation would mean that the aesthetic features of the design are protected per se. The question is then whether the CJEU should follow this approach.

71 It must be first noted that the case is not binding on EU courts. Moreover, the interpretation of the national court seems to directly follow from the tradition in German jurisprudence to conceive – pre-harmonisation – design protection as derivative of copyright (the *Kleines Urheberrecht* doctrine)[159].

72 Nonetheless, in 2015 the German Federal Court of Justice explicitly overruled this long-standing doctrine by recognizing that, after the implementation of the Design Directive, Design law was to be considered as hermeneutically independent of copyright law[160]. In doing so, the Federal Court weakened the ratio decidendi of the *Deutsche Bahn* case. It is unlikely that the CJEU in the future will ever consider the decision to be a persuasive authority in the determination of the scope of design protection.

II. Nintendo v. BigBen: towards a judicial recognition of the "abstract" protection theory at the European level?

73 In Joined Cases C-24/16 and C-25/16 *Nintendo*, the CJEU held that the inclusion on a website of images of goods corresponding to a registered design constitutes an act of reproduction for the purpose of making citations[161]. In confirming the applicability of the limitation in Article 20(1)(c), this judgement is the first explicitly recognition that the mere reproduction of an image of a design on a webpage may fall within the concept of "use of a design" under Article 19(1). This seems to constitute an endorsement of the Abstract view, insofar as it implicitly extends the scope of protection to cover both material and immaterial reproductions of a design. In its most extreme interpretation, it follows from this judgement that any form of reproduction would be covered by the design right.

74 The decision's importance should however not be overstated. After all, the literal text of the provision that the CJEU was asked to interpret referred to an "act of reproduction for the purpose of making citations"[162]. The conclusion of the court was to the same extent predetermined by the inclusion of a citation exception in the legislation. As it will be discussed later, it is difficult to justify its existence unless design rights could be infringed by bidimensional reproductions – whether digital or printed. Any other interpretation would render the scope of this exception incredibly narrow, raising the question of why it was included in the first place.

75 In other words, it appears that the judgement merely confirms the literal reading of the Regulation without really engaging with the underlying conceptual tensions between Article 19 – referring to "use of a product" and thus supporting the Concrete view – and Article 20 – which seemingly assumes the possibility that design rights may be infringed simply by reproducing the design. A textual and systematic analysis of these provisions is inconclusive, making it necessary to focus on the drafting history of the Regulation.

76 For present purposes, it suffices to say that the CJEU simply accepted the Abstract view without spending much time considering the issue. However, what the judgement does not clarify – therefore remaining a contentious issue moving forward – is how broadly the concept of reproduction should be interpreted, a point that was briefly touched upon in the Advocate-General's Opinion. The discussion is limited to a few paragraphs, where the AG cites a publication by Kaesmacher and Stamos to support an interpretation "as broad as possible" of the concept of reproduction[163]. The AG then concludes his Opinion by treating the matter as obvious: the publication of

158 France is another example of a jurisdiction where reproduction of a design of an umbrella was deemed to infringe rights in the registered design; see Paris Court of Appeal, pôle 5, ch. 2, 27 Nov. 2015, *S.A.S. Piganiol c/S.A.S. Publicis Conseil et al.*, No. 13/21612, JurisData No. 2015-029315

159 Design rights as *kleines Urheberrecht*: "*... zwischen dem Urheberrecht und dem Geschmacksmusterrecht kein Wesensunterschied, sondern nur ein gradueller Unterschied bestehe*" (unoffical translation: "[...] there is no difference in essence between copyright law and design law, but only a difference in degree", in *Geburtstagszug* (n 6) para 18. See also Kur (n 22); Kur and Levin (n 18) 53.

160 *Geburtstagszug* (n 6) para 33-40; discussed in Ansgar Ohly, 'The Case for Partial Cumulation in Germany' in Estelle Derclaye (ed), *The copyright/design interface: past, present and future* (Cambridge: Cambridge University Press, 2018).

161 *Nintendo* (n 7) para 86.

162 Regulation (1) art 20(1)(c).

163 Dominique Kaesmacher and Theodora Stamos, *Brevets, Marques, Droits d'auteur ... Mode d'emploi* (Liège : Edipro 2009) 164.

images of the design on packages as well as on the website amounts to an act of reproduction[164].

77 The AG's reliance on Kaesmacher and Stamos' statement is problematic and likely misplaced. The source of the assertion is an intellectual property textbook and, crucially, it appears in the section of the book discussing the interpretation of the concept of reproduction under copyright law, not design law; such a broad interpretation is fully supported by the definition of reproduction found in the Info Soc Directive[165]. On the contrary, the Regulation includes the act of reproduction within the rights conferred by a design only as an "afterthought"[166] and without providing a definition.

78 In addition, from reading the text of the source cited by the AG it emerges that the two authors were working under the assumption that the use of a design necessarily involves the use of a product[167]. The AG appears oblivious to this, or at least fails to make explicit why a literal interpretation of Article 19 is ignored without argument.

79 Alternatively, it is also possible to regard the AG's Opinion as implicitly supporting that the right of reproduction under the Regulation should be consistently interpreted with Article 3 of the Info Soc Directive – notwithstanding that the very broad interpretation in the Info Soc Directive stems from a very specific wording which leaves no doubt as to its wide application.

80 It is not possible to know whether the CJEU endorsed the AG's reasoning when holding that the use of "images of goods corresponding to such designs" amounts to "an act of reproduction"[168]; yet it is undeniable that the inclusion of the term "reproduction" in the wording of Article 20(1)(c) further strengthens the case for the Abstract view. For this reason, an analysis of the legislative and drafting history of Article 20(1)(c) is necessary to assess whether such an inclusion reflects a commitment of Design law to the Abstract view – in other words, whether Design law should include immaterial uses of the design.

III. An analysis of the legislative history of Article 20(1)(c)

81 In the original MPI proposal – considered the "blueprint" or the doctrinal foundation of EU Design law – there is interestingly no mention of an exception to design rights for the purpose of teaching or citation; on the contrary, the precursor to Article 20[169] consisted in only a general exclusion for acts done in private for non-commercial purposes, in addition to a more detailed list of specific acts referring to typical limitations in patent law (e.g., exceptions for installation on craft – e.g., ships – temporarily entering the Member States' territory)[170]. It is therefore safe to assume that this controversial provision was not part of the architecture of Design law as initially conceived by its founders.

82 The first traces of what was to become Art 20(1)(c) can be found in the Green Paper[171], where a provision was included to exclude from liability acts of reproduction of a design "for the purpose of teaching"[172]. Limiting this exception to the right of reproduction – whatever it may mean – is a peculiar choice, especially when considering that this term could have more naturally been subsumed under the concept of "use of a design"[173].

83 There is no exhaustive description of the acts falling under the concept of reproduction, although in the text of the Green Paper the term "reproduction" is often employed as synonymous with "manufacture" of a design product, thus most likely excluding instances of immaterial uses of a design (e.g., reproduction in a book)[174].

84 A more interesting note on the semantic use of "reproduction" can be gleaned from section 6.4 of

164 Joined cases C-24/16 and 25/16 Nintendo v. BigBen (Opinion of Advocate General Bot) ECLI:EU:C:2017:146 (AG's Opinion).

165 Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, OJ L 167 (Info Soc Directive).

166 The idea of the right of reproduction as an "afterthought" is reflected in its legislative history, as later discussed.

167 Kaesmacher and Stamos (n 163) 165.

168 *Nintendo* (n 7) para 86.

169 Then Article 23.

170 Ritscher (n 20) 528.

171 Green Paper (n 3).

172 ibid para 6.4.7.2.

173 A more natural wording could have been: "use of a design for the purpose of teaching".

174 An example of this semantic use of 'reproduction' can be found in the Green Paper's Introduction: "*Reproduction* of design products does not, in many cases presuppose know-how as regards sophisticated manufacturing process". Ibid 2.

the Green Paper, where the term suggests a specific meaning: to make a derivative copy of the protected design; it thus incorporates a subjective requirement of either fraud "or at least of negligence"[175]. It follows that "reproduction", as used in this section, implies a requirement of derivation – yet again this does not necessarily cover immaterial reproductions. If this interpretation is correct, then the use of the term "reproduction" in the teaching exception may be solely attributable to the drafters' assumption that "teaching" a design necessarily implies having prior knowledge about the design, which may suggest derivation.

85 An alternative explanation for the use of the term "reproduction" is provided by Musker, who argues that the historical origins of the provision are to be traced back to the Directive on Semiconductor Topographies 1986[176]. The similarities in the text point to the influence of this Directive on the drafting of the 1990 Green Paper[177]. Under this theory, no particularly specific meaning should be attached to it.

86 More problematic would be to explain the rationale behind the addition in the 1993 Regulation Proposal[178] of a further purpose under which acts of reproduction may be excluded from liability: the purpose of *"making citations"*[179]. The inclusion of a citation exception severely impairs the argument in favour of limiting design protection to the use of a product; for this reason, its origin should be carefully considered.

87 The amendment is most likely the result of the Commission's hearings with interested parties which were held throughout 1992[180]. Although there is no record confirming exactly when and why this provision was introduced, this is a reasonable inference based on the review of the procedural history of the Regulation.

88 What we can however glean from the available documentary evidence is that since its introduction this amendment to the original text proved to be confusing and controversial; most delegees who participated in the proceedings of the Economic and Social Committee proposed to remove the words "making citations" altogether, with three delegations commenting that the provision was not needed and would be likely to create difficulties in the interpretation of the text[181]. There is unfortunately no evidence of the ensuing discussions; the following documents report that delegees removed all their reservations within a year of raising them[182], while the amended proposal for the Community Design Regulation still reported Article 22(1)(c) [now Article 20(1)(c)] in an unaltered form[183] and no further amendments nor discussions followed.

89 It is also possible that the inclusion of a citation exception may be the result of a translation error during the drafting process, as suggested by Musker[184]. First, he notes how this limitation has no analogues in other IP rights; despite this, it does not appear to have ever been discussed in any policy document of the time. This is surprising considering its potential controversial nature, raising the suspicion that its inclusion may have been unintentional. His main argument then rests on a consideration of potential drafting mistakes in the transposition of the wording of the Article from other legislative instruments. He notes for example how both Article 10 of the Berne Convention and Article 5(3)(a) of the Info Soc Directive include an exception for the purpose of "illustration *for teaching*". In the French version, this provision would be translated as "illustration *de* l'enseignement". It is therefore easy to imagine how a small drafting mistake – replacing *de* with *ou* – would result in the following text version: "illustration *ou* enseigement" (unofficial translation: citation or teaching), thus substantially altering the meaning of the exception by giving both purposes independent standing. In its English version, it would then be possible to translate "illustration" as citation, accounting for the current wording to be found in Article 20(1)(c). Albeit quite complex and lacking strong supportive evidence, this theory offers an interesting perspective, cautioning against over-reliance on the wording of the Article. It is further reinforced by evidence of several drafting and translating errors reproduced

175 ibid section 6.4.2.

