AI4media
ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

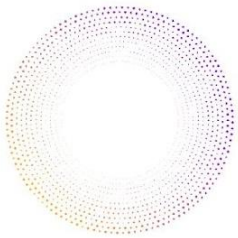
D7.3

From platform liability to platform responsibility - Analysis of the shifting policy approach, guidelines for the AI-on-Demand-Platform and policy recommendations

Project Title	AI4Media - A European Excellence Centre for Media, Society and Democracy
Contract No.	951911
Instrument	Research and Innovation Action
Thematic Priority	H2020-EU.2.1.1. - INDUSTRIAL LEADERSHIP - Leadership in enabling and industrial technologies - Information and Communication Technologies (ICT) / ICT-48-2020 - Towards a vibrant European network of AI excellence centres
Start of Project	1 September 2020
Duration	48 months

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu



Deliverable title	From platform liability to platform responsibility - analysis of the shifting policy approach, guidelines for the AI-on-Demand-Platform and policy recommendations
Deliverable number	D7.3
Deliverable version	1.0 – public version
Previous version(s)	-
Contractual date of delivery	31 August 2023
Actual date of delivery	20 October 2023
Deliverable filename	D7.3_AI4MEDIA_Final
Nature of deliverable	Report
Dissemination level	Confidential (a public version will also be made available in the project website, focusing on Section 3 of this deliverable)
Number of pages	74
Work Package	WP7
Task(s)	T7.5
Partner responsible	KUL
Author(s)	Noémie Krack (KUL), Lidia Dutkiewicz (KUL), Sven Becker (FhG-IAIS)[1]
Editor	Noémie Krack (KUL)*, Lidia Dutkiewicz (KUL)*,
EC Project Officer	Evangelia Markidou
Abstract	The deliverable presents an analysis of the shifting approach from platform liability to platform responsibility. It provides an overview of the new Digital Services Act (DSA) due diligence and accountability obligations, that the providers of intermediary services, including online platforms, need to comply with. It also provides guidelines on platform liability and accountability for the AI-on-Demand platform.
Keywords	Platform, Liability, Accountability, Digital Services Act, Content, AI-on-Demand, Artificial Intelligence, Media

Copyright

[1] The work and research of Noémie Krack and Lidia Dutkiewicz in AI4Media is conducted under the supervision of Prof. Peggy Valcke and Dr Aleksandra Kuczerawy in KU Leuven.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

Contributors

NAME	ORGANISATION
Noémie Krack	KUL
Lidia Dutkiewicz	KUL
Sven Becker	FhG-IAIS

Peer Reviews

NAME	ORGANISATION
Atia Cortés	BSC
Samuel Almeida	F6S
Filareti Tsalakanidou	CERTH

Revision History

VERSION	DATE	REVIEWER	MODIFICATIONS
0.1	14/03/2023	Noémie Krack & Lidia Dutkiewicz	Initial table of content
0.2	27/03/2023	Noémie Krack, Lidia Dutkiewicz & Sven Becker	New table of content
0.3	28/04/2023	Sven Becker	Contribution of FhG-IAS
0.4	12/05/2023	Noémie Krack, Lidia Dutkiewicz & Sven Becker	Edit of the content and revision of the deliverable's table of content
0.5	20/09/2023	Lidia Dutkiewicz, Noémie Krack, Sven Becker	Contribution and version ready for review
0.6	28/09/203	Atia Cortés, Samuel Almeida, Filareti Tsalakanidou	Internal review
1.0	16/10/2023	Noémie Krack, Lidia Dutkiewicz, Sven Becker	Version ready for submission to EC (confidential)
2.0	9/11/2023	Noémie Krack, Lidia Dutkiewicz	Public version for project website

The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf.

Table of Abbreviations and Acronyms

Abbreviation	Meaning
AI	Artificial Intelligence
AIoD	AI-on-Demand platform
API	Application Programming Interface
Art.	Article
AVMSD	Audiovisual Media Service Directive
B2C	Business-to-Consumer
CDSM	Copyright in the Digital Single Market Directive
CJEU	Court of Justice of the European Union
CMS	Content management System
Council	Council of the European Union
CSAM	Child Sexual Abuse Material
D.	Deliverable
Dir.	Directive
DMA	Digital Markets Act
DSA	Digital Services Act
DSC	Digital Service Coordinators
EC	European Commission
ECD	e-commerce Directive
EECC	European Electronic Communication Code
EP	European Parliament
EU	European Union
GDPR	General Data Protection Regulation
HSP	Hosting Service Provider
IP	Intellectual Property
IP	Internet Protocol

Abbreviation	Meaning
IT	Information Technology
JSON	JavaScript Object Notation
MS	Member State
OCSSP	Online content-sharing service providers
ODS	Out-of-court dispute settlement
OECD	Economic Co-operation and Development
OJ	Official Journal
OSP	Online Services Providers
Rec.	Recital
Reg.	Regulation
REST	Representational State Transfer
T&C	Terms & Conditions
TERREG	Regulation on addressing the dissemination of terrorist content online
TOS	Terms of Services
UN	United Nations
UGC	User Generated Content
URL	Uniform Resource Locator
US	United States
VIP	Very Important Person
VLOPs	Very Large Online Platforms
VLOSE	Very Large Online Search Engines
VSP	Video Sharing Platform

Index of Contents

Index of Tables

Index of Figures

1 Executive Summary

Deliverable 7.3 focuses on the evolution of the EU intermediary liability regime and the impact it has on the AI-on-Demand platform. This deliverable is a follow-up to D6.2 "*Report for Policy on Content Moderation*".[2] Deliverable D6.2 provided an overview of the EU policy initiatives on content moderation and the future trends and alternative approaches to content moderation by online platforms. Building on the research conducted in D6.2, the deliverable D7.3 goes more in depth into the Digital Services Act (DSA) obligations for online platforms.

More specifically, Section 3 analyses the shifting approach from an intermediary liability regime for intermediary services providers towards an accountability regime. Sub-section 3.1 first investigates the roles of online platforms and their identification according to the law and their business models. It concludes that an "online platform", as defined in the DSA is a special category of a hosting service: not only does it host third-party content, but also disseminates it to the public. As the intermediary liability regime focuses on illegal content, this notion is being analysed in sub-section 3.2.

The evolution of the liability regime under the EU legal framework is then described in sub-section 3.3. It explains the rationale of exempting certain intermediary providers from the liability for third-party content. The analysis of the e-commerce directive shows that the directive does not provide with a general liability regime for online intermediaries but instead carves out specific rules under which those intermediaries are not held liable under EU law. This sub-section also explains the main conditions for liability exemptions: platforms and other intermediaries are not liable for users' unlawful behaviour unless they are aware of illegal acts and fail to remove them. It analyses how the concept has changed due to the case law of the Court of Justice of the European Union, and how it is clarified in the DSA.

Sub-section 3.4 looks at the evolution of the accountability regime which took place in the DSA. The main novelty of the DSA is the shifted regulatory focus from the liability for content to regulating platforms' responsibility for what they do. The DSA places the focus much more on the platforms' *conduct*, and not content, through the regime of due diligence obligations. These include: transparency provisions, clear notice-and-action processes, internal complaint mechanisms, and risk assessment measures. After explaining the scope of these obligations for each type of an intermediary service (e.g. hosting, online platforms, very large online platforms), the deliverable takes the perspective of an end-user to provide guidelines of what obligations intermediary services have toward their users. Then, the sub-section 3.4 provides an overview of the lex specialis accountability obligations, as provided by the Audiovisual Media Services Directive (AVMSD), terrorist content regulation, Child Sexual Abuse Material (CSAM) and copyright legislation.

[2] Krack N, Dutkiewicz L and Yildirim EO, 'AI4Media Report on Policy for Content Moderation (D6.2)' <https://www.ai4media.eu/reports/report-on-policy-for-content-moderation-d6-2/> accessed 29 September 2023

Section 4 presents the AI-on-Demand (AIoD) platform, its history, rationale and main functionalities. Section 5 explains how guidelines on platform liability and accountability for the AI-on-Demand platform were conducted. Finally, Section 6 offers conclusions.

2 Introduction

In an era marked by the exponential growth of digital content and the ever-expanding influence of online platforms, the dynamics of cyberspace are undergoing a profound transformation. The traditional paradigms of platform liability are gradually giving way to a more encompassing notion: platform responsibility. This progressive shift, driven by an imperative to address the challenges related to the online content landscape, has significant implications for the various and numerous online players.

The internet, seen as a limitless space for creativity and free speech, has turned into a vast space filled with an enormous amount of content created by users. Social media, e-commerce, and content-sharing platforms have become integral facets of our daily lives, serving as conduits for communication, commerce, and entertainment. However, this exponential growth has also engendered a plethora of challenges, from the spread of illegal content and harmful content to concerns about privacy, hate speech, and digital security.

In response to these challenges, the global conversation surrounding the accountability of online platforms has intensified. A fundamental shift is occurring, as societies and policy makers ask for a greater degree of responsibility for platforms and the content they host. The shift from platform liability to platform responsibility represents a pivotal moment in the evolution of the digital age. This research effort studied this transformation, with a specific focus on its ramifications for the AI on-demand platform to which guidelines were formulated on an internal level

The ramifications of this shift are profound and multifaceted. Questions about content moderation, transparency, ethical AI development, and user empowerment come to the forefront. In this context, it becomes crucial to analyse the evolving landscape, identify the key aspects and provisions behind this paradigm shift. The deliverable will be first structured around theoretical sections with legal framework analysis and then the knowledge gathered has been used in the practical assessment and guidance communicated to FhG, the partner working on and liaising with the AI-on-Demand platform.

Disclaimer: Some sections from this deliverable are taken or built from the analysis conducted in D6.2 "*Report on Policy for Content Moderation*" as the two topics are closely interlinked.

3 Analysis of the shifting approach from platform liability to platform responsibility

This section will investigate the shifting approach from platform liability to platform responsibility for third party infringing and/or illegal content.

3.1 The role of platforms and their identification

3.1.1 Their definition and role

The term "online platform" has been used in public discourse to describe a range of services available on the Internet such as online marketplaces, search engines, social media, communication and payment services and so on. The Organisation for Economic Co-operation and Development (OECD) defines an online platform as "a digital service that facilitates interactions between two or more distinct but interdependent sets of users (whether firms or individuals) who interact through the service via the Internet".[3]

Since the adoption of the Digital Services Act (DSA)[4], the term "online platform" has a now a specific, legal meaning. An online platform falls under the broader category of **intermediary services**. Intermediary services are defined in the DSA as the provision of any of the three following services:

- **'mere conduit'** service that consists of the transmission of information provided by a recipient of the service in a communication network, or the provision of access to a communication network; such as internet service providers, direct messaging services, virtual private networks, domain name systems, voice over IP, top level domain name registries;
- a **'caching' service** that consists of the transmission of information provided by a recipient of the service in a communication network, involving the automatic, intermediate and temporary storage of that information, for the sole purpose of making the information's onward transmission to other recipients more efficient upon their request; such as content delivery networks, content adaptation proxies or reverse proxies;
- a **'hosting' service** that consists of the storage of information provided by, and at the request of, a recipient of the service; such as cloud service providers, online

[3] OECD, *An Introduction to Online Platforms and Their Role in the Digital Transformation* (Organisation for Economic Co-operation and Development 2019) <https://www.oecd-ilibrary.org/science-and-technology/an-introduction-to-online-platforms-and-their-role-in-the-digital-transformation_53e5f593-en> accessed 8 August 2023.

[4] 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), OJ L 277, 27.10.2022, p. 1–102. For a general overview about the DSA, see Deliverable D6.2 "Report on Policy for Content Moderation".

marketplaces, social media, app stores. An online platform is a subset of a hosting service: not only does it store information, but also disseminates it to the public.[5]

3.1.2 Their identification

An important question to investigate is how to distinguish an online platform from a 'mere conduits', a 'caching services' or a 'hosting services'. This is of key importance when it comes to the liability regime and accountability obligations. Some criticisms were made about the unclear scope of the definitions of online intermediaries, particularly in the case of services such as video-sharing sites or social networking sites. In particular, the concept of what constitutes a 'communication to the public' has been subject to debates. In *YouTube, C-683/18 Cyando* case, the Court of Justice of the European Union (CJEU) concludes that video-sharing or file-hosting platforms do not themselves make a communication to the public by merely providing the platform infrastructure.[6] Such conclusion has consequences for the question of liability (see Section 3.3).

It is also worth mentioning, that the DSA concerns the *services* and not the *provider* as such. Wilman explains that if a provider offers several services, it is perfectly possible that some of them are covered by the DSA and others are not.[7] It is also noteworthy that it is not relevant whether the provider is established in the EU or in a third country. The applicability of the DSA instead depends on whether the service is offered in the EU[8].

3.1.3 Their business models

On many platforms the **cross-group network effect** is present, meaning that users benefit from the presence of other user groups on the platform. This is the case of, for example, eBay[9] or Airbnb[10], where customers get more choices and potentially a better price if more sellers are present on the platform.[11] **Within-group network effects** can be beneficial as well for the service providers. Whatsapp[12] users, for instance, get an advantage of seeing many of their contacts on the service.[13] Since platforms generate revenues from their intermediation service, they must incentivize the users to join the platform. A platform's income can be generated in different ways:

[5] Art. 3(i) DSA.
[6] Joined Cases C-682/18 and C-683/18: Judgment of the Court (Grand Chamber) of 22 June 2021 Frank Peterson v Google LLC, YouTube LLC, YouTube Inc, Google Germany GmbH (C-682/18) and Elsevier Inc v Cyando AG (C-683/18).
[7] Folkert Wilman, 'The Digital Services Act (DSA) - An Overview' [2022] SSRN Electronic Journal <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4304586>.
[8] Ibid.
[9] https://www.ebay.com/
[10] https://www.airbnb.com/
[11] Miriam C Buiten, Alexandre de Streel and and Martin Peitz, 'Rethinking Liability Rules for Online Hosting Platforms' (2020) 28 International Journal of Law and Information Technology 139.
[12] https://www.whatsapp.com/
[13] Ibid.

- **A fee**. For instance, Vinted[14] is a free app but takes a fee when a deal is concluded between a buyer and a seller. The more users are on the platform, the more exchanges/transactions are concluded, and the more fees can the platform collect.

- **A subscription.** On Spotify[15], you can listen to the music you like without interruption, and get access to additional services such as downloading a playlist for a monthly fee. Even though the access to the content is for free, you pay a fee to listen freely to it.

- **An advertisement.** Some platforms, such as Meta's Facebook[16], are "free" to use, but users are paying with their data. The more users are on the platform, the more lucrative it is for a company to place its ad on the platform. The advertisements are then competing in a real-time bidding process for the best advertising spot, hence generating a profit for a platform.

Platforms can of course combine the different business models to generate more income.

S. Etlinger suggests the following classification of online platforms and their business models (Figure 1):

Type of Platform	Example	Business Model
Advertising platforms	Google, Facebook	Extracting user data and monetizing on ad space
Cloud platforms	Salesforce, Microsoft Azure, Amazon Web Services	Owning and renting out hardware and software
Industrial platforms	GE, Siemens	Building the necessary infrastructures "to transform traditional manufacturing into internet-connected processes"
Product platforms	Rolls-Royce, Spotify	Making use of other platforms "to transform a traditional good into service"
Lean platforms	Uber, Airbnb, DoorDash	Operating on a business model of minimal asset ownership

Figure 1 Classification of online platforms and their business models [17]

[14] https://www.vinted.com/
[15] https://open.spotify.com/
[16] https://www.facebook.com/
[17] Etlinger, Susan. 2021. 'The Next Wave of Platform Governance'. Centre for International Governance Innovation. Accessed 8 August 2023. https://www.cigionline.org/articles/next-wave-platform-governance/ based on Nick Srnicek, Platform Capitalism (Wiley, 2016), 48-49.

3.2 Illegal Content

In the fast-paced digital age, the internet has become an integral part of our daily lives, providing a platform for communication, information-sharing, and entertainment on a global scale. However, this immense connectivity has also given rise to a pressing concern – the proliferation of illegal content online. The European Union, in its commitment to upholding fundamental values and protecting its citizens, has taken significant strides to regulate and combat this issue within its borders.

Illegal content on the internet encompasses a wide range of harmful materials, from hate speech and incitement to violence to copyright infringement, child exploitation, and more. The impact of such content can be profound, threatening the safety, well-being, and integrity of individuals and society as a whole. European legislation has recognized the urgency of addressing these challenges and has implemented a comprehensive framework composed of different legislations and instruments to mitigate their adverse effects.

This deliverable will also explore the balance that European legislation strives to strike between safeguarding the digital space from harm and preserving the open and democratic nature of the internet and the freedom to conduct business from the online digital players.

Ultimately, the regulation of illegal content within the European Union (EU) represents a dynamic and evolving landscape. However, illegal content is always defined as such in the legislation. Indeed, illegal online content refers to material that violates both European and national regulations. The DSA provides a broad definition of what illegal content is: "'illegal content' means any information that, 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 which is in compliance with Union law, irrespective of the precise subject matter or nature of that law".[18] Recital 12 of the DSA states that the "concept of 'illegal content' should broadly reflect the existing rules in the offline environment. In particular, the concept of 'illegal content' should be defined broadly to cover information relating to illegal content, products, services and activities." This includes "the sharing of images depicting child sexual abuse, the unlawful non-consensual sharing of private images, online stalking, the sale of non-compliant or counterfeit products, the sale of products or the provision of services in infringement of consumer protection law, the non-authorised use of copyright protected material, the illegal offer of accommodation services or the illegal sale of live animals".[19]

For the exact content matching the definition of illegal content, there is a need to have a look into the specific/sectoral/vertical content moderation legislations: terrorist, copyright protected, child abuse sexual material,…

The principle is that what is illegal offline should also be illegal online. The classification of such content as unlawful serves various purposes, including safeguarding broader societal interests

[18] Art. 2, h) DSA.
[19] Recital 12 DSA.

and public order (for instance, content promoting terrorism), protecting fundamental personal rights and privacy (e.g., unauthorized handling of personal data), or upholding individual economic rights of significant societal importance (e.g., distributing content in violation of intellectual property rights).[20]

It is important to point the differences between illegal and harmful content. Harmful content is not *per se* illegal but its uses, spread or consumption can generate harms to society and individuals. Harmful content is more difficult to regulate and frame by as it is harder to define.[21]

3.3 Evolution of the liability regime

It is certain that the primary liability lies on the end-users creating and uploading the illegal content or the ones re-uploading or disseminating illegal content. However, in light of internet scale, rapidity and its role as a content dissemination facilitator, online platforms own a peculiar position when it comes to liability for illegal content. In addition, with the Internet "identifying the author seems increasingly complex due to these successive reuploads but also due to increasingly sophisticated anonymisation and obfuscation techniques"[22]. It is therefore important to safeguard end-user and society to have another means of action.

Prior to the EU legislation on the topic, the issue of intermediary liability was addressed differently in the EU Member States. The EU, aware that the lack of a unified approach to intermediary liability created legal uncertainty for online service providers and hindered the growth of e-commerce and online services across the EU, decided to establish a legal framework addressing the topic. The scope and evolution of this liability regime from the e-commerce Directive to the Digital Services Act (DSA) will be the focus of the following sub-sections.

3.3.1 The e-commerce directive liability exemption regime

The e-commerce Directive[23], adopted in 2000, is one of the cornerstones of the Digital Single Market. The e-commerce Directive (ECD) aimed at harmonising minimum standards of liability for internet (online) intermediaries across the EU.[24]

---

[20] European Parliament. Directorate General for Parliamentary Research Services., *Liability of Online Platforms.* (Publications Office 2021) <https://data.europa.eu/doi/10.2861/619924> accessed 24 January 2023.

[21] ibid.

