Computer Law & Security Review 51 (2023) 105893

Contents lists available at ScienceDirect

Computer Law & Security Review: The International Journal of Technology Law and Practice

journal homepage: www.elsevier.com/locate/clsr

Making the private public: Regulating content moderation under Chinese law

Baiyang Xiao [a,b]

[a] PhD Candidate, University of Szeged, Institute of Comparative Law and Legal Theory, Szeged 6722, Hungary
[b] Scholarship Holder, Max Planck Institute for Innovation and Competition, Munich 80539, Germany

ARTICLE INFO

Keywords:
Content moderation
China
Monitoring obligations
Digital Service Act

ABSTRACT

With the expansion of digital economy, tackling illegal online content is an increasingly challenging task. China implemented a dual-track legal mechanism on content moderation, whereby it exempts general monitoring obligations of intermediaries under private law while imposing monitoring obligations under public law. In recent years, major platforms exercise much stronger control over flow of information, regardless of more serious consequences that impact the fundamental rights of users. Meanwhile, a series of Chinese court rulings have shown that these divergent attitudes towards monitoring obligations under public and private law have given rise to legal conflicts that may deprive intermediaries of their legitimate immunity, undermining the stability and efficiency of the safe harbor rule. Furthermore, the lack of adequate legal safeguards against the risk of abusing automatic content filtering technology might transform the internet into a digital panopticon. To redraw boundaries between monitoring obligations under private and public law, future Chinese legislation should not only provide clearer clarification on the scope of monitoring, but also include a provision prohibiting general monitoring obligations in private law. To provide legal predictability for affected parties and flexibility for future technological developments, a Good Samaritan clause should be introduced in *Cybersecurity Law* by learning from the substance of Article 7 of the DSA.

1. Introduction

In the past decades, safe harbor provisions serve as an essential legal foundation to shield intermediaries from legal liability in moderating and managing content posted by users.[1] As is widely recognized, the genesis of these safe harbor provisions is situated within Section 512 of the Digital Millennium Copyright Act (DMCA).[2] The copyright-specific safe harbor provisions, centered around the 'notice and takedown' mechanism as well as the principle of prohibition on general monitoring obligations,[3] quickly became a legislative blueprint for the allocation for liability of online platforms in other nations.[4] Particularly, it is conventional wisdom that the European safe harbor scheme set forth in Article 12 to 14 of the E-Commerce Directive,[5] which provides mere conduit, caching, and hosting exemptions for intermediaries subject to

E-mail address: baiyang.xiao@ip.mpg.de.

[1] Anupam Chander, 'How law made Silicon Valley' (2013) 63 Emory Law Journal 639.

[2] See 17 U.S.C. § 512(a) to (d). See Jonathan Band and Matthew Schruers, 'Safe Harbors against the Liability Hurricane: The Communications Decency Act and the Digital Millennium Copyright Act' (2002) 20 Cardozo Arts & Entertainment Law Journal 295, 303.

[3] Chander (n 1). Section 512(m) specifies the principle of prohibition of general monitoring obligations. ('Nothing in this section shall be construed to condition the applicability of subsections (a) through (d) on……a service provider monitoring its service or affirmatively seeking facts indicating infringing activity……') See also (1998) R. REP. NO. 105-551, pt. 2, at 53 ('a service provider need not monitor its service or affirmatively seek facts indicating infringing activity.')

[4] Although scholars and officials argue that the DMCA safe harbor 'went far beyond treaty requirements' and may undermine freedom of speech, other countries subsequently followed the lead of the US. See Directive 2000/31/EC; Canadian Copyright Act (R.S.C., 1985, c. C-42).

[5] Article 12(3), 13(2), 14(3) of 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.

https://doi.org/10.1016/j.clsr.2023.105893

Available online 31 October 2023

specific requirements, is deeply shaped by the US approach.[6] The EU safe harbor legislation does not require intermediaries to monitor the information that they transmit or store, or to actively seek facts or circumstances indicating illegal activity.[7] However, the DMCA safe harbor seems to be vertical as it limits liability arising from copyright infringement alone, as Section 230 of the Communication Decency Act (CDA) does not require a 'notice-and-takedown' mechanism but provides intermediaries by far the strongest unconditional immunity for online speech.[8] While the EU safe harbor scheme aims to judge intermediary liability in a horizontal approach that applies to various categories of illegal content under the same criteria,[9] and it leaves room for injunctions and duties of care at the national level with respect to illegal content.[10]

Instead of reinventing the wheel, China transplanted and incorporated safe harbor provisions within the 2006 *Regulation on the Protection of the Right of Communication to the Public on Information Networks* by referring to Section 512 of the DMCA and Article 14 of the E-Commerce Directive.[11] Subsequent amendments to the 2006 *Regulation,* the *Tort Law* (2009) (coded in the *Tort Chapter of the Civil Code* (2020))[12] and the *E-Commerce Law* (2018)[13] have not only further refined and improved the joint liability of ISPs for contributory infringement, but also gradually expanded the applicability of the notice-and-takedown mechanism to all civil law issues, including intellectual property rights, defamation, unfair competition, and other types of infringement.[14] Noteworthy, the above legal transplant of safe harbor rules remains incomplete, as the prohibition of general monitoring obligations is absent from the relevant provisions.[15]

However, online tech powerhouses are often in the eye of the storm as they have amassed unprecedented power to proactively control the flow of information within society.[16] The radical paradigm shift in the digital services landscape has not only fundamentally changed the supply chain ecosystem, but opened the door to the unprecedented massive spread of illegal and harmful content, resulting in potential damage to market growth and industry sustainability.[17] Setting an effective and prompt regulatory framework to combat the dissemination of illegal and harmful content online without violating fundamental rights or disrupting innovation, is an inevitable but challenging task for regulators all across the globe.[18] One attractive idea is to redefine the intermediary liability conundrum and lift the monitoring obligation ban, thus requiring ISPs to take on the role of gatekeepers to proactively monitor and control the dissemination of illegal content on the Internet.[19] Policy makers on both sides of the Atlantic have joined the debate over whether platforms should be expelled from first-generation safe harbors and expected to take enhanced liability.[20] The latest endeavor, encapsulated in the controversial Article 17 of the Copyright Directive of the Digital Single Market (CDSM),[21] imposes a proactive

[6] Miquel Peguera, 'The DMCA safe harbors and their European counterparts: a comparative analysis of some common problems' (2008) 32 Columbia Journal of Law & Arts 481. However, scholars also suggest that the E-Commerce Directive does not actually provide a 'notice and takedown' scheme as it merely implies it through its conditions for liability exemption. See Aleksandra Kuczerawy, 'Intermediary liability & freedom of expression: Recent developments in the EU notice & action initiative' (2015) 31 Computer Law & Security Review 46.

[7] Article 15 and recital 47 of Directive 2000/31/EC. The abundant connotation of this principle is construed and explore by the CJEU through considerable decisions. See Case C-324/09, L'Oréal v. eBay, ECLI:EU:C:2011:474; Case C-70/10, Scarlet Extended SA (2011), ECLI:EU:C:2011:771; Case C-360/10, SABAM v. Netlog, ECLI:EU:C:2012:85; Case C-484/14, McFadden v. Sony, ECLI:EU:C:2016:689. See also 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 Journal of Intellectual Property, Information Technology & Electronic Commerce Law 176.

[8] The safe harbor in DMCA remains 'vertical' as it offers a specific coverage of copyright infringements. Joris van Hoboken and Daphne Keller, 'Design Principles for Intermediary Liability Laws' (Transatlantic Working Group, 8 October 2019) < https://cdn.annenbergpublicpolicycenter.org/wp-content/uploads/2020/05/Intermediary_Liability_TWG_van_Hoboken_Oct_2019.pdf > accessed 1 May 2023.

[9] The exemptions in the E-Commerce Directive have a general and horizontal scope, covering all types of illegal content. See 47 U.S.C. § 230(c)(1) (2000) ('No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.'); See Eric Goldman, 'An Overview of the United States' Section 230 Internet Immunity' in Giancarlo Frosio (ed.), *Oxford Handbook of Online Intermediary Liability* (OUP 2020), at pp.167-68. <https://doi.org/10.1093/oxfordhb/9780198837138.013.8> accessed 1 May 2023.

[10] Article 15(2) of E-Commerce Directive.

[11] *Regulations for the Protection of the Right of Communication through the Information Network* (信息网络传播权保护条例), enacted on 2006, amended in 2013. Specifically, this Regulation defines four types of services, including automatic access and transmission, automatic storage, information storage, search and link services, and imposes different obligations depending on the type of service provided.

[12] Article 1194-1197 of *Civil Code* (2020). Especially, the Article 1195 of *Civil Code* explicitly stipulates that all kinds of ISPs are eligible to enjoy the safe harbor protection.

[13] Article 42-45 of *E-Commerce Law* (2019).

[14] Jie Wang, 'How to utilize notice-and-takedown procedures in IP enforcement on e-commerce platforms–a lesson from China' (2021) 29 Asia Pacific Law Review 243.

[15] Dong Zhu, 'Transplantation and Variation of Secondary Liability of ISPs (网络服务提供者间接侵权责任的移植与变异)' (2019) 31 Peking University Law Journal 1340.

[16] Kapczynski Amy, 'The law of informational capitalism' (2020) 129 Yale Law Journal 1460.

[17] Giancarlo Frosio, 'From horizontal to vertical: an intermediary liability earthquake in Europe' (2017) 12 Journal of Intellectual Property Law and Practice 565; Annemarie Bridy and Daphne Keller, 'US Copyright Office Section 512 Study: Comments in Response to Notice of Inquiry' (2016) Available at SSRN 2757197; Daphne Keller, 'Amplification and its discontents: why regulating the reach of online content is Hard' (2021) 1 Journal of Free Speech Law 227.

[18] For example, over the past years the EU has adopted various initiatives to address illegal online content, including sector-specific legislation, non-binding guidelines for platforms and self-regulatory cooperation initiatives. See European Commission, Recommendation of 1.3.2018 on measures to effectively tackle illegal content online (C (2018) 1177 final); See DSA Inception Impact Assessment 2020 < https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12417-Digital-Services-Act-deepening-the-internal-market-and-clarifying-responsibilities-for-digital-services_en> accessed 1 May 2023.

[19] Network Enforcement Act, or NetzDG <https://netzpolitik.org/2020/hass-im-netz-oesterreich-soll-ein-netzdg-erhalten/#vorschaltbanner> accessed 1 May 2023; Rachel Griffin, 'New school speech regulation as a regulatory strategy against hate speech on social media: The case of Germany's NetzDG' (2022) 46 Telecommunications Policy 1.

[20] In the US, there are laws that create exceptions for certain types of content. In 2018, the U.S. Senate voted 97-2 to pass Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA, H.R. 1865) <https://www.congress.gov/bill/115th-congress/house-bill/1865> accessed 1 May 2023; 'Online Platforms and the Digital Single Markets' (Communication) COM (2016) 288 final, 9; Commission, 'Mid-term Review on the implementation of Digital Single Market Strategy' (Staff Working Document) SWD (2017) 155.

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

obligation upon OCSSPs to identify and block access to content that is identical to works claimed by copyright holders.[22] Moreover, the Digital Service Act (DSA), to a certain extent aimed at complementing the E-Commerce Directive, sets clear responsibilities for online platforms, encouraging content moderation and due diligence obligations to protect users' rights while preserving the key pillars of the E-Commerce Directive.[23]

Considering the fact that the EU has been a *de facto* global regulatory superpower characterized by extensive market importance and stringent regulatory capacity, the ramifications of this newly instituted regulation will definitely reverberate far beyond the EU's geographical confines.[24] This is particularly relevant for non-EU online content-sharing service providers because the EU's regulatory approach may shape international standards and practices of content moderation.[25] In light of the distinct characteristics of China's internet industry, direct transplantation EU regulations may not be an optimal choice, while recent EU developments on content moderation could serve as good references for potential revisions to intermediary liability laws in China. Such a decision necessitates a meticulous examination and comparative investigation of the legal and industrial underpinnings as well as the pragmatic context within the Chinese landscape.

By focusing on monitoring obligations, this paper aims to explain what legal measures China adopted to serve the needs of content control and why the current ambiguous and overlapping regulations on content moderation inevitably fail to safeguard the legitimate rights and interests of users through a comparison with the regulatory approach of the EU. Moreover, public policing and private removal concretely employed by legal and technological mechanisms are presented with details. It also revisits the EU legislative initiatives on content moderation, particular the DSA, and draws on the EU experiences and provides implications for future Chinese regulations.