176 Musker (n 62) 834.

177 See, for example, Regulation art 13(1)(c).

178 1993 Regulation Proposal (n 97).

179 Regulation art 20(1)(c).

180 Detailed minutes of the hearing have been submitted by Commission services (III/F/5252/92) July 1992.

181 Summary of Proceedings of Working Party on Intellectual Property (Designs) (20 May 1994) (7298/94) 6.

182 Summary of Proceedings of Working Party on Intellectual Property (Designs) (9 October 1995) (10486/95) 6.

183 Amended proposal for a Council Regulation (EC) on Community Design, 21 June 1999, (COM (1999) 310 final) 28.

184 David Musker, "Making Citations'—Mystery or Mistranslation? The Opinion of Advocate General Bot in Nintendo v BigBen' (2017) 12 Journal of Intellectual Property Law & Practice 834.

in other provisions of EU Design law[185].

90 Unfortunately, the lack of access to public documents shedding light on the drafting process make any attempt to conclusively resolve these questions impossible. For this reason, the existence of a "citation exception" within Design law remains theoretically confusing, with much uncertainty revolving around its scope of application. Whether the existence of this provision is sufficient to warrant a broad interpretation of the scope of design rights as covering digital reproductions remains unresolved. What is however clear is the important role it played in shaping our current understanding of the scope of protection, supporting arguments in favour of extending protection to mere digital reproductions. Arguably, this copyright-like interpretation of design rights is made possible by the existence of this exception. It is therefore surprising that its discussion in the recent Commission's evaluation of the liability arising from the peer-to-peer sharing of DD files has been very limited.

91 In the final section of this article, and despite the inevitable uncertainty currently pervading design law, we will attempt a fresh assessment of the liability for the sharing of DD files in online platforms, questioning whether the Commission Proposal satisfactorily addresses the inconsistencies likely to result from the application of the existing framework. As it will be shown, the answer is negative; for this reason, possible ways forward to solve these inconsistencies will be canvassed, making direct reference to the reform proposal by the Commission[186].

F. Assessment of the liability for the peer-to-peer sharing of Digital Design files – a coherent framework?

I. The Commission's position on the liability for sharing Digital Design files

92 The Commission study analyses the question of liability for the sharing of a DD file[187]. For the purpose of the discussion at hand, the act of sharing a DD file can be characterised as the uploading of a DD file to a publicly accessible website (e.g., by a user or by an online platform). The view of the Commission seems to be that the scope of protection of the current liability regime is sufficiently flexible to cover such acts[188].

93 The Commission's analysis however fails to address –at least explicitly – the thorny question of whether digital reproductions fall within the concept of use of a design (the Abstract view)[189], providing no account of what "use of a design" means more generally. Instead, the study assesses the extent to which acts of "uploading" and "hosting" a DD file may be conceptualised under any of the rights of "use of a design" already explicitly listed in Art 19 of the Regulation.

94 The study finds that the notion of "offering a product made to the design" is sufficiently flexible to encompass both acts – namely, uploading and hosting a DD file[190]. However, it is submitted that by extending the concept of "offering" to a purely digital context, this approach exacerbates the doctrinal confusion. First, the Commission's interpretation is inconsistent with the text of the Regulation, which refers to the *offering* and *stocking* of a *product*. Secondly, the Commission's reasoning is self-contradictory: it maintains that offering means "proposing to a third party the *transfer of physical control* of the design-infringing products" while at the same time arguing that the design-infringing product does not need to exist at the time of offer[191]. This obviously begs the question of what "transfer of control" could mean in a purely digital context (e.g., a design product used in the Metaverse), especially considering the non-rivalrous nature of digital consumption.

95 Even accepting the Commission's premise, which predicates the notion of offer on the potential exercise of *physical control* imports in the legislation a requirement of "an intention to bring the object, as represented in the DD file, into existence" (e.g., 3D printing). Incidentally, this seems to be the approach

185 See for example Art 110 CDR as discussed in *BMW v Round & Metal* [2012] EWHC 2099 (Pat), [2013] Bus LR D30, and the very un-aligned versions of Art 11 CDR. These examples were provided in Musker (183).

186 Commission Proposal (n 11).

187 Commission study (n 4) para 4.4.2.1.

188 Ibid 140-2.

189 It could however be argued that this point is taken for granted, especially as the report accepts that digital uses of a design may in principle give rise to liability. As discussed in this article, such an assumption is problematic.

190 Commission study (n 4) 141-2

191 ibid para 4.4.2.1.

taken in the Commission Proposal[192], where a new provision is included whereby digital uses of a design – e.g., sharing a design – are deemed within the scope of design protection only if carried out for the "purpose of reproducing a product that infringes the design"[193].

II. "Use of a design" as "use of the appearance of a product": is the current regime of liability coherent?

96 In contrast to the approach taken by the Commission's study, this article argues that to understand the scope of protection of Design law it is first necessary to recognise the crucial role played by the "appearance" of a design in the legal framework.

97 A systematic reading of Article 3, 10, and 19 of the Regulation reveals that "use of a design"[194] presupposes the use of the *appearance of a product*. The argument goes as follows: a design is defined in the Regulation as "the appearance of a product"[195]; in addition, the test for infringement also heavily relies on the "appearance" – the overall impression produced by the *appearance or visual features* of a design[196]. Consequently there cannot be a "use of a design" if the design is not visible at any point in time[197]. For this reason, it is submitted that "use of the appearance of a product" is a necessary condition for design infringement.[198]

98 This seems to be confirmed by C-23/99 *Commission c/France*, where the CJEU observes that the physical transportation of a product in which the infringing design is incorporated cannot amount to an act of infringement as it does not involve "use by a third party of the *appearance* of the product"[199]. The AG's Opinion further reiterates that for "the purposes of the transport operation, *the appearance of the goods* transported is of *no importance* and *has nothing to do with the benefits which the carrier derives* from providing the transport service"[200].

99 Applying this doctrine to the act of sharing a DD file leads to an interesting result. In fact, the act of sharing or uploading a DD file on a peer-to-peer website merely provides access to information, without any visual element. It is only the running of the file on the computer of the recipient that will provide the visual element to constitute the infringement – an analytically separate and independent act of use of the design.

100 The argument is reinforced by the separation of preparatory acts from the concept of "use of a design"[201]. The acts preceding the visible reproduction of the design (e.g., the download of the design file) should therefore be classified as preparatory acts, thus removing any potential liability[202]. The sharing of a DD file online cannot per se infringe any design right; the real act of infringement is rather the reproduction of the design (e.g., in the form of JPEG). This is problematic as it makes liability depend on a contingent factor[203]: whether, in addition to providing a link to download the file, the platform's user has also uploaded a reproduction of the design[204].

101 In the digital environment, protection of the appearance per se provides only a limited safeguard to the interests that design rights are meant to protected. This leads to the conclusion that, in its present condition, the current regime of liability is conceptually capable of applying to the peer-to-peer sharing of DD files in the platform ecosystem, yet it does so in an inconsistent and unprincipled

192 Commission Proposal (n 11).

193 ibid recital 11 and art 19(2)d.

194 Regulation (n 1) art 19(1).

195 Ibid art 3(a).

196 Article 10(1).

197 This is reinforced by the centrality of the requirement of visualisation of design features, Article 36(1) and (6) CDR.

198 This generally justifies the exclusion of verbal description from design protection. See Anna Tischner, 'Lost in Communication: A Few Thoughts on the Object and Purpose of the EU Design Protection', *The Object and Purpose of Intellectual Property* (Edward Elgar Publishing 2019).

199 *Commission c/France* (n 147) para 42.

200 C-23/99 *Commission c/France* (Opinion of Advocate General Mischo) ECLI:EU:C:2000:212 para 83.

201 Franzosi (n 144) 131.

202 This classification relies on the correctness of our treatment of the digital file as medium or mere information, as distinct from the design that it incorporates.

203 It is contingent to the point of view of the purpose of design law, namely the protection of the economic value of the design. See Green Paper (n 3) para 2.1.2 and 5.4.7.1.

204 From a practical point of view, this inconsistency will not be a problem. Most often, unless the design is so famous that a verbal description suffices, a digital reproduction will accompany the download link.

manner. Most importantly from a practical point of view, it also risks making design protection easily circumventable. For example, a would-be infringer could in fact avoid liability by ensuring that at no point the design is ever reproduced, replacing instead such a reproduction with an accurate description of the design.

102 It appears intuitively correct that the sharing of DD files is an activity against which Design law should afford protection, given the economic relevance of such acts. Not only could they be considered functionally equivalent to the transfer and sale of physical designs. They may arguably also be even more prejudicial to the interests of rightsholders[205]. The problem highlighted in this article is that the current system is ineffective in affording such protection. Recent proposals for reform of Design law partly address this issue by providing a right to authorise the "downloading ... and sharing or distributing to others any medium or software recording the design" (e.g., a DD file) but only for the purpose of enabling a product to be made[206]. Although this is a positive development, the creation of a purpose-oriented produces considerable uncertainty that will have to be ultimately resolved by the judiciary[207]. For example, extending protection beyond uses of the "appearance" of a design is a considerable transformation of what we currently understand as the scope of design rights; it also stands in contrast with the new articulation of the "object of protection" of Design law in Art 18a of the Commission Proposal: "the *features of the appearance* of a design shown *visibly* in the application for the registration". In other words, this reform demonstrates how nebulous and undefined the identity of this right is in its current form[208].

103 At a time when the overall framework is being reassessed, it is important to face these conceptual challenges lest they will be exacerbated by the new developments in technology and social practices. Potential solutions will be sketched out in the final section of this article. In the conclusion, the Commission Proposal will also be briefly commented to determine whether it sufficiently addresses the issue outlined.

III. Proposal for a consistent and coherent application of Design law online – possible ways forward

104 A possible solution to the issues discussed could be to amend the current Regulation by adding that the notion of use of a design includes the "making or distributing a design document for any of those purposes" [namely – the purpose of making, offering, putting on the market ... a product in which the design is incorporated/applied – see Art 19(1)]. This option – albeit conceived in a different context – was recommended by Malaquias[209], drawing inspiration from Section 226(1)(b) CDPA 1998[210], and considered by the Commission in its 2016 review[211]. Interestingly, the new Commission Proposal opted for a very similar solution[212]. The merits of this amendment will now be assessed.