[22] Noémie Krack and others, 'AI in the Belgian Media Landscape. When Fundamental Risks Meet Regulatory Complexities', Artificial Intelligence and the Law, vol 13 (Second Revised Edition, Jan De Bruyne and Cedric Vanleenhove (eds), Intersentia 2023) <https://intersentia.com/en/artificial-intelligence-and-the-law-2nd-edition.html>.

[23] 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'), OJ L 178, 17.7.2000, p. 1–16.

[24] European Parliamentary Research Service, 'Reform of the EU Liability Regime for Online Intermediaries.' (2020) <https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/649404/EPRS_IDA(2020)649404_EN.pdf>.

The Directive applies to any kind of illegal or infringing content. It sets out the framework for the liability regime for intermediary services concerning third-party infringing or illegal content present on their services. Actually, "the directive does not provide with a general liability regime for 'online intermediaries' but instead carves out specific rules under which those intermediaries are not held liable under EU law".[25] Until the DSA becomes fully applicable, the ECD remains a 'lex generalis' for the intermediary liability regime.

This legal regime has four objectives[26]:

1) To share the responsibility for a safe Internet between all the private actors involved and a good cooperation with public authorities.
2) To encourage the development of e-commerce in the EU by increasing legal certainty on the role of each actor and by ensuring that the hosting platforms do not have an obligation to monitor the legality of all material they store.
3) To strike a fair balance between different fundamental rights of the several stakeholders (freedom of expression, right to privacy, freedom to conduct business, property rights for intellectual property).
4) To strengthen the digital single market by adopting a common EU standard for a liability exemption to solve the discrepancies between national rules and divergent caselaw.

Rationale

The ECD provides for horizontal **liability exemptions**. Liability for hosting illegal content, goods or services on the platform's service can be waived according to certain conditions which were harmonised at the EU level thanks to the ECD. The idea behind this regime is that imposing liability on platforms for all illegal activity or content related to their services would constitute a considerable burden and prevent e-commerce development.[27] The exemptions have a broad scope as - if all applicable conditions are met - the providers are exempt "from a wide array of liabilities including contractual liability, administrative liability, tortious (delictual) or extra-contractual liability, penal liability, civil liability or any other type of liability, for all types of activities initiated by third parties, including copyright and trademark infringements, defamation, misleading advertising, unfair commercial practices, unfair competition, publications of illegal content, etc."[28].

[25] European Parliamentary Research Service, 'Reform of the EU Liability Regime for Online Intermediaries.' (2020) <https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/649404/EPRS_IDA(2020)649404_EN.pdf>.

[26] European Commission, 'Explanatory Memorandum of the Commission Proposal for a Directive on Certain Legal Aspects of Electronic Commerce in the Internal Market' COM (1998) 586 final.

[27] Christina Angelopoulos and Martin Senftleben, 'An Endless Odyssey? Content Moderation Without General Content Monitoring Obligations' <https://papers.ssrn.com/abstract=3871916> accessed 1 February 2023.

[28] European Parliamentary Research Service (n 24).

Each liability exemption is attached to one of the intermediary service categories and is therefore governed by a separate set of conditions enabling the benefits of the exemptions.[29]

Mere conduit

Under Article 12 of the ECD, **mere conduit service providers**[30] are exempt from liability when the service provider is only passively involved in the transmission of data. The provider must not:

- Initiate the transmission
- Select the receiver of the transmission
- Select or modify the information contained in the transmission.

Caching

Under Article 13 of the ECD, **caching providers** are "exempt from liability when they temporarily and automatically store data in order to make transmission more efficient (e.g. proxy server) and if several technical conditions for storing the information are met (e.g. local copy identical to original)."[31] The provider must:

- Not modify the information
- Comply with the conditions on access to the information:
    - comply with the rules regarding the updating of the information
    - not interfere with the lawful use of technology to obtain data on the use of the information
    - upon obtaining such knowledge or awareness, act expeditiously to remove or to disable access to the information it has stored.

Hosting

Under Article 14 of the ECD, **hosting providers** can benefit from a liability exemption provided that:

- They do not have actual knowledge of illegal activity or information;
- Upon obtaining such knowledge or awareness, they act expeditiously to remove or to disable access to the information. The provider of a hosting service can obtain knowledge about the illicit character of hosted content through their own activities or through notifications submitted from recipients to take down the content in question (so-called notice and takedown procedure). As a result, it becomes the provider's task to assess whether the complaint is justified and to decide about the illegal or infringing character of the content. The provider can either leave the content on its platform and risk liability for it, or relieve themselves of the problem altogether by simply removing the content.[32]

[29] Christina Angelopoulos and Martin Senftleben, 'An Endless Odyssey? Content Moderation Without General Content Monitoring Obligations' <https://papers.ssrn.com/abstract=3871916> accessed 1 February 2023.

[30] For a definition and example of this type of provider, see Section 3.1.1.

[31] European Parliamentary Research Service (n 24).

[32] Aleksandra Kuczerawy, 'Safeguards for Freedom of Expression in the Era of Online Gatekeeping' (20180914) 2017 Auteurs en Media 292.

The scope of hosting exemptions is quite broad as the case law of the CJEU confirmed its applicability to marketplaces and social media.[33]

Prohibition of general monitoring obligation

Article 15 of the Directive prohibits EU Member States to impose on intermediary service providers a general obligation to monitor content that they transmit or store. Member States cannot introduce a general obligation to actively look for facts or circumstances indicating illegal activity. The prohibition of monitoring obligations does not concern monitoring obligations in a specific case.

Critics

Legitimacy shift

One fundamental criticism comes from the fact that the ECD grants hosting providers the power to decide which content can remain online and which should be removed. They may be considered private 'gatekeepers', who are able to regulate the behaviour (and speech) of their users. By providing conditional liability exemptions for third parties' illegal content or activities, the States enlist the intermediaries to enforce the public policy objectives (i.e., to remove unlawful content).[34] Tambini calls it 'the first settlement on internet content': huge economic benefits during the internet boom made the governments tackle problems of hate speech, piracy, and harm to children by self-regulation. As the author puts it: "whilst the immense public benefits of free speech over the internet were clear, the framework also permitted a new reality of media freedom to open: net neutrality neutered the ability of networks to control speech, even to protect the public from harm."[35]

Legal uncertainty

The bulk of the analysis focuses on issues of fragmentation and legal uncertainty. There is a lack of uniform rules for notice and action procedures across the EU. The details of these national obligations vary from member state to member state. This led to a fragmented EU landscape where some member states decided to only obligate hosting service providers to remove

[33] CJEU, Aleksandra Kuczerawy, 'The Power of Positive Thinking'.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) [2010] CJEU Joined cases C-236/08 to C-238/08; Belgische Vereniging van Auteurs, Componisten en Uitgevers CVBA (SABAM) v Netlog NV [2012] CJEU Case C-360/10; L'Oréal SA et autres contre eBay International AG et autres [2011] Cour de justice Affaire C-324/09.

[34] Aleksandra Kuczerawy, The Power of Positive Thinking: Intermediary Liability and the Effective Enjoyment of the Right to Freedom of Expression, 8 (2017) JIPITEC 226 para 1.

[35] Damian Tambini, Media Freedom (Polity 2021)

content when the notification contains certain information and/or is made by a competent authority.[36]

Lack of safeguards for fundamental freedoms

Moreover, Kuczerawy points to the lack of sufficient safeguards to prevent violations of fundamental rights, in particular freedom of expression.[37] The directive does not include provisions, which would provide for effective mechanisms to avoid and/or resolve incorrect removals of content, such as, for example, out-of-court dispute settlement. This lack of safeguards "leads to over-notification by notifiers, over-removal by providers, and under-assertion of rights by affected users."[38] As noted by Kuczerawy, "the absence of any incentive to conduct a thorough assessment, together with a risk of being held liable, results in a situation where the contested information is often removed or blocked by service providers without giving it a second thought. This leads to situations when legitimate content, for example, criticism in academic discussion or research, political speech, parody or tribute suffers from such risk-averse behaviour by intermediaries."[39]

Lack of technical safeguards

Additionally, Article 14 of the ECD does not explicitly impose an obligation on hosting service providers to respond to such notifications (and the subsequent takedowns). The content can therefore be removed before the content providers have a chance to contest the notification, without an opportunity to answer to the allegations of illegality of their content. Several EU countries have introduced 'counter-notification' measures in their national procedures, but it has not become a standard part of the procedure across the EU.[40] Moreover, once a notice has been issued, the hosting provider is expected to react 'expeditiously'. What constitutes an 'expeditious' reaction and what timeframe is foreseen for this action are not specified, and opinions differ as to when this timeframe starts running.[41] Furthermore, the directive does not

[36] Raphaël Gellert and Pieter Wolters, 'The revision of the European framework for the liability and responsibilities of hosting service providers. Towards a better limitation of the dissemination of illegal content.' (Ministerie van Algemene Zaken 2021) rapport <https://www.rijksoverheid.nl/documenten/rapporten/2021/04/07/the-revision-of-the-european-framework-for-the-liability-and-responsibilities-of-hosting-service-providers> accessed 1 September 2023.

[37] Aleksandra Kuczerawy, Intermediary Liability and Freedom of Expression in the EU: From Concepts to Safeguards (Intersentia 2018).

[38] Alexandre de Streel and Martin Husovec, 'The E-Commerce Directive as the Cornerstone of the Internal Market' [2020] SSRN Electronic Journal <https://www.ssrn.com/abstract=3637961> accessed 24 January 2023

[39] Aleksandra Kuczerawy, 'Intermediary Liability & Freedom of Expression: Recent Developments in the EU Notice & Action Initiative' (2015) 31 Computer Law & Security Review 46.

[40] Kuczerawy, 'Intermediary Liability & Freedom of Expression'

[41] Kuczerawy, 'Intermediary Liability & Freedom of Expression'

envisage that notifications may be sent by bots and fails to incentivise the quality of sent and reviewed notifications.[42]

Vagueness of the prohibition of general monitoring

Some criticism emerged about the prohibition of general monitoring. This concept was not defined in the ECD, and the limits of the concept were subject to hot debates throughout the years leading to different interpretations.[43]

Often related to the Good Samaritan paradox[44], the Directive actually discourages the platforms to proactively monitor the legality of the content, goods or services they are hosting as it may cause the loss of their liability exemption.

But, the question of determining permissible monitoring obligations (specific) and the prohibited ones (general) has been clarified by the CJEU case law. However, the concept seems to vary in the case law depending on the type of content incriminated: copyright infringing content[45] or defamation content[46].

Indeed, the Glawischnig (defamation) case has set a turn in the CJEU's constant interpretation of the prohibition of general monitoring obligation on intermediary service providers. It widened the scope of permissible specific monitoring.[47] In this case, the CJEU ruled that a national court can issue an injunction against a hosting provider to detect and remove an illegal message, as well as any equivalent message with an essentially unchanged message without this constituting a general monitoring obligation. This creates uncertainty about the use of AI tools to moderate content as in this case the intermediary had no other option than to monitor all information uploaded by all users, which is contradictory to the previously established CJEU case law on the

[42] Alexandre de Streel and Martin Husovec, 'The E-Commerce Directive as the Cornerstone of the Internal Market' [2020] SSRN Electronic Journal <https://www.ssrn.com/abstract=3637961> accessed 24 January 2023.

[43] To dive deeper into these different interpretations, we recommend the paper of Christina Angelopoulos and Martin Senftleben, An Endless Odyssey? Content Moderation Without General Content Monitoring Obligations previously cited in this report.

[44] "The disincentives for hosting providers to take a proportional approach against such infringements in fear of losing safe harbor protection is described as the Good Samaritan Paradox" in Oscar Daniel Del Valle Salinas and others, 'A User-Centered Approach To Content Moderation' (Hertie School Centre for Digital Governance 2020) <https://digitalservicesact.eu/wp-content/uploads/2020/06/DSA-Policy-Brief-Content-Moderation.pdf>; Patrick Eecke, 'Online Service Providers and Liability: A Plea for a Balanced Approach' (2011) 48 Common Market Law Review 1455.

[45] *L'Oréal SA et autres contre eBay International AG et autres*; *Belgische Vereniging van Auteurs, Componisten en Uitgevers CVBA (SABAM) v Netlog NV*.

[46] *Eva Glawischnig-Piesczek v Facebook Ireland Limited* [2019] CJEU Case C-18/18.

[47] Toygar Hasan Oruç, 'The Prohibition of General Monitoring Obligation for Video-Sharing Platforms under Article 15 of the E-Commerce Directive in Light of Recent Developments: Is It Still Necessary to Maintain It?' (2022) 13 JIPITEC <http://www.jipitec.eu/issues/jipitec-13-3-2022/5555>.

topic.[48] This shows the importance of the case context (defamation) and the evolution of the online environment. Later on, this interpretation was re-iterated but in a copyright case this time and also in an annulment action against the Copyright in the Digital Single Market (CDSM) Directive.[49] However, it was a case with a notified copyright infringement and where the provider received a sufficiently substantiated notice of the specific infringement or relevant and necessary information regarding the copyright-protected work. These elements must enable the service provider to identify the unlawful content without conducting legal assessment.

The prohibition of general monitoring obligation can lead to paradoxical situations. For instance, a platform carrying out some ex-ante moderation practices to spot illegal content would therefore lose its liability exemption because of this general monitoring prohibition and obligation of conducting a passive role. This passive role would be lost when AI systems are used to proactively search for some content. However, a "refusal or unwillingness to use filters could also be considered as a form of negligence on the part of the intermediary".[50] These contradictory guidelines seem to find a solution in the recent encouragement from the European Commission (EC) for intermediary services providers to adopt a more proactive approach to content moderation.[51] Nevertheless, this is not corroborated by the CJEU case law and therefore uncertainty remains despite some clarity brought by the DSA (see section 3.3.3). So far, content moderation obligations remain confusing, and have not received a full answer in legal texts or case-law up until now. There is a call to have regulatory explicit clarification stating that "the mere fact that providers use AI-technologies does not automatically preclude the exemption of responsibility".[52] For now, this is very implicitly mentioned in case law, and without a proper framework and clear rules, there are concerns about fundamental rights and respect.

3.3.2 The road towards the DSA

On top of all these criticisms, it appeared clear that the ECD needed a refreshment and a complementary framework to address the evolutions of the digital landscape in citizens' lives.

The review process of the ECD started in 2011, with a **public consultation on the future of electronic commerce in the internal market,**[53] having received 420 responses. Amongst the

[48] Ibid.

[49] *Republic of Poland v European Parliament and Council of the European Union* [2022] CJEU Case C-401/19; *Joined Cases C-682/18 and C-683/18: Judgment of the Court (Grand Chamber) of 22 June 2021*.

[50] Emma Llansó and others, 'Artificial Intelligence, Content Moderation, and Freedom of Expression' 30.

[51] 'Commission Recommendation (EU) 2018/334 of 1 March 2018 on Measures to Effectively Tackle Illegal Content Online', vol 063 (2018) <http://data.europa.eu/eli/reco/2018/334/oj/eng> accessed 1 February 2023

[52] Alexandre De Streel and others, *Study on Potential Policy Measures to Promote the Uptake and Use of AI in Belgium in Specific Economic Domains* (FPS Economy 2022).

[53] European Commission, 'Archive - E-Commerce Directive - What Happened before and since Its Adoption | Shaping Europe's Digital Future' (24 March 2017) <https://digital-

responses to obstacles to the development of e-commerce, the divergence in the application of the EU acquis in the Member States was one of the issues raised. "The majority of respondents argued that a revision of the ECD's liability regime would be unnecessary, but thought the existing rules require clarification."[54]

Later, in May 2015, the European Commission (EC) announced a plan to assess the role of online platforms in the **Communication on a Digital Single Market Strategy for Europe.**[55] The work towards re-imagining how digital services work, had started.

In July 2019, the new EC presented its **Political Guidelines** for the new mandate.[56] The President presented 6 priorities and one is a Europe fit for the digital age.

A couple of months later, the EC released a **Communication entitled 'Shaping Europe's Digital Future'**.[57] The communication was organised around 3 pillars: (1) technology that works for people, (2) a fair and competitive digital economy, and (3) an open, democratic and sustainable society. In the same communication, the EC made a commitment to update the horizontal rules that define the responsibilities and obligations of providers of digital services, and online platforms in particular. The Council's Conclusions[58] welcomed the EC's announcement of a Digital Services Act, emphasised 'the need for clear and harmonised evidence-based rules on responsibilities and accountability for digital services that would guarantee internet intermediaries an appropriate level of legal certainty', and acknowledged 'the need to address the dissemination of hate speech and disinformation online'. It also stressed 'the need for effective and proportionate action against illegal activities and content online (...) whilst ensuring the protection of fundamental rights, in particular the freedom of expression, in an open, free and secure internet.'

strategy.ec.europa.eu/en/library/archive-e-commerce-directive-what-happened-and-its-adoption> accessed 4 September 2023.

[54] European Commission, 'Summary of the Results of the Public Consultation on the Future of Electronic Commerce in the Internal Market and the Implementation of the Directive on Electronic Commerce (2000/31/EC)' (2011) <https://ec.europa.eu/information_society/newsroom/image/document/2017-4/consultation_summary_report_en_2010_42070.pdf>.

[55] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: on a Digital Single Market Strategy for Europe 2015.

[56] Directorate-General for Communication (European Commission) and Ursula von der Leyen, *Political Guidelines for the next European Commission 2019-2024 ; Opening Statement in the European Parliament Plenary Session 16 July 2019 ; Speech in the European Parliament Plenary Session 27 November 2019* (Publications Office of the European Union 2020) <https://data.europa.eu/doi/10.2775/101756> accessed 4 September 2023.

[57] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Shaping Europe's digital future 2020.

[58] Council conclusions on shaping Europe's digital future Brussels, 9 June 2020 (OR. en) 8711/20.

The ambition of the EC was to release a package containing two legislative instruments to provide new rules for the online digital space. In 2020, the EC launched a **consultation on the Digital Services Act package** and the results were published on the 15th of December, date of the publication of the two legislative proposals part of the package.[59] Namely the proposal for a Digital Services Act and the proposal for a Digital Market Act. The consultations takeaways point that there is a consensus from the respondents to get a uniform set of rules. There was a strong request for transparency, getting a clear notification scheme for content to be flagged and taken down. The consultation also showed that respondents were in favour of having different rules depending on the size of the digital players and to have a broad territorial and material scope similar to the one of the General Data Protection Regulation (GDPR).[60] The consultation highlighted the need for better enforcement and cross-border cooperation. Another aspect of the results was that users massively encountered harmful and illegal content, goods or services and that there was dissatisfaction with platform's responses when notified. The results pointed out that platforms were taking too long to take down illegal content or were removing a lot of legal content. Famous examples from this era are the following:

Rupi Kaur's photograph

The Canadian poet and artist Rupi Kaur explored, for a University project, the taboo around menstruation through her photographs. Her picture depicting a woman fully clothed with period spots on her trousers and bedsheet got censored twice by Instagram.[61] The picture got removed 24 hours after it was uploaded.[62] The artist explained that she didn't get contacted by the platform nor did she receive explanations about the removal. She reuploaded the picture which got removed a second time. The artist decided to raise awareness and complain in a post about this and her story went viral, Instagram apologised and re-instated the picture.

[59] European Commission, 'Summary Report on the Open Public Consultation on the Digital Services Act Package | Shaping Europe's Digital Future' (15 December 2020) <https://digital-strategy.ec.europa.eu/en/library/summary-report-open-public-consultation-digital-services-act-package> accessed 4 September 2023.

[60] 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 persona data and on the free movement of such data, and repealing Directive 95/ 46/EC (General Data Protection Regulation) 2016 [Regulation (EU) 2016/679] 88.

[61] Tsjeng Zing, 'Why Instagram Censored This Image of an Artist on Her Period' (*DAZED*, 27 March 2015) <https://www.dazeddigital.com/artsandculture/article/24258/1/why-instagram-censored-this-image-of-an-artist-on-her-period> accessed 5 September 2023.