2. Monitoring obligations within public/private distinction under Chinese law

Monitoring obligations are not uncommon for ISPs to oversee and regulate content on their service.[26] In general, monitoring obligations may emanate from explicit legislative mandates, such as Article 17 of the CDSM, or from the imposition of strict liability for user-generated content by judicial authorities, effectively necessitating that intermediaries actively monitor and moderate illegal content to circumvent liability.[27] Regarding the monitoring obligations of ISPs, Chinese law adopts a dual-track approach that emphasizes the public and private distinction.[28] ISPs are exempted from monitoring obligations in private law, while public law explicitly imposes statutory requirements on the monitoring obligations of ISPs, requiring them to take on the role of gatekeepers who have a responsibility towards the public interest.[29]

The dual approach is well reflected in the Guiding Opinions formulated by the Beijing Higher People's Court.[30] Article 17 of the Guiding Opinions provides that 'ISPs […] generally are not obliged to conduct proactive review and monitoring of others' use of their services to disseminate content to determine whether they infringe on copyrights. If monitoring is required according to relevant laws and regulations, it shall be conducted.'[31] The first sentence reiterates the general monitoring obligations ban, while the second implies that ISPs still have to perform public law monitoring obligations stipulated in relevant legislation.

2.1. No general monitoring obligations under private law

It is worth noting that the prohibition of general monitoring obligations constitutes a critical complement to safe harbor immunity for ISPs,[32] as it prevents conscripting intermediaries to act as unofficial censors.[33] For example, Section 512(m) of the DMCA specifically clarifies that an ISP shall not be required to '[monitor] its service or affirmatively [seek] facts indicating infringing activity' to maintain their safe harbor immunity.[34] In a similar manner, Article 15(1) of the E-Commerce Directive explicitly states that ISPs are not subject to a general monitoring obligation 'to monitor the information which they transmit or store,' nor 'to seek facts or circumstances indicating illegal activity.'[35]

However, the principle of prohibition of general monitoring obligations is absent in Chinese private law legislation. In the third draft of the amendment to the Chinese Copyright Law, Article 73 explicitly provides that 'ISPs are not subject to monitoring obligations related to copyright or related rights, when providing mere technical services such as storage, searching and linking services to users.'[36] However, this draft brought about significant controversies by favoring the internet industry and encouraging copyright infringements, and the proposed provision

[22] Marcin Rojszczak, 'Online content filtering in EU law' (2022) 47 Computer Law & Security Law Review 1, 10; Jane Ginsburg, 'A United States Perspective on Digital Single Market Directive Art. 17' in Irini Stamatoudi and Paul Torremans (eds.) *EU COPYRIGHT LAW: A COMMENTARY*, (2nd edn, Edward Elgar 2020).

[23] João Pedro Quintais and Sebastian Felix Schwemer, 'The Interplay between the Digital Services Act and Sector Regulation: How Special is Copyright?' (2022) 13 European Journal of Risk Regulation 191.

[24] Anu Bradford, 'The Brussels effect.' (2012) 107 Northwestern University Law Review 1; See also Anu Bradford, *The Brussels effect: How the European Union rules the world* (OUP 2020).

[25] Ally Boutelle and John Villasenor, 'The European Copyright Directive: Potential impacts on free expression and privacy' (Brookings, 2 February 2021) <https://www.brookings.edu/blog/techtank/2021/02/02/the-european-copyright-directive-potential-impacts-on-free-expression-and-privacy/> accessed 1 May 2023.

[26] Sunimal Mendis and Giancarlo Frosio, 'Monitoring and filtering: European reform or global trend?' in Giancarlo Frosio (ed.) *Oxford Handbook of Online Intermediary Liability* (OUP 2020), 544-565.

[27] Giancarlo Frosio, 'Why keep a dog and bark yourself? From intermediary liability to responsibility' (2018) 26 International Journal of Law and Information Technology 1.

[28] A telling case referring the private/public distinction on monitoring obligation is the WeChat mini-program case. The court held that defendant Tencent, being the provider of internet access services in this case, did not assume the obligation of 'notice and takedown'; however, the court still emphasized that Tencent should proactively monitor undesirable information related to pornography, terrorism, gambling, and other illegal activities. In other words, Tencent must assume the monitoring obligation under public law even if it does not have a private law obligation. See *Hangzhou Internet Court [2019] Zhe 01 Min Zhong No.4286 Civil Judgement* (2019)浙01民终4286号民事判决书.

[29] Jonathan Zittrain, 'A history of online gatekeeping' (2005) 19 Harvard Journal of Law & Technology 253; Tarleton Gillespie, *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media* (Yale University Press 2018).

[30] *Guiding Opinions on Several Issues Concerning the Trial of Copyright Disputes in the Network Environment* (北京高院关于审理涉及网络环境下著作权纠纷案件若干问题的指导意见), issued by Beijing Higher People's Court on 19 May 2010.

[31] Ibid.

[32] Kuczerawy (n 6) 47.

[33] Marcelo Thompson, 'Beyond Gatekeeping: The Normative Responsibility of Internet Intermediaries' (2020) 18 Vanderbilt Journal of Entertainment and Technology Law 783, 785.

[34] See 17 U.S.C. § 512 (c).

[35] Article 15(1) of E-Commerce Directive.

[36] 'Revised Draft for Amendment to Copyright Law of the People's Republic of China' (State Council, 10 June 2014) <http://www.gov.cn/xinwen/2014-06/10/content_2697701.htm> Accessed 1 May 2023.

on the monitoring obligation ban was deleted after several rounds of revision.[37]

Article 36 of the Tort Law (codified into the Civil Code), which addresses online infringement, is a manifestation of the legal transplantation of the safe harbor rules delineated in Section 512 of the DMCA. Although this provision does not explicitly require ISPs to bear monitoring obligations, the Legislative Affairs Commission referred to international conventional wisdom and clarified that 'ISPs that provide technical services are not subject to general monitoring obligations.'[38] After seven years, the legislative Affair Commission reiterated the same principle in its authoritative interpretations of Article 1197 of the Civil Code.[39] Obviously, the term 'general monitoring obligation' is an imported lexicon from EU safe harbor legislation. However, the E-Commerce Directive exempts ISPs from general monitoring obligations, but leaves the discretion to national laws to provide for monitoring obligations 'in a specific case.'[40] Particularly in cases of alleged infringement of IP rights, the CJEU allowed specific monitoring measures when a fair balance between the fundamental rights of the different stakeholders was achieved.[41] In the same vein, the Chinese jurisprudence also recognizes the prohibition of general monitoring obligations under private law but does not preclude the possibility of monitoring obligations of a specific nature.[42]

According to Article 8(2) of the *[2020] Judicial Interpretation No.19*, the Supreme People's Court clarifies that the court shall not determine an ISP is at fault where it fails to conduct proactive monitoring regarding a user's infringement.[43] Article 8(3) further states that 'where an ISP can demonstrate that it has employed reasonable and efficacious technical measures, yet remains unable to identify a user's infringement […], the court shall ascertain that the ISP is not at fault.' In another Guiding Opinion, the Supreme People's Court explicitly stated that '[courts shall] not impose a general obligation of prior review and a relatively high degree of duty of care upon the ISPs […].'[44] Courts all across the country also confirm the principle of no general monitoring obligations in numerous cases.[45]

In conclusion, the Chinese jurisprudence has reached consensus that the principle of prohibition on general monitoring obligations applies in private sphere and leaves certain room for monitoring obligations in cases of specific natures.[46] Yet it is worth noting that this consensus only extends to the prohibition of general monitoring obligations in private law, not those under public law.

2.2. Proactive general monitoring obligations under public law

Considering the significant risks involved in content moderation, governments struggle to determine the proper oversight of digital platform companies.[47] Due to limited technical capabilities and enforcement resources, administrative agencies tend to impose obligations on platforms to urge them to carry out internal regulations to regulate illegal content.[48] Unlike the sector-specific approach in the EU, the Chinese regulatory framework of content moderation consists of a vertical approach combining public intervention and self-regulation.[49] ISPs are required to review, monitor, and inspect information prohibited from being disseminated by laws and administrative regulations.[50] When they 'discover' illegal content disseminated on their services, they must fulfill their proactive monitoring obligations, by taking certain measures to prevent the transmission of such content, namely to stop transmission, remove disputed content, prevent dissemination, preserve records, and report to relevant departments. In addition to technical

[37] Keli, 'PKU and Stanford Conference-The Development of the Internet Industry in Light of Article 69 of the Copyright Law Amendment Act' (Tencent Research Institute 10 August 2014) <https://www.tisi.org/436> accessed 11 September 2023; 'Explanation on the Amendment to the Copyright Law of the People's Republic of China (Draft)' (National People's Congress, 26 April 2020) <http://www.npc.gov.cn/npc/c30834/202011/f254003ab9144f5db7363cb3e01cabde.shtml> accessed 1 May 2023.

[38] Shengming Wang, *Interpretation of the Tort Liability Law of the People's Republic of China* (中华人民共和国侵权责任法》释义) (2nd ed, Law Press China 2013), 218.

[39] *Annotations to the Civil Code of the People's Republic of China* (中华人民共和国民法典释义) (1st ed, Law Press China 2020), p.695.

[40] Giancarlo Frosio, 'Reforming intermediary liability in the platform economy: A European digital single market strategy' (2017) 112 Nw. UL Rev. Online 18, 41.

[41] Article 15 (2) of the E-Commerce Directive provides that the immunities shall not affect the possibility of a court or administrative authority … of requiring the service provider to terminate or prevent an infringement. See Case C-314/12 UPC Telekabel Wien v Constantin Film Verleih GmbH (2014) EU:C:2014:192 (elaborating detailed requirements for 'specific nature'); Case C-18/18, Glawischnig-Piesczek v. Facebook Ireland, ECLI:EU:C:2019:821.

[42] See *Suzhou Intermediate Court [2019] Su 05 Min Zhong No.4709 Civil Judgement* (2019)苏05民终4709号民事判决书 (ISPs are required to monitor copyright infringing content uploaded by third parties through targeted measures under specific circumstances); *Shandong Higher People's Court [2008] Lu Min San Zhong Zi No.8 Civil Judgement* (2008)鲁民三终字第8号民事判决书 (concludes that ISPs are not subject to an ex ante general motoring obligation, but should bear certain ex post monitoring obligation).

[43] *Provisions by the Supreme People's Court on Several Issues Concerning the Application of Law in Hearing Civil Dispute Cases Involving Infringement of the Right of Communication to the Public on Information Networks* (最高人民法院关于审理侵害信息网络传播权民事纠纷案件适用法律若干问题的规定), issued by Supreme People's Court on 29 December 2020.

[44] Supreme People's Court, *Notice of the Supreme People's Court on Issuing the Opinions on Issues concerning Maximizing the Role of Intellectual Property Right Trials in Boosting the Great Development and Great Prosperity of Socialist Culture and Promoting the Independent and Coordinated Development of Economy* (关于充分发挥知识产权审判职能作用推动社会主义文化大发展大繁荣和促进经济自主协调发展若干问题的意见), issued on 16 December 2011.

[45] See *Liaoning Higher People's Court [2013] Liao Min San Zhong Zi No. 178 Civil Judgement* (2013)辽民三终字第178号民事判决书 (denying proactive monitoring obligation of ISPs); *Hangzhou Intermediate People's Court [2019] Zhe 01 Min Zhong No.4268 Civil Judgement* (2019)浙01民终4268号民事判决书(denying ex ante monitoring obligation of ISPs) ; *Beijing Internet Court [2019] Jing 0491 Min Chu No.22238 Civil Judgement* (2019)京0491民初22238号民事判决书(ISPs are not subject to proactive, general monitoring obligation); *Shanghai Xuhui District People's Court [2020] Hu 0104 Min Chu No.8302 Civil Judgement* (2019)沪0104民初8302号民事判决书(platforms are not subject to general proactive monitoring obligation, and it is practically difficult to conduct comprehensive and active monitoring of a large number of short videos, or to block keywords in advance).

[46] However, no detailed elaboration on the 'specific' nature of such monitoring obligation is provided in relevant legislative documents or court decisions.

[47] Julia Pohle and Daniel Voelsen, 'Centrality and power. The struggle over the techno-political configuration of the Internet and the global digital order' (2022) 14 Policy & Internet 13.

[48] The Chinese authorities adopted a 'regulate the Internet by Internet' strategy, a multifaceted approach that leverages various stakeholders, technology, and self-regulation to control and govern the online content in accordance with the government's objectives and regulations. See 'Accelerate the establishment of a comprehensive Internet governance system and comprehensively improve the level of Internet management and control capabilities.' (Cyberspace Administration of China 9 June 2022) <http://www.cac.gov.cn/2022-06/08/c_1656303130339484.htm.> accessed 1 May 2023,

[49] Yong Shan,' Digital Gatekeeper and Governance of Crimes on Mega Platforms (数字看门人与超大平台的犯罪治理)' (2022) 2 Legal Science 74, 82-85.