105 It must be first noted that this new ground of liability would significantly alter the current nature of Article 19, which does not cover any form of indirect infringement of design rights. In other words, once it is accepted the design need to be visible in some form in order for an act to constitute a (direct) infringement of a design, the distribution of a design document could be construed as a supply of the means to infringe such a design[213] – an act having all the hallmarks of indirect infringement – and be considered foreign to the spirit of that Article.

106 It would however be effective in ensuring consistency, being applicable to all cases of sharing of a DD file regardless if there is any reproduction of the design, and would increase legal certainty. More concerns, however, exist about the possible divergent interpretations of "making a design document". This term could be interpreted as extending to the automatic creation of a document by a computer machine, thus requiring the creation of a new exception to design rights similar to Article

205 See for a similar analysis C-263/18 *Nederlands Uitgeversverbond and Groep Algemene Uitgevers (Tom Kabinet)* ECLI:EU:C:2019:1111 para 57-58.

206 See Commission Proposal (n 11) art 19.

207 A Kur and T Endrich-Laimböck and M Huckschlag, 'Position Statement of the Max Planck Institute for Innovation and Competition of 23 January 2023 on the 'Design Package" (2023) Max Planck Institute for Innovation and Competition Research Paper No. 23-05, p. 12.

208 See Commission Proposal (n 11) art 18a.

209 Malaquias (n 26).

210 Copyright, Designs and Patents Act 1988.

211 Legal Review (n 14) 133.

212 Commission Proposal (n 11).

213 Martin Mengden, '3D-Druck – Droht eine "Urheberrechtskrise 2.0"? Schutzumfang und drohende Rechtsverletzungen auf dem Prüfstand' (2014) 17(2) *MultiMedia und Recht*, p. 80.

5(1) of the Info Soc Directive[214].

107 Another potential issue is the compatibility of the new provision with the definition of design in Article 3(a)[215]. The price for consistency would be therefore to abandon "appearance" as the kernel of design protection, moving into a new territory where designs gain protection independently of their visibility[216]. While this could be formally fixed by defining in Article 3 what a "design document" is, the prospect that an infringement of a design right may occur without at any point the design's appearance being visible raises the question of whether a particular mental state should be required before the act may attract any liability.

108 Finally, protecting DD files as design documents could potentially violate the exclusion of computer programs from the scope of design protection. This assessment is made particularly difficult by the absence of a positive definition of what a computer program is[217]. It is important however to keep in mind that this exclusion only applies to the definition of a product. As the introduction of the concept of "design document" would be independent of either the concept of "design" or "product", it is possible to argue that the exclusion simply does not apply at all. It is worth looking at other possible scenarios in case this may prove to be incorrect.

109 Malaquias compares "the sharing of a DD file" to "the sale of a computer program" on the basis that they both enable hardware (e.g., 3D printer) – to carry out an auction – (e.g., produce an object)[218]. It can however be argued that the ability to "enable" a printer to operate is not a sufficient condition. Considering the question of the copyrightability as software of CAD files under US law, Rideout maintains that since CAD files do not control the way 3D printers operate, they are not equivalent to software; rather, they function as a blueprint[219] and should be considered more akin to a graphical work than a literary work[220].

110 Since the exclusion of computer programs from the scope of protection serves the purpose of ensuring there is no overlap between Design law and copyright law in protecting software, it is also useful to assess whether the DD file could fall within the scope of the Software Directive. Although we defined the DD file as comprising the source code[221], protecting it as a computer program would be inconsistent with the requirement that the program is a literary work[222]; the author's intellectual creation does not go towards writing the source code and arguably does not involve programming at all. Protection of a DD file as a computer program seems therefore inappropriate, a conclusion reinforced by the judgement of the CJEU in *SAS Institute*[223].

214 Info Soc Directive (n 165).

215 Regulation (n 1) art 3(a). Discussed in T-494/12 *Biscuits Poult v OHMI – Banketbakkerij Merba (Biscuit)* (GC) ECLI:EU:T:2014:757.

216 The role of the 'appearance of a design' as a constitutive element of design infringement was discussed in art 96-7. Not discussed in this article is how the 'appearance of a design' may be translated into a visibility requirement applicable for all type of products – contrary to the current position, where a visibility requirement during normal use applies only to components of complex products. See Regulation art 4(2), as interpreted in 11/08 Kwang Yang Motor (n 119); Third BoA Lindner Recyclingtech v. Franssons Verkstäder (2009) R 690/2007–3; and T-494/12 Biscuits Poult v. Banketbakkerij Merva (GC) EU:T:2014:757.

217 It is preferred to avoid an ontological argument on whether data (e.g., CAD files) could be classified as computer programs; after all, courts are unlikely base their judgements on such discussions.

218 Malaquias (n 26) para 3.1.1.1.

219 This is further confirmed when we consider that an argument in favour of protecting a DD file as a computer program would also most likely apply to Word Doc and other file formats.

220 Brian Rideout, 'Printing the Impossible Triangle: The Copyright Implications of Three-Dimensional Printing' (2011) 5 *J. Bus. Entrepreneurship & L.* 161, 168.

221 The present discussion assumes that the DD file can be expressed as source code. it is important to note that this is not always the case: in AutoCAD, for example, designs are created by interactive modelling without a human-readable source code (just a binary file). This difference does not affect our conclusions: if no written language is used in the creation of the design, then it would seem even more inappropriate to protect under the Software Directive.

222 Following Case C– 5/08 *Infopaq International* ECLI:EU:C:2009:465 , an act to fall within the concept of 'reproduction' has to reproduce the elements which are the expression of the intellectual creation of the author. Arguably, the designer intellectual creation is expressed in the design itself – which may be protected as an artistic work – but not the 'source code', protected as a literary work. David Nickless, 'Functionality of a Computer Program and Programming Language Cannot Be Protected by Copyright under the Software Directive' (2012) 7 *Journal of Intellectual Property Law & Practice* 709, 709.

223 In the judgement, the CJEU held that 'neither the *functionality* of a computer program nor the *programming language* and the *format of data files* used in a computer program in order to exploit certain of its functions constitute a form of expression of that program for the purposes of Article

111 Regarding the question of how a design document is to be defined, a good starting point is once again Section 263(1) CDPA 1988. According to this provision, a design document consists of "any *record of a design*, whether in the form of a drawing, a written description, a photograph, *data stored in a computer* or otherwise". This definition is extremely wide, and sufficient to cover digital files stored on a computer and even on the cloud[224]. The requirement of visibility is somehow retained by the condition that the design document "corresponds to a record which clearly shows a visual representation of the design"[225]. DD files should be able to comply with this condition if they are capable of reproducing the design visually – e.g., should be machine-readable and produce a clear image of the design containing all its distinctive features.

112 The concept of distribution should also be interpreted as broadly as possible to ensure technological neutrality and guarantee its application to online peer-to-peer sharing of DD files. A good blueprint could be the right of distribution in the Software Directive, which covers "*any form* of distribution to the public"[226]. Despite that "distribution" is commonly understood only to apply to physical transfers, the CJEU in *UsedSoft* (2012) has extended its scope of application to digital distribution in circumstances where there is no tangible medium involved.[227]

113 It is important to stress that an essential premise of the solution proposed above is that the mere reproduction of a design constitutes a "use of a design" and can therefore give rise to liability (as stipulated by the Abstract view). As this article intended to demonstrate, this conclusion is not inevitable. For this reason, an alternative possible solution is to formally recognise in the legislation that the existence of a physical product is a necessary precondition for the infringement of a design right. Not only would this approach solve much of the conceptual uncertainty described in this article, but it would still leave open the option to extend the scope of protection of design rights to target specific factual scenarios: e.g., sharing DD files for the purpose of 3D printing.

114 This solution is not currently reflected in the Commission Proposal; on the contrary, the Commission Proposal gives further support to the Abstract view – see, as an example, the inclusion of an exception for the purpose of "comment, critique or parody"[228] – while at the same time, it includes a limited-in-scope extension of design rights to address the threat of illegal 3D printing incorporating registered designs.

115 Adding to the confusion, Article 19 of the Commission Proposal confers the exclusive right to "creating, downloading, copying and sharing or distributing to others any medium or software recording the design" but only when these acts are carried out "for the purpose of enabling a product [incorporating the design] to be made", mostly using 3D printing technology. While an in-depth criticism of this provision is beyond the scope of this article, it is apparent how this solution is likely not increasing legal certainty. Especially when considering that the most recent Commission study treated the right to "offer a design" as covering both the "sharing and offering" of a DD file, it is not clear whether the Commission Proposal will reduce rights – by extending protection to sharing only if done with the purpose to print the product – or whether it leaves the previous framework intact. If the latter, then framing Art 19(d) as a purpose-limited right is redundant and likely to increase the already existing doctrinal confusion. Finally, in light of the increasing economic importance of purely Digital Designs, the future Regulation may be outdated soon after its enactment. A more general reconceptualisation and reflection of what the "design approach" means in today's context is required. Unfortunately, the current Commission proposal falls short of offering a "protection system fit for purpose in the digital age"[229] and leaves unaddressed most of the important issues outlined in this article.

G. Conclusion

116 What the above analysis shows is that the extension of design protection to forms of immaterial exploitation of the appearance of a product (e.g., sharing of a DD file) causes several doctrinal problems which should be urgently addressed. Such an extension however should not be considered as a *fait accompli* or inevitable; in other words, it

1(2) of Directive 91/250 [Software Directive]'. C-406/10 *SAS Institute Inc. v World Programming* ECLI:EU:C:2012:259 para 39. Similarly, protection as a computer program of the DD file seems inappropriate and extend beyond the mere protection of the source code.

224 David I Bainbridge, *Intellectual Property* (10th edn Pearson 2018) 497.

225 John Sykes, *Intellectual Property in Designs* (LexisNexis Butterworths 2005) 240.

226 Software Directive (n 98) art 4(1)(c).

227 C-128/11 *UsedSoft GmbH v Oracle International* EU:C:2012:407.

228 Commission Proposal (n 11) art 20(e).

229 ibid 2.

is still possible to recognise that "use of a design" necessarily requires an interaction with a physical product. The extension of design protection to mere reproductions of a design seems to receive support from the jurisprudence and the wording of the Regulation itself; however, a careful analysis of its drafting history suggests that several explanations exist that would prompt us to recognise how the introduction of a right to authorise reproductions of a design may have been in reality an unintended consequence of the drafting process.