[62] Sarah Cascone, 'Instagram Censors Period Blood, Enrages Artist' (*Artnet News*, 31 March 2015) <https://news.artnet.com/art-world/instagram-slammed-for-censoring-period-283123> accessed 5 September 2023.

Figure 2 Photo from Rupi Kaur

19th Century painting of Gustave Courbet

In 2018, a French court had to decide over the allegation of censorship by Facebook over a user's account.[63] At the origin of the case, in 2011, the user, a teacher, had posted on social media the famous painting of Gustave Courbet 'The origin of the world'. However, Facebook had shut down the user's account because his post was contradictory to the ban on nude images from the platform. The Facebook user claimed that his account got deactivated without any warning or justification. Since then, Facebook has made changes to its policies allowing for nudity in artwork. The Court ruled that Facebook failed to fulfil its contractual obligations by closing without prior notice the account of a user who posted a photo of a famous 19th century nude painting.[64]

Figure 3 Painting from Gustave Courbet entitled "L'Origine du Monde".

[63] AFP, 'Facebook to French Court: Nude Painting Did Not Prompt Account's Deletion' *The Guardian* (1 February 2018) <https://www.theguardian.com/technology/2018/feb/01/facebook-nude-painting-gustave-courbet> accessed 5 September 2023.
[64] Philippe Sotto, 'French Court Issues Mixed Ruling in Facebook Nudity Case' (*Courthouse News Service*, 15 March 2018) <https://www.courthousenews.com/french-court-issues-mixed-ruling-in-facebook-nudity-case/> accessed 5 September 2023.

The question of **liability was a hot debate** before the DSA package release. Many were wondering whether the liability exemption mechanisms which helped platforms grow for 20 years was still appropriate in today's online world.[65] The consultation showed that the public wanted platforms to do more against illegal content and the EU legislator came up with several EU legislations. On a horizontal level, the DSA proposal and the Digital Market Act proposal were released on 15 December 2020. On a sectoral level, legislations specific to terrorist content, CSAM content, content infringing intellectual property rights and so forth got adopted in recent years. They came revising the existing measures in place and asked for more from platforms in the fight against illegal content online.

3.3.3 The DSA and the liability regime

The Digital Service Act was adopted in October 2022. The objective about the revision of the intermediary liability rules was to reinforce and further clarify the conditions for liability exemptions: platforms and other intermediaries are not liable for users' unlawful behaviour unless they are aware of illegal acts and fail to remove them. Indeed, the liability regime of the ECD has been through the years further clarified by the case-law of the Court of Justice of the European Union (CJEU), this legal construction brought legal security to the digital players and it has been deemed important to keep the original framework while improving it and enshrining the clarifications in law.[66]

The DSA is a central piece for content moderation and for the intermediary liability regime as it is the first "overhaul of the horizontal rules of the e-commerce Directive".[67] The shift from a Directive towards a regulation will now harmonise and uniformize the liability rules across the EU thanks to a directly applicable Regulation.

Preserving the intermediary liability exemption regime

The articles from the ECD (Art. 12-15) were transplanted in the DSA almost literally. Therefore, the rationale applicable to the ECD's conditional horizontal liability exemptions is also applicable for the DSA.

Rec. 17 of the DSA provides that "the rules on liability of providers of intermediary services set out in this Regulation should only establish when the provider of intermediary services concerned cannot be held liable in relation to illegal content provided by the recipients of the service. (…) Furthermore, the exemptions from liability established in this Regulation should apply in respect of any type of liability as regards any type of illegal content, irrespective of the precise subject matter or nature of those laws." The DSA therefore upholds the pre-existing

[65] Miriam C Buiten, Alexandre de Streel and and Martin Peitz, 'Rethinking Liability Rules for Online Hosting Platforms' (2020) 28 International Journal of Law and Information Technology 139.
[66] Recital 16 DSA.
[67] Sebastian Felix Schwemer, 'Digital Services Act: A Reform of the e-Commerce Directive and Much More' (10 October 2022) <https://papers.ssrn.com/abstract=4213014> accessed 1 December 2022.

regime for illegal content. The DSA does not include harmful content in the scope of the liability exemptions.

The personal scope is still the same. The provisions on liability applies to all intermediary services: caching, mere conduit and hosting. The DSA clarifies in its recital 28 that "providers of services establishing and facilitating the underlying logical architecture and proper functioning of the internet, including technical auxiliary functions, can also benefit from the exemptions from liability set out in this Regulation, to the extent that their services qualify as 'mere conduit', 'caching' or 'hosting' services".

Clarifying the liability exemptions regime

However, even if the regime remains the same, some clarifications were brought by the DSA to solve the unclarities and enshrine the CJEU case-law. The articles have almost been literally transposed from the Directive. These tiny changes have nevertheless an impact on the liability exemption regime. Much has been clarified through the use of recitals.

The concept of active and passive role is something that has been discussed significantly considering the automation of content moderation tasks. However, besides transposing the interpretation of the CJEU on this aspect in Recital 18 of the DSA, the concept isn't further specified.

However, the concept of actual knowledge is now further clarified with the recital 22 of the DSA. It states that the "actual knowledge or awareness cannot be considered to be obtained solely on the ground that the provider is aware, in a general sense, of the fact that its service is also used to store illegal content."

Updating the liability regime

Figure 4 shows the comparison between the ECD Articles and the DSA articles.

e-Commerce Directive	Digital Services Act
Article 12 – "Mere conduit"	Article 4 – 'Mere conduit'
Article 13 – "Caching"	Article 5 – 'Caching'
Article 14 – Hosting	Article 6 – Hosting
	Article 7 – Voluntary own-initiative investigations and legal compliance
Article 15 – No general obligation to monitor	Article 8 – No general monitoring or active fact-finding obligations
	Article 9 – Orders to act against illegal content[45]
	Article 10 – Orders to provide information

Figure 4 Comparison between the e-Commerce Directive Articles and the DSA articles (from Sebastian Felix Schwemer, 'Digital Services Act: A Reform of the e-Commerce Directive and Much More).

Voluntary own-initiative investigations and legal compliance

The DSA introduces in its Art. 7 a provision focusing on the voluntary own-initiative investigations and legal compliance from intermediary services. This provision has been compared to a Good Samaritan provision.[68] Article 7 clarifies the role of intermediary services in situations involving voluntary own-initiative investigations or the adoption of other measures designed to detect, identify, and remove illegal content or ensure compliance. The question was important in terms of legal certainty as it would provide the reassurance to the providers when it comes to the benefits of liability exemption. Indeed, if the providers would lose the passive role status, they could no longer benefit from the liability exemption regime. Art. 7 means that "voluntary measures taken by intermediaries on their own initiative should not be the sole reason for the loss of immunity. (…) taking voluntary actions in good faith neither guarantees not precludes neutrality."[69] The provision does not specify what could qualify as a voluntary action. Some wonder whether this provision would not lead providers to conduct general monitoring through voluntary basis.

In relation to the use of technology for content moderation purposes and its impact on intermediary liability rules, recital 26 of the DSA states that: "the providers concerned should, for example, take reasonable measures to ensure that, where automated tools are used to conduct such activities, the relevant technology is sufficiently reliable to limit to the maximum extent possible the rate of errors." S. Schwemer explains that "a low percentage of false-positives and false-negatives, seems to be the guiding benchmark for voluntary actions. The DSA, however, remains silent as to what error rate would be acceptable for society."[70]

Cooperation with national, judicial and administrative authorities

What is also new in the DSA are the Art. 9 and 10. They introduce rules about the national orders coming from judicial or administrative authorities requesting intermediary services providers' cooperation. They clarify what are the conditions associated with these orders to bring clarity and certainty to the providers but also against State interference. Indeed, the orders need to contain a statement of reasons indicating specific information. These provisions enshrine previous CJEU case-law rulings establishing that the orders must be specific enough to avoid any

[68] In the US, Section 230 of the US Code provides that online intermediaries should not be held liable for any voluntary actions taken in good faith against certain types of objectionable content. A. Kuczerawy wrote that "Section 230 essentially protects intermediaries when they act with good intentions to restrict access to or availability of content. But it also protects intermediaries when they do not act against such content, regardless of whether they have knowledge of it or not". Aleksandra Kuczerawy, 'The Good Samaritan that wasn't: voluntary monitoring under the (draft) Digital Services Act' [2021] Verfassungsblog <https://verfassungsblog.de/good-samaritan-dsa/> accessed 5 September 2023.
[69] Aleksandra Kuczerawy, 'The Good Samaritan that wasn't: voluntary monitoring under the (draft) Digital Services Act' [2021] Verfassungsblog <https://verfassungsblog.de/good-samaritan-dsa/> accessed 5 September 2023.
[70] Schwemer (n 67).

general monitoring. In that sense, these provisions do not grant any new power, but they harmonise the framework.[71]

Art. 9 focuses on the orders to act against illegal content. The orders "should balance the objective that the order seeks to achieve, in accordance with the legal basis enabling its issuance, with the rights and legitimate interests of all third parties that may be affected by the order, in particular their fundamental rights under the Charter."

Art. 10 focuses on the orders to provide information. Rec. 37 specifies that there is the need to identify specifically the individuals targeted by the orders.

In addition, other DSA articles under the due diligence obligations Chapter of the DSA prove to be updating the liability regime. These obligations will be further detailed in section 3.4 They include:

- Art. 16 harmonising the notice-and-action procedures.
- Art.17 providing end-user a statement of the reasons following a content moderation decision.
- Art. 20 and 21 DSA setting out an effective internal complaint mechanism and the option to rely on an out-of-court dispute settlement body for complaint about a content moderation decision.
- Art. 22 establishing the status of trusted flaggers, whose notices must be handled in priority by the online platforms.
- Art. 23 containing a prohibition of abuse of right, where the platform can act following the frequent submission of manifestly unfounded notices or complaints.

3.3.4 Sector specific legislations

As pointed out earlier, the ECD and now the DSA set up horizontal and general rules applicable for intermediary providers and providing the liability exemption regime for illegal content on their services. This general framework which can also be called a baseline framework is complemented by vertical rules, some lex specialis addressing specific types of content deserving specific attention, rules, and processes (see Figure 5). They cover terrorist content, child sexual abuse material (CSAM), copyright infringing content, racist and xenophobic content, disinformation, and hate speech. Given the various sensitivity or degrees of the illegality of this content, a one size fits all approach would be detrimental to freedom of expression; therefore, specific rules have been adopted. These lex specialis rules are often complemented by self-regulatory initiatives. Lex specialis means that when there is a conflict of laws of equal

[71] Miriam C Buiten, 'The Digital Services Act: From Intermediary Liability to Platform Regulation' (2022) 12 JIPITEC <http://www.jipitec.eu/issues/jipitec-12-5-2021/5491>.

importance in the hierarchy of norms, the preference/applicability shall be given to the most specific, the one that approaches most nearly to the subject at hand.[72]

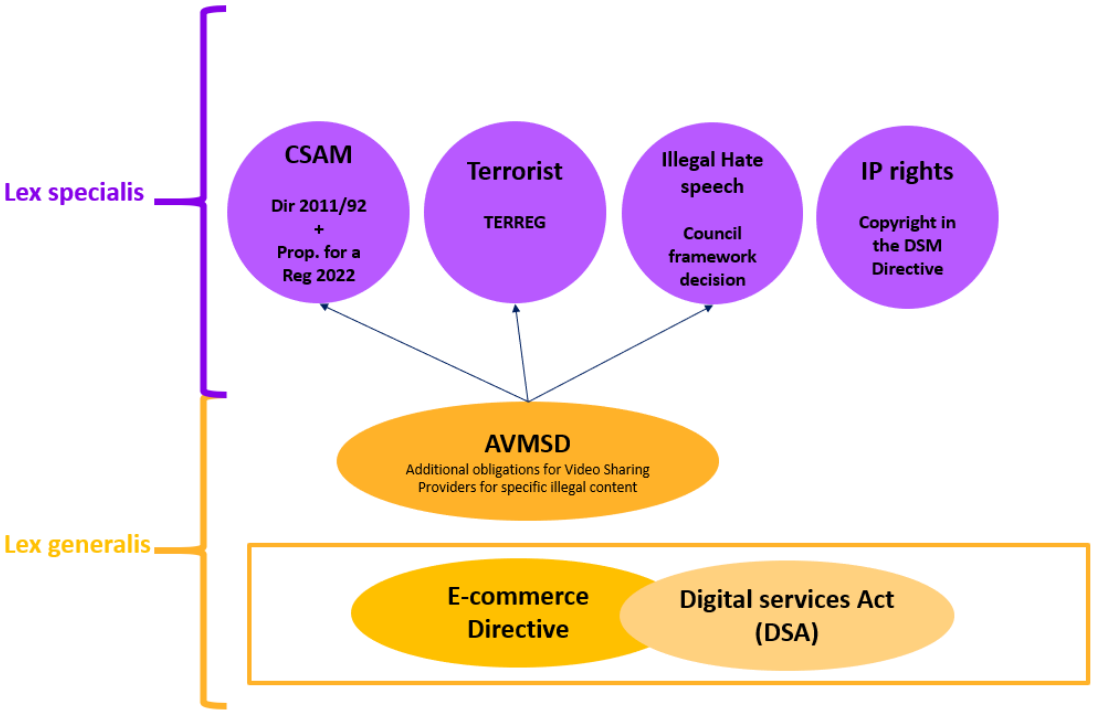


Figure 5 Overview of the lex generalis and lex specialis for illegal content moderation.

3.4 Evolution of accountability regime

The DSA's main novelty is the shifted regulatory focus from the liability for content to regulating platforms' responsibility for what they do. As explained above, prior to the DSA, the regulation of e-commerce in the EU focused primarily on platforms' ***liability*** for user generated content. Now, societal responsibility of online services to provide a trustworthy, fair and safe digital space has been recognized. This recognition comes from the realisation that content which is not illegal, may nonetheless be harmful to individuals and society as such. *"What responsibility do online platforms' algorithms play in the distribution and promotion of online content? And what are the commercial incentives that guide the development of those algorithms?"* – these are some of the questions which influenced the thinking process about the new rules. [73]

[72] Hugo Grotius, De Jure Belli Ac Pacis. Libri Tres; Anja Lindroos, 'Addressing Norm Conflicts in a Fragmented Legal System: The Doctrine of Lex Specialis' (2005) 74 Nordic Journal of International Law 27.

[73] Sebastian Becker Castellaro and Jan Penfrat, 'The DSA fails to reign in the most harmful digital platform businesses – but it is still useful' [2022] Verfassungsblog <https://verfassungsblog.de/dsa-fails/> accessed 10 August 2023.

The DSA places the focus much more on the platforms' ***conduct***, and not content, through the regime of **due diligence obligations**. Transparency provisions, clear notice-and-action processes, internal complaint mechanisms, and risk assessment measures are thought to strengthen the users' position vis-à-vis the platforms.

Importantly, the due diligence obligations apply **irrespective of the question of liability** for dissemination of illegal content. Even providers who are not liable for users' content remain accountable for their failings to be diligent.[74] In addition, the failure from an intermediary service provider on the due diligence obligations does not imply that the liability exemptions are not available to the provider.[75] The two need to be assessed separately.[76]

Chapter III of the DSA is the core chapter of the legislative instrument containing most of the provisions. It brings the biggest innovations of the text. Importantly, the obligations under the DSA are cumulative: for example, a provider of very large online platforms must comply not only with the specific rules for such services, but also with those for online platforms, hosting services and intermediary services. A hosting service, on the contrary, must comply with the rules applicable to all intermediary services, as well as with the specific rules applicable to hosting services.

3.4.1 Obligations for all intermediary services (mere conduit, caching and hosting)

Except for annual reporting, the obligations apply irrespective of the size of the firm:

A single point of contact or legal representative

All digital services in scope of the regulation must have a single point of contact (Article 10). The single point of contact is required to be easily contactable and responsive to users' requests. Intermediary services that are not legally established in the EU must appoint a legal representative in a Member State (Article 11). This representative is responsible for communicating directly with Member States' authorities, the EC, and the European Board for Digital Services.

Explanation in the terms and conditions whether and how they moderate third-party content

The DSA provides, for the first time, a legal definition of content moderation. Article 3(t) defines 'content moderation' as the activities, whether automated or not, taken by providers of intermediary services, that are aimed, in particular, at detecting, identifying, and addressing illegal content or information incompatible with their terms and conditions. These activities include any measures that affect the availability, visibility, and accessibility of that content, such as demotion, demonetisation, disabling of access to, or removal, or that affect the ability of the recipients of the service to provide that information, such as the termination or suspension of a

[74] Martin Husovec and Irene Roche Laguna, 'Digital Services Act: A Short Primer' [2022] SSRN Electronic Journal <https://www.ssrn.com/abstract=4153796> accessed 14 October 2022.
[75] Schwemer (n 67).
[76] Recital 41 of the DSA.

recipient's account. Importantly, this broad definition of content moderation includes remedies that go beyond content removal.

All intermediary services must publish in their terms and conditions, in a "clear and unambiguous language", information on any policies, procedures, measures and tools used for the purpose of content moderation, including about "algorithmic decision-making" and human review (art. 14). Moreover, they need to apply their content moderation policies in a diligent, objective, and proportionate manner. When enforcing these rules, they are under an obligation to do so "with due regard to" the rights and legitimate interests of all parties involved. This is an important obligation, as so far, in practice, platforms often applied restrictions based on a vague and undefined terms and conditions. Now, the DSA requires providers to clarify in advance what they allow on their platforms.

Annual transparency reports

Except for micro or small firms[77], intermediary services must publish detailed yearly reports about their content moderation (Article 15). Specifically, such reports must include:

- The number of orders received from Member States' authorities including categorised by the type of illegal content;
- Meaningful and comprehensible information about the content moderation on its own initiative, including the use of automated tools;
- The measures taken to provide training and assistance to human moderators;
- The number and type of measures taken that affect the availability, visibility and accessibility of content and other related restrictions of the service categorised by the type of illegal content or violation of the terms and conditions;
- The number of complaints received through the internal complaint-handling systems;
- Any use made of automated means for the purpose of content moderation, including a qualitative description, a specification of the precise purposes, indicators of the accuracy and the possible error rate.

3.4.2 Obligations for hosting services

Notice and action procedures and a statement of reasons

Providers of hosting services are required to introduce mechanisms to allow users to notify them of the presence of allegedly illegal content (so-called notice and action procedure). Moreover, under Article 17, providers of hosting services have an obligation to provide a clear and specific statement of reasons to the affected user in which they must explain the facts and circumstances

[77] Microenterprise employs fewer than ten persons and its annual turnover and/or annual balance sheet total does not exceed €2 million. A small enterprise employs fewer than 50 persons and its annual turnover and/or annual balance sheet total does not exceed €10 million.

relied on in taking a content moderation decision, information on the use of automated means, the grounds and reasons behind this action. Such justification must be properly explained and must contain information about potential redress.

Detailed transparency reports

Providers of hosting services must also report the number of notices submitted under the notice and action procedures, categorized by the type of alleged illegal content, as well as the number of notices submitted by so-called "trusted flaggers" (a definition of "trusted flaggers" can be found in section 3.4.3, below). It should also be specified whether the action in response to a notice was taken on the basis of the law or the terms and conditions. The number of notices processed by using automated means and the median time needed for taking the action are also part of the reporting obligation.

Reporting criminal offences

Under Art. 18, providers of hosting services are obliged to inform the national law enforcement or judicial authorities of the relevant EU Member State of any information that gives rise to suspicions of criminal offences involving a threat to the life or safety of persons.

3.4.3 Obligations for online platforms

Online platforms are a special category of hosting services: not only do they provide a technical function of storage of third-party information, but also, they disseminate the information to the public.