[50] Article 47 of *Cybersecurity Law* provides that 'If any operator finds any information of which the release or transmission is prohibited by any law or administrative regulation, it shall immediately cease the transmission of such information, take deletion or any other handling measure to prevent the information from spreading, preserve relevant records, and report it to the competent department.'

filtering mechanisms, platforms must also employ trained personnel to conduct human reviews of uploaded content. Otherwise, they will face penalties such as warnings, fines, suspension of services, and cancelation of permissions or licenses for business operation, for their failure to perform their monitoring obligations.[51]

2.2.1. 'Eleven boundaries': overinclusive general monitoring obligations

In general, the scope of public law monitoring obligations primarily encompasses illegal information pertaining to political matters, explicit, violent and terrorist-related content, and ethnic and religious issues.[52] The 'Nine Prohibitions' specified in *Administrative Measures for Internet Information Services*, addresses nine types of information that ISPs shall not produce, copy, publish or distribute.[53] Chinese scholars also refer to the 'Nine Prohibitions' as nine bottom lines that ISPs shall not step over.[54] The long list of prohibited content best sums up the primary targets of authorities when it comes to illegal and harmful content online.[55]

The *Cybersecurity Law*, which took effect in 2017, does not provide a precise definition of illegal information, but it outlines in the general provisions the prohibited illegal and harmful online content, which are similar in scope to the 'Nine Prohibitions,' albeit with some slight variations in phrasing.[56] In 2019, the *Provisions on the Ecological Governance of Network Information Content* adds two further types of illegal and harmful content to the prohibitions, namely content demeaning or denying the deeds and spirit of heroes and martyrs, and content promoting terrorism or extremism, thus turning 'Nine Prohibitions' into "Eleven Boundaries.'[57]

Furthermore, other than providing an inclusive list of prohibited illegal and harmful content, the 2019 *Provisions* also requires ISPs to take measures to 'prevent and resist' nine types of 'undesirable content,'[58] Through a contextual interpretation, the nine categories of 'undesirable content' clearly fall outside the scope of 'Eleven Boundaries,' and thus are not subject to a general monitoring obligation.

2.2.2. Jigsaw puzzles: fragmented administrative regulations regarding content moderation

Furthermore, administrative authorities launched dozens of regulatory projects tackling online illegal content under a multipronged regulatory strategy. Based on the diverse nature and characteristics of different ISPs' services, a series of fragmented administrative regulations were issued to target the dissemination of illegal content and activities on the internet. To comply with the above regulations, the ISPs are required to proactively monitor illegal content related to food safety,[59] online posts and comments,[60] group chat service,[61] artificial intelligence,[62] online cultural activities,[63] mobile applications,[64] live streaming,[65] online searching,[66] audio-visual programs,[67] and so forth. Unsurprisingly, the scope of monitoring can be considered comprehensive, as the ISPs are required to monitor almost all online content in accordance with various laws, administrative regulations, and even 'relevant state provisions.'[68] By employing unrelated laws and regulations as a pretext to block content, authorities thus hold a powerful tool to filter unfavorable content through soft censorship.[69]

[51] See Chapter 3, 4 and 5 of *Cybersecurity Law*. The massive dissemination of illegal information due to failure to monitor illegal content may lead to the crime of refusing to fulfill the obligations of information network security management under Article 286 of the *Chinese Criminal Law*.

[52] Xiangwen Kong, 'A Public Law Perspective on Reflection of the Structure of Information Content Regulation for Online Platforms (网络平台信息内容规制结构的公法反思)' (2022) 2 Global Law Review 133.

[53] *Administrative Measures for Internet Information Services* (互联网信息服务管理办法), enacted by State Council on 8 January 2011.

[54] Kong (n 52) 137.

[55] Article 15 provides that internet information service providers shall not produce, copy, publish or distribute information having the following contents: (1) violates cardinal principles set forth in the Constitution; (2) endangers national security; (3) damages national honor and interests; (4) undermines the state's religious policies; (5) propagates cults and feudal superstitions; (6) disseminates rumors that disrupts social order and stability; (7) disseminates obscenity, pornography, brutality and terror or crime-abetting elements; (8) infringes upon the legitimate rights and interests of others through insults and defamation; and (9) involves other information that violates laws and regulations.

[56] Article 12 of *Cybersecurity Law*.

[57] Article 6 of *Provisions on the Ecological Governance of Network Information Content* (网络信息内容生态治理规定).

[58] 'Undesirable content' refers to (1) content using exaggerated titles that are seriously inconsistent with the contents; (2) content hyping gossips, scandals, bad deeds, and so forth; (3) content making improper comments on natural disasters, major accidents or other disasters; (4) content containing sexual innuendo, sexual provocations, and other information that easily leads to sexual fantasy; (5) content showing bloodiness, horror, cruelty, and other scenes that cause physical and mental discomfort; (6) content inciting discrimination among communities or regions; (7) content promoting indecency, vulgarity, and kitsch; (8) content that may induce minors to imitate unsafe behavior, violate social morality, or induce minors to indulge in unhealthy habits; and (9) other contents that adversely affect network ecology.

[59] Article 14 of *Measures for the Investigation and Punishment of Illegal Acts Related to Online Food Safety* (网络食品安全违法行为查处办法) (2021), enacted by State Administration for Market Regulation.

[60] Article 4 (5) of *Provisions on the Administration of Internet Comments Posting Services* (互联网跟帖评论服务管理规定) (2022), enacted by Cyberspace Administration of China. Article 5 of *Provisions on the Administration of Internet Forum and Community Services* (互联网论坛社区服务管理规定) (2017), enacted by Cyberspace Administration of China.

[61] Article 5 of *Provisions on the Administration of Internet Group Information Services* (互联网群组信息服务管理规定) (2017), enacted by Cyberspace Administration of China.

[62] Article 7 of *Provisions on the Administration of Deep Synthesis of Internet-based Information Services* (互联网信息服务深度合成管理规定) (2022), enacted by Cyberspace Administration of China, Ministry of Industry & Information Technology, and Ministry of Public Security; Article 7 of *Provisions on the Administration of Algorithm-generated Recommendations for Internet Information Services* (互联网信息服务算法推荐管理规定) (2022), enacted by Cyberspace Administration of China, the Ministry of Industry and Information Technology, the Ministry of Public Security, and the State Administration for Market Regulation.

[63] Article 18 of *Interim Provisions on the Administration of Internet Culture* (互联网文化管理暂行规定) (2017), enacted by Ministry of Culture (dissolved).

[64] Article 5 of *Provisions on the Administration of Information Services of Mobile Internet Apps* (移动互联网应用程序信息服务管理规定) (2022), enacted by Cyberspace Administration of China.'

[65] Article 7(2) of *Provisions on the Administration of Internet Live-Streaming Services* (互联网直播服务管理规定) (2016), enacted by Cyberspace Administration of China.

[66] Article 6 of *Provisions on the Administration of Internet Information Search Services* (互联网信息搜索服务管理规定) (2016), enacted by Cyberspace Administration of China.

[67] Article 20 of *Provisions on the Administration of Private Network and Targeted Communication Audiovisual Program Services* (专网及定向传播视听节目服务管理规定) (2021), issued by National Radio and Television Administration.

[68] The highly inclusive and problematic term 'relevant state provisions' was adopted in *Provisions on the Administration of Internet Forum and Community Services* (n 60), as well as *Provisions on the Administration of Internet Group Information Services* (n 61). It reveals that the vague provisions of higher-level laws provide opportunities and space for lower-level regulations to continuously expand the scope of content monitoring.

[69] PEN America, *FORBIDDEN FEEDS: Government Controls on Social Media in China*, (PEN American Center, 2019), pp.21-22. <https://policycommons.net/artifacts/1736566/forbidden-feeds/2468203/> accessed 1 May 2023.

3. Making the private public: expansion of public law monitoring obligations

Effective management of illegal content depends heavily on internal platform regulation in addition to state interventions. Law enforcement agencies fully utilize the advantages of platforms in discovering, identifying, and handling illegal content, and entrust ISPs to proactively engage in collateral censorship through private ordering.[70] Thus, house rules, consisting of substantive norms voluntarily adopted by companies to regulate content and activities on their services,[71] act as a critical supplement to state legislation by restricting otherwise-legal content or activities based on their idiosyncratic editorial policies.

Usually, the house rules that determine which content can be published and disseminated on the platforms are not established by users but rather unilaterally decided by the platforms.[72] Platforms often present users with a 'take it or leave it' option, essentially forcing them to accept the terms or refrain from using their services.[73] In practice, online platforms usually further expand the scope of illegal and harmful content, indicating that the concentrated and pervasive power that corporations hold over online content might arguably surpass state power within its sphere.[74] Telling examples can be found in the terms and conditions of three exceptionally mega platforms that dominate online content in China, namely Tencent,[75] Weibo[76] and Douyin.[77]

3.1. Beyond 'Eleven boundaries': house rules regarding content moderation

Without proper content moderation, the internet would drown in spam and disturbing imagery, which deteriorates the user experience and risks losing users to more trustworthy competitors.[78] Platforms are incentivized to provide content governance, aiming to maintain their credibility and reputation.[79] Therefore, from the perspective of business operations, the scope and measures of content governance may inadvertently serve as a competitive asset among different platforms.

Section 8.1.2 of the WeChat T&Cs clearly prohibits the dissemination, transmission, storage, and publication of five types of illegal content.[80] In addition, WeChat further provides a detailed list of 12 types of content prohibited in its Community Guidelines.[81] Compared to the scope of 'Eleven Boundaries,' it seems that the house rules of WeChat are not only pervasive but more subtle.[82] Essentially, these provisions contain considerable unclear concepts such as 'national interests,' 'legitimate interests,' 'social morality' and 'public order,' thus making platforms prone to abuse their power in the interpretation of such terms. If a user violates the house rules, WeChat may take actions such as restricting the visibility of content, deleting the non-compliant content, restricting the accounts from accessing some or all of the WeChat features, or blocking the user accounts.

Content prohibited on WeChat
content that(1) violates the laws and regulations (2) infringes upon others' reputation rights, portrait rights, intellectual property rights, trade secrets and other legitimate rights; (3) contains others' privacy, personal information or materials; (4) contains harassment and advertising information, over-marketing information, spam or any information containing any sexual content or sexual connotation; (5) violates laws, regulations, policies and public order, that contradicts social morality, or interferes with the normal operation of WeChat or infringes the legitimate rights and interests of other users or third parties.

Section 8.1.2 of T&Cs of Tencent WeChat

The Weibo platform has developed its own content regulation system, which categorizes targeted online information into four types, namely sociopolitical information, illegal information, undesirable information, and the protection of minors. In this guideline, Weibo contends that it discovers non-complaint content by proactively detecting online content or receiving reports from users. Users violating house rules will be prohibited from posting and commenting, from being followed by others, from modifying account information, and from having access restricted until the account is canceled. Moreover, Weibo may adopt more diverse measures to handle prohibited content, including tagging, removing, blocking, restricting visibility, banning from using features, banning from monetizing content, and so forth.[83]

Content prohibited on Weibo	
Sociopoliticalinformation	'Eleven Boundaries'+content that incites illegal assemblies, associations, processions, demonstrations and gatherings to disrupt social order.
Illegal information	(1) content containing content that disturbs public order, obstructs public safety, infringes on personal rights and property rights, and obstructs social administration;(2) content involving pornography;(3) content on the sale or trafficking of prohibited or restricted items as defined by various laws and regulations;(4) fraudulent content;

(*continued on next page*)

[70] Kate Klonick, 'The new governors: The people, rules, and processes governing online speech.' (2017) 131 Harvard Law Review 1598; Jack M. Balkin, 'Free speech is a triangle' (2018) 118 Columbia Law Review 2011; Felix T. Wu, 'Collateral Censorship and the Limits of Intermediary Immunity' (2011) 87 Notre Dame Law Review 293.

[71] Eric Goldman, 'Content moderation remedies' (2021) 28 Michigan Technology Law Review 1, 8.

[72] Luca Belli and Jamila Venturini, 'Private ordering and the rise of terms of service as cyber-regulation' (2016) 5 Internet Policy Review 1.

[73] Péter Mezei and István Harkai, 'End-user flexibilities in digital copyright law–an empirical analysis of end-user license agreements' (2022) 5 Interactive Entertainment Law Review 2; see also Frago Kourandi, Jan Krämer, and Tommaso Valletti, 'Net neutrality, exclusivity contracts, and internet fragmentation' (2015) 26 Information Systems Research 320.

[74] Kyle Langvardt, 'Regulating online content moderation' (2017) 106 Georgetown Law Journal 1353.

