

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/rjml20>

# How equalitarian regulation of online hate speech turns authoritarian: a Chinese perspective

Ge Chen

To cite this article: Ge Chen (2022): How equalitarian regulation of online hate speech turns authoritarian: a Chinese perspective, Journal of Media Law, DOI: [10.1080/17577632.2022.2085013](https://doi.org/10.1080/17577632.2022.2085013)

To link to this article: <https://doi.org/10.1080/17577632.2022.2085013>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 21 Jun 2022.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

# How equalitarian regulation of online hate speech turns authoritarian: a Chinese perspective

Ge Chen

Durham Law School, Durham University, Durham, UK

## ABSTRACT

This article reveals how the heterogeneous legal approaches of balancing online hate speech against equality rights in liberal democracies have informed China in its manipulative speech regulation. In an authoritarian constitutional order, the regulation of hate speech is politically relevant only because the hateful topics are related to regime-oriented concerns. The article elaborates on the infrastructure of an emerging authoritarian regulatory patchwork of online hate speech in the global context and identifies China's unique approach of restricting political contents under the aegis of protecting equality rights. Ultimately, both the regulation and dis-regulation of online hate speech form a statist approach that deviates from the paradigm protective of equality rights in liberal democracies and serves to fend off open criticism of government policies and public discussion of topics that potentially contravene the mainstream political ideologies.

**KEYWORDS** Authoritarianism; China; equality rights; free speech; online hate speech

## Introduction

The universal adoption of digital technologies has made it more convenient to spread various types of speech, including hate(ful) speech that impinges on racial, religious and gender equality, all of which constitute themes of socio-political significance.<sup>1</sup> Due to broad cultural discrepancies and political disparities, however, countries around the world endorse tremendously different regulatory systems to deal with unpopular speech that may be branded as hateful.<sup>2</sup> To be fair, most of these legal rules are rooted in conventional speech law supposed to be largely applicable in digital context.<sup>3</sup> But the idiosyncratic feature of the digital environment, together with the

**CONTACT** Ge Chen  [ge.chen@durham.ac.uk](mailto:ge.chen@durham.ac.uk)

<sup>1</sup>See Iginio Gagliardone and others, *Countering Online Hate Speech* (UNESCO 2015) 33–52 <<https://unesdoc.unesco.org/ark:/48223/pf0000233231>> accessed 27 March 2022.

<sup>2</sup>Alexander Brown, *Models of Governance of Online Hate Speech* (Council of Europe 2020) 44 <<https://rm.coe.int/models-of-governance-of-online-hate-speech/16809e671d>> accessed 27 March 2022.

development of international politics and economic globalisation, attaches interdisciplinary considerations to the regulation of Internet speech.<sup>4</sup> Indeed, online speech regulation is an integral part of global Internet governance captured by the tension between the state, multi-stakeholders and the Internet itself.<sup>5</sup> Whereas the initial stage of Internet governance was characterised by the decentralisation of technological control and the marginalisation of the role of the state,<sup>6</sup> the beginning of the second decade of this new century seems to be bidding farewell to that entirely liberal yet at times anachronistic regulatory mode. This is a trend duly visible in a multitude of recently enacted laws and regulations to combat online hate speech around the world.<sup>7</sup>

Thus, the regulation of online speech falls, inevitably, within the purview of the intransigent debates on the division of roles between state and society and is increasingly subject to the narrative of the confrontation between authoritarianism and constitutionalism.<sup>8</sup> In today's world, the governance models of Internet speech can be discerned roughly in a taxonomy of 'three empires of Internet': China, the EU and the US.<sup>9</sup> As such, more profound features are identifiable in the distinctions between speech regulation in authoritarian countries and that in democratic countries.<sup>10</sup> Amidst this turmoil, China often plays a leading role in exemplifying how authoritarian governments regulate Internet speech, in sharp contrast to liberal democracies.<sup>11</sup>

<sup>3</sup>United Nations General Assembly, Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, para 57, 74th session, UNDOC A/74/486, 9 October 2019.

<sup>4</sup>Laura DeNardis, 'The Emerging Field of Internet Governance' in William H Dutton (ed), *The Oxford Handbook of Internet Studies* (OUP 2013) 556, at 572.

<sup>5</sup>Jack Balkin, 'Old School/New School Speech Regulation' (2014) 127 Harvard Law Review 2296, 2308.