Internal complaint-handling system

As explained above, hosting services, including online platforms, have to provide a statement of reasons as to why the content has been removed, disabled, suspended, or a user's account terminated and to implement an internal complaint-handling system that enables users to lodge complaints. Moreover, providers of online platforms are obliged to inform complainants without undue delay of their reasoned decision and of the possibility of out-of-court dispute settlement and other available possibilities for redress.

Trusted flaggers

Providers of online platforms shall take the necessary technical and organisational measures to ensure that notices submitted by so-called "trusted flaggers" are given priority and are processed and decided upon without undue delay. The status of a trusted flagger will be given to an entity upon its application to the Digitals Services Coordinator (DSC)[78] in their Member State. A trusted flagger shall have particular expertise and competence for the purposes of detecting, identifying and notifying illegal content; be independent from any provider of online platforms; and shall carry out their activities for the purposes of submitting notices diligently, accurately and objectively.

Transparency reporting obligations

Online platforms also have additional transparency reporting obligations, which include:

- The number of disputes submitted to the out-of-court dispute settlement bodies, the outcome of such proceedings and the median time needed for completing the dispute settlement procedures and whether or not the provider implemented the decision;
- The number of service suspensions of those users who frequently provide manifestly illegal content.

Fair design of services

Under Art. 25 DSA providers of online platforms shall not design, organise or operate their online interfaces in a way that deceives or manipulates the recipients of their service or in a way that otherwise materially distorts or impairs the ability of the recipients of their service to make free and informed decisions. Moreover, providers of online platforms accessible to minors shall put in place appropriate and proportionate measures to ensure a high level of privacy, safety, and security of minors, on their service (Art. 28(1)).

Advertising

The providers of online platforms must provide users with information relating to any online advertisements on their platform. The recipients of the service must be able to identify, in a clear, concise and unambiguous manner and in real time the information displayed is an advertisement, the person on whose behalf the advertisement is displayed, and the main parameters used to determine the recipient to whom the advertisement is displayed to. Providers of online platforms are also prohibited from presenting targeted advertisements based on profiling using special category of data (as defined in the GDPR) and they shall not present ads based on profiling to a minor.

[78] DSC are the national competent authorities responsible for the application and enforcement of the DSA Regulation. Art. 38 of the DSA.

Recommender system transparency

The DSA also regulates recommender system transparency. The DSA explains that recommender systems can have a significant impact on the ability of recipients to retrieve and interact with information online. Consequently, providers of online platforms are expected to set out in their terms and conditions in plain and intelligible language the main parameters used in their recommender systems and the options for users to modify or influence them (Art. 27). The main parameters shall explain why certain information is suggested, and include, at least, the criteria which are most significant in determining the information suggested, and the reasons for the relative importance of those parameters.

3.4.4 Additional provisions applicable to providers of online platforms allowing consumers to conclude distance contracts with traders (B2C)

Additionally, the DSA provides for special rules for the providers of platforms allowing for distance contracts to be concluded between consumers and traders (B2C online marketplaces). The DSA provides for a so-called 'know your business customer' obligation, which is intended to provide greater traceability of traders and thus, legal certainty for a consumer (Article 30). Prior to concluding a contract, a customer should have the trader's contact details, the payment account details, information where the trader is registered and a self-certification by the trader committing to only offer products or services that comply with the applicable rules. Providers of B2C online platforms shall moreover ensure that its online interface is designed and organised in a way that it provides: (i) clear and unambiguous identification of the products or the services promoted or offered to consumers; (ii) signs identifying the trader such as the trademark, symbol or logo; and, (iii) where applicable, the information concerning the labelling and marking in compliance with rules product safety regulations (so-called 'compliance by design'). Lastly, providers shall make reasonable efforts to check whether the products or services offered have not been identified as illegal. If providers become aware that such illegal content was purchased, they must inform consumers about this fact and about relevant redress mechanisms.

3.4.5 Obligations for very large online platforms (VLOPs) and very large search engines (VLOSEs)

On 25 April 2023, the EC adopted the first designation decisions, designating 17 Very Large Online Platforms (VLOPs) and 2 Very Large Online Search Engines (VLOSEs) that reach at least 45 million monthly active users.[79] Following their designation, companies will now have to comply, within four months, with the full set of new obligations under the DSA. Apart from the obligations explained above, the additional obligations for the VLOPs and VLOSEs include:

Risk assessment and mitigation

A true novelty of the DSA is a risk assessment mechanism. VLOPs and VLOSEs are required to conduct an annual assessment on any systemic risks stemming from the functioning and use of

[79] The list is available here: https://ec.europa.eu/commission/presscorner/detail/en/IP_23_2413

their services, including algorithmic systems. The DSA mentions a wide array of systemic risks, in particular: (i) dissemination of illegal content; (ii) any actual or foreseeable negative effects for the exercise of fundamental rights (such as respect for private and family life, freedom of expression and information, the prohibition of discrimination, and the rights of the child as set out in the EU's Charter of Fundamental Rights); (iii) any actual or foreseeable negative effects on the protection on civic discourse and electoral processes, and public security; and (iv) any actual or foreseeable negative effects in relation to gender-based violence, the protection of public health and minors and serious negative consequences to the person's physical and mental well-being.

It is not entirely clear when an online content or conduct risk becomes "systemic". Critics foresee that a substantial work is needed to define systemic risk, as for now there lacks clarity about what this term means, and how it should be understood and operationalized in DSA risk assessments.[80] Critics point out that the variety of types of potential risk may make it difficult to define appropriate and coordinated assessment tools.[81] Although there are some specific tools to counter illegal content and methodologies assessing the impact on human rights (Human Rights Impact Assessments, UN Guiding Principles on Business and Human Rights), how to assess negative effect on civic discourse and electoral process is more complex. Allen makes an important point about the need to adopt an intersectional methodology[82] to assess negative effects in relation to gender-based violence[83].

Such risks will have to be **mitigated** by implementing tailored, reasonable, proportionate, and effective mitigation measures. When conducting risk assessments, VLOPs and VLOSEs shall consider, in particular, whether the design of their recommender systems and their content moderation systems influence any of the systemic risks. If so, they must put in place mitigation measures, such as testing and adapting their algorithms (Art. 35). The risk mitigation measures will be subject to an independent audit and oversight by the EC.

---

[80] David Sullivan, 'Unpacking "Systemic Risk" Under the EU's Digital Service Act' (*Tech Policy Press*, 19 July 2023) <https://techpolicy.press/unpacking-systemic-risk-under-the-eus-digital-service-act/> accessed 9 August 2023.

[81] Alessandro Mantelero,Alessandro Mantelero, 'Fundamental Rights Impact Assessments in the DSA' (*Verfassungsblog*, 1 November 2022) <https://verfassungsblog.de/dsa-impact-assessment/> accessed 10 August 2023.

[82] Intersectionality is "*an analytical framework for understanding how aspects of a person's social and political identities combine to create different modes of discrimination and privilege. Concretely, the method looks at the interconnected nature of social categorisations such as race, class, and gender, which can create overlapping and interdependent systems of discrimination or disadvantage.*" in Allen, Asha: *An Intersectional Lens on Online Gender Based Violence and the Digital Services Act*, VerfBlog, 2022/11/01, https://verfassungsblog.de/dsa-intersectional/, DOI: 10.17176/20221101-215626-0.

[83] Asha Allen, 'An Intersectional Lens on Online Gender Based Violence and the Digital Services Act' [2022] Verfassungsblog <https://verfassungsblog.de/dsa-intersectional/> accessed 10 August 2023.

Crisis management

The question of how to deal with crises has become an important question also for the DSA to consider, in light of the recent COVID-19 pandemic and the war in Ukraine. The DSA crisis response mechanism gives the EC significant powers: the EC, acting upon a recommendation of the Board may require VLOPs and VLOSEs to: (i) assess whether, and if so to what extent and how, the functioning and use of their services significantly contribute to a serious threat to public security or public health in the EU or significant parts thereof; (ii) to identify and apply specific, effective and proportionate measures to prevent, eliminate or limit any such contribution; (iii) to report to the EC on the assessment, the measures taken. This mechanism can be seen as a kind of "*ad hoc* risk management regime" which needs to be put in place during an unexpected and urgent situation.[84]

External and independent auditing

VLOPs and VLOSEs will also have to submit annual independent audits to confirm their compliance with various obligations under the DSA. They will also have to implement the remedial measures in case of receiving an audit report that is not 'positive'.

Additional transparency obligations

VLOPs and VLOSEs will need to provide an easily understandable, plain-language summary of their terms and conditions, in the languages of the Member States where they operate. Moreover, there are additional requirements imposed on the providers of VLOPs and VLOSEs to provide at least one option for their recommendation systems which is not based on profiling. Advertising carried by VLOPs must be accompanied by a publicly accessible archive – a repository. Such a repository should include the content of the advertisement, the person on whose behalf the advertisement is presented, who paid for the advertisement, the period during which the advertisement was presented; who was the ad intended to be presented to, the main parameters used for that purpose and the total number of recipients of the service reached. In addition to other transparency obligations, VLOPs must specify what human resources they put into the content moderation per each official language of Member States, the qualifications of the persons carrying out content moderation, and the indicators of accuracy of such processes.

Access to data for researchers

The DSA requires providers of VLOPs and VLOSEs to grant 'vetted researchers'[85] access to data, subject to certain conditions. Data can be provided "for the sole purpose of conducting research that contributes to the detection, identification and understanding of systemic risks (…) and to the assessment of the adequacy, efficiency and impacts of the risk mitigation measures (…)"(Art. 40(4)). Vetted researchers must meet certain criteria and procedural requirements in the

[84] Wilman (n 7).

[85] Researchers are awarded this status if they comply with several conditions detailed in this section. This status gives them access to data from very large online platforms that are necessary to monitor and assess compliance with this Regulation.

application process. Importantly, they must be affiliated to a research organization or a not-for-profit body, organization or association (Art. 40(12)). Many details around researchers' access to data through the DSA will be decided in delegated acts that have yet to be adopted by the European Commission.

Figure 6 provides a summary of the asymmetric due diligence obligation for intermediary services providers.

	VLOPs and VLOSEs	Online Platforms	Hosting Services	All intermediaries
Transparency reporting	•	•	•	•
Requirements on terms of service due account of fundamental rights	•	•	•	•
Cooperation with national authorities following orders	•	•	•	•
Points of contact and, where necessary, legal representative	•	•	•	•
Notice and action and obligation to provide information to users	•	•	•	
Reporting criminal offences	•	•	•	
Complaint and redress mechanism and out of court dispute settlement	•	•		
Trusted flaggers	•	•		
Measures against abusive notices and counter-notices	•	•		
Special obligations for marketplaces, e.g. vetting credentials of third party suppliers ("KYBC"), compliance by design, random checks	•	•		
Bans on targeted adverts to children and those based on special characteristics of users	•	•		
Transparency of recommender systems	•	•		
User-facing transparency of online advertising	•	•		
Risk assessment and mitigation	•			
External and independent auditing	•			
Compliance officer- internal compliance function	•			
User choice not to have recommendations based on profiling	•			
Enhanced advertising transparency	•			
Data sharing with authorities and researchers (access and scrutiny)	•			
Codes of conduct	•			
Crisis response cooperation	•			

Figure 6 Summary of the asymmetric due diligence obligation for intermediary services providers, adapted from a previous design by the EC.

3.4.6 New end-users' rights granted by the DSA

Reporting illegal content / All hosting services

As rightly pointed out by Kuczerawy, the right to an effective remedy comes into play on two separate occasion: when a user attempts to stop an infringing expression (e.g. by requesting its removal); and, in a case of successful removal, when the author tries to contest the removal and asks for the reinstatement.[86] As already mentioned, the *notice and action mechanism* allows users to report illegal content. After the submission of a notice, the service provider should get back to them without delay and inform about redress possibilities. Under 44 of the DSA, the EC shall promote the standardization of the electronic submission of Article 16 notices. As pointed

[86] Aleksandra Kuczerawy, 'Remedying Overremoval: The Three-Tiered Approach of the DSA' [2022] Verfassungsblog <https://verfassungsblog.de/remedying-overremoval/> accessed 8 March 2023.

out by Ortolani, such standardized notices could enhance the usefulness of notice-and-action mechanisms as a tool for access to justice and they may help avoid dark patterns, and ensure that affected parties have equal access to the mechanism, irrespective of the type of illegality they are reporting.[87] In his view, this may help overcome the current status quo, in which platforms facilitate the reporting of certain categories of illegal content, while failing to do the same for others.[88]

Statement of reasons / All hosting services

Hosting services will have to provide a *statement of reasons* to any affected recipient of the service on any restrictions on visibility of the content, restrictions of monetisation of the content, suspension, or termination of the service in whole or in part, also of the recipient's accounts. The statement must also provide user-friendly information on redress mechanisms available. Overall, this shall bring more transparency to the users about the decisions that platforms take in their content moderation. Understanding the reasons behind such decision is a first step towards a successful complaint.

Internal complaint handling system / All online platforms except micro and small ones

All online platforms except micro and small ones will need to have a system allowing users to complain about platforms' content moderation decisions, both in cases of over-moderation (take downs) and under-moderation (when platforms decide not to act on reported content). Specifically, platforms need to provide users with access to an effective internal complaint-handling system that enables them to lodge complaints, electronically and free of charge, against the decision whether or not to:

- disable access to or restrict visibility of the information;
- suspend or terminate the provision of the service, in whole or in part, to the recipients;
- suspend or terminate the recipients' account;
- suspend, terminate or otherwise restrict the ability to monetise information provided by the recipients.

If a complaint contains sufficient grounds for the platform to consider that its decision not to act upon the notice is unfounded or that the information to which the complaint relates is not illegal and is not incompatible with its terms and conditions, it should reverse its decision without undue delay. Moreover, the user will have to be informed of the possibility of out-of-court dispute settlement provided and other available possibilities for redress.

Challenging the outcome of the decision / All online platforms except micro and small ones

Users can challenge the platform's decision with yet another tool: *out-of-court dispute settlement (ODS),* provided under Art. 21 DSA. Any effected user is entitled to file an ODS

[87] Pietro Ortolani, Pietro Ortolani, 'If You Build It, They Will Come' (*Verfassungsblog*, 7 November 2022) <https://verfassungsblog.de/dsa-build-it/> accessed 10 August 2023.
[88] ibid.

complaint even if they did not initiate an internal complaint. Member States shall create independent bodies through which users could settle disputes with the platform. These bodies would be certified by the Digital Service Coordinators (DSC) of the Member States. To be certified, the ODS body must demonstrate its impartiality and independence from providers, their users, and notifiers, an expertise in the subject matter area, such as area of illegal content, terms and conditions of specific services or platforms (there is no requirement of legal expertise), its ability to settle disputes in a "swift, efficient and cost-effective manner" and in at least one of the official languages of the institutions of the EU, and it must adopt "clear and fair rules of procedure that are easily and publicly accessible". It follows, that two main types of ODS bodies are likely to emerge: (i) area-specific ODS bodies that focus on specific types of illegal content, such as hate speech, child abuse material, or copyright infringements; (2) platform specific ODS bodies, such as Facebook-related, Twitter-related.[89] Moreover, since ODS bodies must be "independent" from providers, their users, and notifiers, existing oversight mechanisms such as Meta's Oversight Board, are unlikely to qualify as ODS in their current form.[90]

The out-of-court dispute settlement bodies cannot, however, impose a binding solution. It does not preclude users to initiate, at any stage, proceedings to contest the decisions by the providers of online platforms before a court. As pointed out by Husovec, *"ODS bodies are not courts, not even "de facto" courts. The DSA's system is probably better described as a system of second opinions that providers cannot easily ignore and must pay for if they lose".* [91] He argues that the *ratio legis* behind this system, is to prevent well-documented over-blocking of content by the platforms: the external body will assess the dispute and impose small costs onto providers for their initial mistakes.

Representation / All online platforms except micro and small ones

Under Article 86 DSA, users have the right to mandate a non-for-profit body, organisation or association to exercise the rights provided for by the DSA on their behalf. Internal complaints under Article 20(1) submitted by these entities on behalf of the users should be processed and decided upon with priority and without undue delay.

Complaint about a platform to Digital Services Coordinator (DSC)

Users and the abovementioned bodies, organisations or associations have the right to lodge a complaint against providers of intermediary services with the Digital Services Coordinator (DSC) about any platform for an alleged breach of the DSA. The DSC in a Member State where the

[89] Martin Husovec, 'Certification of Out-of-Court Dispute Settlement Bodies under the Digital Services Act' [2023] SSRN Electronic Journal <https://www.ssrn.com/abstract=4501726> accessed 16 August 2023.
[90] Aleksandra Kuczerawy, 'Social Media Councils under the DSA: A Path to Individual Error Correction at Scale? By Aleksandra Kuczerawy :: SSRN' <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4497877> accessed 3 August 2023.
[91] Husovec (n 74).

complainant is located should assess the complaint and submit it to the DSC of the Member State that is overseeing the regulation of the platform.

3.4.7 Lex specialis accountability obligations and end-user rights

Indeed, the approach towards more accountability is not only in the DSA but is also a visible trend in content moderation lex specialis. It should be noted that the following section is not intended to provide an in-depth exploration and comprehensive analysis of the accountability obligations. The purpose of this section is to demonstrate the common trend towards more accountability in the various instruments and in their latest revisions.

3.4.7.1 AVMSD

The Audiovisual Media Service Directive (AVMSD) is the cornerstone of audiovisual media regulation in the EU. The EC proposed a revision of the old Audiovisual Media Services Directive in May 2016 to include a new approach to online platforms disseminating audiovisual content. The revision of the AVMSD was concluded in November 2018, and Member States had until September 2020 to transpose the AVMSD into their national legislation.[92] The revised AVMSD introduced major changes with regard to the broadening of the scope to include video-sharing platforms (VSPs). The new Directive includes more accountability obligations, including the selection below:

- Video sharing platforms (VSP) that host content for which they have no editorial responsibility, such as videos posted by users (Art. 28b) must :
    - Take measures with regards to harmful content in the areas of terrorist and racist subject matters, child pornography and hate speech to the general public.
    - Adopt appropriate measures to deal with the different type of content such as flagging systems, effective complaint systems, age, verification, and transparency obligations.
    - Enhance transparency on their services: VSP must establish and operate transparent and user-friendly mechanisms for users of a video-sharing platform to report or flag the content, easy-to-use and effective procedures for the handling and resolution of users' complaints. Kuklis explains that this provision serves both the user who complained and the user against whose content the complaint was directed.[93] This provision is thus potentially a useful tool in protecting the rights of users, especially those who are actively uploading content. The user whose content is taken down by a platform provider usually receives only a generic explanation of the reasons why it happened.

[92] 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, OJ L 303, 28.11.2018, p. 69–92.

[93] Lubos Kuklis, 'Media Regulation at a Distance: Video-Sharing Platforms in AVMS Directive and the Future of Content Regulation'.

- Media service providers must put in place a rating systems or symbols that indicate the presence of violence, nudity, and adult language about the content offered.
- Accountability: The AVMSD obliges Member States to ensure that users have easy and direct access at any time to information about the media service provider (Art. 5).

3.4.7.2 Terrorist content online

The Counter-Terrorism Directive (2017) obliges Member States to take the necessary measures to ensure the prompt removal of, or with appropriate safeguards block access to, online content constituting a public provocation to commit a terrorist offence. Member States implemented these obligations via two main types of measures: notice-and-takedown measures and criminal measures.[94] However, as the directive addresses Member States, the measures did not target directly the platforms, which are in the best position to address the topic. The Directive was complemented by a voluntary system for tackling terrorism online based on guidelines and recommendations, but it was deemed insufficient to deal with terrorist content online.[95]

This is the reason why, already in September 2018, the EC submitted a proposal for a Regulation on preventing the dissemination of terrorist content (TERREG).[96] Similarly to the DSA shift, a regulatory shift is operated in the fight against terrorist content by choosing a regulation as an instrument. This instrument imposes directly on Hosting Services Providers duties of care and proactive measures to remove terrorist content including by deploying automated detection tools. The Regulation was published at the Official Journal (OJ) in May 2021, entered into force on 6 June 2021, and applies as of 7 June 2022.[97]

There are some similarities between the DSA and TERREG when it comes to national orders and due diligence obligations.