[75] Tencent WeChat Software License and Service Agreement (腾讯微信软件许可及服务协议) <https://weixin.qq.com/agreement?lang=zh_CN> accessed 1 May 2023. WeChat is a free messaging and calling application with more than 1.3 billion monthly active users in 2022. See 'Number of monthly active WeChat users from 2nd quarter 2011 to 4th quarter 2022' (Statista) <https://www.statista.com/statistics/255778/number-of-active-wechat-messenger-accounts/> accessed 1 May 2023.

[76] Community Guidelines of Weibo (微博社区公约) < https://service.account.weibo.com/h5/roles/gongyue?ua=iPhone11%2C8_weibo_12.8.2_iphone_os15.6&from=10C8293010>;Terms & Conditions of Weibo (微博服务使用协议) <https://weibo.com/signup/v5/protocol> accessed 1 May 2023. Weibo is a Chinese microblogging website with more than 580 million monthly active users in 2022. See Number of monthly active users of Weibo Corporation from 1st quarter of 2014 to 4th quarter of 2022, (Statista) <https://www.statista.com/statistics/795303/china-mau-of-sina-weibo/> accessed 1 May 2023.

[77] Community Self-Discipline Convention of Douyin (抖音社区自律公约) <https://www.douyin.com/rule/policy> accessed 1 May 2023. Douyin is the Chinese version of TikTok, with more than 730 million monthly active users in 2022. See 'Number of monthly active users of popular short video apps in China in November 2022' (Statista) <https://www.statista.com/statistics/910633/china-monthly-active-users-across-leading-short-video-apps/> accessed 1 May 2023.

[78] Langvardt (n 74).

[79] Rotem Medzini, 'Enhanced self-regulation: The case of Facebook's content governance' (2022) 24 New Media & Society 2227.

[80] T&Cs of WeChat (n 75).

[81] Standards of Weixin Account Usage (updated: 28 October 2022) <https://weixin.qq.com/cgi-bin/readtemplate?&t=page/agreement/personal_account&lang=en_US&head=true> accessed 1 May 2023.

[82] Another list of 24 subtypes of 'content that violates the laws and regulations' is provided in section 8.1.2.1. Even the section 8.1.2.1.(24) is presented in the form of miscellaneous provision.

[83] Terms & Conditions of Weibo (n 76).

(*continued*)

Content prohibited on Weibo	
Undesirable information	(1) content containing malicious marketing;(2) content containing promotion of hatred;(3) content containing other undesirable information;
Child protection	(1) content that is sexually suggestive, sexually seductive, or other elements that may easily evoke sexual associations;(2) content displaying gore, horror, or cruelty, which may cause physical and mental discomfort;(3) content that promotes promoting vulgar or pandering material;(4) content that may induce minors to imitate unsafe behavior, violate social ethics, or develop unhealthy habits;(5) Other contents that may affect, harm, or endanger the safety and mental and physical health of minors.

Community Guidelines of Weibo

Based on the 'Eleven Boundaries,' Community Self-Discipline Convention of Douyin further specifies ten categories with an immensely complicated list of 35 types of prohibited content and activities. Moreover, it provides the most detailed definition and explanation for each type of content and activity by enumerating and providing examples. Users who violate this community guideline will be imposed appropriate penalties, including but not limited to removal or blocking of prohibited content, banning or blocking of non-complaint accounts. Moreover, the scope of content monitoring further extends in practice as platforms might moderate content that negatively impacts their interests and infrastructural values.[84]

Content and activities prohibited on Douyin	
Violence and criminal behaviors	1. Incitement and perpetration of violence 2. Prohibited and controlled substances 3. Terrorism and extremism 4. Dangerous persons and organizations 5. Display or promotion of criminal activities 6. Aiding and abetting the commission of a crime
Harmful and inaccurate information on current affairs	7.Harmful information on current affairs that endangers national and social security 8. Inaccurate information on current affairs that damages the image of the nation and the social order
Violation of personal rights	9. Suicide, self-injury 10. Cyber violence 11. Violation of personal freedom 12. Dangerous behavior 13. Invasion of privacy and personal information 14. Other violations of personal rights and interests
Illegal and undesirable content	15. Pornographic and obscene content 16. Hate and discriminatory speech 17. Vulgar content 18. Bloody and gory content 19. Excessively horrifying content 20. Brutal and apathetic content
Misinformation	21. Rumor and other types of misinformation
Violation of social morality	22. Content that is contrary to social ethics 23. Content that disseminates negative value orientations 24. Content that seriously hurts national sentiments
IPRs Infringement	25.Content that infringes upon IPRs of others
Infringement of the rights of minors	26. Sexual abuse of minors 27. Content related to sexual abuse of minors 28. Improper sexual exploitation of minors for profit 29. Content that endangers or affects the physical or mental health of minors 30. Other criminal activities against minors
False and dishonest conduct	31. Cheating and spamming 32. Improper Marketing and Misrepresentation

(*continued on next column*)

(*continued*)

Content and activities prohibited on Douyin	
	33. Malicious traffic diversion 34. Deceptive behavior
Jeopardize the order and safety of the platform	35. Content that threaten the security of the platform

Community Self-Discipline Convention of Douyin

Noteworthy, these house rules classify all the illegal, harmful and undesirable content as prohibited content, and ignore the distinction between prohibited content and undesirable content made in relevant administrative regulations. Apparently, these platforms adopted a crafty approach by introducing more blurred and abstract concepts to explain the ambiguous language of legislation, thus worsening the predictability of house rules. Although commentators voice concerns about legal uncertainty deriving from ambiguous rules, the platforms seem willing to regard them as 'flexibility.' By embracing an expansive scope of monitoring and an erratic and opaque decision-making process, mega platforms may exercise much stronger control over the flow of information, regardless of more serious consequences that impact the fundamental rights of users.[85] Smaller platforms may outsource their moderation to third-party services via the same software and human teams. Nevertheless, these standards deployed for content moderation often share a high level of similarity to house rules phrased by US-domiciled mega platforms.[86]

3.2. *Putting public law monitoring obligations into practice*

When lacking systematic and institutional constraints, the constantly expanding content moderation practices are characterized by quasi-legislative (T&Cs and Community Guidelines), quasi-executing (content moderation measures), and quasi-judicial (determination of illegal and harmful) natures. Evidently, under the top-down collateral censorship mechanism, platforms try to adopt various stricter content moderation measures and further extend the scope of monitoring to eliminate potential uncertainties and risks.[87] Such practices can fully empower themselves with greater control over content and information on the internet from the perspectives of moderation technology and norm-making.[88]

3.2.1. *Diverse toolkits for content moderation*

On the one hand, in the overly inclusive T&Cs and Community guidelines, a vast space is left for platforms to apply alternative mechanisms, which are often not transparent and not subject to external oversight, to moderate content.[89] Platforms adopt more diverse measures to conduct content moderation, both preventive (ex ante) and reactive (ex post). Reactive measures such as region- and service-specific methods are employed to control the availability, visibility and accessibility of certain content, or restrict users' ability to provide

[84] Community Self-Discipline Convention of Douyin (n 77).

[85] 'Surveillance giants: How the business model of Google and Facebook threatens human rights', (Amnesty International, 21 November 2019) <https://www.amnesty.org/en/documents/pol30/1404/2019/en/> accessed 1 May 2023.

[86] Tarleton Gillespie et al, 'Expanding the debate about content moderation: scholarly research agendas for the coming policy debates' (2020) 9 Internet Policy Review 1, 5.

[87] Jack M. Balkin, 'Old-school/new-school speech regulation' (2013) 127 Harvard Law Review 2296, 2309-2310.

[88] Rebecca Tushnet, 'Power without responsibility: Intermediaries and the First Amendment' (2007) 76 George Washington Law Review 986.

[89] Klonick (n 70).

information, independently or in response to government mandates.[90] Meanwhile, preventive content moderation, which aims to make content contingent on the prior consent of a designated public authority, usually takes the form of automated content filtering of unpublished content. Among them, two types of measures, automated content filtering (ex ante)[91] and visibility remedies (ex post),[92] need to be highlighted and further analyzed.

Major platforms implement ex ante algorithm-based filtering mechanisms as a regular weapon to define the scope of visibility of content on their services.[93] While the increased danger of false positives and false negatives is the most evident drawback of automated content filtering.[94] The facilitation of large scale and effortless removal of allegedly infringing content, is an extensively examined consequence of the traditional 'notice and takedown' process, ultimately resulting in a substantial chilling effect on users' freedom of expression.[95]Furthermore, algorithm-based automated content moderation systems amplify such an outcome, because the indifferent nature of the online intermediary is translated and coded into the design of the decision-making algorithm by setting the defaults.[96] In the context of increased responsibilities for illegal content, platforms are incentivized to expand the scope of monitoring and flag controversial marginalized content as illegal in order to avoid liability and minimize the compliance cost, resulting in a rising number of false positives.[97]

Moreover, platforms adopt 'shadow banning' to set an output-based form of visibility restriction on user content, which gives the user the false impression that the content can still be posted, while in fact it is not visible to other users.[98] Leerssen succinctly suggests that shadow banning is used to manage new controversies which often fall short of violating established laws.[99] Shadow banning usually takes a subtler form as the complement to conventional moderation practices, making affected users struggle to ascertain whether or not they have been sanctioned.[100] Even though shadow banning appears less restrictive than removal and blocking, it may have a greater impact on users' freedom of expression and privacy due to a lack of transparency and proportionality.[101] The shadow banning not only challenges the predictability of the procedures of content moderation, but also practically precludes possibilities for individual or collective resistance.[102]

3.2.2. *Constantly widening scope of monitoring*

Platforms extend the scope of content moderation with the substantial quasi-legislative power obtained from house rules. By introducing more uncertain concepts to elaborate on vague terms in public law, the predictability and transparency of house rules are further diminished. What is more, platforms may encode infrastructural values in both house rules and content moderation enforcement.[103] For instance, Weibo removed all misleading posts supporting a celebrity named Kris Wu in 2021, who was detained under suspicion of rape then.[104] The removal of such posts is neither based on a determination of the illegality of the content posted nor in accordance with any specific provision of the community guidelines, but driven by the platform's self-interest and the eagerness to appease popular public sentiments.[105] Under this parental state, other types of political heterodox speeches,[106] legal speeches that violate widely held social norms and moral beliefs,[107] or infrastructural values of platforms,[108] are removed or

[90] 'NCAC and Other Three Authorities Launched 'Jianwang 2020' Campaign' (National Copyright Administration of China, 16 July 2020) <https://en.ncac.gov.cn/copyright/contents/10373/339825.shtml> accessed 1 May 2023.

[91] Niva Elkin-Koren, 'Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence' (2020) 7 Big Data & Society 2053951720932296; Tarleton Gillespie, 'Content moderation, AI, and the question of scale' (2020) 7 Big Data & Society: 2053951720943234.

[92] Sarah Myers West, 'Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms' (2018) 20 New Media & Society 4366.

[93] Jeffrey Knockel and others, '(Can't) Picture This: An Analysis of Image Filtering on WeChat Moments' (The Citizen Lab, 14 August 2018) <https://tspace.library.utoronto.ca/bitstream/1807/94801/1/Report%23112-Can%27t%20Picture%20This.pdf> accessed 1 May 2023.

[94] Evan Engstrom and Nick Feamster, 'The limits of filtering: A look at the functionality & shortcomings of content detection tools' (Engine, March 2017) <https://static1.squarespace.com/static/571681753c44d835a440c8b5/t/58d058712994ca536bbfa47a/1490049138881/FilteringPaperWebsite.pdf> accessed 1 May 2023.

[95] Jennifer M. Urban, Joe Karaganis, and Brianna Schofield, 'Notice and takedown in everyday practice' (2017) UC Berkeley Public Law Research Paper 2755628.

[96] Orit Fischman-Afori, 'Online Rulers as Hybrid Bodies: The Case of Infringing Content Monitoring' (2021) 23 University of Pennsylvania Journal of Constitutional Law 351, 368; Jonathon Penney, 'Privacy and Legal Automation: The DMCA as a Case Study' (2019) 22 Stanford Technology Law Review 412; Martin Husovec, 'The Promises of Algorithmic Copyright Enforcement: Takedown or Staydown: Which Is Superior: And Why' (2018) 42 Columbia Journal of Law & Arts 53.

[97] Thomas Riis and Sebastian Felix Schwemer, 'Leaving the European safe harbor, sailing towards algorithmic content regulation' (2019) 22 Journal of Internet Law 1; Sabine A. Einwiller and Sora Kim, 'How Online Content Providers Moderate User-Generated Content to Prevent Harmful Online Communication: An Analysis of Policies and Their Implementation' (2020) 12 Policy & Internet 184.

[98] Paddy Leerssen, 'An End to Shadow Banning? Transparency rights in the Digital Services Act between content moderation and curation' (2023) 48 Computer Law & Security Review 105790.

[99] Ibid, at p.4.