<sup>6</sup>Jeanette Hofmann, 'Internet Governance: A Regulative Idea in Flux' in Ravi Kumar Jain Bandamutha (ed), *Internet Governance: An Introduction* (Icfai University Press 2007) 74, at 75–76.

<sup>7</sup>Above all, Europe is the leader in these legislative activities. The best known example is Germany's Network Enforcement Act ('NetzDG') which came into effect on 1 January 2018, see *Gesetz zur Verbesserung der Rechtsdurchsetzung in Sozialen Netzwerken* (*Netzwerkdurchsetzungsgesetz* – 'NetzDG'), 1 September 2017, BGBl I S 3352. It was followed by the UK's Online Harms White Paper, a legislative proposal that came out on 8 April 2019 (see HM Government, *Online Harms White Paper*, CP 354, April 2019), and France's Law no 2020–766 of 24 June 2020 ('*Loi Avia*', '*loi visant à lutter contre les contenus haineux sur internet*' – 'Act aiming to fight against online hatred'). The European Union also proposed its non-binding *Code of Conduct on Countering Illegal Hate Speech Online* on 7 October 2021. Beyond the EU, the legislative model of NetzDG found an echo in legislations in at least Australia, Belarus, Honduras, India, Kenya, Malaysia, Philippines, Russia, Singapore, Venezuela and Vietnam. For an overview of these online hate speech laws, see Jacob Mchangama and Joelle Fiss, *The Digital Berlin Wall: How Germany (Accidentally) Created a Prototype for Global Online Censorship* (Justitia 2019).

<sup>8</sup>Blayne Haggart, Jan Aart Scholte and Natasha Tusikov, 'Introduction: Return of the State?' in Blayne Haggart, Natasha Tusikov and Jan Aart Scholte (eds), *Power and Authority in Internet Governance: Return of the State?* (Routledge 2021) 1, at 3–4.

<sup>9</sup>Jack Balkin, 'How to Regulate (and Not Regulate) Social Media' (Knight First Amendment Institute at Columbia University, 25 March 2020) <<https://knightcolumbia.org/content/how-to-regulate-and-not-regulate-social-media>> accessed 27 March 2022.

<sup>10</sup>Mark Tushnet, 'Authoritarian Constitutionalism' (2015) 100 Cornell Law Review 391.

<sup>11</sup>Harriet Moynihan and Champa Patel, *Restrictions on Online Freedom of Expression in China* (Chatham House 2021) <<https://www.chathamhouse.org/2021/03/restrictions-online-freedom-expression-china>> accessed 27 March 2022.

Nevertheless, the borderline between authoritarian states and democratic states in their specific regulation of online speech is, on the surface, somewhat blurring, although the fundamental distinctions in the regulatory regimes remain intact. On the one hand, even democratic states need to stretch out their hands occasionally into a domain born to be free from regulation in order to find a compromise between free speech and other constitutional values.<sup>12</sup> On the other hand, authoritarian countries are also seized by capitalism, so they must take market factors into consideration when regulating the Internet, albeit to different degrees.<sup>13</sup> Therefore, certain measures that impose restrictions on online speech in both types of countries, when perceived independently of their underlying institutional, social and political structures, may be overlapping. This is the contemporary global context in which authoritarian states may well usurp international legal rules to reach overt or covert goals of changing the international legal order gradually across the borders.<sup>14</sup>

It is against such a backdrop that this article seeks to explore certain features of the regulation of online hate speech that China is developing in conjunction with its overall target of online speech regulation. To be clear, hate speech, as it is so termed also in China, is situated only peripherally in the spectrum of political speech that the Chinese government aims to regulate wholeheartedly by means of various types of laws and regulations. This is because, as an authoritarian state, China lays its regulatory focus on traditional, core political speech that could be deemed subversive, inciteful and seditious to the ruling Chinese Communist Party (CCP), whereas racial, religious and gender-oriented speech are often regarded as less relevant in China's domestic political context. This is so even though ethnic tension and suppression of religious freedom often constitute a major topic in China's foreign relations, and, thus, a potential catalyst for political upheaval.<sup>15</sup> Therefore, not only does China's regulation of online hate speech exist in an emerging patchwork of legal rules that are politically oriented, but this regime also borrows extensively from what is known as the regulation of online hate speech in international and comparative perspectives. However, the implications of the application of similar legal techniques in a completely different setting are not to be ignored, since they may sometimes generate similar, if not unexpected and undesirable, effects on the daily use of such

<sup>12</sup>Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power* (Profile Books 2019).