- The new regulation foresees that a competent authority of a Member State can issue a removal order requiring hosting service providers to remove terrorist content or to disable access to such content in the whole European Union. The competent authority is in this case as well not necessarily a judicial body.

---

[94] Directorate-General for Internal Policies of the Union (European Parliament) and others, *Online Platforms' Moderation of Illegal Content Online: Laws, Practices and Options for Reform* (Publications Office of the European Union 2020) <https://data.europa.eu/doi/10.2861/831734> accessed 23 January 2023.

[95] Flavia Giglio, 'The New Regulation on Addressing the Dissemination of Terrorist Content Online: A Missed Opportunity to Balance Counter-Terrorism and Fundamental Rights?' (*CITIP blog*, 14 September 2021) <https://www.law.kuleuven.be/citip/blog/the-new-regulation-on-addressing-the-dissemination-of-terrorist-content-online/> accessed 3 February 2023.

[96] Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online A contribution from the European Commission to the Leaders' meeting in Salzburg on 19-20 September 2018 2018 [COM/2018/640 final].

[97] Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online (Text with EEA relevance) 2021 (OJ L)

- There is a fast lane, similar to the fast lane for the trusted flaggers under the DSA, with the expeditious assessment of the Content referrals sent from either a national competent authority or an EU body such as Europol.
- The matter is so serious, that the window time for action upon receipt of an order requires terrorist content to be removed within one hour from the receipt of the removal order.
- A competent authority can also decide that a certain Hosting Service Provider (HSP) is particularly exposed to terrorist content. They can oblige the hosting service to adopt measures to prevent the dissemination of terrorist content on its services.

The new duties of care can be considered as a push towards a more proactive role from hosting providers and hence towards algorithmic moderation with its risks for freedom of expression. There is fear among civil society that "important public debates that help society understand, monitor and deal with terrorism get categorised incorrectly and could just be censored"[98].

3.4.7.3 CSAM

The Child Sexual Abuse Directive (2011)[99] requires member states to ensure that intermediaries promptly remove websites that contain or distribute child pornography.[100] However, since the expansion of the notion of electronic communication services in the European Electronic Communication Code (EECC)[101], e-privacy now includes interpersonal communication services in its scopes such as WhatsApp, Instagram, and Messenger. The detection and reporting of CSAM by these services have clashed with the protection granted under the e-Privacy Directive.[102]

To fix this issue, the EC has adopted an interim CSAM regulation in July 2021 which will last until August 2024.[103] This is the reason why the 2022 proposal has been released and is now in the EU policy-making pipeline to make sure to reach an agreement before the end of the interim text.

---

[98] EDRI, 'European Parliament Confirms New Online Censorship Powers' (*European Digital Rights (EDRi)*, 29 April 2021) <https://edri.org/our-work/european-parliament-confirms-new-online-censorship-powers/> accessed 7 September 2023.

[99] 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.

[100] Buiten (n 65).

[101] Directive (EU) 2018/1972 of the European Parliament and of the Council of 11 December 2018 establishing the European Electronic Communications Code (Recast) (Text with EEA relevance)Text with EEA relevance 2018.

[102] Charlotte Somers, 'The Proposed CSAM Regulation: Trampling Privacy in the Fight against Child Sexual Abuse?' (*CITIP blog*, 3 January 2023) <https://www.law.kuleuven.be/citip/blog/the-proposed-csam-regulation-trampling-privacy-in-the-fight-against-child-sexual-abuse/> accessed 20 January 2023.

[103] Regulation (EU) 2021/1232 of the European Parliament and of the Council of 14 July 2021 on a temporary derogation from certain provisions of Directive 2002/58/EC as regards the use of technologies by providers of number-independent interpersonal communications services for the processing of personal and other data for the purpose of combating online child sexual abuse (Text with EEA relevance) 2021 (OJ L).

The proposal for a new CSAM regulation aims to replace the current system based on voluntary detection and reporting by companies. The proposal suggests imposing qualified obligations on providers of hosting services, interpersonal communication services, and other services concerning the detection, reporting, removing, and blocking of known and new online child sexual abuse material, as well as solicitation of children. New accountability obligations are imposed on the providers of services, including the collaboration with Europol, law enforcement authorities and the data protection authorities.

However well intended, the proposal has been subject to criticisms from scholars, EU co-legislators, and civil society. The private companies would have a very broad margin of appreciation to ensure compliance with the obligations contained in the AVMSD, leaving the door open to potential abuse and legal uncertainty on how to balance the rights at stake in each case.

The obligations are such that the providers will have to rely on detection technologies in interpersonal communication services. The recourse to AI and technology is here as well pushed by the legislation in light of the strict conditions and obligations to react swiftly to the illegal content. Critics pointed that technology false positives could expose innocent people to the screen of inspectors.[104] In general, they believe that the proposal should operate a better balance between freedom of expression, right to privacy, and data protection and the fight against CSAM in order to meet societal needs such as having secure and private communication channels.

3.4.7.4 Copyright

The 2019/790/EC Directive on Copyright in the Digital Single Market (CDSM) requires that online content-sharing service providers use their best efforts to obtain licences for content posted by their users and holds them liable for copyright or related rights infringement if they do not remove the material after notification and prevent its reappearance. Art. 17 of the CDSM is of particular interest for the push towards more accountability obligations for providers. The provision imposes direct liability on online content-sharing service providers (OCSSPs) for copyright- protected works or other protected subject-matter uploaded by users.[105]

In addition, there are notice-and-takedown and notice-and-stay-down. Under the notice-and-takedown process, platforms are required to make 'best efforts' to takedown copyright infringing user generated content (UGC) upon receiving notice from rightsholders. When it comes to the notice-and-stay-down process, platforms are obliged to make the 'best efforts' to prevent future uploads of works that have been taken down after notice from rightsholders or had previously been flagged as infringing. By fear of facing liability, the Directive incentivizes platforms to use ex-ante upload filters to remove or block content before it even has a chance to be made available to the public. However, as already mentioned the use automated filtering

[104] EDRi, 'News from Ireland Question Effectiveness and Lawfulness of Online Scanning for Tackling Child Sexual Abuse: Lessons for the EU' (European Digital Rights (EDRi), 15 October 2022) <https://edri.org/our-work/breaking-irish-story-shows-that-eus-csam-proposal-can-never-work/> accessed 28 September 2023.
[105] Art. 17(1), the Copyright in the Digital Single Market Directive.

technologies comes with their limitations, over-blocking, freedom of expression concerns and so forth. Relying on filters could only be justified in the case of removing or blocking manifestly infringing content.

3.5 Policy recommendations on the shift from platform liability to responsibility

3.5.1 General considerations

The question of liability and accountability has only grown similarly to the growth of the use of the Internet, social media and all the other services offered by intermediary services providers. The topic is delicate as it is at the intersection of different interests, fundamental rights and vision about innovation.

Originally, advocates for a free and open internet argued that minimal regulation promotes free speech, fosters innovation, and allows for a diversity of voices and ideas to flourish online. At the start of the "Internet boom", having excessive regulation would likely hinder the development of new online services and platforms. However, as the internet and its massive use became a fertile ground for various harmful activities, such as hate speech, misinformation, cyberbullying, harassment, and illegal content, the need for some level of regulation was deemed necessary by policy makers to protect individuals from these harms and to maintain a safe and civil online environment. In view-of-these "impacts of social media on society, no surprise that in the past years, instruments to counteract these possible undesirable effects have been considered around the globe"[106].

In addition, regulating intermediary services providers impacts the basic tension between the platform autonomy and platform accountability. According to the freedom to conduct business, the platform autonomy advocates argued for intermediary services providers to be free to set their own content moderation policies and decide on what type of content is allowed on their platforms, also to prevent any government outreach on their services. But on the other side, massive disinformation or hate speech campaigns and the spread of CSAM online really pushed for hard regulation. As the platforms became indeed a major facilitator of online communication, it was deemed that they should bear a degree of responsibility for the content they host and external oversight is necessary to hold them accountable. Platforms now do more than just transmit content as they moderate, rank, prioritise content, impacting the public discourse and the prominence of certain content online.[107]

[106] Judit Bayer, Lorna Woods and Bernd Holznagel, 'Introduction to Perspectives on Platform Regulation' (Nomos Verlagsgesellschaft mbH & Co KG 2021) <https://www.nomos-elibrary.de/10.5771/9783748929789-9/introduction?page=1> accessed 8 September 2023.
[107] Judit Bayer, 'Rights and Duties of Online Platforms', *Perspectives on Platform Regulation* (Nomos Verlagsgesellschaft mbH & Co KG 2021) <https://www.nomos-elibrary.de/10.5771/9783748929789-23/rights-and-duties-of-online-platforms?page=0> accessed 8 September 2023.

In order to balance these different visions and tensions, the ECD was a first EU step on intermediary liability regulation. The section above already details the main components of this regime and how the shift from liability exemption towards accountability obligations operated.

Despite the improvements, the liability and content moderation regimes are certainly not perfect and can be improved. Already in 2010; the US scholar T. Gillepsie was saying that platform companies successfully positioned themselves in a sweet spot that allows them to seek protection for facilitating user expression, yet also seek limited liability for what users say.[108] Some speak about regulatory deficit when it comes to platform governance, "where there is a well-founded societal desire for governmental response to a social need, as yet unsatisfied"[109]. The authors underline the challenge that it is to address the demand for regulation that is consistent with international human rights norms and laws.

M.C. Buiten et al. also underlined the negative effect of higher requirements to disproportionately burden small entrants.[110] This would only benefit the solid and big market players creating bigger barriers to entry to the digital single market and reinforcing the powers of those already detaining it. The same authors were putting forward that there was a need to split the responsibility among all actors diffusing illegal material.

Now that the DSA has been adopted, some also wonder whether the DSA's regulatory design will allow this Regulation to be as future proof as the ECD was.[111] They claim that the DSA focuses on the most pressing issues which were targeting the VLOP's to the detriment of more technological neutral framework.

3.5.2 Remaining challenges

Platforms' rule-making power

The DSA leaves the "rule-making" power in the hands of service providers. In short, they are free to decide what they allow on their platforms, but are limited in how they interpret and enforce their own (content moderation) towards individuals.[112] With the DSA proposal, in 2020, some scholars[113] were warning policy makers to be cautious "in entrusting—and burdening— private parties with such an extensive 'policing' role". The rule-making power of platforms is only vaguely limited. Whereas a non-binding Recital 47 considers that providers of intermediary services should consider the rights and legitimate interests of the recipients of the service,

[108] Tarleton Gillespie, 'The Politics of "Platforms"' (1 May 2010) <https://papers.ssrn.com/abstract=1601487> accessed 8 September 2023.
[109] Monroe E. Price, 'Foreword for Perspectives on Platform Regulation' (Nomos Verlagsgesellschaft mbH & Co KG 2021) <https://www.nomos-elibrary.de/10.5771/9783748929789/perspectives-on-platform-regulation?page=1> accessed 8 September 2023.
[110] Miriam C Buiten, Alexandre de Streel and and Martin Peitz, 'Rethinking Liability Rules for Online Hosting Platforms' (2020) 28 International Journal of Law and Information Technology 139.
[111] Schwemer (n 67).
[112] Husovec (n 74).
[113] Miriam C Buiten, Alexandre de Streel and and Martin Peitz, 'Rethinking Liability Rules for Online Hosting Platforms' (2020) 28 International Journal of Law and Information Technology 139.

including fundamental rights, when "designing, applying and enforcing" the restrictions, Art. 14(4) only refers to "applying and enforcing". How the rules are designed is left in the hands of service providers.

Husovec provides two examples of what such rule-setting power could mean in practice. First, online platforms will be still allowed to carry on a VIP no-moderation list of users. If such a practice is clearly described in platforms' terms and conditions, it will still be allowed.[114] It should be mentioned that such practices have already sparked some controversies. In 2021, the Wall Street Journal revealed, that Facebook was using a program is known as "cross check" or "XCheck," that whitelists millions of VIP Facebook and Instagram users from the company's standard content moderation practices.[115] Second, online platforms will also be allowed to introduce a so-called "Pay2Say" service: for a small monthly fee, one will be allowed to say anything on the platform (if it is not illegal). In other words, none of the restrictions provided for in the terms and conditions will apply – such as prohibition of spreading mis- and dis-information, nudity, harmful content etc.

To conclude, one can wonder whether platforms' power in the rule-making sphere should remain unconstrained. Of course, platforms' freedom to conduct business remains their fundamental right, and any restriction of this right would have to be carefully assessed. Nonetheless, this point should be subject to further discussion about platforms' accountability.

Responsibility for the dissemination of news and media content

Under the DSA, online platforms can delete legal news content they do not want to host on their platform if it does not comply with their terms and conditions (so-called "*private ordering*"[116]). The so-called 'media exemption' in the Digital Services Act (amendments 511 and 513 to article 12(1) and recital 38) caused heated discussions around an important issue: once the media content is already subject to editorial responsibility, should it be subject to additional scrutiny by online platforms? Some media organisations pointed out that the 'non-interference principle' would ensure that platforms do not undermine the independence of media publishers by arbitrarily deleting legal, public-interest content.[117]

---

[114] Martin Husovec, 'How Does the EU's Digital Services Act Regulate Content Moderation? And Will it Work?' A conversation with Martin Husovec, moderated by Daphne Keller of the Program on Platform Regulation. Available at: https://www.youtube.com/watch?v=np05wM3h2mc&ab_channel=StanfordCyberPolicyCenter.

[115] 'The Facebook Files' *Wall Street Journal* (1 October 2021) <https://www.wsj.com/articles/the-facebook-files-11631713039> accessed 18 January 2023.

[116] Luca Belli and Jamila Venturini, 'Private Ordering and the Rise of Terms of Service as Cyber-Regulation' (2016) 5 Internet Policy Review <https://policyreview.info/node/441> accessed 18 March 2022.

[117] EBU, 'Protecting Media Content Online: A Decisive Moment' <https://www.ebu.ch/news/2021/10/protecting-media-content-online-a-decisive-moment> accessed 19 September 2023.

As already mentioned, the idea of imposing additional accountability and due diligence obligations on online platforms is a major force behind the DSA. Under the adopted text, there is no special protection or any obligation of a prior notice to media organizations. Media organizations and journalists can invoke the same procedural rights which apply to all users of online platforms (see above): under Article 14(4) DSA online platforms are required to have "*due regard*" to the fundamental rights of all users, including "*freedom and pluralism of the media*".

The DSA leaves the impression that the assessment of what is a "freedom and pluralism of the media" is an easy task. However, there are doubts whether online platforms are qualified to make such an assessment. The assessment of whether a particular news outlet is a media service provider benefiting from media freedom is much more complex. Recent Twitter's saga labelling US public broadcaster NPR as "state-affiliated media"[118] – just as Russia's RT – or BBC as a "government-funded" body[119], demonstrates the challenge of classifying media. Moreover, the DSA considers the relationship between media and platforms by obliging the VLOPS to identify the risks stemming from their service, to carry out risk assessments and to put in place mitigation measures. One of the possible systematic risks is any actual or foreseeable negative effects on the exercise of fundamental rights, including freedom and pluralism of the media, such as, potentially, wrongful takedowns of content or suspensions of journalists' accounts. Again, such assessment is far more complex and requires a media expertise. Clarifying the DSA platforms' responsibilities with regard to content moderation towards media organisations, i.e. how platforms should deal with lawful content under the editorial control and legal liability of the publisher (or broadcaster) has been one of the recommendations of D6.2.

Individual remedies vs systemic remedies

Another paradigm shift which comes with the DSA is moving from individual remedies to systemic remedies for platforms' content moderation decisions. Previously, an individual remedy mechanism which allows users to contest platforms' decisions, has been the main way in which users could contest platforms' behaviour. Notably, the enforcement mechanism of the GDPR has put a lot of power (and responsibility) onto users' proactive approach in enforcing their rights. However, the structural deficits of individual remedy consist, in their limited scope: individual remedy does not address the algorithmic infrastructure, amplification of content, the design of recommender systems or platforms' impacts on civic discourse, electoral processes or public health and security.[120] "Individual remedy only empowers users of platforms, although

[118] Taylor Hatmaker, 'Twitter Singles NPR out with Misleading State-Backed Media Label' (TechCrunch, 5 April 2023) <https://techcrunch.com/2023/04/05/twitter-singles-npr-out-with-misleading-state-backed-media-label/> accessed 28 September 2023.

[119] Paul Glynn, 'Elon Musk: Twitter Owner Changes BBC Account's "government Funded" Label' BBC News (12 April 2023) <https://www.bbc.com/news/entertainment-arts-65248554> accessed 28 September 2023.

[120] Niklas Eder, 'Making Systemic Risk Assessments Work: How the DSA Creates a Virtuous Loop to Address the Societal Harms of Content Moderation by Niklas Eder :: SSRN' <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4491365> accessed 12 September 2023.

non-users equally suffer the detrimental consequences of content moderation."[121] As the US scholar, Evelyn Douek argues, "because it focuses on the merits of individual speech decisions, it leads to endless and irresolvable arguments (…) whether [platforms' rules] have been correctly and impartially applied in particular cases, and whether platforms have afforded due process to individual users".[122] She argued for a toolset for content moderation which expands beyond individual error correction.[123]

Systemic risk assessments in the DSA offer such a new regulatory approach. The individual remedy options are complemented with a systemic approach which promises to capture platforms' influence on societal interests. Articles 34 and 35 of the DSA establish obligations for platforms to assess and mitigate these systemic risks. Although it remains to be seen how the systemic risk assessment mechanism will work in practice, it may be concluded that "a systemic approach to regulate content moderation could be a huge leap forward, allowing to hold platforms accountable for their societal impact."[124] In addition, the DSA brings also a welcomed novelty with the system trusted flaggers also brings the possibility for NGOs, hotlines or rightsholders representative to act on behalf of societal or a group of individuals interests and see their notices treated with priority.

Beside this systemic and collective aspect, the DSA for the rest, contains rights for individuals. In this scenario, where part of the enforcement lies on the shoulders of users and individuals, massive communication and education campaigns must accompany the legislations. There must be investments in these initiatives using the modern arsenal to raise awareness and incentivize to individuals to enforce their rights.

Enforcement challenges

For the new accountability rules to function and deliver their promises, the enforcement of the DSA must be on point. The legislation sets up an ambitious enforcement structure composed of the various EU Member States DSC (national regulatory authorities), a European Board for Digital Services and the EC's enforcement team. To solve the shortcomings of the GDPR enforcement structure with its country of origin principle, the DSA provides a great deal of oversight power to the EC for the VLOPS and VLOSES. The new powers to investigate and sanction platforms

For these purposes, numerous cooperation agreements must be signed before the 24 February 2024 between the EC and national regulatory authorities, EU agencies and competence

[121] ibid.
[122] Evelyn Douek, 'Content Moderation as Systems Thinking - Harvard Law Review' <https://harvardlawreview.org/print/vol-136/content-moderation-as-systems-thinking/> accessed 19 September 2023.
[123] ibid.
[124] Niklas Eder, 'Making Systemic Risk Assessments Work: How the DSA Creates a Virtuous Loop to Address the Societal Harms of Content Moderation', SSRN' (n 106).

centres.[125] However, G. Miller reports that only few Members States are far ahead in their DSC designation and selection process.[126]

The DSA being a regulation, will be directly applicable and has to be uniformly applied. However, this last point risks to be a challenge given the variety of stakeholders involved in the enforcement. The cooperation and harmonious application will be key for a successful enforcement. To achieve this, adequate financial and human resources at the EU and national level and an efficient information exchange system are the basis for a proper functioning.[127] In addition, the delegated acts complementing and clarifying some DSA need to be "be developed quickly and in a transparent, inclusive process."[128] Another aspect still unclear is the overlap with other content moderation legislation and the cooperation with enforcement structure of other accountability rules contained in lex specialis.