117 While it is certain that the Concrete view would avoid much of the conceptual confusion, the broader reappraisal of Design law by the EU Commission offers the opportunity to decide whether design legislation should be applicable to forms of digital value-creation, distribution, and consumption.

118 Several options are available to implement such a policy, and all of them require some forms of amendment of the existing regime. For example, and as recommended by Malaquias, it could be possible to include in the list of exclusive rights conferred by a Design the "making or distributing of a design document"[230], thus ensuring that DD files attract protection without any visibility requirement. Another possibility is offered by the recent Commission Proposal: extend design protections to digital uses of the design (e.g., sharing) but only when it is done for the purpose of "making a product" (e.g., 3D printing)[231].

119 It is nonetheless submitted that without a clear spelling and elucidation of what is the "function of Design law", coupled with a clarification of its broader conceptual architecture, such an amendment would risk raising more questions than it can answer. It is also evident how the newly proposed Article 19 – arguably a legislative-driven foray of Design law into the digital ecosystem – is an ad hoc response to a specific threat: in the words of the Commission, "the challenges brought by the increased deployment of 3D printing technologies"[232]. As a result, the intervention may reveal itself to be short-sighted in so far as it ignores other forms of digital exploitations (e.g., in-game and purely digital consumptions of designs) and does not increase the inherent conceptual flexibility of Design law.

120 In conclusion, it is likely that the broader conceptual uncertainties identified in this article will not be resolved by the introduction of legislative amendments to the Regulation; a broader reconceptualization of EU Design law is called for.

230 Malaquias (n 26).

231 Commission Proposal (n 11) Art 19(2)(d).

232 ibid 8.

Role of White Paper in smart contract formation within ICO (IEO, IDO)

by **Sergey Kasatkin***

Abstract: One of the primary issues for blockchain's widespread adoption in society is the issue of applying traditional contract law to smart contracts. This is because certain elements of contract law are not fully adapted to the formation of agreements with the blockchain. White Paper, which is widely used in other procedures for the placement of digital assets (ICO, IEO, IDO) can serve as an appropriate instrument to explain blockchain code in valid legal manner .

This article investigates the interaction between law and software by means of White Paper. According to the author, approaching a White Paper and program code as a unified concept could solve many practical problems, including the creation of a clear model of ICO smart contract formation, determination of the time of ICO smart contract conclusion and ensuring consistency between the White Paper and contract law requirements for the proper structure of contracts.

Keywords: ICO, white paper, smart contract, blockchain

Recommended citation: Sergey Kasatkin, Role of White Paper in smart contract formation within ICO (IEO, IDO), 14 (2023) JIPITEC 484 para 1.

A. Introduction

1 A White Paper is an informative document that helps professionals figure out a technical issue, solve a difficult problem, or make an important management decision. A White Paper is a widely used tool of content marketing, which is based on the assumption that clients need not just a supplier of products and services, but a trusted advisor.

2 Graham defines this document as a "persuasive essay that uses facts and logic to promote a certain product, service, or viewpoint."[1] There are three main types of White Papers in marketing: 1. A descriptive document that explains the technical features and benefits of a product or service; 2. A numbered list provides a light and lively set of points or concerns about some topic; 3. A problem/solution with recommendations for a new, improved solution for an important business or technical problem.[2]

3 The main goals of a White Paper are to increase sales, enhance trust and loyalty for the company, promote a new product or service or increase brand awareness etc. A White Paper solves these problems through

* Chief Compliance Officer at 'NORCHEM LIMITED' (Malta). Address: Valletta Waterfront No. 6, 45-46 Forni Complex, Pinto Wharf, Floriana, Malta FRN1913. E-mail: s.kasatkin@norchem.mt.

1 Graham Gordon, 'White Paper FAQ (Frequently Asked Questions)' (28 September 2022) <https://thatwhitepaperguy.com/white-paper-faq/#what_is> accessed 10 December 2022.

2 Graham Gordon 'White Papers For Dummies' (April 8, 2013).

the use of expertise, that is, by demonstrating a way to solve a practical problem or with a deep analysis of a particular technical or business issue. Thanks to the expert and professional approach, the reader becomes interested in the proposed product or service.

4 Initially, White Papers were widely used mainly in the field of IT. However, at the moment, this document is read by many B2B executives who are interested in purchasing an expensive product for business purposes. The White Paper is addressed to IT managers and staff, financiers, corporate executives (decision-makers) and other managers. Thus, the format, content and structure of the White Paper is generally aimed at the professional and qualified reader from the field of B2B.

5 Since 2017, the White Paper has become widely used in blockchain projects. Early versions explained the procedures for placing digital assets, called an "Initial Coin Offering" (ICO): "Initial coin offerings are a mechanism to raise external funding through the emission of cryptocurrency tokens, which conceptually are entries on a blockchain."[3] "ICOs take place as a form of investment, or fundraising events, for a project that may involve a product or a service. The entity creating an ICO offers participants unique 'coins' or 'tokens' in exchange for consideration."[4] "ICO can be generally defined as a kickstarter-style crowdfunding campaign that allows the public to participate in an early-stage project and a project team to raise financial capital to support the development of its project across the globe."[5] "Thus, ICOs have become a venture capital-raising tool for start-ups developing projects and applications on the blockchain and trying to escape the constraints of regulation."[6]

6 However, due to the large number of fraudulent ICO projects, new forms of digital assets placement have been developed. They leverage crypto exchanges as intermediaries. Thus, "Initial Exchange Offering" (IEO) and "Initial Dex Offering" (IDO), have appeared and are now widespread.

7 An IEO is a more reliable type of token sale, as each project is verified by a centralized crypto exchange which in turn provides user identification, controls transactions, and ensures security. IDO is very similar to IEO, but instead of a centralized exchange, the organizer of IDO is a DEX platform (i.e. a decentralized exchange). The decentralized exchange does not store or control the digital assets of anonymous users. At the same time, a decentralized exchange serves merely as a trading platform that allows sellers and buyers to find each other in order to exchange digital assets. All operations on DEX-platforms occur through smart contracts, without the participation of a centralized management body.

8 White Papers are widely used to explain procedures for the placement of digital assets (ICO, IEO, IDO)The conclusions of this study could be applied equally to ICO, IEO and IDO. Therefore, for convenience, all mentioned procedures will be referred in this article by one term: "ICO".

9 The ICO White Paper, as well as the White Papers of non-blockchain projects, is in most cases a fairly complex document designed for a professional reader. The ICO White Paper describes the technical features of the project, and contains many special terms and concepts. However, during the ICO process, such a document is also offered to consumers who, in most cases, do not have professional knowledge and skills in the field of blockchain and digital assets. Thus, from the perspective of an ICO, an apparent contradiction arises regarding the initial designation of the White Paper. Therefore, the format, content, and functions of the ICO White Paper should be thoroughly analyzed and aligned through legal regulation..

10 Undoubtedly, the structure and content of an ICO White Paper has many unique features compared to a more traditional marketing document. An ICO White Paper is always a descriptive document that explains the features and benefits of the ICO project. The main purpose of an ICO White Paper is to attract the investors and make them want to invest . For that, it is necessary to present all information about the project in an appealing manner, emphasize its strengths, describe prospects and benefits of investing.

3 Florysiak David, Schandlbauer Alexander, 'The Information Content of ICO White Papers' (December 23, 2019) <https://ssrn.com/abstract=3265007> accessed 10 December 2022.

4 Stylianou Theodoros, 'An Investigation into the Utility and Potential Regulation of Initial Coin Offerings and Smart Contracts in Selected Industries and Jurisdictions' (November 1, 2018) <https://ssrn.com/abstract=3276822> accessed 10 December 2022.

5 Burilov Vladislav, 'Regulation of Crypto Tokens and Initial Coin Offerings in the EU' (May 30, 2019) <https://ssrn.com/abstract=3400705> accessed 10 December 2022.

6 Barsan Iris M., 'Legal Challenges of Initial Coin Offerings (ICO)' (November 2, 2017) <https://ssrn.com/abstract=3064397> accessed 10 December 2022.

B. White Paper as a link between contract law and ICO smart contract

11 When thinking about how to qualify blockchain in legal terms, one of the most important problems is the issue of the applicability of traditional contract law to ICO smart contracts. On the one hand, modern civil law is based on legal concepts that have proven their effectiveness over many centuries. But on the other hand, traditional regulatory approaches can hold back the introduction of innovative blockchain technology into social and legal practice.

12 This problem is clearly reflected in relation to ICO smart contracts. All ICOs are based on smart contracts, defined as a computerized transaction protocol that fulfills the provisions of a contract, or in other words, a program that enforces the contract.[7] Smart contracts guarantee automatic placement of tokens to investors after the transfer of cryptocurrency to the designated wallets of the ICO issuer. Smart contracts thus enable the exchange of cryptocurrencies for the newly issued tokens.[8] The program determines everything: from how parties will transact with the cryptocurrency, how the transactions will be recorded, and how the new coins may be created and released.[9]

13 The use of the blockchain makes it possible to seamlessly structure a multi-level chain of rules, conditions and consequences all implemented in a smart contract. All transactions happening on the blockchain may be programmed in and executed by smart contracts. In this regard it is very important that there is a fundamental similarity between the linguistic structure of code, and content of the contract: computer code is based on statements like "if *x* then *y*". In some sense such an approach correlates with contractual terms and conditions.[10]

14 Meanwhile, it is unclear how contract law would deal with an ICO smart contract. Traditional contract law is not fully adapted to the formation of agreements that exist only in a programming language. In such cases, the White Paper accompanying an ICO has special importance and performs three important functions: 1) providing an appropriate insight of ICO smart contacts in terms of contract law; 2) ensuring the applicability of existing legislative provisions to ICO smart contacts; 3) providing the interpretation of ICO smart contacts.

I. White Paper as a way of providing an appropriate insight of ICO smart contacts in terms of contract law

15 There are several approaches to the legal qualification of smart contracts. Some researchers define a smart contract as a classic legal contract. De Caria states that smart contracts can be considered as actual contracts in their legal meaning, or at the least some form of self-help technology which ensures compliance with contractual obligations.[11] Independently from being digitally expressed, every contract is ruled and guaranteed by the law and the parties have the right to file a court-case for compensation in case of agreement violation.[12]

16 In my opinion, this approach implies the recognition at the legislative level of agreements formed only in the language of the program code. However, at the present time, contract law of most countries does not provide for such recognition. In addition, the place of smart contracts in the system of concepts of traditional contract law are still not defined. Therefore, it could be argued that the recognition of a smart contract as a classic legal contract is premature.