<sup>13</sup>Ashutosh Bhagwat, 'Free Speech without Democracy' (2015) 49 U C Davies Law Review 59, at 70 and 72–74.

<sup>14</sup>Tom Ginsburg, 'Authoritarian International Law' (2020) 114(2) American Journal of International Law 221, at 255–56.

<sup>15</sup>For an explanation of such ethnic tension in China, see Mark Elliott, 'What Is the Source of Ethnic Tension in China?' in Jennifer Rudolph and Michael Szonyi (eds), *The China Questions: Critical Insights into A Rising Power* (Harvard University Press 2018) 33, at 34–37.

legal infrastructure elsewhere. For these reasons and those stated above, it makes sense to look into the Chinese case as a primary model in contradistinction to what is understood as the regulation of online hate speech universally.

This article identifies a trend in Chinese law in which regulators take the protection of equality rights as a legitimate claim to regulate online hate speech but, in fact, subject such speech regulation to the broader regulation of core political speech. To engage with this unique approach, it is necessary to grasp its roots and paths in international and comparative perspectives. First, this article builds on the transition from traditional regulation of hate speech and addresses the new technological, legal, political and constitutional challenges posed by online hate speech to the protection of equality rights. Second, the article analyses the heterogeneous approaches by which European and US courts deal with those challenges and discusses differing roles of the state, private actors, and media organisations. Third, the article explains how those approaches have informed Chinese law while being used in different ways within an authoritarian constitutional order. Finally, the article develops a case study of the dis-regulation of ‘White-Left’ that commentators deem to represent Western liberalism in China, questioning both theoretical and practical aspects of that statist approach in light of the tension between free speech and equality rights.

### The conventional imbroglio of hate speech and equality rights

Traditional hate speech law in Europe and the US focuses on the challenges to racial, religious and gender equality, which necessitates nuanced constitutional protection of equality rights.<sup>16</sup> A comparison of the judicial approaches in these two major systems belies a dichotomy between content-based and content-neutral speech. On the one hand, European courts apply a systemic interpretation of Articles 10, 14 and 17 of the European Convention on Human Rights (ECHR), highlighting an equalitarian and dignitarian reading of the protection of speech.<sup>17</sup> The Strasbourg jurisprudence ranging from *Glimmerveen* to *Jersild* well illustrates this.<sup>18</sup> More recently, the *Lilliendahl* case shows that the Strasbourg court stresses an ‘assessment of the content of the expression’ in determining the degree of protection granted to hate speech.<sup>19</sup> On the other hand, US courts have almost always held on to a ‘content-neutral’ stance regarding the regulation of hateful speech. Apart from the earlier deviation from this standard in

<sup>16</sup>Eric Barendt, *Freedom of Speech* (2nd edn, OUP 2005), 177–86.

<sup>17</sup>Dieter Grimm, ‘Freedom of Speech and Human Dignity’ in Adrienne Stone and Frederick Schauer (eds), *The Oxford Handbook of Freedom of Speech* (OUP 2021) 107, at 108.

<sup>18</sup>*Glimmerveen and Hagenbeek v the Netherlands* App nos 8348/78 and 8406/78 (11 October 1979); *Jersild v Denmark* (ECHR, 23 September 1994) App no 15890/89.

<sup>19</sup>*Lilliendahl v Iceland* App no 29297/18 (ECHR, 12 May 2020) [36].

*Beauharnais*,<sup>20</sup> both *RAV* and *Virginia*,<sup>21</sup> the two landmark cases, are testament to the rule that US courts do not endorse the regulatory approach of distinguishing between different types of speech based on their content or topic. The more recent *Matal* case reiterates the fact that the First Amendment protection would extend to hateful speech ‘that demeans on the basis of race, ethnicity, gender, religion, age, disability, or any other similar ground’.<sup>22</sup>

Therefore, there is a clear borderline between the two classical approaches: US courts would only consider whether they *may* prohibit certain hateful speech to stop imminent violence regardless of the discriminatory elements in speech, whereas European courts are obliged to draw a red line where they *must* protect equality rights from hateful and extremist speech. Despite such a dichotomy of judicial approaches in terms of the content of speech, both European courts and US courts would engage with a more profound debate between those who stand for an absolutist reading of free speech rights and those who strive for the equality rights of minorities.<sup>23</sup>