3.5.3 Recommendations

Table 1 below provides policy recommendations to tackle the above-mentioned challenges.

Table 1 Policy recommendations on the shift from platform liability to responsibility

Policy recommendations on the shift from platform liability to responsibility
• **Reflect on a question of legitimacy** Making online intermediaries responsible for identifying and addressing societal harms constitutes an indirect delegation of public functions. Although not problematic *per se*, the questions of online platforms' legitimacy to make the normative judgements about the availability and accessibility of online content should be part of a societal discussion. • **Balance stricter regulation for providers with fundamental rights requirements and protection** As platforms acquire an increasing influence on the enforcing and balancing of users' fundamental rights, they should be subject to substantial checks and oversights.

[125] Théophane Hartmann, 'Challenges Mount for European Commission's New DSA Enforcement Team' www.euractiv.com (30 August 2023) <https://www.euractiv.com/section/law-enforcement/news/challenges-mount-for-european-commissions-new-dsa-enforcement-team/> accessed 2 October 2023.
[126] Gabby Miller, 'Who's Afraid of the DSA?' (Tech Policy Press, 25 August 2023) <https://techpolicy.press/whos-afraid-of-the-dsa/> accessed 2 October 2023.
[127] Julian Jaursch, 'Barriers to Strong DSA Enforcement – and How to Overcome Them' (Tech Policy Press, 5 December 2022) <https://techpolicy.press/barriers-to-strong-dsa-enforcement-and-how-to-overcome-them/> accessed 2 October 2023.
[128] Julian Jaursch, 'Barriers to Strong DSA Enforcement – and How to Overcome Them' (Tech Policy Press, 5 December 2022) <https://techpolicy.press/barriers-to-strong-dsa-enforcement-and-how-to-overcome-them/> accessed 2 October 2023.

- **Ensure that the accountability rules are not undermining the small and not-for-profit providers access to the digital single market**
  Although no one-fit-all approach in the DSA is undoubtedly beneficial, some uncertainty remains about digital services which have purely non-profit character. The enforcement of the DSA should hinder not-for-profit scientific and educational repositories, digital archives, and libraries access to the digital single market.

- **Investigate the risks and impacts of new, less known and alternative online platforms and services**
  Much focus is currently put on preventing the harms to fundamental rights coming from the VLOPS (e.g. TikTok, X, Meta etc.). However, it is equally important to make sure that other, perhaps less known online phenomena do not fall through the cracks of DSA enforcement.

- **Support individual remedy mechanisms**
  Whereas the DSA provides users with new rights, the remedy mechanisms would only work if users are made aware of these options. An effort is needed to encourage users to take an active role in the private enforcement of the DSA.

- **Provide methodologies for the DSA systemic risks assessments to further develop a collective and societal dimension to platform's accountability.**
  There is a need for unified methodology on how to identify and address online harms, especially those faced by different (marginalized) groups and how these harms intersect.

- **Clarify platforms' responsibilities with regard to content moderation towards media organisations' content**
  The DSA does not directly tackle the relationship between the VLOPS and media organizations' content. The European Media Freedom Act proposal[129] however, proposes the provide media service providers specials procedural rights vis-à-vis VLOPS content moderation practices. The harmonious relationship between the two procedures needs to be ensured.

- **Ensure a harmonious DSA enforcement**
  The DSA puts an obligation on member states to establish Digital Services Coordinators which are independent from political and business and have certain investigatory and enforcement powers as well as adequate resources. It is crucial to make sure there is a well-functioning collaboration between the national coordinators. The European Board for Digital Services, which is made up of all Digital

[129] European Commission, Proposal for a Regulation of the European Parliament and of the Council establishing common framework for media services in the internal market (European Media Freedom Act) and amending Directive 2010/13/EU 2022 [COM(2022) 457 final].

Services Coordinators, should also play a key role in providing recommendation to the Commission.

4 The AI-on-Demand platform (AIoD)

4.1 The AI-on-Demand Platform

The AI-on-Demand Platform (AIoD)[130] is part of the European AI strategy and "will act as a community resource for the research community, facilitating experimentation, knowledge sharing and the development of state-of-the-art solutions and technologies", as the European Commission lines out in the objective of the currently ongoing project AI4Europe[131].

The development of the platform, the technical operation and the moderation of content contributions is entrusted to a series of successively implemented projects funded by the EC. The platform is currently in a development state in which existing subsystems are already in productive use, but further components are still under development. Over time the platform "will mature to add AI assets and tools to be used by the broader community to upskill and transfer knowledge to innovation sectors. It will supply new services and a marketplace for non-experts, so that they can experiment and deploy AI solutions in their own workplaces".[132]

The content that is accessible on the AI-on-Demand Platform is the result of a joint effort of several European research projects. To date, these include AI4EU, multiple supporting projects funded under H2020 calls ICT-48 (AI4Media[133], ELISE[134], HumanE-AI-Net[135], TAILOR[136]) and ICT-49 (AIPlan4EU[137], AI4Copernicus[138], BonsAPPs[139], DIH4AI[140], I-NERGY[141], StairwAI[142]) and AI4Europe (see Figure 7).

Today's vision for the further development of the AI-on-Demand Platform is described in the call for proposals DIGITAL-2022-CLOUD-AI-B-03[143]. In the future, the AIoD shall include among others:

- A visible and secure catalogue of trustworthy AI resources made in Europe.
- A one-stop shop to access AI tools for European industry and public administrations.
- A reference and trusted marketplace for trustworthy AI resources.
- A platform embedding a quality stamp recognised and accepted by the European industry and public administrations.

[130] https://www.ai4europe.eu/, last visited 24/05/2023
[131] https://cordis.europa.eu/project/id/101070000, last visited 24/05/2023
[132] https://www.ai4europe.eu/about/ai-on-demand-platform, last visited 24/04/2023
[133] https://www.ai4media.eu/, last visited 24/05/2023
[134] https://www.elise-ai.eu/, last visited 24/05/2023
[135] https://www.humane-ai.eu/, last visited 24/05/2023
[136] https://tailor-network.eu/, last visited 24/05/2023
[137] https://www.aiplan4eu-project.eu/, last visited 24/05/2023
[138] https://ai4copernicus-project.eu/, last visited 24/05/2023
[139] https://bonsapps.eu/, last visited 24/05/2023
[140] https://www.dih4ai.eu/, last visited 24/05/2023
[141] https://i-nergy.eu/, last visited 24/05/2023
[142] https://stairwai.nws.cs.unibo.it/, last visited 24/05/2023
[143] https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/digital/wp-call/2022/call-fiche_digital-2022-cloud-ai-b-03_en.pdf, last visited 24/05/2023

- Interconnections to computing resources, data spaces and testing and experimentation facilities.
- A coordinated governance mechanism, integrating all projects taking part on the AI-on-Demand ecosystem (H2020, Horizon Europe, Digital Europe).
- And more.

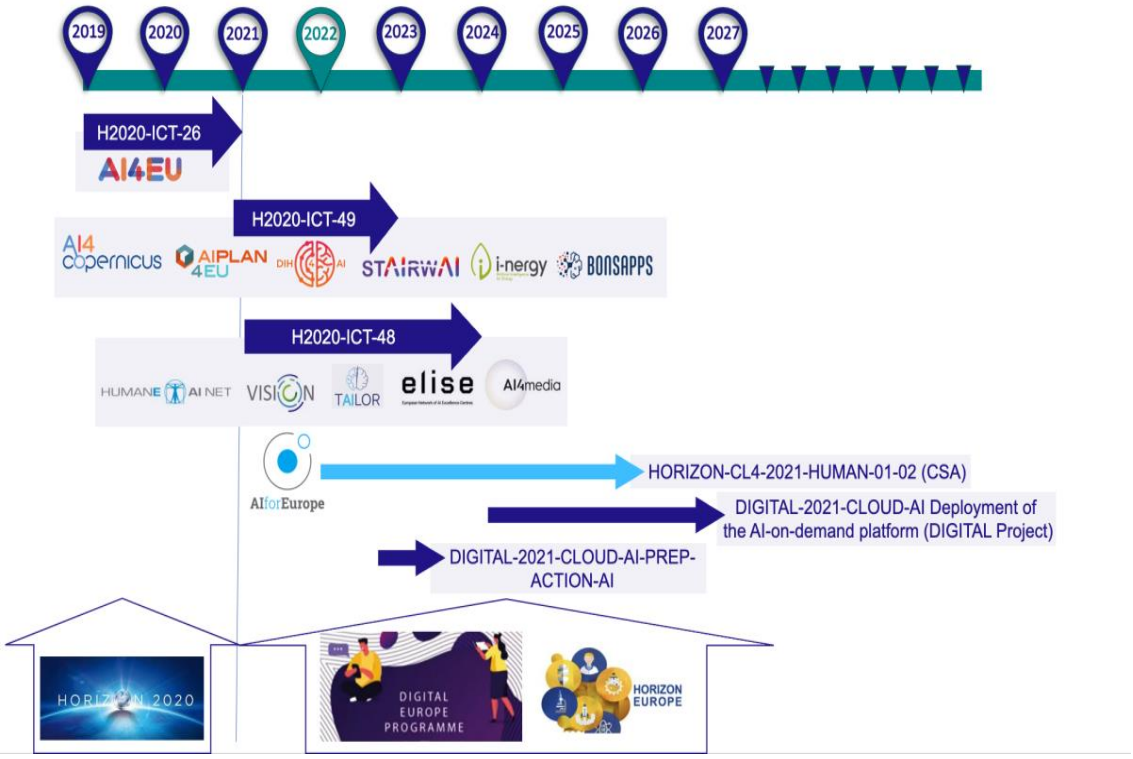


Figure 7 EU projects supporting the development of the AIoD.[144]

4.2 History of the platform

"The first iteration of the AI-on-Demand platform was developed under the leadership of the AI4EU project (ICT26-2019). This initial version of the platform put in place the basic functionalities of the platform and initiated many community efforts to further support the necessary administration and technical development. The initial investment by the EC was followed by the funding of six supporting projects (ICT49-2020), each of which is committed to building new innovative services for the platform. In addition, each of Europe's Network of AI Excellence Centres (ICT-48-2020) committed to further enrich the platform with the outputs of their activities".[145]

With the completion of the AI4EU project at the end of 2021, the AIoD included two publicly available subsystems:

1. AIoD Website, including the AI Assets Catalog and other content, and

[144] https://www.ai4europe.eu/about/ai-on-demand-platform, last visited 24/04/2023
[145] https://www.ai4europe.eu/about/ai-on-demand-platform, last visited 24/04/2023

2. AI4EU Experiments for the publication of AI solutions and the creation of AI pipelines.

In January 2022, after the end of AI4EU project, the technical operation and the continuation of the organisation of the platform was transferred to the AI4Media project[146]. In July 2022 the AI4Europe project[147] started and took over these tasks. AI4Europe is "further advancing the technical development of the platform and developing a supporting community to create additional value for the AI Research community".[148]

Simultaneously with the further development of the platform in the AI4Europe project "in early 2022, under the DIGITIAL EUROPE programme, there was a dedicated call focused on putting in place a Preparatory Action that would be responsible for analysing the features and requirements of the platform as they applied to innovation, industry, and the public sector".[149] The Preparatory Action project Pre-PAI[150] will set the foundation for future developments of the platform, complementing the work in the AI4Europe project and continuing this work after its end.

In addition to the projects funded by the European Commission, the KI.NRW competence platform[151] has contributed the AI-Lab Playground[152] as a new subsystem to the AI-on-Demand Platform. The AI-Lab Playground is an execution environment for Docker containers designed for a very simple deployment of pipelines created in AI4EU experiments. It allows experimenting with AI technologies without the need to have experience with the technical deployment of Docker containers.

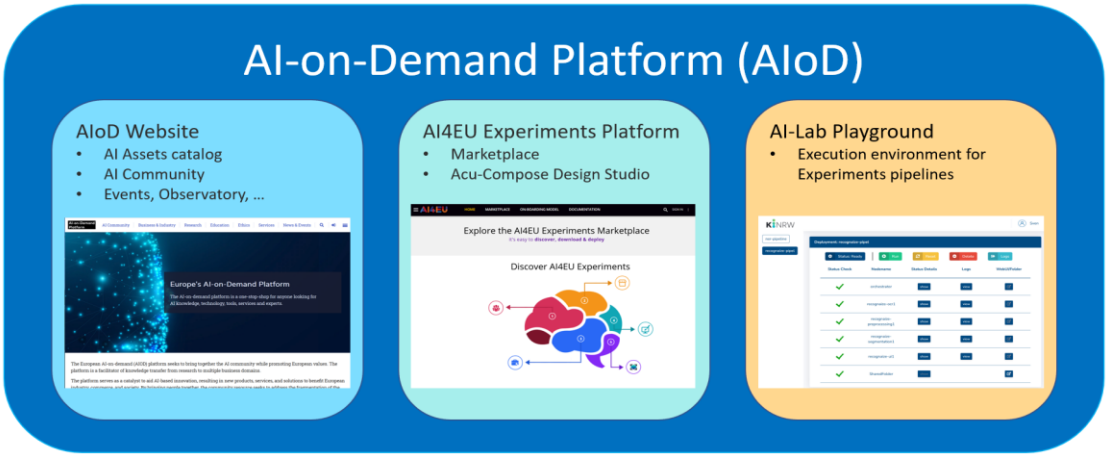


Figure 8 AIoD subsystems as of April 2023

[146] see AI4Media deliverable D7.2 "*Extended version of the integration result with the AI-On-Demand-Platform*"
[147] https://cordis.europa.eu/project/id/101070000, last visited 24/05/2023
[148] https://www.ai4europe.eu/about/ai-on-demand-platform, last visited 24/04/2023
[149] https://www.ai4europe.eu/about/ai-on-demand-platform, last visited 24/04/2023
[150] https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/how-to-participate/org-details/999999999/project/101083674/program/43152860/details, last visited 28/04/2023
[151] https://www.ki.nrw/en/, last visited 24/04/2023
[152] https://www.ai-lab.nrw/#/home

4.3 Content on the AI-on-Demand Platform

The AI-on-Demand Platform's technical subsystems make content of various kinds and from various creators available to the public. In order to be able to analyse the importance of content moderation for the different types of content, this section takes a closer look at the different types of content and the existing content contribution, publication and moderation processes.

4.3.1 AIoD website

The content available on the AIoD website can be roughly divided into two categories.

On the one hand, content was created by the editorial team of the AIoD website. Over time different projects have been responsible for the provision of an editorial team for the website. Accordingly this content comes from different people belonging to different organisations. On the other hand, content is available that was supplied via the Contribution Gateway[153], which provides a form for the submission of content for registered users. This content can be one of the following types (see Figure 9):

- Organisation
- AI Asset
- News
- Event
- Project
- Case Study
- Open Call
- Education content
- Research Bundle

The Contribution Gateway can be used by anyone to contribute content to the platform, including members of the editorial team, collaborators in funded projects whose obligation is to publish project results on the platform, and anyone else.

Content of both categories is seamlessly integrated on the website. Therefore, it is not easy to tell whether a post was created by the editorial team or by a member of the community.

For some types of content, the date of the last update is visible: AI Asset, News, Education content and Research Bundle.

For News content, the contributor is mentioned with their login ID.

[153] https://www.ai4europe.eu/contribute

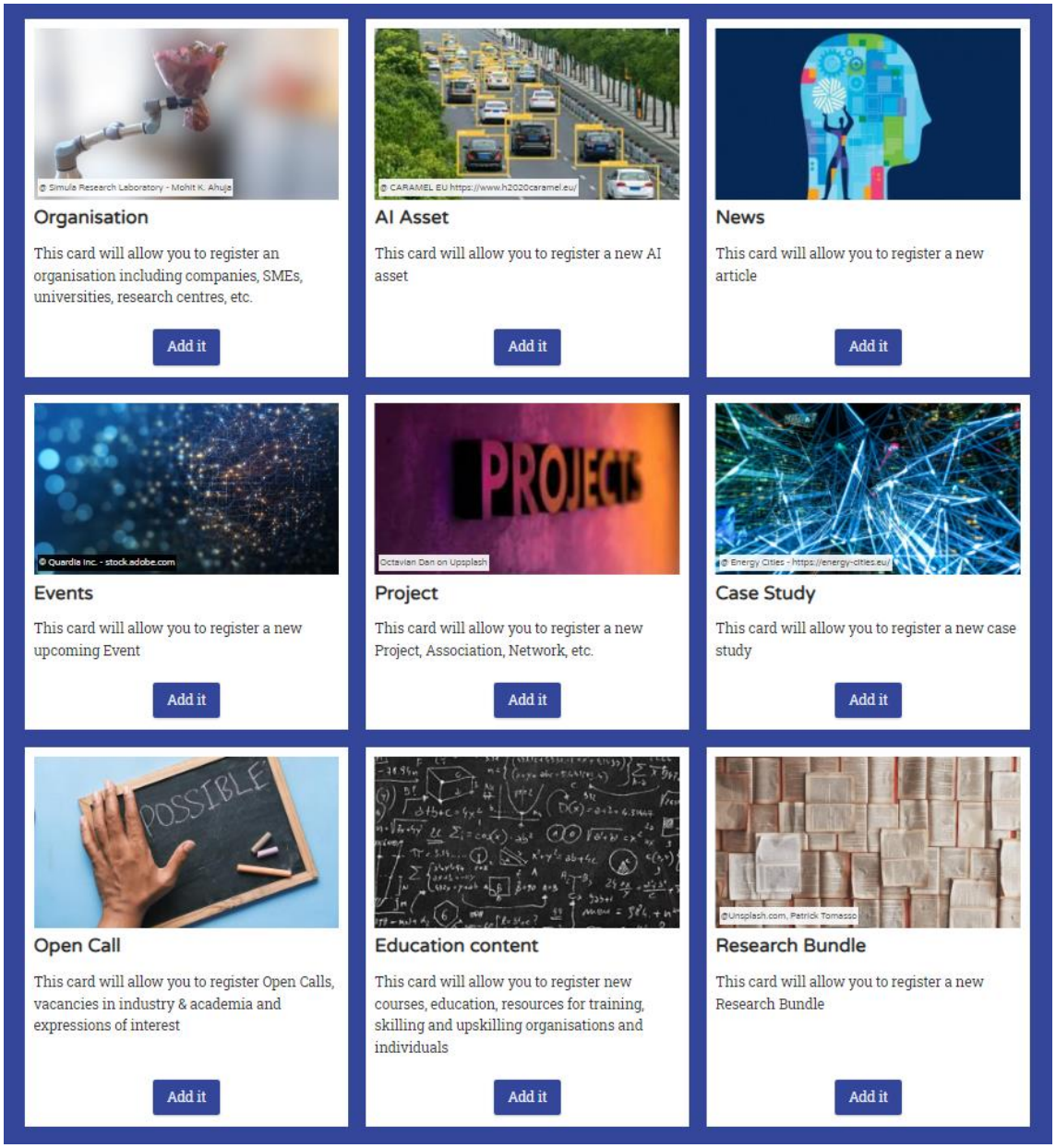


Figure 9 Content type selection in the AIoD Contribution Gateway (as of April 2023)

4.3.1.1 Publication and content moderation workflow on the AIoD website

The AI-on-Demand Platform supports a simple workflow for content moderation for all contributions provided using the Contribution Gateway.

As a prerequisite for the publication of content on the AIoD via the Contribution Gateway only a valid EU Login[154] is required for authorization. The Contribution Gateway provides the contributing user with a form to publish different content depending on the type selected:

[154] https://ecas.ec.europa.eu/cas/about.html, last visited 24/05/2023

- text,
- images,
- weblinks, and
- downloadable files.