[100] Tarleton Gillespie, 'Do not recommend? Reduction as a form of content moderation' (2022) 8 Social Media+ Society 20563051221117552.

[101] Sebastian Felix Schwemer and Jens Schovsbo, 'What is Left of User Rights? – Algorithmic Copyright Enforcement and Free Speech in the Light of the Article 17 Regime' (2019). In Paul Torremans (ed), *Intellectual Property Law and Human Rights*, (4th edn, Wolters Kluwer 2020), pp. 569-589 <http://dx.doi.org/10.2139/ssrn.3507542> Accessed 1 May 2023; See Amélie Heldt, 'Borderline speech: Caught in a free speech limbo' (2020) 15 Internet Policy Review 1; Zeng, Jing, and D. Bondy Valdovinos Kaye, 'From content moderation to visibility moderation: A case study of platform governance on TikTok' (2022) 14 Policy & Internet 79.

[102] Cobbe Jennifer, 'Algorithmic censorship by social platforms: Power and resistance' (2021) 34 Philosophy & Technology 739.

[103] Blake Hallinan, Rebecca Scharlach and Limor Shifman, 'Beyond Neutrality: Conceptualizing Platform Values' (2022) 32 Communication Theory 201; See Community Self-Discipline Convention of Douyin (n 77).

[104] Mandy Zuo, 'Kris Wu removed from Chinese social media and nearly 1,000 supporters' accounts meet the same fate following rape allegations' (South China Morning Post, 2 August 2 2021,) <https://www.scmp.com/news/people-culture/china-personalities/article/3143525/kris-wu-removed-chinese-social-media-and> accessed 1 May 2023.

[105] Langvardt (n 74).

[106] 'China to Cleanse Online Content That 'Bad-Mouths' Its Economy' (Bloomberg, 28 August 2021) <https://www.bloomberg.com/news/articles/2021-08-28/china-to-cleanse-online-content-that-bad-mouths-its-economy?leadSource=uverify%20wall> accessed 1 May 2023.

[107] Officials ordered to target celebrity fan groups and money worship in drive to clean up online content. See Ryan McMorrow, 'China launches internet 'purification' campaign for lunar new year' (Financial Times, 25 January 2022) <https://www.ft.com/content/285059f7-3f0e-4083-b01f-02e48eccff88> accessed 1 May 2023.

[108] Xinmei Shen, 'Weibo will cap share counts to fight fake traffic (but government accounts are exempt)' (South China Morning Post, 9 January 2019) <https://www.scmp.com/abacus/culture/article/3029087/weibo-will-cap-share-counts-fight-fake-traffic-government-accounts> accessed 1 May 2023.

blocked in practice.

Most platforms ignore due process and transparency since no laws or regulations mandate them to disclose how they put their content moderation policies and procedures into everyday practice.[109] Contending that platform rules are inequitable under the doctrine of unconscionability and abusive clauses in relation to standard terms, or asserting that platform sanctions are unwarranted due to excessive contractual breach liabilities, is generally improbable to garner legal backing.[110]

3.3. Expansive public law monitoring obligations in private law judicial practices

The above analysis shows that, practically, the monitoring obligations under public law conflict with 'no general monitoring obligations' principle under private law when ISPs conduct content moderation on their services. Consequently, the scope of platforms' liability for illegal content uploaded by third parties might be affected in both 'positive' and 'negative' ways.[111]Specifically, the courts misinterpreted the monitoring obligation set by an explicit statutory requirement of public law as a duty of care, thus turning the safe harbor into an empty shell. In addition, fulfilling public law monitoring obligations may expose platforms to civil liability due to their actual knowledge concerning the existence of infringing content.

3.3.1. Conflict between public and private law

In China, the legislative and judicial authorities have reached a consensus on the prohibition of general monitoring obligations in private sphere.[112] Yet, monitoring obligations established in public law conflict with the 'no general monitoring obligations' principle in private law. Public law monitoring obligations encompass not only content that violates public law norms, but also content that violates private law norms.[113] Under private law, infringing content is subject to notice-and-takedown mechanism, it may, however, violate 'Eleven Boundaries' stipulated in administrative regulations and thus fall within the scope of the public law monitoring obligation.

In fact, the overinclusive monitoring obligations under public law have given rise to legal conflicts that unfairly distorted the knowledge-based standards establishing secondary liability. In judicial practice, courts directly interpreted the public law monitoring obligation into a duty of care, and determined that ISPs failed to fulfill its duty of care where they failed to perform public law monitoring obligations against online illegal content.[114] The logic behind such legal reasoning indicates that, by virtue of their public law monitoring obligation, ISPs are presumed to have a corresponding monitoring obligation under private law. More importantly, courts implied that platforms should bear civil liability if they failed to perform their monitoring obligations. Such unreasonable decisions not only imposed unduly heavy-headed burdens on platforms but also eroded the distinction between public law monitoring obligations and private law monitoring obligations.

3.3.2. A higher duty of care arising from public law monitoring obligations

In certain exceptional circumstances, the level of duty of care for ISPs may be significantly elevated, resulting in constructive knowledge with regard to potential infringements.[115] For example, an ISP providing the information storage space service has constructive knowledge of a user's infringement of the right of communication to the public on information networks[116], if the ISP substantially accesses the disputed content of popular movies and TV series or establishes a dedicated ranking for them on its own initiative.[117] When performing their public law monitoring obligations, whether an ISP would be considered to have substantially accessed third-party content by monitoring or reviewing it, and thus be required to assume a higher level of duty of care[118], remains unanswered in this judicial interpretation.

However, Chinese courts have held that, when reviewing the legality of uploaded contents, the human reviewer can make preliminary judgments on whether the content infringes on the rights of others by drawing upon their common sense and professional expertise.[119] The Beijing Internet Court ruled that, in order to comply with the monitoring obligation set in administrative regulations, the defendant, a video sharing provider, is obliged to monitor and review the uploaded content to prevent the dissemination of illegal content. The court further explained that, 'although such monitoring does not directly target copyright infringing content, it is not difficult for a professional video sharing provider, to be aware that uploading a whole movie to its website has a high risk of infringing upon others' copyright.'[120] Therefore, the courts held the defendant liable as it had constructive knowledge of the infringement and failed to perform its duty of care. The legal reasoning in this decision implies that, since ISPs must fulfill their public law monitoring obligations by monitoring illegal content, they should also be aware of potential copyright infringement within the content being monitored.

Therefore, platforms are faced with the dilemma that, if they fail to fulfill their monitoring obligation set by public law, they are deemed to

[109] Douyin did publish its transparency report every month, the content, however, only revealed the rough number of removed content and suspended account, rather than the specific standard or criteria of monitoring. See Transparency Report of Quarter 3 of 2022, <https://www.douyin.com/transparency> accessed 1 May 2023; While Weibo and Tencent do not publish either transparency reports or standard of monitoring. Weibo only publicizes decisions on disputed content involving 'misinformation,' 'personal attack,' and 'infringement of personal rights.'

[110] *Announcement of the State Administration for Industry and Commerce on Issuing the Guidelines for Regulating the Standard Terms of Online Trading Platform Contracts* (工商总局关于发布网络交易平台合同格式条款规范指引的公告), issued by State Administration for Industry & Commerce on 30 July 2012.

[111] Positive approach refers to the standards establishing secondary liability; negative approach refers to immunity provisions precluding liability. See Graeme B. Dinwoodie, '*A comparative analysis of the secondary liability of online service providers*' (Springer International Publishing, 2017).

[112] See Section 2.1.

[113] See Article 15 of *Administrative Measures for Internet Information Services* (n 53).

[114] *Suzhou Intermediate People's Court [2004] Su Zhong Min San Chu Zi No.098 Civil Judgement* (2004)苏中民三初字第098号民事判决书; *Guangzhou Intermediate People's Court [2008] Sui Zhong Fa Min San Zhong Zi No.119 Civil Judgement* (2008)穗中法民三终字第119号民事判决书.

[115] Supreme People's Court (n 43).

[116] Article 12(10) of *Chinese Copyright Law* 2020.

[117] In practice, courts have included all professionally produced film and TV series within the scope of the 'red flag.' *See Beijing IP Court [2021] Jing 73 Min Zhong No.220 Civil Judgement* (2021)京73民终220号民事判决书.

[118] The higher level of duty of care is a problematic term. For example, in the first NFT related case, the Hangzhou Internet Court ruled that NFT trading platforms should bear a higher level of duty of care to introduce ex-ante monitoring mechanisms to conduct a preliminary review of the authenticity and legality of NFTs traded and prevent potential infringement at source. Apparently, the higher level of duty of care interpreted by Hangzhou Internet Court refers to a de facto general monitoring obligation. See *Hangzhou Internet Court [2022] Zhe 0192 Min Chu No. 1008 Civil Judgement* (2022)浙0192民初1008号民事判决书. See also Baiyang Xiao, 'Copyright law and non-fungible tokens: experience from China' (2022) 30 International Journal of Law and Information Technology 444.

[119] *Shanghai Higher People's Court [2008] Hu Gao Min San (Zhi) Zhong Zi No.62 Civil Judgement* (2008) 沪高民三(知)终字第62号民事判决书.

[120] *Beijing Internet Court [2019] Jing 0491 Min Chu No.16240 Civil Judgement* (2019)京0491民初16240号民事判决书.

have committed a fault that contributes to the occurrence of the infringement, for which they must assume administrative liability;[121] while they need to conduct ex ante monitoring of content uploaded to fulfill the monitoring obligation set by public law, which means they have had constructive knowledge of the existence of infringing content and thus may bear a higher level of duty of care. Upon the existence of infringing content on a platform, there is a high probability that it will be considered to have constructive knowledge regarding the existence of such content and thus be held liable. That said, platforms risk losing their safe harbor protection if they take proactive measures to address illegal and harmful content.

3.4. Successful public law monitoring obligation, but at what cost?

Currently, the regulation of moderate content serves as a 'policy lever' used by public authorities to obtain control over tech powerhouse.[122] At the same time, platforms are vested with a potent power, which has substantially mitigated illicit online content to a large extent.[123] However, this has accelerated the fragmentation of online law enforcement and generated the need for algorithmic recommendation and filtering systems.[124] In the long run, excessively vague rules, inconsistent enforcement, together with excessive reliance on algorithms will render the expansive collateral censorship of online content an inevitable failure, since it burdens ISPs with significant compliance costs and impacts freedom of expression, access to information and media pluralism at large.[125]

According to Article 47 of the *Cybersecurity Law*, ISPs bear the obligation to tackle illegal information once 'discovered.' However, the absence of a clear explanation on how to perform the obligation of 'discovery' brings legal uncertainty to ISPs, because adopting different standards may have different impacts on the cost from platform operation. Illegal content can be primarily 'discovered' through notices or complaints of users and rights holders, monitoring activities by ISPs, and orders from competent administrative agencies.[126] Having said that ISPs have to invest significant resources in implementing effective mechanism to not only discover illegal and harmful content on their services but respond to notifications from users and orders from authorities. The significant cost might not be a huge burden for giant tech powerhouses like Tencent, Weibo and Douyin, who 'are willing to' invest in monitoring measures to amass more control over the flow of online information.[127] Yet, such a high threshold will keep median, small and start-up platforms from competing in the market.[128]

Additionally, to perform their public law monitoring obligations, platforms have to inevitably conduct a preliminary assessment of the legality of information before taking any further action when faced with specific information. Indeed, given the highly abstract and over-inclusive language used in relevant norms and the vast quantity and diverse nature of information posted by users, platforms, especially smaller ones and startups, often lack the professional expertise and capabilities to evaluate the legality of the vast array of content they encounter.[129] For example, without further guidance or clarification, when faced with content related to local beliefs, an ISP can hardly make a decision on whether the specific content propagates feudal superstitions within the 'Eleven Boundaries,' or promotes national cultures.[130]

Furthermore, scholars observed that, in practice, law enforcement agencies are prone to fall into 'results-oriented' reasoning due to the lack of clear explanation and guidance on performing monitoring obligations.[131] That is, they often presume that ISPs failed to fulfill monitoring obligations based on the result of illegal content existing on their platform without further investigation.[132] For instance, in 2017, irrespective of considerations such as due process and proportionality assessments, Cyberspace Administration of Guangdong directly decided that Tencent failed to perform its public law monitoring obligations based on the existence of content on its service.[133] Under the significant pressure of outcome-oriented reasoning, platforms would tend to over-block content in an attempt to avoid any possible suggestion of liability.[134]

Finally, platforms are imposed with excessive monitoring obligations under public law on the one hand, and entrusted with massive power to govern content on their services on the other hand.[135] Without properly designed procedural safeguards and complaint mechanisms, the power to monitor and the right to report can be easily abused. However, neither

[121] Article 20 of *Provisions on the Administration of Private Network and Targeted Communication Audiovisual Program Services* (n 67) requires ISPs to monitor legality of all programs before dissemination. That said, if ISPs fails to establish an ex-ante monitoring mechanism to monitor and review uploaded content, they will be deemed to have violated their administrative obligations and will be held legally accountable for illegal content on their services.