This imbroglio of hate speech and equality rights has become the litmus test of international standards now enshrined in the prevalent conventions such as the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights (ICCPR) and the Race Convention, all of which are corollaries of rounds of earlier international negotiations that reflected conflicting values in dealing with hateful speech.<sup>24</sup> At the end of the day, the US’ standpoints emanating from the jurisprudence of the First Amendment helped dilute the proposals led by the USSR that represented overall suppression of hate speech prone to incite national, racial and religious resentment and culminated in the criminalisation of war propaganda or genocide.<sup>25</sup> Apart from Article 19 of the ICCPR, a reconcilable approach adopted by most countries is hallowed in Article 20 of the ICCPR, which accommodates not only content-based elements such as national, racial and religious hatred but also relatively content-neutral standards such as discrimination, hostility and violence.<sup>26</sup>

As a result, prevalent international jurisprudence is characterised by considerations that espouse both the criminalisation of hate crimes and legitimate caveats against misuse of speech-repressive laws.<sup>27</sup> This mentality is

---

<sup>20</sup>*Beauharnais v Illinois* 343 US 256 (1952).

<sup>21</sup>*RAV v St Paul*, 505 US 377 (1992); *Virginia v Black* 538 US 343 (2003).

<sup>22</sup>*Matal v Tam*, 582 US \_\_\_, 137 S. Ct. 1744 (2017).

<sup>23</sup>See e.g. Jeremy Waldron, *The Harm of Hate Speech* (Harvard University Press 2012), 105–43.

<sup>24</sup>Sarah H Cleveland, ‘Hate Speech at Home and Abroad’ in Geoffrey R. Stone and Lee C. Bollinger (eds), *The Free Speech Century* (OUP 2018) 210, at 216–25.

<sup>25</sup>*Ibid* 211–12.

<sup>26</sup>United Nations Human Rights Committee (UNHRC), General Comment No 34 (2011) on the Freedoms of Opinion and Expression, paras 50 and 52.

profoundly influenced by, and associated with, the post-war European political context that precludes holocaust denial. In its leading case *Faurisson*, the United Nations Human Rights Committee (UNHRC) upheld the French court's conviction of anti-Semitic speech regardless of the speaker's intent under the three-step test of Article 19.<sup>28</sup> In another enlightening case, *Rabbae*, the UNHRC established Article 20 as a bridge between free speech and equality rights – where, however, courts will need to draw a fine line of distinction between substantial harms and criticism of religious leaders or tenets such as that allowed under Dutch law.<sup>29</sup> Indeed, the leading regional jurisdiction that has developed a variety of judicial approaches toward fine-tuning of such international standards is the European Court of Human Rights (ECtHR), which grants member states of the ECHR a 'margin of appreciation' in making such decisions.<sup>30</sup> However, the ECtHR and domestic European courts are less tolerant of hateful speech than are US courts, where discriminatory considerations may be triggered.

### New challenges of online hate speech to the protection of equality rights

With the development of the Internet, artificial intelligence (AI) and big data, the free flow of information and speech grows at such an exponential speed that hate speech in social media becomes an inevitable theme in Internet governance.<sup>31</sup> Arguably, various types of hateful speech in the virtual environment contribute to the distribution of political, national, racial, religious and gender hatred that, in turn, could be reflected in the physical world in ubiquitous forms of hostility, discrimination and violence.<sup>32</sup> Online hateful speech is common both in liberal democracies that promote free speech and in authoritarian regimes that impose ruthless censorship. The harms of political hateful speech that supporters of Trump caused to ethnic groups during the 2016 US presidential campaign are no less illegitimate than the harms of cultural hateful speech that Chinese bloggers who mocked the Hui minority roasting pettitoes caused to members of that ethnic nationality.<sup>33</sup> Moreover, the impacts of online hateful speech are no

<sup>27</sup>RA Duff and SE Marshall, 'Criminalizing Hate?' in Thomas Brudholm and Birgitte Schepelern Johansen (eds), *Hate, Politics, Law: Critical Perspectives on Combating Hate* (OUP 2018) 116, at 139–41.

<sup>28</sup>UNHRC, *Faurisson v France*, Communication No 550/1993, UN Doc CCPR/C/58/D/550/1993, 8 November 1996.

<sup>29</sup>UNHRC, *Rabbae v The Netherlands*, Communication No 2124/2011, UNDOC CCPR/C/117/D/2124/2011, 29 March 2017.