17 Other researchers point out that a smart contract is a programming concept. A smart contract is an executable program, running automatically (i.e. without human intervention) on a blockchain, embodying and enforcing a pre-existing agreement between the contracting parties.[13] Smart contracts

7 Thibault Schrepel, 'Collusion by Blockchain and Smart Contracts' (14 January 2019) <https://ssrn.com/abstract=3315182> accessed 10 January 2023.

8 Philipp Maume, Mathias Fromberger, 'Regulation of Initial Coin Offerings: Reconciling US and EU Securities Laws' (15 June 2018) <https://ssrn.com/abstract=3200037> accessed 10 January 2023.

9 Sam Waxenbaum, 'The SEC and ICOs: Connections Between Digital Assets and Citrus Groves' (8 May 2019) <https://ssrn.com/abstract=3385064> accessed 10 December 2022.

10 Alexander Savelyev, 'Contract Law 2.0: 'Smart' Contracts As the Beginning of the End of Classic Contract Law' (14 December 2016) <https://ssrn.com/abstract=2885241> accessed 10 January 2023.

11 Riccardo De Caria, 'The Legal Meaning of Smart Contracts' (2018) <https://kluwerlawonline.com/journalarticle/European+Review+of+Private+Law/26.6/ERPL2018052> accessed 10 January 2023.

12 Perugini Maria Letizia, Dal Checco Paolo, 'Smart Contracts: A Preliminary Evaluation' (8 December 2015). <https://ssrn.com/abstract=2729548> accessed 10 December 2022.

13 Yongfeng Huang, Yiyang Bian, Renpu Li, J. Leon Zhao and

are "self-executing electronic instructions drafted in computer code."[14]

18 I believe that the main disadvantage of this approach is that a smart contract is considered as a code isolated from human relations and actions. In practice, the conclusion of any ICO smart contract includes the provision by the developers of information about the project on the website or in a special document, user interaction with the software and website (including clicking an OK button), subsequent feedback to the user, etc. The importance of these actions cannot be ignored in the process of smart contract legal conceptualizing. In other words, the context in which an ICO smart contract is used is crucial to its legal qualification. Moreover, it also means that it is not possible to consider the legal qualification of the smart contract in isolation, but only as a part of the "real" contract, which includes smart contract and appropriate context.

19 The most reasonable approach seems to be the point of view of researchers, who argue that smart contract has a dual nature, combining legal and non-legal features. Schuster states that smart contracts fuse contracts and computer programs together by envisioning computer programs written in a way that reflects what two or more parties agree to in a contract.[15] In this case, smart contracts could be defined as "programs that perform part of the contractual obligations, and may contain and execute contractual conditions, as well as invoke physical remedies."[16]

20 Indeed, from a legal point of view, the promises by the developers and the justified expectations of investors are just as important as the actual results of the execution of the smart contract. Expectation and reality are two integral parts of contract realm in social and legal practice, both for traditional and smart contracts. In traditional civil law relations, the actual will of one of the parties may not coincide with the expression of will and with the actual legal results in terms of rights and obligations. Likewise, a smart contract can lead to unexpected results for various reasons (due to the intent of the developers or a technical problem). Thus, in my view, a smart contract should be considered in connection with a set of related documents and actions by the parties that together constitute the agreement between the parties. The code is very important, but it is not the only component of ICO agreement of the parties. Without informational and organizational measures, no one would even know about the existence of the code.

21 Within the ICO, the will, statements and promises of the smart contract developers (token issuers) are expressed in the White Paper. This document allows investors to align their expectations with the goals and objectives of the project. The White Paper defines the relationship and interaction between a regulated social reality and a program code, which is by its nature in itself not susceptible to legal regulation. The White Paper and website information, taken together with the program code, make up the agreement between the parties. Traditional contract law (including special rules for e-commerce transactions) would apply to this agreement (consisting of White Paper and program code). The combination of smart contractcode and context (White Paper) will be referred to hereinafter as the "ICO-contract".

22 This approach reflects the concept that is embodied in contract law called "*consensus ad idem*" which translates as "a meeting of minds". *Consensus ad idem* in contract law means that there is an agreement of all parties involved and everyone has accepted the offered contractual terms and conditions of each party. This is the main principle that underlies enforceable contracts because for contracts to be enforceable, agreement of all parties is necessary.

23 Besides, the concept of *consensus ad idem* states that parties should have an identical or similar mindset regarding the details of the contract they conclude. In other words they could not have entered into a contract where they had no knowledge (nor could they have had knowledge) of its conditions. Nowadays the test for *consensus ad idem* has evolved into an objective standard of a reasonable observer and a requirement of reasonable availability and notice of contractual terms has been formulated in this context.[17]

24 An ICO White Paper provides an appropriate insight of ICO smart contacts in terms of contract law. This document ensures reasonable availability and notice of contractual terms. The consequence of this is a

Peizhong Shi 'Smart Contract Security: A Software Lifecycle Perspective' (October 2019) <https://www.researchgate.net/publication> accessed 10 January 2023.

14 O'Shields Reggie, 'Smart Contracts: Legal Agreements for the Blockchain' (2017) <https://ssrn.com/abstract=2985764> accessed 10 January 2023.

15 Schuster Edmund-Philipp, 'Cloud Crypto Land' (November 21, 2018) <https://ssrn.com/abstract=3476678> accessed 10 December 2022.

16 E. Tjong Tjin Tai, 'Force Majeure and Excuses in Smart Contracts' (4 May 2018) <https://ssrn.com/abstract=3183637> accessed 10 January 2023.

17 Wijayasriwardena Dasuni, 'Consent in Online Contracts - Mindless or Mindful?' (May 24, 2016) <https://ssrn.com/abstract=2783793> accessed 10 December 2022.

consensus ad idem between token issuers and ICO participants. This meeting of minds is determined by both the smart contract program code and ICO White Paper.

II. Applicability of contract law to ICO smart contracts by means of White Paper

25 In most jurisdictions there is no special legal regulation dedicated to ICO smart contracts and White Papers. Therefore the general provisions of contract law apply to smart contracts. The will of the parties to be bound by the terms of the contract is the essential requirement for a valid contract – this does not change simply because the execution of the contract is automated, as is the case in smart contracts.

26 Even if the legislation of any state does not explicitly recognize digital assets as potential objects of civil rights, the law cannot ignore the relationships that are formed on a distributed ledger. The main reason is that digital assets are important from an economic point of view, since they give rise to investment activity in the state and can become an enabler of economic growth. Moreover, in all states, regardless of existing legislation, consumers can purchase digital assets by paying for them with real (fiat) currency. Therefore the legislator is obliged to intervene in such transactions and protect consumers by developing new, special legislation or by applying the provisions of the traditional civil law (in particular consumer protection law) to smart contracts.

27 Meanwhile, by their legal nature, ICO-contracts are quite similar to other agreements concluded by electronic means, and, most importantly, they ensure the achievement of the same goals - the legal formation of agreements between the parties in electronic form (hereafter called e-contracts). E-contracts as well as smart contracts presume that the parties can agree to reach their agreements and to document their transactions only through the exchange of computer-generated messages.[18] E-contracts and smart contracts are agreements formed, specified, fulfilled and deployed by a software system.[19]

28 The legislation of most countries recognizes that contracts can be concluded by electronic means. Moreover, in most cases, the only requirement for a contract in electronic form is the above-mentioned *consensus ad idem*. Smedinghoff reckons that almost all transactions can be done by electronic means. The challenge is to define the electronic-specific requirements that should be met to comply with electronic transaction laws.[20]

29 Meanwhile, the legislation of many countries contains special provisions for contracts formed by electronic means. Traditionally, there are specific kinds of contracts that the law requires to be concluded in writing. The most obvious example is the sale of real estate. In these cases, there are special provisions for how this requirement of a contract being in written form can be met when the contract is concluded electronically.

30 For instance, according to article 1174 of the French Civil Code, "When a writing is required for the validity of a contract, it can be drawn up and kept in electronic form under the conditions provided for in Articles 1366 and 1367..."

31 In accordance with article 1366 of the French Civil Code, "An electronic document has the same probative force as a written document on paper, provided that the person from whom it emanates can be duly identified and that it is drawn up and kept in conditions such as to guarantee its integrity." Article 1367 of the French Civil Code states that "the signature necessary for the perfection of a legal act identifies its author. It expresses the consent on the obligations resulting from this act. When it is electronic, it consists of the use of a reliable identification process guaranteeing its link with the act to which it is attached."

32 According to article 6:227a of Dutch Civil Code, "If a statutory provision implies that an agreement can only be formed validly and inviolably (unchallengeable) in writing, then this formal requirement will be met as well if the agreement is entered into by electronic means and: (a) the agreement is and remains accessible for the parties; (b) the authenticity of the agreement is sufficiently guaranteed; (c) the moment on which the agreement was formed, can be determined with sufficient certainty, and (d) the identity of the parties can be assessed with sufficient certainty."

18 Martin Charles H., 'The Electronic Contracts Convention, the CISG, and New Sources of E-Commerce Law' (2008) <https://ssrn.com/abstract=1120333> accessed 10 January 2023.

19 Jain Sankalp, 'Electronic Contracts: Nature, Types and Legal Challenges' (May 26, 2016) <https://ssrn.com/abstract=2786438> accessed 10 January 2023.

20 Smedinghoff Thomas J., 'The Legal Challenges of Implementing Electronic Transactions' (September 28, 2008) <https://ssrn.com/abstract=1275108> accessed 10 December 2022.

33 Undoubtedly, due to the literal interpretation of the legal provisions, the mentioned additional requirements do not directly apply to ICO-contracts, since no mandatory written form exists. However, I would argue that these requirements should be considered as a standard and necessary for any contracts in electronic form, because they provide the clarity that is required for the meeting of the minds of the parties.

34 Meanwhile, in the absence of a White Paper, an ICO smart contract really does not comply with the above-mentioned legislative requirements for specific kinds of E-contracts: it is not accessible for the parties in clear form; the authenticity of the agreement is not guaranteed; the moment of ICO smart contract completion is unclear; the identity of ICO participants can also be hidden.