[122] Robert Gorwa, Reuben Binns, and Christian Katzenbach, 'Algorithmic content moderation: Technical and political challenges in the automation of platform governance' (2020) 7 Big Data & Society 2053951719897945.

[123] See Annual Online Content Governance Research Group, 'Element-based Governance and Relationship Coordination: Online Content Governance Report 2021 (要素治理与关系协调—2021年网络内容治理报告)' (2022) <https://jil.nju.edu.cn/DFS//file/2022/01/25/20220125152613006 2qxi84.pdf> accessed 1 May 2023.

[124] Tianxiang He, 'Online content platforms, copyright decision-making algorithms and fundamental rights protection in China' (2022) 14 Law, Innovation and Technology 71.

[125] Wu (n 70) 296 ('The unique harm of collateral censorship, as opposed to self-censorship, lies in the incentives that intermediaries have to suppress more speech than would be suppressed by original speakers'); Gorwa, Binns, and Katzenbach (n 122); Daphne Keller, 'Facebook Filters, Fundamental Rights, and the CJEU's Glawischnig-Piesczek Ruling' (2020) 69 GRUR International 616; Pamela Samuelson, 'Pushing Back on Stricter Copyright ISP Liability Rules' (2020) 27 Michigan Technology Law Review 299.

[126] Article 14, 28, 47 of *Cybersecurity Law*.

[127] Julie E. Cohen, 'Law for the Platform Economy' (2017) 51 U.C. Davis Law Review 133, 175 ('compromises that involve voluntary filtering shift much day-to-day authority over interdiction of information flows to platforms and at the same time make interdiction decisions more difficult to contest.').

[128] As Tianxiang He suggests that 'once internet users find a platform that is lenient or insensitive towards certain illegal content and they switch platforms, then censorship problems will soon become obvious in that platform.' See He (n 124) 88.

[129] Aleksandra Kuczerawy, 'General monitoring obligations: a new cornerstone of Internet regulation in the EU?' in KU Leuven Centre for IT and IP law (ed.), *Rethinking IT and IP Law* (Wellington, Intersentia, 2019).

[130] Especially in contemporary China, many local beliefs that were stigmatized as 'feudal superstitions', now are promoted as China's national intangible cultural heritage. See Ziying You, 'Conflicts over Local Beliefs' (2020) 79 Asian ethnology 137.

[131] Zhiwei Yao, 'Technical Monitoring: the dilemma of public law monitoring obligation of ISPs (技术性审查-网络服务提供者公法审查义务困境之破解)' (2019) 1 Journal of Legal Studies 31.

[132] Ibid.

[133] Tencent was accused of hosting content containing violence, terrorism, false information, obscenity and pornography that endangered national security, public safety and social order 'Cyberspace Administration of Guangdong ordered Administrative Penalties for Tencent's Violation of the Cybersecurity Law (广东省网信办对腾讯公司违反《网络安全法》作出行政处罚)' (China Daily, 25 September 2017) <http://china.chinadaily.com.cn/2017-09/25/content_32467839.htm> accessed 1 May 2023.

[134] Zittrain (n 29) (Over-removal incentives are likely to be greatest when platforms fear high regulatory attention that can lead to other costs or business impact, or business-altering injunctions).

[135] Steven L. Schwarcz, 'Private ordering' (2002) 97 Northwestern University Law Review 319; Luca Belli and Jamila Venturini, 'Private ordering and the rise of terms of service as cyber-regulation' (2016) 5 Internet Policy Review 1.

the *Cybersecurity Law* nor the administrative regulations provide a redress mechanism to restrict the platform's power to monitor voluntarily. Even affected parties are absent in the negotiation stage during the making of house rules. Consequently, platform monitoring denies the affected users the due process safeguards and remedies to which they are entitled in the administrative legal process. In addition, stricter content regulation by platforms cannot only disproportionately silence lawful speeches,[136] but also lead to self-censorship as users may limit their expression to avoid any potential negative consequences.[137] Nonetheless, the content reporting system established in administrative regulations is excessively biased towards complainants. No provisions concerning liability for erroneous or even malicious complaints are provided in the relevant regulations, rendering the report and complaint process available at no cost. Due to the lack of a substantial negotiation stage in 'notice and counternotice' mechanism provided in the *Civil Code*, complaints can hardly be fully delivered to the content publishers, thereby indirectly encouraging malicious users to abuse the right to report illegal content.

4. Implications from the EU: a critical analysis of content moderation regulation in the EU

To sufficiently protect individual rights and achieve overall security goals, the EU seeks to play an active role in steering and influencing the implementation of content moderation measures through various regulations.[138] Those sector-specific legislative initiatives are characterized by a limited scope, targeting specific types of illegal content within a specific sub-set of services.[139] However, those initiatives pose a challenge to the long-established principle of prohibition on general monitoring obligations as they indicate a supportive attitude toward monitoring the entirety of uploaded content to fight against illegal online content.[140]

In contrast to the Chinese regulatory landscape, EU regulations pay more attention to due process and transparency regarding the protection of fundamental rights.[141] Most regulations require an evaluation of the adequacy of the safeguards against power abuse and arbitrary decisions when implementing preventive content moderation systems to address illegal content online.[142] For example, Article 12 of DSA also requires that platforms have 'due regard' to the 'fundamental rights' of users under the EU Charter of Fundamental Rights in the enforcement of T&Cs that restrict user-generated content,[143] which targets platforms' abuse of power that the Chinese regulatory approach fails to address.

4.1. How good intentions make bad laws: recent EU regulations on content moderation

Before the introduction of Article 17 of the CDSM, the Commission had proposed broader policy and legislative developments related to a shift towards proactive measures against online illegal content.[144] Acknowledging that a strict and narrow interpretation of the prohibition of general monitoring obligations could be a barrier to effectively tackling illegal online content,[145] regulators repeatedly emphasized the adoption of 'effective proactive measures to detect and remove illegal content online' in multiple policy documents.[146] Moreover, the CJEU departed from the earlier broad interpretation of the concept of general monitoring obligations,[147] rather acknowledged that preventive measures targeting illegal content are ineffective without prior monitoring of all the content transmitted.[148] Besides, various national-level initiatives have imposed more stringent obligations on platforms, requiring them to combat the spread of specific types of illegal content.[149] However, they further add normative fragmentation and legal uncertainty to the already complex EU regulatory landscape, particularly impeding small providers' ability to effectively compete in the market.[150]

In response to the controversial discussion on the need for proactive monitoring obligations,[151] the European lawmakers have introduced several sector-specific rules and guidelines for hosting platforms, most recently the introduction of specific liability rules on video-sharing

[136] Christophe Geiger and Bernd Justin Jütte, 'Platform liability under Article 17 of the Copyright in the Digital Single Market Directive, automated filtering and fundamental rights: An impossible match' (2021) 70 GRUR International 517.

[137] Leslie Kendrick, 'Speech, intent, and the chilling effect.' (2012) 54 William & Mary Law Review 1633; Monica Youn, 'The Chilling Effect and the Problem of Private Action' (2013) 66 Vanderbilt Law Review 1473.

[138] Rocco Bellanova and Marieke de Goede, 'Co-Producing Security: Platform Content Moderation and European Security Integration' (2022) 60 Journal of Common Market Studies 1316. See João Pedro Quintais, Péter Mezei and others, 'Copyright Content Moderation in the EU: An Interdisciplinary Mapping Analysis' (2022). <http://dx.doi.org/10.2139/ssrn.4210278> accessed 1 May 2023.

[139] Ilaria Buri and Joris van Hoboken, 'The Digital Services Act (DSA) proposal: a critical overview' (Digital Services Act (DSA) Observatory, 8 October 2021), at p.8. <https://dsa-observatory.eu/wp-content/uploads/2021/11/Buri-Van-Hoboken-DSA-discussion-paper-Version-28_10_21.pdf> accessed 1 May 2023.

[140] Kuczerawy (n 129); See also Martin Senftleben and Christina Angelopoulos, 'The Odyssey of the Prohibition on General Monitoring Obligations on the Way to the Digital Services Act: Between Article 15 of the E-Commerce Directive and Article 17 of the Directive on Copyright in the Digital Single Market' (2020) Amsterdam/Cambridge <https://ssrn.com/abstract=3717022> accessed 1 May 2023.

[141] João Pedro Quintais, Naomi Appelman, and Ronan Fahy, 'Using Terms and Conditions to Apply Fundamental Rights to Content Moderation' (2022) Forthcoming in German Law Journal <http://dx.doi.org/10.2139/ssrn.4286147> accessed 1 May 2023.

[142] Rojszczak (n 22) 12.

[143] Article 12 of the DSA. See also Quintais, Appelman, and Fahy (n 141).

[144] Sebastian Felix Schwemer, 'Article 17 at the Intersection of EU Copyright Law and Platform Regulation' (2020) Nordic Intellectual Property Law Review, p.30.

[145] (COM(2018)640, Recital 19. The Commission suggested allowing for derogations from the prohibition of general monitoring obligations in certain cases.

[146] COM(2017) 555 final, p.3. Commission Recommendation (EU) 2018/334, pp.50–61.

[147] Imposing an obligation on online intermediaries to monitor and filter all information from all service users for potential infringements is considered to be within the scope of prohibition, as it represents an undue encroachment on the fundamental rights of both online intermediaries and internet users. See *L'Oréal, Scarlet Extended, Netlog, McFadden* case.

[148] See Eva Glawischnig-Piesczek; Peterson/Elsevier v Youtube/Cyando, Opinion of the AG Saugmandsgaard Øe (55) 221; Peterson/Elsevier v Youtube/Cyando (33); Poland v. European Parliament and Council. See also Clara Rauchegger and Aleksandra Kuczerawy, 'Injunctions to remove illegal online content under the e-Commerce Directive: Glawischnig-Piesczek' (2020) 57 Common Market Law Review 1495, 1505

[149] See German Network Enforcement Act (NetzDG) of 30 June 2017; French "Avia" Law 2020-766 of 24 June 2020 on online hateful content.

[150] Buri and Hoboken (n 139) 5.

[151] Thomas Spoerri, 'On upload-filters and other competitive advantages for Big Tech companies under Article 17 of the Directive on Copyright in the Digital Single Market' (2019) 10 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 173, 174; Christina Angelopoulos and João Pedro Quintais, 'Fixing copyright reform: a better solution to online infringement' (2019) 10 Journal of Intellectual Property, Information Technology & Electronic Commerce Law 147.

platforms in cases of hate speech,[152] terrorist content,[153] and copyright.[154] Those scattered regulations echo the hardly contested prohibition of general monitoring obligations,[155] and introduce a *lex specialis* model to general requirements of the E-Commerce Directive.[156] The above legal instruments could constitute a solid ground for the introduction of various preventive content moderation measures to monitor specific or even the entirety of users' activities and uploaded content. One could not help but wonder if the prohibition of general monitoring obligations only exists beneath legal texts. The vagueness, complexity and opaqueness inherent to the wordings of regulations bring more legal uncertainty to the effective protection of fundamental rights throughout the process of moderating illegal content, especially in terms of obligations, responsibilities and regulatory oversight.[157] After all, the goal of all initiatives indicates a good intention to protect online users; the result, however, is rather bad to some extent, particularly with regard to the fundamental rights of users.