Once submitted content is saved, it is visible to the user as a draft in their personal "My content" page of the platform (Figure 10). The platform also creates a URL alias derived from the content title (clean URL). This URL is very handy when referencing content. The contributing user can edit and save draft contributions and finally submit the draft for review.

The review is conducted by a member of the AIoD governance team, which is located at one or more beneficiaries of the project in charge of the operation of the platform. The reviewer can publish the contributed content, reject the publication, or return the contribution to the contributor for revision[155]. Once contributed content has been published, it can no longer be edited or deleted by the contributing user.

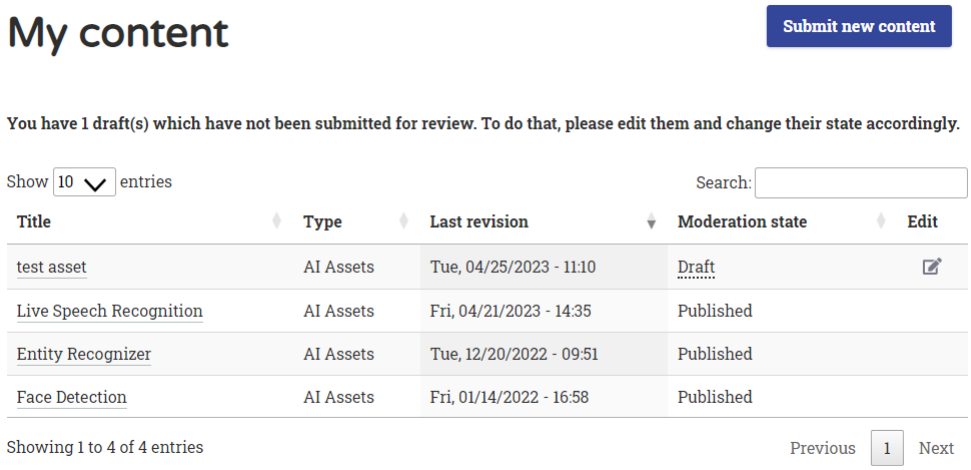


Figure 10 "My content" page of the AIoD (as of April 2023)

4.3.2 AI4EU Experiments Marketplace

The AI4EU Experiments Marketplace is a publicly accessible repository for AI technologies.[156] Content items in the marketplace are called "models". A model consists of a few descriptive information elements, depending on the type of the model.

Models for Docker containers

[155] Roles and publication process see https://aiondemand.readthedocs.io/en/latest/03Editorial.html#roles-and-publication-process, last visited 16/10/2023

[156] https://aiexp.ai4europe.eu/#/marketPlace#marketplaceTemplate, last visited 24/05/2023

[156] https://aiexp.ai4europe.eu/#/marketPlace#marketplaceTemplate, last visited 24/05/2023

[156] https://aiexp.ai4europe.eu/#/marketPlace#marketplaceTemplate, last visited 24/05/2023

[156] https://aiexp.ai4europe.eu/#/marketPlace#marketplaceTemplate, last visited 24/05/2023

The AI4EU Experiments Marketplace makes AI technologies accessible in a uniform format. Therefore, all contributed AI technologies are wrapped into Docker containers. A Docker container is a standardized package of files containing software and possibly data that can be executed by a Docker engine. A Docker container can also contain a dataset that can be used for training or inference of AI technologies.

Models for Docker containers are composed of a description (text, image) of the AI technology included and associated artifacts like license information, link to the Docker container, a signature defining the interface of the model in Protocol Buffers (Protobuf) syntax, author information and other optional documents (Figure 11-Figure 13). All artifacts of a model are provided for download (Figure 14).

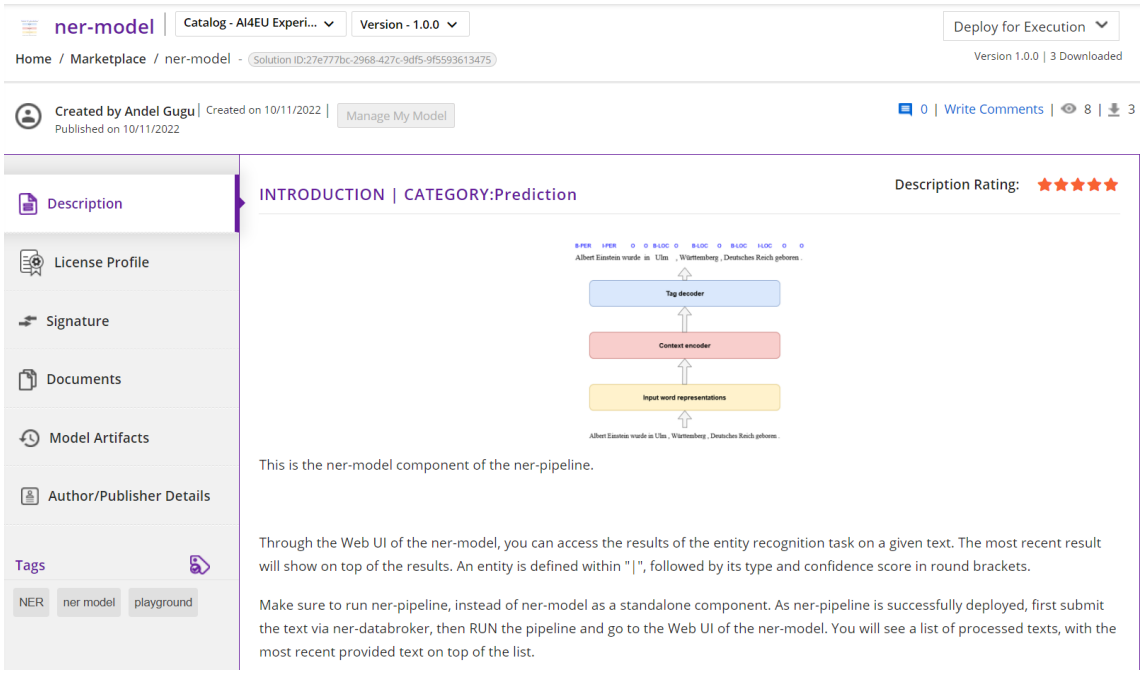


Figure 11 Description of a model in AI4EU Experiments Marketplace

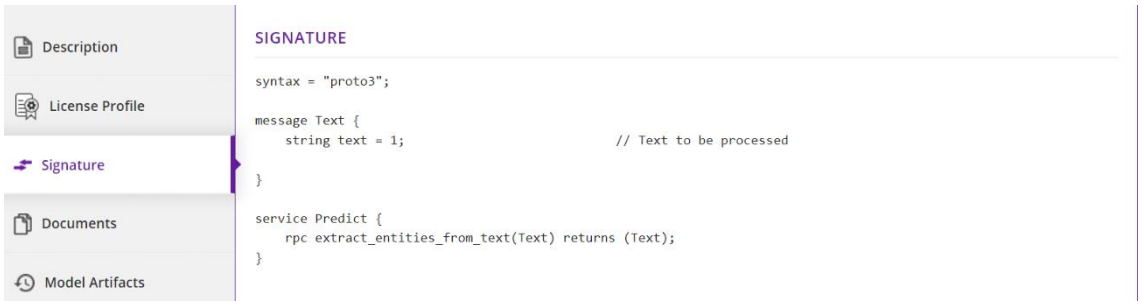


Figure 12 Protobuf signature of a model in AI4EU Experiments Marketplace

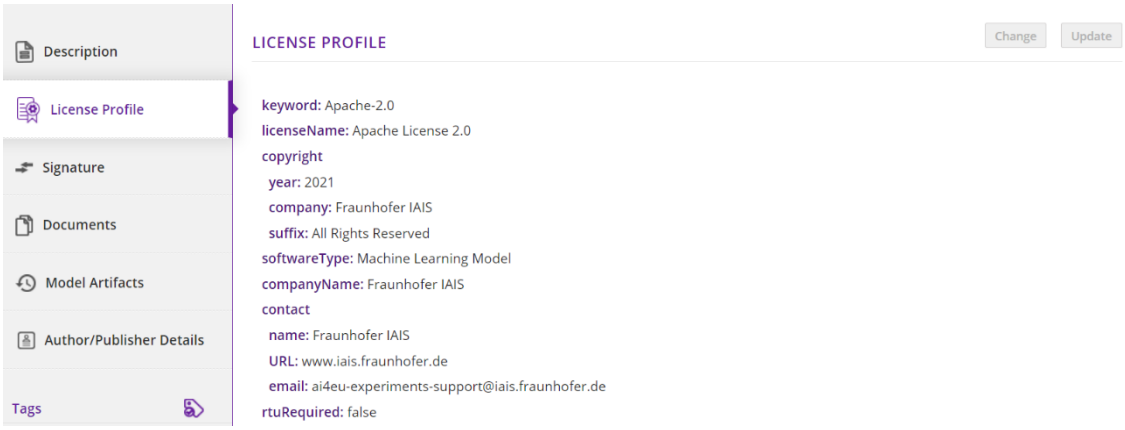


Figure 13 License information of a model in AI4EU Experiments Marketplace

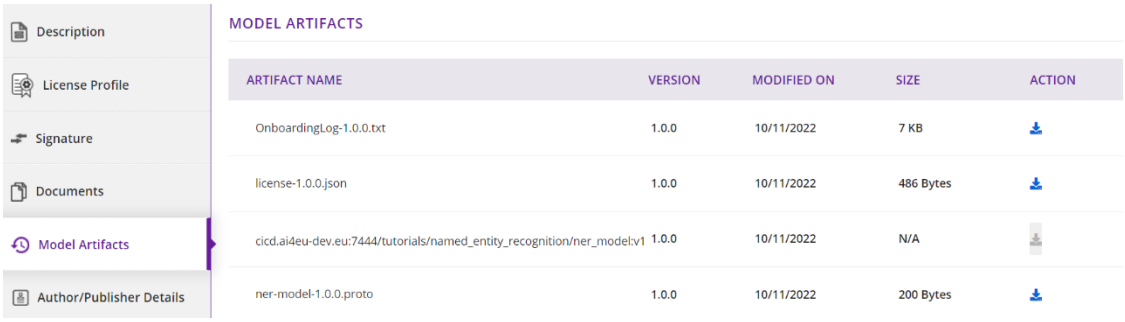


Figure 14 Artifacts of a model in AI4EU Experiments Marketplace

It is important to note that the actual Docker container is not stored in the marketplace. The marketplace only provides metadata and access information for these containers. The Docker containers themselves can be either publicly accessible or protected. The decision not to host any data or software in the marketplace itself, but only to offer links to external resources, was made consciously when designing the platform. This enables both freely accessible and commercially offered data and solutions to be managed equally in the marketplace without having to worry about access or intellectual property protection.

Models for pipelines

Multiple models can be joined together to build AI pipelines. Therefore, the AI4EU Experiments platform provides the Design Studio application[157], which is an interactive application that runs in the web browser. The Design Studio verifies the connectivity of models according to the provided Protobuf signatures. In this way, every user can compose AI pipelines and save and edit them in a personal workspace in the Design Studio application. Completed pipelines can be deployed for execution, either to a local execution environment or to the AI-Lab Playground.

Models for pipelines can also be published in the marketplace. Published models for pipelines are composed of a description (text, image) of the pipeline and associated artifacts like license information, author information and other optional documents (Figure 15). In addition, two text

[157] https://aiexp.ai4europe.eu/#/designStudio (only visible for logged in users), last visited 16/10/2023

files in JSON syntax are provided, which carry all information about the composition of the required models and their connections using the Protobuf interfaces of the models.

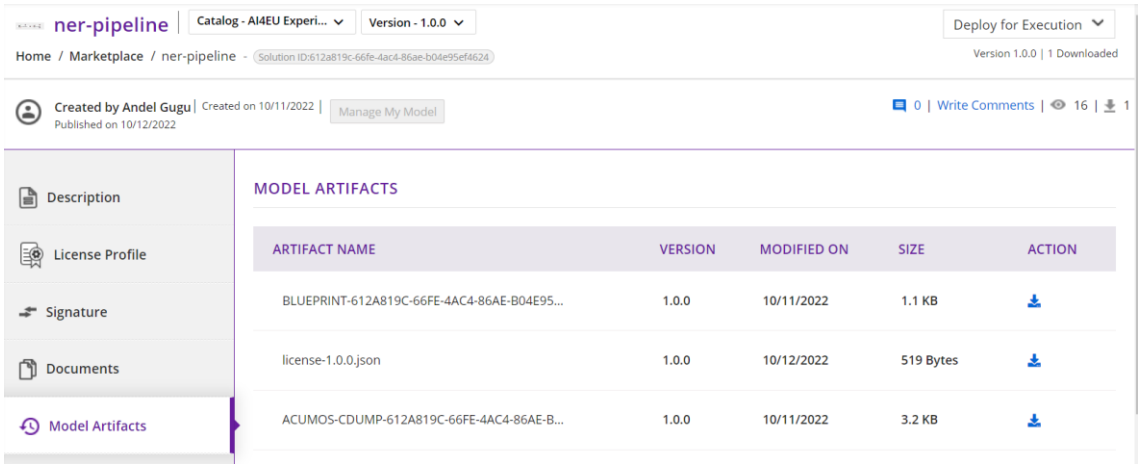


Figure 15 Artifacts of a pipeline in AI4EU Experiments Marketplace

4.3.2.1 Publishing models to AI4EU Experiments Marketplace

Models can be published by anybody to the AI4EU Experiments Marketplace. As a prerequisite for the publication only a valid EU Login is required for authorization.

Model publishing follows a simple workflow using a web form that ensures all required information is provided by contributing users. Each contribution is reviewed by a member of the AI4EU Experiments Marketplace governance team[158], which is located at one or more beneficiaries of the project in charge of the operation of the marketplace. The review looks at the information to be published in the marketplace, there is no review of the referenced Docker containers.

4.3.3 AI-Lab Playground

The AI-Lab Playground is a web-based platform for conducting experiments using AI technology. It allows the deployment and execution of models from the AI4EU Experiments Marketplace. Every user is assigned a personal workspace, and a valid EU Login is required for authorization. "The AI-Lab Playground is operated by KI.NRW and Fraunhofer IAIS as the German node of the international AI4EU initiative, which was initiated by the European Commission and is supported by the Ministry of Economics in North Rhine-Westphalia" [159].

The personal dashboard of the AI-Lab Playground allows access to the deployed Docker containers (Figure 16). In addition to read-only status information and logs, the dashboard

[158] Email address for support is provided on the website: https://aiexp.ai4europe.eu/#/home, last visited 16/10/2023
[159] https://www.ai-lab.nrw/#/home, last visited 16/10/2023

allows you to invoke a web-based user interface, if the containers provide one. This can be used to enable any interaction with the container provided by the developer.

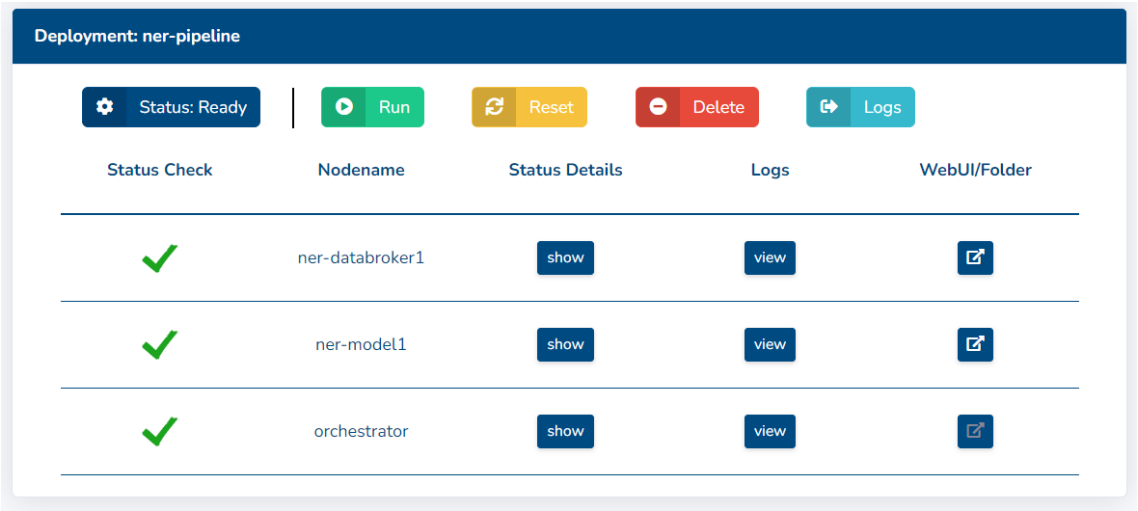


Figure 16 Dashboard for a pipeline in AI-Lab Playground

The AI-Lab Playground also feature its own marketplace and the Design Studio application to build AI pipelines. These features are provided by the same source code base like the comparable features of AI4EU Experiments. In terms of assessing responsibility and liability, it is also important to consider that by means of the linked Docker containers, untested software is downloaded from external sources and executed on the servers that are the responsibility of beneficiaries of the respective EU funded project. The docker container itself may contain any kind of data and access any resource on the internet.

The AI-Lab Playground takes no responsibility for the data processed on the platform in the experiments. It prohibits processing of personal data or other real data in the terms of use[160] and privacy information.[161] The platform may only be used with anonymized test data. Use of the platform for productive purposes is not permitted.

4.4 Developments in AI4Europe

The AI4Europe project is developing an infrastructure to make contributed content available to different systems. In January 2023, a preview of the technical architecture for the systems to be built under the AI4Europe project was presented to the community.[162] The architecture foresees a central meta-data store (database) for the storage of the information about AI assets and all crucial metadata. The central database is complemented by an interface (AIoD REST API) that will allow access to the stored information for various AIoD client applications.

[160] https://www.ai-lab.nrw/index.html#/termsCondition, last visited 28/04/2023
[161] https://www.ai-lab.nrw/index.html#/dataProtection, last visited 28/04/2023
[162] Presentation: https://github.com/ai4eu/Technical-Governance-Board/blob/main/Meeting-2023-01-27/meta-data-service.pdf, last visited 27/04/2023

"The AIoD REST API will allow any kind of service to interact with the AIoD portal to discover, retrieve, and share AI resources. It forms the connection between user-facing components, such as the AIoD website or Python Client API, and the backend. The metadata for datasets, models and other resources can be accessed, added, updated and deleted through this API."[163]

One of the first client applications could be a new AIoD website that uses the AIoD REST API to access community-contributed information from the central meta-data store, such as AI assets, rather than from a database integrated into a CMS. The plan is to incorporate the content contributed by the community into the new meta-data store.

[163] (https://github.com/aiondemand/AIOD-rest-api/, last visited 25/04/2023)

5 Guidelines on platform liability and accountability for the AI-on-Demand platform

The AI-on-Demand platform is in its development stage and during exchanges and meetings with FhG and other stakeholders involved in the platform, the aim of KUL was to better understand the functioning of the AI-on-Demand platform to fine tune the analysis and guidance elaborated on the subject covered by this deliverable. Guidelines and information on the relevant liability and accountability considerations attached to AI-on-Demand services were provided. The aim of these guidelines was to provide valuable insights and guidance, to help the stakeholders and different partners involved in the platform to navigate the complex landscape of platform liability and accountability.

An analysis was conducted to identify the overarching challenges related to the management of the AI-on-Demand platform. The KUL's team further researched and analysed the three different components of the platform namely the AIoD Website, the AI4EU Experiments Platform and the AI-lab Playground. They investigated into the different type of content hosted on the platform, who could publish content and what was the content moderation processes. The analysis continued by assessing in which category of intermediary service would the different components of the AI-on-demand platform fall. Based on this assessment the liability and accountability considerations associated to this status were presented. Guidelines and recommendations on these aspects and the legal documentation necessary were also delivered.