35 However, the use of ICO smart contracts is quite similar to the application of the so-called "automated message systems for contract formation". According to the article 12 of the United Nations Convention on the Use of Electronic Communications in International Contracts "a contract formed by the interaction of an automated message system and a natural person, or by the interaction of automated message systems, shall not be denied validity or enforceability on the sole ground that no natural person reviewed or intervened in each of the individual actions carried out by the automated message systems or the resulting contract."[21] In fact, transactions made on the blockchain (such as when accepting the offer by sending an amount of money to an ICO smart contract) are also the result of the exchange of automated messages, and such operations may be performed without the intervention of natural persons.

36 Moreover, the principle of technological neutrality mentioned in the preamble to the United Nations Convention on the Use of Electronic Communications in International Contracts cannot be ignored, namely "uniform rules should respect the freedom of parties to choose appropriate media and technologies, taking account of the principles of technological neutrality."[22] This principle means that "legislation should define the objectives to be achieved and should neither impose, nor discriminate in favor of, the use of a particular type of technology to achieve those objectives."[23] "The same regulatory principles should apply regardless of the technology used. Regulations should not be drafted in technological silos."[24]

37 In accordance with the principle of technology neutrality, maximum efforts should be made to attain full legal recognition of agreements concluded on the blockchain. At the same time, one should agree with the scholars, who note that smart contracts do not need any special new laws or regulations. Instead, existing legal principles of contract law should be adapted or modified, either statutorily or judicially, to deal with smart contracts.[25] Indeed, to include ICO smart contracts in the scope of E-contracting, it is only necessary to bring them in line with the concepts and principles of existing legislation, the most important of which is the concept of consensus ad idem.

38 In particular, the content of the ICO-contract must be clear and accessible to both parties who know each other's identity. In addition, a clear procedures for concluding an ICO-contract should be established, in accordance with modern E-contracting rules.

39 I believe that the White Paper is able to subordinate ICO smart contracts to the norms of national legislation (including E-commerce rules) by complementing the code with what is necessary to truly reach consensus ad idem. The next sections will show how the White Paper helps to eliminate possible inconsistencies and contradictions between the existing legislation and ICO smart contracts. In particular, the White Paper allows us to build a clear model for concluding an ICO smart contract from the point of view of modern civil law; determine the moment of conclusion of an ICO contract and establish essential and other terms of an ICO contract. Clarity and harmonization on these issues will enable ICO contracts to be included in the scope of E-contracting in accordance with the current legislation.

21 United Nations Commission On International Trade Law 'United Nations Convention on the Use of Electronic Communications in International Contracts' (November 23, 2005) <https://uncitral.un.org/en/texts/ecommerce/conventions/electronic_communications> accessed 10 January 2023.

22 Ibid

23 Commission of the European communities 'Towards a new Framework for Electronic Communications Infrastructure and Associated Services: the 1999 Communications Review COM' (November 10, 1999) <https://aei.pitt.edu/5978> accessed 10 December 2022.

24 Maxwell Winston, Bourreau Marc, 'Technology Neutrality in Internet, Telecoms and Data Protection Regulation' (November 23, 2014) <https://ssrn.com/abstract=2529680> accessed 10 December 2022.

25 O'Shields Reggie (n 14).

III. Providing interpretation of ICO smart contacts through White Paper

40 Since smart contracts embody and enforce the nature of a parties' agreements in the language of the program code, contradictions between the legal obligations of the parties and transactions that occur according to the rules of a computer program may arise. Schuster describes this situation quite accurately as a "synchronisation conflict": when compliance with the law would yield a state of affairs that is different from the state of affairs that is created by and reflected in the distributed ledger.[26]

41 The reason for such contradictions lies not only in the distinction between the rules of legal and technical (software) regulation, but also in the obvious differences between natural language and software code. "Ambiguity is celebrated in human language. It is a central feature of literature, poetry, and humor. However, ambiguity is anathema to computer language. An ambiguous computer language is a nonsensical concept because the predictability of computers is what gives part of their value."[27] In other words, smart contracts replace vague natural language with precise computer code.[28]

42 In the process of an ICO, a White Paper allows the expression of the will of the developers, which is decisive from a contract law perspective. Thus, the White Paper legally obliges the developers to follow the announced plan of project implementation.

43 Undoubtedly, the contents of a White Paper may not correspond to the program code due to a mistake or the intention of the developers or discrepancy between meanings of words and program code lines. In this regard, it is necessary to determine the priority between the natural language contained in the White Paper and the program code of the smart contract. Obviously, the decision on the priority of the code over the language is unacceptable, since this will make the content of the agreements of the parties inaccessible to the absolute majority of persons who do not have special skills and special education in the field of programming.

44 Thus, White Paper, rather than the program code, should be decisive in determining the will of the parties. In other words the program code should be interpreted in the light of the White Paper, where, in case of any differences, the White Paper is prioritized. White Paper creates the certainty and clarity which are necessary for interpretation and regulation of ICO-contracts.

C. ICO White Paper as an invitation to conclude an ICO-contract

45 Creation of a clear model for concluding an ICO-contract is the most important prerequisite for including such contracts in the scope of civil law regulation. The parties must accurately understand the legal meaning and consequences of all actions that they take before the ICO-contract comes into force. This is very important in terms of protecting the rights of investors, as well as to ensure the stability of the implementation of ICO procedures. At the same time, it seems necessary to correctly understand the meaning of White Paper in the ICO-contract concluding process. There are many practical problems associated with this issue, including the possibility to change or revoke a White Paper after it has been published.

46 In most cases, ICO-contract formation includes the following actions: After the publication of information about the project in a White Paper and on the project website, all interested investors get the opportunity to participate in the ICO. To do this, investors must register on the ICO website and identify themselves. This requirement is primarily related to compliance with KYC (Know Your Customer) rules. In most cases, in order to fulfill the registration, the investor needs to indicate name, country of residence and the planned amount of investment. After that, the investor will be included in the so-called "White list" - a list of approved participants of ICO procedure. To complete the purchase of digital assets, the investor sends the appropriate amount of cryptocurrency to the address of the ICO smart contract, which is listed on the official website of the project. In exchange for the received cryptocurrency, the smart contract automatically sends the appropriate amount of digital assets to the investor in accordance with the program code.

47 The correct legal qualification of these actions in terms of civil law theory is the first step towards the elimination of contradictions between the existing legal systems and ICO smart contracts. In addition, such a qualification will make it possible to determine the role of White Paper in ICO smart contract formation.

48 Traditionally, in civil law systems there are two stages in a contract formation: first an offer, that is, a proposal to enter into a contractual relationship,

26 Schuster Edmund-Philipp (n 15).

27 Raskin Max, 'The Law and Legality of Smart Contracts' (September 22, 2016) <https://ssrn.com/abstract=2959166> accessed 10 December 2022.

28 Schuster Edmund-Philipp (n 15).

and second, an acceptance of this offer. The contract comes to existence by the offer and its acceptance, which together constitute the *consensus ad idem*.

49 In accordance with the requirements of common law systems, in addition to the offer and acceptance for contract formation, consideration is required. According to the classical doctrine of consideration it may consist either in some benefit to the promisor or in some detriment to the promisee.[29] Usually, the full or partial payment under the contract is considered as consideration. Offer, acceptance and consideration form *consensus ad idem* which is supposed to demonstrate the fact that the parties have a similar mindset and corresponding intentions regarding the details of the contract.[30]

50 Meanwhile, the application of these classical models of contract formation to agreements in electronic form (including ICO-contracts) is a rather difficult task. There is even a point of view that the orthodox rule, which demands concurrence of offer and acceptance for contract formation becomes inapplicable in all cases of e-contracting.[31]

51 Indeed, an important peculiarity of the ICO-contracts and many other E-contracts , is that the meeting of the parties wills is ensured by the interaction of the user with the website, and not by the exchange of separate messages. At the same time, the difference between these methods of contract formation is quite significant. In case of contract conclusion through a website a variety of complex functions are fulfilled on the computer or on the site's server. The website can send and receive messages, display media, alter and rearrange information, and communicate with other sites and devices.[32] Thus, in the case of E-contracting by means of website, it is important to determine exactly which actions should be considered as an offer and an acceptance.

52 For the above reasons, legal science has created new approaches to describing the procedure for concluding contracts in electronic form. In particular, the concept of click-wrap agreements have become widespread. These agreements are concluded electronically on the website by clicking on the "I agree" button that accompanies the text of the agreement. Currently, click-wrap agreements are widely used in many areas of E-contracting, including the purchase of software and access to services on the Internet.

53 At the level of the European Union, click-wrap agreements were recognized primarily due to the Directive of the European Parliament and of the Council (EC) 2000/31/EC on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market. Following the provisions of article 9(1) of "Directive on electronic commerce", "Member States shall ensure that their legal system allows contracts to be concluded by electronic means. Member States shall in particular ensure that the legal requirements applicable to the contractual process neither create obstacles for the use of electronic contracts nor result in such contracts being deprived of legal effectiveness and validity on account of their having been made by electronic means."[33]

54 Depending on the peculiarities of the legal system, the display by websites of their goods (or services) qualifies as an invitation to offer (invitation to treat). In some countries, so as to ensure consumer interests, the offer of goods on a website is considered as a public offer. The seller cannot refuse to sell the goods if the buyer accepts an offer containing all the essential terms of the contract. For example, in accordance with article 1114 of the French Civil Code, the offer, made to a determined or indeterminate person, includes the essential elements of the envisaged contract and expresses the will of its author to be bound in the event of acceptance. Otherwise, there is only an invitation to enter into negotiations. According to article 1127-1 of the French Civil Code, anyone who professionally offers, by electronic means, the supply of goods or the provision of services, makes available the applicable contractual provisions in a manner that allows their conservation and reproduction. The author of an offer remains committed by it until it is accessible electronically by him.

55 In other countries, on the contrary, the website reflects only an invitation to offer to conclude a contract. For example, under English law, the offer of a product or service on a website is an invitation

29 Lorenzen Ernest, 'Causa and Consideration in the Law of Contracts' (May 1919) <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/11443/55_28YaleLJ621_1918_1919_.pdf?sequence=2> accessed 10 January 2023.

30 Wijayasriwardena Dasuni (n 17).

31 Gebrehiwot Entehawu Desta, 'Enforceability of electronic contracts in light of the Ethiopian General Contract Law: appraising the issues' (December, 2018) <https://www.researchgate.net/publication> accessed 10 January 2021.

32 Sklaroff Jeremy, 'Smart Contracts and the Cost of Inflexibility' (September 18, 2017) <https://ssrn.com/abstract=3008899> accessed 10 January 2023.