As a result, there was an urgent need for new legislation to upgrade the liability rules for intermediary services while effectively protecting the fundamental rights enshrined in the Charter in the EU's internal market.[158] Pursuing to consolidate various separate pieces of EU legislation and self-regulatory practices addressing online illegal and harmful content, the DSA retains the conditional immunity and the prohibition of general monitoring obligations, but further lays down horizontal rules on wide-ranging transparency and due diligence obligations for platforms.[159]

4.2. How to make bad laws into good ones: regulating content moderation under the DSA

Aiming to modernize the existing legal framework for digital services laid down by the E-Commerce Directive, the DSA introduces a general framework for the provision of intermediary services.[160] It adopts a tiered structure with four horizontal layers,[161] and targets different types of obligations on different types of service providers, namely intermediaries, hosting providers, online platforms, and very large online platforms (VLOPs).[162]

For the widest subcategory, all intermediaries are subject to general due diligence obligations, including establishing a single point of contact or designating a legal representative,[163] incorporating certain information in the provider's terms and conditions[164] as well as complying with transparency reporting duties.[165] Notably, Article 12 allows powerful intermediaries to suppress legal content based on their T&Cs, thereby vesting the power of formulating adequate rules for online communication in the intermediaries.[166] The DSA also positions platforms at a 'gordian knot' of fundamental rights and public interest pertaining to various affected stakeholders, namely users, content providers, intermediaries, and states.[167] Particularly, Article 12(2) requires intermediaries to apply the above restriction 'in a diligent, objective and proportionate way' that respects the 'fundamental rights of the recipients of the service as enshrined in the Charter.'[168]

In addition, Article 14 requires providers of hosting services, including online platforms, to implement an easily-accessible and user-friendly notice-and-action mechanism, that allows any individual or entity to notify them of the presence on their service of specific items of information that the individual or entity considers to be illegal content. Moreover, regarding additional obligations applicable to online platforms, the DSA upgrades the internal complaint-handling mechanism and reporting obligations to supervisory authorities.[169] Article 18 introduces out-of-court dispute resolution mechanisms, including the introduction of trusted flaggers and precautions against the abuse of complaints. Noteworthy, a carve-out exception is provided for micro and small enterprises, which means these additional obligations shall not apply to them. For VLOPs, they have to not only undertake the above-mentioned obligations, but also obligations with regard to risk management, data access, compliance, and transparency, as well as the implementation of an independent audit.[170]

Safe harbors for intermediaries and the prohibition of general monitoring obligations laid down in the E-Commerce Directive remain unaffected, even though the corresponding provisions (Article 12–15) are slightly amended and incorporated into the DSA instead.[171] Again, Article 7 of the DSA confirms the prohibition of general monitoring obligations and active fact-finding obligations, and Recital 28 of the DSA confirms that obligations imposed on providers to monitor in specific cases are not against the general monitoring obligations ban. This provision also connects the case law of the CJEU regarding general monitoring obligations: obligations to monitor all content for an indefinite period of time qualifies as a prohibited general obligation,[172] while an obligation to detect and remove specific identical or equivalent content that contains specific elements pre-identified by a national court is not covered by the prohibition.[173]

Article 6 of the DSA incorporates a Good Samaritan provision, promising that intermediaries will not automatically lose immunity from liability 'solely' because they carry out voluntary measures aimed at detecting and removing illegal content in good faith, or take the necessary measures to comply with the requirements of Union law.[174]

[152] Article 28b of Directive 2018/1808.

[153] Directive 2017/541; See also Bellanova and Goede (n 138); Aleksandra Kuczerawy, 'The proposed Regulation on preventing the dissemination of terrorist content online: safeguards and risks for freedom of expression' (2018) For Center for Democracy and Technology, <https://ssrn.com/abstract=32968644> accessed 1 May 2023.

[154] Felipe Romero Moreno, 'Upload filters and human rights: implementing Article 17 of the Directive on Copyright in the Digital Single Market' (2020) 34 International Review of Law, Computers & Technology 153; Senftleben and Angelopoulos (n 140).

[155] Article 17(8) of Directive (EU) 2019/790; Article 5(8) of Regulation (EU) 2021/784.

[156] Rojszczak (n 22) 9.

[157] Ibid.

[158] Article 1(1) of DSA.

[159] See Giancarlo Frosio and Christophe Geiger, 'Taking Fundamental Rights Seriously in the Digital Services Act's Platform Liability Regime' (2022) forthcoming in the European Law Journal. <https://ssrn.com/abstract=3747756> accessed 1 May 2023. See also Quintais, Appelman, and Fahy (n 141).

[160] Folkert Wilman, 'The Digital Services Act (DSA)-An Overview' (2022) <http://dx.doi.org/10.2139/ssrn.4304586> accessed 1 May 2023.

[161] The distinction of four different categories of intermediary services is based on the size of services, namely number of monthly active users. Once the monthly active users reach 45 million, online platforms thus fall within the scope of VLOPs under this regulation. See Article 33 and Recital 76 of DSA.

[162] Alexander Peukert, Martin Husovec, Péter Mezei and others, 'European Copyright Society–Comment on Copyright and the Digital Services Act Proposal' (2022) 53 IIC-International Review of Intellectual Property and Competition Law 358.

[163] Article 10 and 11 of DSA.

[164] Article 12 of DSA.

[165] Article 13, 23, 33 of DSA.

[166] Ruth Janal, 'Eyes Wide Open' (Verfassungsblog, 7 September 2021) <https://verfassungsblog.de/power-dsa-dma-15/> accessed 1 May 2023.

[167] Ibid.

[168] Article 12(2) of DSA.

[169] Article 17 of DSA.

[170] Article 25-33 of DSA.

[171] Article 7 of DSA.

[172] Scarlet Extended v. SABAM Case C-70/10; SABAM v. Netlog, Case C-360/10.

[173] Case C-18/18; See Oruç (n 7).

[174] Miquel Peguera, 'The Platform Neutrality Conundrum and the Digital Services Act' (2022) 53 IIC-International Review of Intellectual Property and Competition Law 681.

Recital 25 further clarifies that 'the mere fact that providers undertake such activities does not lead to the unavailability of the exemptions from liability, provided those activities are carried out in good faith and in a diligent manner.'[175] The Good Samaritan protection also applies to 'measures taken to comply with the requirements of Union law, including those set out in this Regulation as regards the implementation of their terms and conditions.'[176]

5. Reposition the gatekeeper: suggestions for future Chinese rulemaking

Even though several concepts concerning crucial obligations adopted remain vague, and guidance on enforcement remains unmentioned,[177] the recent EU legislative initiatives, the DSA in particular, offer thought-provoking and practical insights to improve content governance in China, including the introduction of the Good Samaritan provision detailed in the DSA and transparency obligation, a human rights-centric regulatory system, tiered obligation regimes for intermediaries, and so forth.

5.1. Legal predictability of monitoring obligations

Although the Chinese private law judicial interpretation and guiding opinions have reached a consensus that provides that they are not subject to a general monitoring obligation, a clause expressly stipulating the prohibition of general monitoring obligations is still missing in private law legislation. The consensus is far less solid than a piece of legislation. Consequently, some courts may implement the judicial interpretation based on interpretations that are different or even opposite to the general monitoring obligations ban, thus leading to misunderstandings and chaotic applications in practice.[178]

On the one hand, to better clarify the standpoint of the Legislative Affairs Commission and lessen legal uncertainty, a clause regarding the prohibition of general monitoring obligations should be explicitly introduced in the form of a judicial interpretation by the Supreme People's Court. On the other hand, monitoring obligations under public law should be further limited to ensure the fundamental rights of users and avoid overly intrusive interference by authorities. Specifically, the scope of monitoring should be refined to the extent that the standards for determining illegality are distinct and practical to meet current available technology.[179] That is to say, the permissible monitoring must not require platforms to assess the legality of content, and should target online content that has been previously identified as illegal by national authorities, or is manifestly illegal for a reasonable person.[180] Considering the distinctive dual-track approach concerning monitoring obligations, private sphere should be excluded from the scope of public law monitoring, while public law monitoring obligations are applicable merely to public law issues, namely the illegal content listed in 'Eleven Boundaries.'

5.2. A conditional 'Good Samaritan' protection

As analyzed in previous chapters, both public law monitoring obligations and voluntary monitoring activities could lead to awareness of facts or circumstances from which an illegal activity or information is apparent, therefore, to obtaining constructive knowledge. Failing to remove content that was reviewed and monitored may still result in administrative and civil liability as platforms 'knew' or 'should have known' about the illegality through monitoring.[181] The absence of Good Samaritan protection results in platforms excessively removing content when monitoring activities and content on their services.[182] Moreover, unsuccessful monitoring under public law will not only block legal content, but also result in civil liability. To provide not only legal predictability for affected parties, but also flexibility for future technological developments, a Good Samaritan clause should be introduced in the *Cybersecurity Law* by referring to Article 7 of the DSA.

Noteworthy, there is a major difference between the Good Samaritan Clause in the CDA and the DSA. The former regulation provides intermediaries with full protection when they do not act against illegal content covered by Section 230(c), regardless of whether they have knowledge of it or not.[183] In another word, Section 230 not only protects platforms from liability for failing to remove harmful or illegal content, it also protects them from liability for engaging in the removal of potentially harmful or illegal content, provided the measures are taken in good faith.[184] With this absolute assurance, platforms are incentivized to adopt voluntary monitoring measures. However, Section 230 is not a perfect piece of legislation, as it may be overprotective in some respects and under-protective in others.[185] By tracing the historical background of CDA, Jeff Kosseff summarized two enduring purposes of Section 230 as 'providing platforms with the flexibility to moderate' and 'promoting free speech and online innovation by helping platforms to flourish.'[186] Scholars also suggest that an overbroad reading of Section 230 gives free passes to ignore abusive Bad Samaritans' illegal activities while ensuring that abusers cannot be identified, thus devaluing the efforts of the latter purpose,[187] and at the same time may result in excessive removal on intermediaries' own initiatives in practice.[188]

In a different way, the European Good Samaritan Clause may also lead to certain disadvantages. Recital 25 states that 'any such activities and measures that a given provider may have taken should not be taken into account when determining whether the provider can rely on an

[175] Recital 25 of DSA.
[176] Ibid.
[177] Aina Turillazzi, Federico Casolari, Mariarosaria Taddeo and others, 'The digital services act: an analysis of its ethical, legal, and social implications' (2023) 15 Law, Innovation and Technology 83; Miriam C Buiten, 'The Digital Services Act From Intermediary Liability to Platform Regulation' (2021) 12 Journal of Intellectual Property, Information Technology & Electronic Commerce Law 361.
[178] See Section 3.3.
[179] See generally João Pedro Quintais, Christian Katzenbach, Péter Mezei and others, 'Copyright Content Moderation in the EU: Conclusions and Recommendations' (reCreating Europe Report. March 2022) <https://ssrn.com/abstract=4403423> accessed 1 May 2023.
[180] Mendis and Frosio (n 25); Opinion of the AG Saugmandsgaard ØE in case C-401/19.
[181] Aleksandra Kuczerawy, 'The Good Samaritan that wasn't: voluntary monitoring under the (draft) Digital Services Act' (Verfassungsblog, 12 January 2021) <https:// verfassungsblog.de/good-samaritan-dsa/> accessed 1 May 2023.
[182] Danielle Keats Citron and Benjamin Wittes, 'The internet will not break: Denying bad samaritans sec. 230 immunity' (2017) 86 Fordham Law Review 401, See also Kuczerawy (n 129).
[183] The Section 230 of CDA established a blanket immunity which covers any kind of content excluding from its scope only federal criminal law, state or federal sex trafficking law, intellectual property law and communication privacy law. See Goldman (n 9).
[184] Ashley Johnson and Daniel Castro, 'How Other Countries Have Dealt With Intermediary Liability' (2021) Information Technology and Innovation Foundation <https://itif.org/publications/2021/02/22/how-other-countries-have-dealt-intermediary-liability/> accessed 1 May 2023.
[185] Jack M. Balkin, 'The Future of Free Expression in a Digital Age' (2008) 36 Pepperdine Law Review 427, 434.
[186] Jeff Kosseff, 'What Was the Purpose of Section 230? That's a Tough Question, a Response to Danielle Citron's How to Fix Section 230' (2023) Forthcoming in Boston University Law Review, <https://ssrn.com/abstract=4388216> accessed 1 May 2023.
[187] Citron and Wittes (n 182) 423.
[188] Benjamin Volpe, 'From innovation to abuse: does the Internet still need section 230 immunity' (2019) 68 Catholic University Law Review 597; Ben Horton, 'The Hydraulics of Intermediary Liability Regulation' (2021) 70 Cleveland State Law Review 201.

exemption from liability.' Having said that, adopting voluntary measures in good faith and in a diligent manner neither guarantees nor precludes neutrality, and they may still lose immunity.[189] The question of whether the unsuccessful outcome of voluntary actions undertaken by providers would fall into the scope of 'diligent manner' under this provision remains unclear and needs to be determined on a case-by-case basis.[190] Furthermore, Recital 22 states that platforms' own-initiative investigations could trigger actual knowledge or awareness of illegal content, thus resulting in losing safe harbor protection.[191] In other words, implementing proactive monitoring measures strengthens providers' capability to discover illegal content, which in turn further increases the probability of their exposure to liability.

Considering the reality of the Chinese internet industry, this paper argues that powerful platforms in Web 3.0 era no longer need strong protectionism in Web 2.0. That is, the reading the of Good Samaritan Clause should not be overbroad. A platform may lose its Good Samaritan immunity status when it engages in bad faith like a Bad Samaritan or fails to conduct diligent self-regulation.[192] Therefore, when providers undertake voluntary monitoring measures or fulfills their public law monitoring obligation in good faith and in diligent manner, its private law duty of care should not be affected and the legitimate safe harbor protection should not be deprived. It should be clarified that platforms should not be liable for good-faith unsuccessful monitoring, either voluntarily or to perform public law monitoring obligations. However, if they intentionally or knowingly promote, endorse, or maintain manifestly illegal content that they actually know or have awareness of, Good Samaritan protection should not be extended to them. Of course, rule-making authorities need to provide more specific details about the connotations of 'good faith' and 'diligence.' Moreover, to strike a fair balance between the interests of platforms and users, the above liability exemption under the Good Samaritan clause should be limited to monetary damages, while affected users could still require platforms to stop infringing activities.