6 Conclusions

The growth of online platforms over the past 20 years showed that the first generation of regulatory solutions, based on liability exemptions for content posted by the users, has reached its limits. Online platforms have come under more pressure to do both less and more to monitor their platforms. In 2018, Facebook was accused of failing to adequately address calls for violence against Muslim minorities in Myanmar[164], while Facebook and Twitter were criticised after permanently suspending Donald Trump's account following his comments about violence at the US Capitol in 2021[165]. As pointed out by Buiten, "these examples illustrate that the debate revolving around platform responsibility reaches beyond the question of platforms' liability in curbing illegal content".[166]

The second wave of digital regulation, focuses on the *risks* posed by the service, and not solely on liability. Although the old rules should still apply, the new regulatory expectations – as to the accountability for the design of the services - overlay the liability regime. It may therefore be said, that the shift of the regulatory approach, is not as much a shift "*from*" liability "*to*" responsibility, but rather adding responsibility "*in addition to*" liability.

These new responsibility/accountability rules prompt providers to rely on technology, including automated tools. It's worth noting that while technology and automated decision-making can be valuable tools in enhancing accountability and content moderation, they come with inherent limitations. These limitations include a lack of contextual understanding, quality assessment, diversity recognition, and inclusivity considerations[167]. Moreover, they raise socio-political challenges and require a careful balance with fundamental rights. As we navigate this evolving regulatory landscape, it is crucial to use technology with an awareness of these limitations and to bolster it with appropriate safeguards.

[164] Barbara Ortutay, 'Amnesty Report Finds Facebook Amplified Hate Ahead of Rohingya Massacre in Myanmar' PBS NewsHour (29 September 2022) <https://www.pbs.org/newshour/world/amnesty-report-finds-facebook-amplified-hate-ahead-of-rohingya-massacre-in-myanmar> accessed 29 September 2023; Tom Miles, 'U.N. Investigators Cite Facebook Role in Myanmar Crisis' Reuters (12 March 2018) <https://www.reuters.com/article/us-myanmar-rohingya-facebook-idUSKCN1GO2PN> accessed 17 March 2023.

[165] Kate Conger, Mike Isaac and Sheera Frenkel, 'Twitter and Facebook Lock Trump's Accounts After Violence on Capitol Hill' The New York Times (6 January 2021) <https://www.nytimes.com/2021/01/06/technology/capitol-twitter-facebook-trump.html> accessed 29 September 2023.

[166] Buiten (n 65).

[167] More information on these limitations in our previous report : Noémie Krack, Lidia Dutkiewicz and Emine Ozge Yildirim, 'AI4Media Report on Policy for Content Moderation (D6.2)' <https://www.ai4media.eu/reports/report-on-policy-for-content-moderation-d6-2/> accessed 29 September 2023.

7 References

AFP, 'Facebook to French Court: Nude Painting Did Not Prompt Account's Deletion' *The Guardian* (1 February 2018) <https://www.theguardian.com/technology/2018/feb/01/facebook-nude-painting-gustave-courbet> accessed 5 September 2023

Allen A, 'An Intersectional Lens on Online Gender Based Violence and the Digital Services Act' [2022] Verfassungsblog <https://verfassungsblog.de/dsa-intersectional/> accessed 10 August 2023

Allen A and Stockem O, 'A Series on the EU Digital Services Act: Tackling Illegal Content Online' (*Center for Democracy and Technology*, 2 August 2022) <https://cdt.org/insights/a-series-on-the-eu-digital-services-act-tackling-illegal-content-online/> accessed 6 September 2023

Angelopoulos C and Senftleben M, 'An Endless Odyssey? Content Moderation Without General Content Monitoring Obligations' (22 June 2021) <https://papers.ssrn.com/abstract=3871916> accessed 1 February 2023

Bayer J, 'Rights and Duties of Online Platforms', *Perspectives on Platform Regulation* (Nomos Verlagsgesellschaft mbH & Co KG 2021) <https://www.nomos-elibrary.de/10.5771/9783748929789-23/rights-and-duties-of-online-platforms?page=0> accessed 8 September 2023

Bayer J, Woods L and Holznagel B, 'Introduction to Perspectives on Platform Regulation' (Nomos Verlagsgesellschaft mbH & Co KG 2021) <https://www.nomos-elibrary.de/10.5771/9783748929789-9/introduction?page=1> accessed 8 September 2023

Belli L and Venturini J, 'Private Ordering and the Rise of Terms of Service as Cyber-Regulation' (2016) 5 Internet Policy Review <https://policyreview.info/node/441> accessed 18 March 2022

Buiten MC, 'The Digital Services Act: From Intermediary Liability to Platform Regulation' (2022) 12 JIPITEC <http://www.jipitec.eu/issues/jipitec-12-5-2021/5491>

Buiten MC, de Streel A and Peitz and M, 'Rethinking Liability Rules for Online Hosting Platforms' (2020) 28 International Journal of Law and Information Technology 139

Cascone S, 'Instagram Censors Period Blood, Enrages Artist' (*Artnet News*, 31 March 2015) <https://news.artnet.com/art-world/instagram-slammed-for-censoring-period-283123> accessed 5 September 2023

Castellaro SB and Penfrat J, 'The DSA fails to reign in the most harmful digital platform businesses – but it is still useful' [2022] Verfassungsblog <https://verfassungsblog.de/dsa-fails/> accessed 10 August 2023

Conger K, Isaac M and Frenkel S, 'Twitter and Facebook Lock Trump's Accounts After Violence on Capitol Hill' The New York Times (6 January 2021)

<https://www.nytimes.com/2021/01/06/technology/capitol-twitter-facebook-trump.html> accessed 29 September 2023

Damian Tambini, *Media Freedom* (Polity, 2021)

Del Valle Salinas OD and others, 'A User-Centered Approach To Content Moderation' (Hertie School Centre for Digital Governance 2020) <https://digitalservicesact.eu/wp-content/uploads/2020/06/DSA-Policy-Brief-Content-Moderation.pdf>

De Streel A and others, *Study on Potential Policy Measures to Promote the Uptake and Use of AI in Belgium in Specific Economic Domains* (FPS Economy 2022)

De Streel A and Husovec M, 'The E-Commerce Directive as the Cornerstone of the Internal Market' [2020] SSRN Electronic Journal <https://www.ssrn.com/abstract=3637961> accessed 24 January 2023

Directorate-General for Communication (European Commission) and Leyen U von der, *Political Guidelines for the next European Commission 2019-2024 ; Opening Statement in the European Parliament Plenary Session 16 July 2019 ; Speech in the European Parliament Plenary Session 27 November 2019* (Publications Office of the European Union 2020) <https://data.europa.eu/doi/10.2775/101756> accessed 4 September 2023

Directorate-General for Internal Policies of the Union (European Parliament) and others, *Online Platforms' Moderation of Illegal Content Online: Laws, Practices and Options for Reform* (Publications Office of the European Union 2020) <https://data.europa.eu/doi/10.2861/831734> accessed 23 January 2023

Douek E, 'Content Moderation as Systems Thinking' (*Harvard Law Review*, 12 December 2022) <https://harvardlawreview.org/print/vol-136/content-moderation-as-systems-thinking/> accessed 19 September 2023

Eder N, 'Making Systemic Risk Assessments Work: How the DSA Creates a Virtuous Loop to Address the Societal Harms of Content Moderation' (26 June 2023) <https://papers.ssrn.com/abstract=4491365> accessed 19 September 2023

EDRi, 'News from Ireland Question Effectiveness and Lawfulness of Online Scanning for Tackling Child Sexual Abuse: Lessons for the EU' (European Digital Rights (EDRi), 15 October 2022) <https://edri.org/our-work/breaking-irish-story-shows-that-eus-csam-proposal-can-never-work/> accessed 28 September 2023

EDRI, 'European Parliament Confirms New Online Censorship Powers' (*European Digital Rights (EDRi)*, 29 April 2021) <https://edri.org/our-work/european-parliament-confirms-new-online-censorship-powers/> accessed 7 September 2023

Eecke P, 'Online Service Providers and Liability: A Plea for a Balanced Approach' (2011) 48 Common Market Law Review 1455

Etlinger S, 'The Next Wave of Platform Governance' (*Centre for International Governance Innovation*) <https://www.cigionline.org/articles/next-wave-platform-governance/> accessed 8 August 2023

European Broadcasting Union (EBU), 'Protecting Media Content Online: A Decisive Moment' (29 October 2021) <https://www.ebu.ch/news/2021/10/protecting-media-content-online-a-decisive-moment> accessed 19 September 2023

European Commission, 'Summary of the Results of the Public Consultation on the Future of Electronic Commerce in the Internal Market and the Implementation of the Directive on Electronic Commerce (2000/31/EC)' (2011) <https://ec.europa.eu/information_society/newsroom/image/document/2017-4/consultation_summary_report_en_2010_42070.pdf>

European Commission, 'Archive - E-Commerce Directive - What Happened before and since Its Adoption | Shaping Europe's Digital Future' (24 March 2017) <https://digital-strategy.ec.europa.eu/en/library/archive-e-commerce-directive-what-happened-and-its-adoption> accessed 4 September 2023

European Commission, 'Summary Report on the Open Public Consultation on the Digital Services Act Package | Shaping Europe's Digital Future' (15 December 2020) <https://digital-strategy.ec.europa.eu/en/library/summary-report-open-public-consultation-digital-services-act-package> accessed 4 September 2023

European Parliament. Directorate General for Parliamentary Research Services., *Liability of Online Platforms.* (Publications Office 2021) <https://data.europa.eu/doi/10.2861/619924> accessed 24 January 2023

European Parliamentary Research Service, 'Reform of the EU Liability Regime for Online Intermediaries.' (2020) <https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/649404/EPRS_IDA(2020)649404_EN.pdf>

Giglio F, 'The New Regulation on Addressing the Dissemination of Terrorist Content Online: A Missed Opportunity to Balance Counter-Terrorism and Fundamental Rights?' (*CITIP blog*, 14 September 2021) <https://www.law.kuleuven.be/citip/blog/the-new-regulation-on-addressing-the-dissemination-of-terrorist-content-online/> accessed 3 February 2023

Gillespie T, 'The Politics of "Platforms"' (1 May 2010) <https://papers.ssrn.com/abstract=1601487> accessed 8 September 2023

Glynn P, 'Elon Musk: Twitter Owner Changes BBC Account's "government Funded" Label' BBC News (12 April 2023) <https://www.bbc.com/news/entertainment-arts-65248554> accessed 28 September 2023

Grotius H, *De Jure Belli Ac Pacis. Libri Tres*

Hartmann T, 'Challenges Mount for European Commission's New DSA Enforcement Team' www.euractiv.com (30 August 2023) <https://www.euractiv.com/section/law-

enforcement/news/challenges-mount-for-european-commissions-new-dsa-enforcement-team/> accessed 2 October 2023

Hatmaker T, 'Twitter Singles NPR out with Misleading State-Backed Media Label' (TechCrunch, 5 April 2023) <https://techcrunch.com/2023/04/05/twitter-singles-npr-out-with-misleading-state-backed-media-label/> accessed 28 September 2023

Husovec M, 'Certification of Out-of-Court Dispute Settlement Bodies under the Digital Services Act' [2023] SSRN Electronic Journal <https://www.ssrn.com/abstract=4501726> accessed 16 August 2023

Husovec M and Roche Laguna I, 'Digital Services Act: A Short Primer' [2022] SSRN Electronic Journal <https://www.ssrn.com/abstract=4153796> accessed 14 October 2022

Jaursch J, 'Barriers to Strong DSA Enforcement – and How to Overcome Them' (Tech Policy Press, 5 December 2022) <https://techpolicy.press/barriers-to-strong-dsa-enforcement-and-how-to-overcome-them/> accessed 2 October 2023

Krack N and others, 'AI in the Belgian Media Landscape. When Fundamental Risks Meet Regulatory Complexities', *Artificial Intelligence and the Law*, vol 13 (Second Revised Edition, Jan De Bruyne and Cedric Vanleenhove (eds), Intersentia 2023) <https://intersentia.com/en/artificial-intelligence-and-the-law-2nd-edition.html>

Krack N, Dutkiewicz L and Yildirim EO, 'AI4Media Report on Policy for Content Moderation (D6.2)' <https://www.ai4media.eu/reports/report-on-policy-for-content-moderation-d6-2/> accessed 29 September 2023

Kuczerawy A, 'Intermediary Liability & Freedom of Expression: Recent Developments in the EU Notice & Action Initiative' (2015) 31 Computer Law & Security Review 46

Kuczerawy A, *Intermediary Liability and Freedom of Expression in the EU: From Concepts to Safeguards* (Intersentia 2018)

Kuczerawy A, 'Remedying Overremoval: The Three-Tiered Approach of the DSA' [2022] Verfassungsblog <https://verfassungsblog.de/remedying-overremoval/> accessed 8 March 2023

Kuczerawy A, 'Social Media Councils under the DSA: A Path to Individual Error Correction at Scale?' (31 January 2023) <https://papers.ssrn.com/abstract=4497877> accessed 19 September 2023

Kuczerawy A, 'The Power of Positive Thinking': Intermediary Liability and the Effective Enjoyment of the Right to Freedom of Expression' (2017) 8 JIPITEC <https://www.jipitec.eu/issues/jipitec-8-3-2017/4623>

Kuczerawy A, 'The Good Samaritan that wasn't: voluntary monitoring under the (draft) Digital Services Act' [2021] Verfassungsblog <https://verfassungsblog.de/good-samaritan-dsa/> accessed 5 September 2023

Kuklis L, 'Media Regulation at a Distance: Video-Sharing Platforms in AVMS Directive and the Future of Content Regulation'

Lindroos A, 'Addressing Norm Conflicts in a Fragmented Legal System: The Doctrine of Lex Specialis' (2005) 74 Nordic Journal of International Law 27

Llansó EJ and others, 'Artificial Intelligence, Content Moderation, and Freedom of Expression' (2020)

Mantelero A, 'Fundamental rights impact assessments in the DSA' [2022] Verfassungsblog <https://verfassungsblog.de/dsa-impact-assessment/> accessed 19 September 2023

Miles T, 'U.N. Investigators Cite Facebook Role in Myanmar Crisis' Reuters (12 March 2018) <https://www.reuters.com/article/us-myanmar-rohingya-facebook-idUSKCN1GO2PN> accessed 17 March 2023

Miller G, 'Who's Afraid of the DSA?' (Tech Policy Press, 25 August 2023) <https://techpolicy.press/whos-afraid-of-the-dsa/> accessed 2 October 2023

Monroe E. Price, 'Foreword for Perspectives on Platform Regulation' (Nomos Verlagsgesellschaft mbH & Co KG 2021) <https://www.nomos-elibrary.de/10.5771/9783748929789/perspectives-on-platform-regulation?page=1> accessed 8 September 2023

OECD, *An Introduction to Online Platforms and Their Role in the Digital Transformation* (Organisation for Economic Co-operation and Development 2019) <https://www.oecd-ilibrary.org/science-and-technology/an-introduction-to-online-platforms-and-their-role-in-the-digital-transformation_53e5f593-en> accessed 8 August 2023

Ortolani P, 'If You Build It, They Will Come' (*Verfassungsblog*, 7 November 2022) <https://verfassungsblog.de/dsa-build-it/> accessed 10 August 2023

Ortutay B, 'Amnesty Report Finds Facebook Amplified Hate Ahead of Rohingya Massacre in Myanmar' PBS NewsHour (29 September 2022) <https://www.pbs.org/newshour/world/amnesty-report-finds-facebook-amplified-hate-ahead-of-rohingya-massacre-in-myanmar> accessed 29 September 2023

Oruç TH, 'The Prohibition of General Monitoring Obligation for Video-Sharing Platforms under Article 15 of the E-Commerce Directive in Light of Recent Developments: Is It Still Necessary to Maintain It?' (2022) 13 JIPITEC <http://www.jipitec.eu/issues/jipitec-13-3-2022/5555>

Schwemer SF, 'Digital Services Act: A Reform of the e-Commerce Directive and Much More' (10 October 2022) <https://papers.ssrn.com/abstract=4213014> accessed 1 December 2022

Somers C, 'The Proposed CSAM Regulation: Trampling Privacy in the Fight against Child Sexual Abuse?' (*CITIP blog*, 3 January 2023) <https://www.law.kuleuven.be/citip/blog/the-proposed-csam-regulation-trampling-privacy-in-the-fight-against-child-sexual-abuse/> accessed 20 January 2023

Sotto P, 'French Court Issues Mixed Ruling in Facebook Nudity Case' (*Courthouse News Service*, 15 March 2018) <https://www.courthousenews.com/french-court-issues-mixed-ruling-in-facebook-nudity-case/> accessed 5 September 2023

Sullivan D, 'Unpacking "Systemic Risk" Under the EU's Digital Service Act' (*Tech Policy Press*, 19 July 2023) <https://techpolicy.press/unpacking-systemic-risk-under-the-eus-digital-service-act/> accessed 9 August 2023

'The Facebook Files' *Wall Street Journal* (1 October 2021) <https://www.wsj.com/articles/the-facebook-files-11631713039> accessed 18 January 2023

Wilman F, 'The Digital Services Act (DSA) - An Overview' [2022] SSRN Electronic Journal <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4304586>

Zing T, 'Why Instagram Censored This Image of an Artist on Her Period' (*DAZED*, 27 March 2015) <https://www.dazeddigital.com/artsandculture/article/24258/1/why-instagram-censored-this-image-of-an-artist-on-her-period> accessed 5 September 2023

Zingales N, 'The DSA as a paradigm shift for online intermediaries' due diligence: Hail to meta-regulation' [2022] Verfassungsblog <https://verfassungsblog.de/dsa-meta-regulation/> accessed 2 October 2023

Case-law

Eva Glawischnig-Piesczek v Facebook Ireland Limited [2019] ECJ Case C-18/18

Google France SARL and Google Inc v Louis Vuitton Malletier SA (C-236/08), Google France SARL v Viaticum SA and Luteciel SARL and Google France SARL v Centre national de recherche en relations humaines (CNRRH) SARL and Others) [2010] ECJ Joined cases C-236/08 to C-238/08

Frank Peterson v Google LLC, YouTube LLC, YouTube Inc, Google Germany GmbH and Elsevier Inc v Cyando AG [2021], ECJ, Joined cases C-682/18 & C-683/18

Republic of Poland v European Parliament and Council of the European Union [2022] ECJ Case C-401/19

Stichting Brein [2017], ECJ, C 527/15, EU:C:2017:300

Stichting Brein v Ziggo BV and XS4All Internet BV [2017], ECJ, C-610/15, ECLI:EU:C:2017:456.

EU policy and legal instruments

European Commission, Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises (Text with EEA relevance) (notified under document number C(2003) 1422) 2003

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committe of the Regions: on a Digital Single Market Strategy for Europe 2015

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committe of the Regions, Shaping Europe's digital future 2020

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

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 [Regulation (EU) 2016/679] 88

Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online A contribution from the European Commission to the Leaders' meeting in Salzburg on 19-20 September 2018 2018 [COM/2018/640 final]

Directive (EU) 2018/1972 of the European Parliament and of the Council of 11 December 2018 establishing the European Electronic Communications Code (Recast) (Text with EEA relevance)Text with EEA relevance 2018

Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online (Text with EEA relevance) 2021 (OJ L)

Regulation (EU) 2021/1232 of the European Parliament and of the Council of 14 July 2021 on a temporary derogation from certain provisions of Directive 2002/58/EC as regards the use of technologies by providers of number-independent interpersonal communications services for the processing of personal and other data for the purpose of combating online child sexual abuse (Text with EEA relevance) 2021 (OJ L)

European Commission, Proposal for a Regulation of the European Parliament and of the Council establishing common framework for media services in the internal market (European Media Freedom Act) and amending Directive 2010/13/EU 2022 [COM(2022) 457 final]

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) (Text with EEA relevance) 2022 (OJ L)

Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) (Text with EEA relevance) 2022 (OJ L).