33 Directive of the European Parliament and of the Council (EC) 2000/31/EC on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce') [2000] OJ L178.

to treat, while the offer is made by the buyer at the time the product is placed in the shopping cart and the payment begins.[34] The offer is not made by website showing the goods for sale at a certain price. This represents an invitation to offer and could be revoked at any time before the acceptance. The offer is made by the customer who places the items in the virtual basket for payment.[35]

56 However, the offer of goods using the website must be distinguished from the proposal expressed in ICO White Paper. Regardless of the legal system and the features of the ICO procedure, a White Paper is always an invitation to offer, since it does not contain the main essential condition of the ICO-contract - the current price of the digital asset, which is reflected only on the site and can be changed during the ICO. In addition, the user is not able to interact directly with the White Paper, and in any case is forced to use the developers' site to conclude an ICO-contract. At the same time, the final stage of ICO-contract formation (acceptance) is in any case performed automatically by a smart contract.

57 Recognition of a White Paper as an invitation to offer to conclude an ICO-contract is fully compliant with Article 11 of United Nations Convention on the Use of Electronic Communications in International Contracts. According to this article, "a proposal to conclude a contract made through one or more electronic communications which is not addressed to one or more specific parties, but is generally accessible to parties making use of information systems, including proposals that make use of interactive applications for the placement of orders through such information systems, is to be considered as an invitation to make offers, unless it clearly indicates the intention of the party making the proposal to be bound in case of acceptance."[36]

58 An analysis of Article 10 of the Directive on E-Commerce also leads to the conclusion that the White Paper cannot be considered as a public offer. According to this article, some important information (including different technical steps to follow to conclude the contract) must be provided "prior to the order being placed by the recipient of the service."[37]

34 Reed Chris, Angel John 'Computer Law: The Law and Regulation of Information Technology' (May 17, 2007) <https://archive.org/details/computerlawlawre0000unse> accessed 10 December 2022.

35 Jain Sankalp (n 19).

36 United Nations Commission On International Trade Law (n 21).

37 Directive on electronic commerce (n 33).

59 Thus, it is assumed that within the process of E-contract concluding, the next step after the disclosure of information about a product or service is the placement of an order by the buyer on the website, but not the final acceptance of the offer. In respect to the procedure for concluding an ICO-contract, this means that an offer to acquire a digital asset is not the placement of information about the project in the White Paper, but the investor's actions necessary to acquire tokens (including registration on the project's ICO website, cryptocurrency payment).

60 The recognition of a White Paper as an invitation to offer to conclude an ICO-contract has some undeniable advantages: It ensures that both parties of the ICO-contract are protected from errors related to the use of the Internet and the blockchain in the contract formation process. In particular, investors have the opportunity to withdraw their offer, before an acceptance. In addition, it should be taken into account that during the ICO process, technical problems arise, which should be eliminated by developers before the ICO-contract comes into force. Since the White Paper is not a public offer, developers have the opportunity to change the White Paper and bring it into line with the changed conditions of the ICO procedure.

61 Thus, modern rules of contract law are suitable for describing the procedure of ICO-contract formation if it is interpreted in this way. The White Paper and information on the project website should be considered as an invitation to conclude an ICO-contract. The offer is the registration of a participant on the ICO project website, as well as sending cryptocurrency to the appropriate ICO electronic wallet. At the same time, from the point of view of Common law systems, such a cryptocurrency payment may be considered as a legal consideration. The acceptance is performed through a smart contract automatically, by exchanging the cryptocurrency for the required digital asset.

D. Determination of a moment of ICO smart contract conclusion through a White Paper

62 Determination of the precise moment of civil law agreement conclusion is one of the key issues in contract law. This legal relationship is expressed in terms of mutual obligations. Therefore, the establishment of clear rules for determining moment of the contract formation contributes to the clarity in legal relations, and also ensures the necessary stability and sustainability of civil transfers. In the literature it is rightly pointed out that the contract

should clearly determine the exact time and way of acceptance of the agreement.[38]

63 In traditional contract law, the moment of contract formation is determined according to long-established and proven rules. For example, in accordance with article 1121 of the French Civil Code, the contract is concluded as soon as the acceptance reaches the offeror.

64 Meanwhile, in the field of E-commerce, the traditional rules of contract conclusion sometimes turn out to be inapplicable. Therefore, there exists a point of view that time of contract formation is another area where technological developments have had an impact on the law due to changes resulted from electronic communication technologies.[39]

65 Indeed, due to the importance of precise determination of the E-contract formation, special rules and requirements have appeared in civil law. In particular, according to article 6:227c of the Dutch Civil Code, if the party has made an announcement by electronic means that may be interpreted by the service provider either as the acceptance of an offer which the service provider has made by electronic means or as an offer in response of an invitation to start negotiations made by the service provider by electronic means, then the service provider will confirm that he has received this announcement as soon as possible by electronic means. The opposite party may rescind the agreement as long as the service provider has not confirmed that he has received an acceptance.

66 Another special e-commerce rule concerning the moment of contract conclusion is set out in Article 10(1)(a) of the Directive on electronic commerce. According to this article, Member States shall ensure that information about different technical steps follow to conclude the contract will be given by the service provider clearly, comprehensibly and unambiguously and prior to the order being placed by the recipient of the service.[40] This information requirement needs to avoid that people are contractually bound without knowing it. A simple click on an OK button may be enough to conclude the contract, but only if the recipient has been given information about this "technical step" before he clicked. Otherwise, the contract is not binding.[41]

38 Jain Sankalp (n 19).

39 Gebrehiwot Entehawu Desta (n 31).

40 Directive on electronic commerce (n 33).

41 Lodder Arno R., 'European Union E-Commerce Directive - Article by Article Comments' (2017) <https://ssrn.com/abstract=1009945> accessed 10 January 2023.

67 Meanwhile, during the ICO, above mentioned E-commerce rules are often not being implemented due to peculiarities of the blockchain functioning. Often, investors are not aware of the meaning of their actions in the ICO-contract formation process. The moment when the ICO-contract comes into force is also sometimes vague.

68 The fact is that the legal provisions that determine the moment of ICO-contract formation may differ significantly from the rules encoded in a smart contract. According to the requirements of the legislation, the receipt by the offeror of an electronic notification from the smart contract or website about the acceptance of an offer may be sufficient for an ICO-contract formation. However, from the point of view of programming, an ICO contract cannot be considered as concluded until the execution of the operation of exchanging cryptocurrency for a digital asset by a smart contract becomes irreversible. Up to this point, investors do not have reliable guarantees for the execution of the ICO-contract by means of program code, since the execution of a smart contract can be stopped or changed by developers.

69 Quite often, investors receive a digital asset in their wallet immediately after cryptocurrency payment. In such cases, there are no problems or difficulties, since the moments of ICO-contract formation from a legal and technical point of view coincide. Moreover, in this case, the instance of concluding an ICO-contract coincides with its full or partial execution (depending on the type and functionality of the digital asset).

70 However, according to the rules of many ICO procedures, a significant period of time passes from the moment an investor's cryptocurrency payment to when he receives a digital asset. In this case, determination of ICO-contract formation moment becomes crucial, since it is important for an investor to understand when he receives reliable guarantees of obtaining the necessary digital asset, in terms of law and program code. In particular, an investor may receive a digital asset later than the cryptocurrency payment occurred under the following models of ICO fundraising: Dutch auction model and Soft cap model.

71 Dutch auction model means an auction in which the auctioneer begins with a high asking price in the case of selling, and lowers it until some participant accepts the price, or it reaches a predetermined reserve price. In the case of such an ICO model, the investor receives a digital asset at the end of the auction results.

72 In addition, the terms of ICO quite often provide for the so-called "soft cap", which means the minimum funding aim of the ICO. In this case, the investor receives a digital asset no earlier than when the

project reaches its soft cap, the specific value of which is programmed in the smart contract and is announced in advance by the developers.

73 Thus, it becomes obvious that in order to determine the moment of ICO-contract formation in the mentioned ICO fundraising models, it is necessary to understand the technical features of blockchain transactions. Researchers reckon that it is important for any electronic transaction to begin with a clear and comprehensive understanding of the process involved and how it will actually work from a technical perspective. In fact, understanding "how it works" from a technical perspective is critical to "making it work" from a legal perspective.[42]

74 I believe that the process of concluding an ICO-contract, as well as a classic civil law agreement, can be represented in the form of message exchange. The offer is accepted in an electronic form, by sending an electronic message signed with a private digital key to the smart contract that accepts the offer. At the same time, it should be noted that the traditional E-commerce rules for receiving and sending messages do not fully correspond to the features of the blockchain. According to the article 15(1) of UNCITRAL Model Law on Electronic Commerce, unless otherwise agreed between the originator and the addressee, the dispatch of a data message occurs when it enters an information system outside the control of the originator or of the person who sent the data message on behalf of the originator.[43]

75 In the scope of blockchain, the interaction between the originator and the addressee is always carried out within the framework of one information system - a distributed ledger. This means that the relations of the parties at the stage of ICO-contract formation are regulated by a programthat has unconditional priority over other sources of regulation.

76 For the above reasons, it could be concluded that regardless of the legislative regulation or the content of the White Paper, the-ICO contract will be formed at the moment when, according to the code of the smart contract, the operation of exchanging cryptocurrency for a digital asset becomes irreversible. At the same time, from a technical point of view, it can be considered that the transaction became irreversible at one of the following moments:

- When an entry is made in the distributed registry about the transfer of the token to the ICO participant;

- When a valid transaction amounting to an acceptance of the offer by the smart contract is sent to it;

- When the transaction is included in a valid block that is added to the blockchain, thereby transferring the token to the ICO participant.

77 Thus, from the point of view of programming, the moment of transaction irreversibility within a smart contract is ambiguous and requires clarification for each distributed ledger and each certain ICO project. In this regard, the investor should be provided with additional information in advance about the moment of ICO smart contract conclusion. This is critical point in terms of compliance with civil law and E-commerce rules. At the same time, the best way to disclose such information is in the White Paper, which allows the investor to comprehensively and systematically evaluate the risks associated with the moment of ICO-contract formation, taking into account all significant factors and circumstances.

78 It should also be taken into account that due to the intent or mistake of the developers, false information about encoded moment of ICO smart contract formation may be included in the White Paper. In this case, the developers should be held fully liable for any losses caused to the investor in connection with such unfair reporting.

79 Thus, in the purpose of determining the moment of ICO-contract formation, the White Paper should contain the following information:

-Technical steps required to conclude an ICO-contract;

- Precise moment when the transaction of exchanging cryptocurrency for a digital asset through smart contract becomes irreversible (and thus the contract is considered to have come into being);

- Provisions on the responsibility of developers for providing false information about the moment of ICO contract formation.