5.3. Reduce platforms' concentrated power over speech

There is plenty of lawful but awful content spreading over the internet, ranging from discriminatory speech to medical misinformation.[193] The DSA did not require platforms to moderate such content by prescribing new content prohibitions, but rather regulated the systems and processes by which platforms enforce their own house rules.[194] That is to say, platforms are regarded as a mini-government assigned with the power to define and moderate harmful content within their house rules.[195] Since substantiated notices constitute actual knowledge for the purposes of the hosting immunity under Article 5, hosting providers have a strong incentive to remove content upon effective notices.

However, entrusting content moderation to private actors with market influence may not always be an optimal choice, given the significant concentration of power over internet users' speech that this entails.[196] Utilizing the power of content moderation vested in public law, the Chinese mega platforms further extend the scope of monitoring from illegal content defined by administrative laws to undesirable content through their house rules, thus leading to a host of legal concerns, such as disproportionately undermining freedom of expression, access to information, and media pluralism.[197] In turn, the Chinese experience may very well serve as a warning for EU regulators that discretion and power over fundamental rights granted to platforms should be limited.

Any regime that imposes liability on speech intermediaries should comply with constitutional and human rights safeguards.[198] Intermediary liability laws' restrictions on core democratic freedoms such as freedom of communication, speech, and association, as well as the right to privacy, must be necessary, proportionate, and provided for by law.[199] Rather than imposing stringent liability on platforms for user-generated content or mandating comprehensive content monitoring, contemporary platform regulation ought to concentrate on establishing norms for platforms' operational procedures, including modifications to terms of service and algorithmic decision-making processes.[200] Accountable governance, such as necessary notifications and disclosures to users whenever platforms change their terms of service, can help reduce the information asymmetry between users and powerful gatekeeper platforms.[201] Meanwhile, users should be empowered to better understand how they can notify platforms about both problematic content and problematic takedown decisions and should be informed about how content moderation works on large platforms.[202] Privacy by default, improved transparency, and procedural safeguards, such as due process and effective redress mechanisms for removal or blocking decisions, can help to ensure the protection of fundamental rights online.[203]

5.4. A 'differentiated' liability regime

Before the DSA, most legislative initiatives to regulate content moderation reasonably targeted large platforms like Facebook or Google, however, in practice, these initiatives apply to all types of platforms

[189] Kuczerawy (n 181).
[190] Ibid.
[191] Joris van Hoboken, João Pedro Quintais, Joost Poort and others, 'Hosting intermediary services and illegal content online: An analysis of the scope of Article 14 ECD in light of developments in the online service landscape. Final report prepared for the European Commission' (2018) <https://op.europa.eu/nl/publication-detail/-/publication/7779caca-2537-11e9-8d04-01aa75ed71a1> accessed 1 May 2023.
[192] Andrew M. Sevanian, 'Section 230 of the communications decency act: A good samaritan law without the requirement of acting as a good samaritan' (2014) 21 UCLA Entertainment Law Review 121.
[193] Gillespie (n 29); See also Eric Goldman and Jess Miers, 'Online Account Terminations/Content Removals and the Benefits of Internet Services Enforcing Their House Rules' (2021) 1 Journal of Free Speech Law 191.
[194] Daphne Keller, 'The EU's new Digital Services Act and the Rest of the World' (Verfassungblog, 7 November 2022) <https://verfassungsblog.de/dsa-rest-of-world/> accessed 1 May 2023; Amélie P. Heldt, 'EU digital services act: The white hope of intermediary regulation' In Terry Flew and Fiona R. Martin (ed.), *Digital Platform Regulation: Global Perspectives on Internet Governance*. (Springer International Publishing, 2022), at pp. 69-84.
[195] Janal (n 166).
[196] Daphne Keller, 'Lawful but Awful? Control over Legal Speech by Platforms, Governments, and Internet Users' (The University of Chicago Law Review Online, 28 June 2022) <https://lawreviewblog.uchicago.edu/2022/06/28/keller-control-over-speech/> accessed 1 May 2023.
[197] PEN America (n 69).
[198] Christoph Schmon and Haley Pedersen, 'Platform Liability Trends Around the Globe: Moving Forward' (8 June 2022, Electronic Frontier Foundation). <https://www.eff.org/deeplinks/2022/05/platform-liability-trends-around-globe-conclusions-and-recommendations-moving> accessed 1 May 2023; See also Schwemer and Schovsbo (n 101).
[199] Evelyn Douek, 'Governing online speech: From "posts-as-trumps" to proportionality and probability' (2021) 121 Columbia Law Review 759.
[200] Gillespie (n 29).
[201] Daphne Keller and Paddy Leerssen, 'Facts and where to find them: Empirical research on internet platforms and content moderation' in Nathaniel Persily and Joshua A. Tucker (ed.), *Social Media and Democracy: The state of the Field and Prospects for Reform* (CUP 2020), at p.224.
[202] Paddy Leerssen, 'The Soap Box as a Black Box: Regulating transparency in social media recommender systems' (2020) 11 European Journal of Law and Technology 1.
[203] Giovanni De Gregorio, 'Democratising online content moderation: A constitutional framework' (2020) 36 Computer Law & Security Review 1. See also Giovanni De Gregorio, 'The rise of digital constitutionalism in the European Union' (2021) 19 International Journal of Constitutional Law 41; Quintais, Appelman, and Fahy (n 141).

and services.[204] The power of the largest platforms will be further consolidated, since only the largest platforms have the resources to meet the requirements crafted.[205] Moreover, large platforms may be able to save on costs of detection, monitoring and removal because of economies of scale.[206] Therefore, it is always a challenging task for regulators to ensure that the rules are both effective in combating illegal content online while remaining achievable by platforms of all sizes.

Given the varying costs and benefits associated with controlling illegal content online across different platforms and types of content, a one-size-fits-all liability rule is untenable. Generally, distinctions in the size, reach, technical design and business model of the platform as well as the type of illegal material necessitate distinct liability guidelines. Theoretically, any meaningful reform of ISP liability rules should consider the interests of a wide range of stake holders.[207] The duty of care ascribed to online platforms should be nuanced, with consideration given to the type of illegal material and the type of harm it generates.[208]

Regarding the size of platforms, the tiered system of obligations adopted in the DSA indicates that, with greater economic power and societal influence, come more additional responsibilities. The future Chinese regulations may follow this approach and adopt tailored obligations on different platforms in accordance with the types and scale of services. Even though it might be a complicated task to figure out which type of platform should bear what obligations, more clearly articulated obligations will prevent abuse of power to a certain degree. With regard to the threshold for classification of platforms, in addition to the reasonable number of monthly active users, other factors that reflect providers' power and influence on flow of information should also be taken into consideration when determining the threshold for large or small providers.

Under the DSA, the detailed procedural steps will waste resources that could better be spent elsewhere, and burden smaller platforms to a degree that effectively sacrifices competition and pluralism goals in the name of content regulation.[209] Moreover, effective content moderation requires more investment in knowledge and expertise, and the spectacular failures of some small platforms and startups suggest that this knowledge is often gained too late, or not at all.[210] Thus, a cost-and-benefit analysis should be adopted when assigning obligations to platforms. For example, those costly responsibilities, including public law monitoring obligations, shall not apply to smaller providers, as they are unable to afford the cost of additional responsibility and might be kept from competing in markets.[211]

Furthermore, the burden of detection and removal of illegal material online should be fairly shared among the different parties involved. Therefore, the optimal level of monitoring obligations should be tailored according to the specific category of illegal content, such as serious crime, highly recognizable information. In terms of serious crimes that may inflict server harm, such as terrorist content that threatens national security or contains child sexual abuse, it is necessary for society as a whole to adopt preventative actions in order to take them down. Additionally, the scope of such monitoring should be confined to highly recognizable content that does not require platforms to conduct independent assessments on its legality. In response to these concerns, both CJEU and ECtHR limit the scope of proactive measures against manifestly illegal content that would not require the online intermediary to conduct any legal assessment. Moreover, they allow the imposition of such measures on financially and technically resourceful intermediaries who have influence over the curation of content, as opposed to simply hosting them.[212]As for harmful but lawful content, platforms are encouraged to adopt less restrictive content moderation practices, such as labeling, providing contextual information in relation to disinformation, and de-monetization.[213]

6. Conclusion

In practice, the distressingly calibrated and heavy-handed public law monitoring obligations can be easily abused, which might not only impose a heavy burden on ISPs, but also bring more profound legal uncertainty and complexity to an already fragmented content regulatory landscape. The DSA sets out great examples for regulators in China. However, it is worth noting that transplanting foreign laws or legal institutions does not take place in a legal cultural vacuum, as the path dependence for legal transplantation might impede the well-functioning of transplanted rules or institutions.[214] As Daphne Keller envisioned, the DSA should be perceived as a starting point, rather than an end point in the process of deliberating potential national legal reform.[215]

In China, the debate revolving around platform responsibility reaches beyond how governments seek to regulate platforms. Rather, the scope and range of research on content moderation could reach the 'rule of platform,' which refers to rules on how platforms *de facto* regulate the availability, accessibility, and visibility of online information.[216] Nonetheless, as Bambauer succinctly observes that, 'it is not clear that censorship should occur; rather, it is clear that it is occurring.'[217] The comprehensive censorship system is deeply rooted in every aspect of cyberspace in China, and this paper contends that analysis of the current rules' goals, as well as where and why they fall short, is important for refuting those who propose increased censorship as a means of cleaning up the internet elsewhere. Moreover, monitoring should be conducted in an open, transparent, and narrowly targeted way with due process safeguards.[218] Moreover, other than technical innovation and economic growth, special attention should be paid to the protection of speech and access to information. After all, in what way do we avoid the abuse of such quasi-state powers held by mega platforms and ensure the fundamental rights of users, necessitates a meticulous and thorough analysis.

Declaration of Competing Interest

The author declare that they have no known competing financial interests or personal relationships that could have appeared to influence

[204] Gillespie et al (n 86); Keller (n 194).
[205] Kosseff (n 186).
[206] Miriam C. Buiten, Alexandre De Streel, and Martin Peitz, 'Rethinking liability rules for online hosting platforms' (2020) 28 International Journal of Law and Information Technology 139, 153; Keller (n 194).
[207] Samuelson (n 125).
[208] Buiten, Streel, and Peitz (n 206) 162.
[209] Daphne Keller, 'The DSA's Industrial Model for Content Moderation' (Verfassungsblog, 24 February 2022) <https://verfassungsblog.de/dsa-industrial-model/> accessed 1 May 2023.
[210] Gillespie et al (n 86), at pp.10-11.
[211] Janal (n 164).
[212] Julia Reda, Joschka Selinger, and Michael Servatius, 'Article 17 of the Directive on Copyright in the Digital Single Market: a fundamental rights assessment' (2020) <http://dx.doi.org/10.2139/ssrn.3732223> accessed 1 May 2023; Geiger and Jütte (n 136).
[213] João Pedro Quintais, Giovanni De Gregorio, and João C. Magalhães, 'How platforms govern users' copyright-protected content: Exploring the power of private ordering and its implications' (2023) 48 Computer Law & Security Review 105792.
[214] Jaakko Husa, 'Developing Legal System, Legal Transplants, and Path Dependence: Reflections on the Rule of Law' (2018) 6 The Chinese Journal of Comparative Law 129.
[215] Keller (n 194, n 196).
[216] Lena Ulbricht and Karen Yeung, 'Algorithmic regulation: A maturing concept for investigating regulation of and through algorithms' (2022) 16 Regulation & Governance 3; Quintais, Gregorio, and Magalhães (n 213).
[217] Derek E Bambauer, 'Orwell's armchair' (2012) 79 University of Chicago Law Review 863, 869.
[218] Ibid. See also John Tehranian, 'The New ©ensorship' (2015) 101 Iowa Law Review 245, <https://ssrn.com/abstract=2514224> accessed 1 May 2023.

the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

The author is very grateful to Prof. Peter Mezei, Prof. Orit Fischman-Afori, and the anonymous reviewers for their valuable feedback and support. The author acknowledges the support of China Scholarship Council (Grant Number: 202008500106), Tempus Public Foundation of Hungarian Government (Grant Number: SHE-08308–004/2020), Max Planck Institute for Innovation and Competition, and University of Szeged Open Access Fund (Grant Number: 6565).