<sup>30</sup>Andrew Legg, *The Margin of Appreciation in International Human Rights Law: Deference and Proportionality* (OUP 2012).

<sup>31</sup>Victoria Nash, 'Analyzing Freedom of Expression Online: Theoretical, Empirical, and Normative Contributions' in William H. Dutton (ed), *The Oxford Handbook of Internet Studies* (OUP 2013) 442, 457.

<sup>32</sup>M Cinelli and others, 'Dynamics of Online Hate and Misinformation' (2021) 11 Science Reports 22083 <<https://doi.org/10.1038/s41598-021-01487-w>> accessed 27 March 2022.

longer confined to national territories, but can be extended to recipients across different jurisdictions.<sup>34</sup> Thus, states would often issue long-arm regulations of cross-border data and online hateful speech that reach remote senders and publications.<sup>35</sup> In that sense, all jurisdictions, democratic or non-democratic, need to deal with the conflict-ridden consequences of such online hateful speech, irrespective of whether the regulatory approaches they adopt are content-based or not. The technicalities of the regulation of online hate speech, although country-specific and region-oriented, are mostly backed up by considerations of discriminatory risks and could be seen as part of global Internet governance at least on an ontological level.<sup>36</sup>

While digital media proffers the general public larger space for free expression, such seemingly equal rights to impart and receive information may become vulnerable to substantive inequality as a result of the marginalisation of certain minority groups at whom hate speech is targeted.<sup>37</sup> This is because minority groups are often positioned in a disadvantageous and hazardous environment such that their speech may not be heard or communicated effectively.<sup>38</sup> Besides, hateful and dangerous speech could be the precursor of physical violence.<sup>39</sup> This situation could be aggravated by the fact that the harms of online hate speech, if not subjected to regulatory and technological measures, are spread widely and exist permanently in the Internet, which deepens certain factual inequalities of different minority groups.<sup>40</sup>

### **Technological risks**

Digital technologies often make it difficult to detect and identify certain types of ideological discrimination and inequality. While formalistic equality between different groups is widely recognised, the harms that online hate speech causes to such equality rights may still obtain in an anonymous and dormant manner that potentially shapes users' way of thinking and behaviour.<sup>41</sup> Although online hate speech could be filtered via technological

<sup>33</sup>Notably, in the latter case, ethnic tension constitutes a fundamental threat to political stability and regime existence; see Elliott (n 15) 40.

<sup>34</sup>Daniel Holznagel, 'Platform Liability for Hate Speech and the Country of Origin Principle: Too Much Internal Market?' (2020) 4 Computer Law Review International 103.

<sup>35</sup>James Bank, 'European Regulation of Cross-Border Hate Speech in Cyberspace: The Limits of Legislation' (2011) 19(1) European Journal of Crime, Criminal Law and Criminal Justice 1.

<sup>36</sup>Tim Wu, 'Network Neutrality, Broadband Discrimination' (2005) 2 Journal of Telecommunications and High Technology Law 141.

<sup>37</sup>Gregory P. Magarian, 'The Internet and Social Media' in Adrienne Stone and Frederick Schauer (eds), *The Oxford Handbook of Freedom of Speech* (OUP 2021) 351, at 359–61.

<sup>38</sup>Brown (n 2) 42.

<sup>39</sup>For example, those who are verbally attacked on Facebook, such as women, may well suffer from offline violence. UN Women, *Eliminating Online Hate Speech to Secure Women's Political Participation* (2022) <<https://asiapacific.unwomen.org/en/digital-library/publications/2021/04/eliminating-online-hate-speech-to-secure-women-s-political-participation>> accessed 27 March 2022.

<sup>40</sup>Brown (n 2) 42.