E. Determination of contractual terms through White Paper

80 Accessibility and clarity of contractual terms are the most important prerequisites for legal protection of

42 Smedinghoff Thomas J. (n 20).

43 United Nations Commission On International Trade Law 'Model Law on Electronic Commerce' (June 12, 1996) <https://uncitral.un.org/en/texts/ecommerce/modellaw/electronic_commerce> accessed 10 December 2022.

participants in civil legal relations. The opportunity to find out the terms of the contract in advance and have constant access to them during the contract's execution is an obvious and natural need of each party of the contract.

81 However, in the field of e-commerce (including ICO contracting), the issue of access to the provisions of the contract is extremely relevant. The fact is that when the order of goods or services is carried out using the interface of the website, the buyer (customer) does not always have real technical access to the terms of the formed agreement. In the literature it is rightly pointed out that if the e-contracts are concluded on web-click-on agreements, only the site owner has an access to the conditions of the contract.[44] Nowadays, a lot of websites do not propose the possibility of contract filing, a stored agreement can be used as an evidence. A filed and accessible contract may give the recipient more confidence in the provider and influence his purchase.[45]

82 For the above reasons, article 10(1)(b) of Directive on electronic commerce provides that Member States shall ensure that information about availability of the concluded contract provisions will be given by the service provider clearly, comprehensibly and unambiguously and prior to the order being placed by the recipient of the service.[46]

83 In the field of ICO, the problem of contractual terms availability is especially relevant. Many terms of the agreement are set out in the language of the program code, so it is rather difficult for the parties to gain access to the programmed terms of the contract.

84 Since the White Paper is the main way to provide information about the ICO project, this document should be used to reflect the contractual terms of the ICO-contract. In other words, the White Paper translates the agreements of the parties from the programming language into one of the human languages, and therefore this document can be considered as an ICO-contract in writing.

85 Undoubtedly, the recognition of the White Paper as a constitutive element of contractual relations is fraught with certain risks. Participants in the digital asset market are accustomed to considering White Paper as a purely informational document. The legal force of the White Paper can confuse even professional crypto investors. Indeed, the ICO White Paper is sometimes called an "anarchist" document, which usually "expressly states that transactions on the system are not intended to create legal relations".[47]

86 However, in accordance with one of the principles developed by the European Law Institute, "smart contracts used for consumers always have to be made available as a translation (and explanation) into natural language so the consumer can read and understand what their rights and duties are."[48] Since consumers often participate in ICOs, the mentioned rule must certainly be followed in every case of placing digital assets.

87 Consumers should be informed about the terms and conditions of their contracts in the most generally common, accessible and convenient way. Obviously, in the case of an ICO, these requirements are fulfilled thanks to the White Paper. This document has been accompanying ICO procedures for many years and is usually considered as the main source of information about the project. In addition, the White Paper is proof of the conclusion of the ICO-contract under certain conditions, since this electronic document can be downloaded and saved by the ICO participant. For the reasons stated, I believe that the terms of the ICO-contract should be determined through the White Paper, despite some of the risks.

88 As a form of expression of an ICO-contract, the White Paper must comply with several contract law requirements for the proper content of contracts. These legal requirements include, firstly, the rules of traditional contract law on the essential terms of contracts, and secondly, the Rules on contractual standard terms and conditions.

I. Contract law requirements on the essential terms of contracts

89 The civil legislation of many countries links the creation of the legal consequences of the contract with the parties agreement on some of the most important terms of the contract. Such a minimum required set of contractual terms consists of so-called essential terms of the contract.

90 Parties will have stipulated the performances to which they have committed themselves. This can be

44 Gebrehiwot Entehawu Desta (n 31).

45 Lodder Arno R. (n 41).

46 Directive on electronic commerce (n 33).

47 UK Law commission 'Smart legal contracts. Advice to Government' (November 2021) <www.gov.uk/official-documents> accessed 15 April 2023.

48 European Law Institute 'ELI Principles on Blockchain Technology, Smart Contracts and Consumer Protection' (September 8, 2022) <https://www.europeanlawinstitute.eu> accessed 15 April 2023.

done extensively or briefly, but at least the essential points of the agreement must be set with mutual consent. These essential points indicate what kind of agreement has been concluded, for instance a sale contract or an employment agreement.[49]

91 According to article 966 of Malta Civil Code, the following are the conditions essential to the validity of a contract: (a) capacity of the parties to contract; (b) the consent of the party who binds himself; (c) a certain thing which constitutes the subject-matter of the contract; (d) a lawful consideration. In accordance with the article 982 of Malta Civil Code, every contract has for its subject-matter a thing which one of the contracting parties binds himself to give, or to do, or not to do.

92 According to Article 6:227 of Dutch Civil Code, the obligations to which parties subject themselves under the agreement, must be determinable.

93 In accordance with the article 1163 of France Civil Code, the object of the obligation is a present or future performance. This must be possible and determined or determinable.

94 Thus, according to the legislation of the most countries, the essential condition of any contract is the subject-matter (object). If the parties do not agree on the subject-matter of the contract, then legal consequences will not arise or the agreement will be considered invalid.

95 The above fully applies to ICO-contracts. Researchers believe that the initial stage of a contractual agreement is not markedly different between smart and traditional contracts. This is due to the fact that before any contract software can operate, two parties must agree to some set of terms that initiates the program.[50]

96 It is obvious that the subject-matter, and, therefore, the essential condition of any ICO-contract is a digital asset, which is fundamentally different from traditional intangible objects of civil rights. Typical intangible assets include, for example, patents, trade secrets, copyrights and trademarks. Obviously, despite their immaterial nature, these assets are able to independently satisfy certain needs of individuals and companies. This feature provides their value. For example, an exclusive right given by a patent entails the capacity to exclude others from commercial exploitation of the object of the exclusive right.

97 Meanwhile, a crypto token per se is merely an entry in a blockchain (transaction ledger).[51] Therefore, the value of crypto tokens is determined by the features of the blockchain itself. Firstly, the value of digital assets depends on what the asset stands for within blockchain. It is about benefits and opportunities that are technically provided by a digital asset to their owner. Secondly, in the system of blockchain, the problem of "double spending" information units has been solved. Unlike crypto tokens any other information objects that exist outside the blockchain technology can be copied and transferred from one owner to another for an unlimited number of times, while remaining both with the transmitting and the receiving party. In other words, the value of digital assets is determined by the consensus reached within the blockchain. "Consensus in a blockchain network refers to the process where the distributed nodes agree on the history and the final state of the data on the ledger, usually referred to as distributed consensus. Since all participants in the network hold the data, they can also be a part of the decision-making".[52] Blockchain essentially consists only of transaction history which includes not only the signature and the amount transferred, but also links to all previous transactions in which the payer received the assets concerned.

98 Thus, since the usefulness and value of a digital asset is not obvious and is determined partly by the rules underlying the blockchain system in question, the White Paper should contain not only the designation and definition of the acquired digital asset, but also detailed description of the rights provided by such asset. These rights are essentially ways of program influencing or controlling the information system by the owner of a digital asset. However, regardless of their content, the rights must be presented in the form of legal claims vis-a-vis the token issuer or project developers. Only such a description of the subject-matter of the ICO-contract should be considered appropriate and satisfying the requirements of civil law on the essential terms of contracts.

II. Applying the rules on contractual standard (general) terms and conditions to ICO-contracts

99 The legislation of most countries contains special rules applicable in cases where one of the parties to the agreement (typically a consumer) can only

49 Dutch Civil Law <http://www.dutchcivillaw.com/civilcodegeneral.html> accessed 10 January 2023.

50 Raskin Max (n 27).

51 Vladislav Burilov (n 5).

52 Lawrence J. Trautman, Mason Molesky, 'A Primer for Blockchain' (28 January 2019) <https://ssrn.com/abstract=3324660> accessed 10 December 2022.

accede to the entire agreement without having the ability to negotiate its terms. For example, according to the article 1119 of France Civil Code, the general conditions invoked by one party only have effect with regard to the other if they have been brought to the attention of the latter and if she has accepted them. In case of discrepancy between general conditions and special conditions, the latter take precedence over the former.

100 According to Article 6: 231 of the Dutch Civil Code, "standard terms and conditions" mean one or more contractual provisions or stipulations, drafted to be included in a number of contracts, with the exception of provisions and stipulations that indicate the essence of the performance under the obligation, as far as these last meant provisions and stipulations have been formulated clearly and unambiguously.

101 Following to the article 6:233 of the Dutch Civil Code, a stipulation from the applicable standard terms and conditions is voidable: a. if it is unreasonably burdensome for the counterparty, having regard to the nature and content of the contract, the way in which these standard terms and conditions have been formed, the interests of each party, as evident to the other, and the other circumstances of the case; b. if the user has not given his counterparty a reasonable opportunity to take knowledge of the content of the applicable standard terms and conditions.

102 Besides, the Dutch Civil Code includes so called "Black list" of stipulations which are always unreasonably burdensome for consumers (Article 6:236) and "Grey list" of stipulations which are presumed to be unreasonably burdensome for consumers (Article 6:237).

103 To sum up, regardless of the type of legal system, a White Paper should contain an essential condition of any civil law contract: a subject-matter expressed in a detailed description of a digital asset and appropriate rights. Moreover, the White Paper should indicate that it contains standard terms and conditions, and that appropriate rules of the particular jurisdiction are complied with. These conditions will bring the White Paper into line with the requirements of civil law.

F. Summary

104 The challenge for modern legal systems is to draft new legislation that is tailored to the specific properties of ICO smart contracts that require regulation. Another way to deal with ICO is to create appropriate mechanisms to ensure that existing legislation can be applied to ICO smart contracts.

105 Indeed, even the existing legal provisions and structures created within contract law make it possible to solve many of the challenges of ICO legal regulation. Therefore, what is needed is a reliable connection and interaction between law and software code. I believe that such a connection can be provided through the White Paper.

106 In general, a White Paper allows the application of classical contract law to smart contracts by converting computer code into understandable legal terms that someone can agree to (or not). In other words, from a legal point of view, a smart contract cannot be considered in isolation from the White Paper. In this regard, it is correct to apply the special term "ICO-contract", denoting the unity of the White Paper and the smart contract.

107 Considering a written document and program code as a unified concept could solve many practical problems, including creation of a clear model of ICO-contract formation, determination of precise moment of ICO-contract conclusion and ensuring consistency between the White Paper and contract law requirements for the proper content of contracts.

jipitec

Journal of
Intellectual Property,
Information Technology,
and Electronic Commerce
Law

www.jipitec.eu