<sup>41</sup>Alexander Brown and Adriana Sinclair, *The Politics of Hate Speech Laws* (Routledge 2019) 21.



measures, such measures do not provide a fundamental solution, as the decentralisation techniques serve only to remove speech from one platform while users may transpose it to another.<sup>42</sup> In addition, certain drawbacks of AI-generated arithmetics prevent hate speech from being identified effectively, because the understanding of human language and speech goes far beyond the literal input of linguistic meanings.<sup>43</sup> Words meant to be entirely neutral can be imbued with hateful connotations in different contexts, and complex human culture makes it easy to develop countermeasures to eschew AI-generated filtering software.<sup>44</sup> Worse still, filtering keywords hinders the general public from exercising their right to free expression.<sup>45</sup>

### **Political and constitutional factors**

Traditional constitutional protection of equality rights presupposes the bipolar confrontation between state power and individual rights and hinges on the constitutional mechanism of preventing state power from encroaching on the civil liberties of individuals. With the revolutionary advent of digital technologies, however, part of the protégé of such constitutional power turns into digital media, a fourth constitutional power,<sup>46</sup> which alters the established model of Montesquieurian *trias politica*.<sup>47</sup> The decentralisation of media power enables a variety of influential distributors to acquire control of information in a way comparable to the governance of speech by public authorities.<sup>48</sup> Where the law fails to respond to such a fundamental change of political and constitutional infrastructure in a timely and effective manner, the multiplication of such public roles in media could further weaken the voice of minorities and endanger the protection of equality rights.<sup>49</sup>

### **Legal challenges**

Online hate speech also poses unprecedented challenges to the protection of equality rights when compared with the classical legal approaches to discriminatory practice in the past. First, corresponding legislation may lag behind

---

<sup>42</sup>Nash (n 31) 449.

<sup>43</sup>*Ibid* 451.

<sup>44</sup>Mihaela Mihai, 'From Hate to Political Solidarity' in Thomas Brudholm and Birgitte Schepelern Johansen (eds), *Hate, Politics, Law: Critical Perspectives on Combating Hate* (OUP 2018) 193, at 197–98.

<sup>45</sup>Nash (n 31) 451–52.

<sup>46</sup>Nic Newman, William H. Dutton and Grant Blank, 'Social Media in the Changing Ecology of News: The Fourth and Fifth Estates in Britain' (2012) 7(1) *International Journal of Internet Science* 6.

<sup>47</sup>For a recent review of this constitutional doctrine, see e.g. Lukas van den Berge, 'Montesquieu and Judicial Review of Proportionality in Administrative Law: Rethinking the Separation of Powers in the Neoliberal Era' (2017) 10(1) *European Journal of Legal Studies* 203.

<sup>48</sup>Magarian (n 37) 355–57.

<sup>49</sup>Brown (n 2) 77–83.

technological progress. Whereas the anonymity, immediacy and interoperability of the Internet makes it extremely convenient to distribute messages across the world, the automatic sharing and volatility of hate speech renders it difficult to provide a timely remedy for the harms to equality rights.<sup>50</sup> National legislation with respect to online regulations could plummet into a standoff because, even in traditional domains, there is no uniformly agreed set of rules of hate speech and equality rights.<sup>51</sup> Moreover, the nature of digital media that carries hate speech brings significant obstacles to direct supervision by government authorities, since it is almost impossible for them to ask even explicit but unidentified sites to delete hateful information thoroughly.<sup>52</sup> Still other websites have a variety of ways, such as using neutral domain names and a deceptive user interface, to cover up their discriminatory purposes.<sup>53</sup> These potentially harmful websites are hardly visible to supervisory bodies where one sets off to search for discriminatory information directly.<sup>54</sup> Besides, government authorities, above all, in the US, may not deal with complaints about hateful speech in a serious manner where there is no imminent danger of violence. Finally, despite the existing international standards, courts around the world would interpret cross-border hate speech in entirely different legal, political and cultural contexts.<sup>55</sup>

### **The equalitarian dilemma: heterogeneous contours of regulating online hate speech**

These new challenges call for concerted efforts of regulation and cautious response to the potential harms of online hate(ful) speech.<sup>56</sup> In the wake of growing terrorism, racial tension and globalisation problems, online hate speech has elicited broad social and political impacts on minority groups such as among refugees and immigrants and impinged on democratic countries that embrace both free speech and equality rights. In 2019, France and New Zealand jointly pledged with the world's Big Tech from 26 countries to fight against online hate speech.<sup>57</sup> Under this global call,

<sup>50</sup>DeNardis (n 4) 571–73.

<sup>51</sup>Jonathan Turley, 'No, the US Does Not Need European-Style Hate Speech Laws' *USA Today* (8 November 2019) <<https://eu.usatoday.com/story/opinion/2019/11/08/no-us-not-need-european-style-hate-speech-laws-column/4157833002/>> accessed 27 March 2022.

<sup>52</sup>Jessie Daniels, 'Race, Civil Rights, and Hate Speech in the Digital Era' in Anna Everett (ed), *Learning Race and Ethnicity: Youth and Digital Media* (MIT Press 2007) 129, at 132.

<sup>53</sup>*Ibid.* 138–46.

<sup>54</sup>*Ibid.* 137–78.

<sup>55</sup>This is best exemplified by the Yahoo case where the potential violation of French criminal law could be trumped by the application of the First Amendment in the US; see *Yahoo! v LICRA*, 433 F 3d 1199 (9th Cir 2006).

<sup>56</sup>Clara Hendrickson and William A Galston, *Big Tech Threats: Making Sense of the Backlash against Online Platforms* (Brookings, 28 May 2019) <<https://www.brookings.edu/research/big-tech-threats-making-sense-of-the-backlash-against-online-platforms/>> accessed 27 March 2022.

many US tech companies including Facebook initiated their own governance codes and adjusted their platform rules so as to curb hateful speech. In conjunction with such efforts, the EU also issued its *Code of Conduct on Countering Illegal Hate Speech Online*, which represents a non-binding call for a non-discriminatory, tolerant and respectful Internet while stressing the importance of free speech.<sup>58</sup> However, such an emerging global regulatory patchwork still derives from the doctrines of protecting free speech and equality rights rooted in the heterogeneous contours of the constitutional and legal traditions of these countries.

In essence, a liberal democracy requires restriction of public power including that of the media with regard to the potential encroachment on individual rights.<sup>59</sup> Yet the power of the media, if it can be dubbed as such, is different from state power in that the media power stems from the civil liberty of free expression, and restricting media in the way restrictions are imposed on public authorities may impair media freedom itself.<sup>60</sup> As a result, digital media often enjoys greater freedom of speech and transfers such freedom to its users.<sup>61</sup> Conversely, the protection of equality rights could be weakened disproportionately if the potential conflicts between these two fundamental rights could not be reconciled in this new context. Based on its stronghold of the protection of equality rights, Germany issued the 2017 Network Enforcement Act (*Netzwerkdurchsetzungsgesetz* – ‘NetzDG’), which requires social media to take down manifestly unlawful content within 24 h of receipt and imposes fines of up to €5 million on sites that violate such rules.<sup>62</sup> Nonetheless, it is this law that has become the focus of debates on the regulation of online hate speech with regard to its potential merits and drawbacks.<sup>63</sup>

Admittedly, post-war European hate speech law relies on the protection of equality rights to mitigate racial and ethnic tension as a socio-political task. Thus, the ECtHR found prosecution of holocaust denial compatible with

<sup>57</sup>Charlotte Graham-McLay, ‘New Zealand and France to Seek Pact Blocking Extreme Online Content’ *New York Times* (24 April 2019) <<https://www.nytimes.com/2019/04/24/world/asia/ardern-social-media-content.html>> accessed 27 March 2022.

<sup>58</sup>The European Commission, *The EU Code of Conduct on Countering Illegal Hate Speech Online*, <[https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online\\_en](https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en)> accessed 27 March 2022.

<sup>59</sup>Eric Heinz, *Hate Speech and Democratic Citizenship* (OUP 2016) 81.

<sup>60</sup>*Ibid.*

<sup>61</sup>*Ibid* 130.

<sup>62</sup>NetzDG (n 7), ss 3.2.2 and 4.2, 1 September 2017.

<sup>63</sup>On the one hand, the NetzDG has caused wide concerns about its chilling effects and its potential influences on legislations in other countries. See, generally, UNHRC, Concluding Observations on the Seventh Periodic Report of Germany, paras 46–47, UNDOC CCPR/C/DEU/CO/7, 11 November 2021. For a more detailed analysis, see Peter Coe, ‘The Draft Online Safety Bill and the Regulation of Hate Speech: Have We Opened Pandora’s Box?’, this issue. On the other hand, the law seems to serve as a model for assuaging the fears of speech harms in different societies. See Uta Kohl, ‘Platform Regulation of Hate Speech: A Transatlantic Speech Compromise?’, this issue.

Article 14 of the ECHR when it upheld German defamation law in prohibiting such speech and sanctioning such an attack on the Jewish community.<sup>64</sup> However, the legal techniques of addressing relevant hate speech in member states could be rather sophisticated. For instance, German courts would not deal with hate speech against Jews directly as a violation of the constitutional right to equality, but tend to address such speech along the lines of human dignity under Article 1 of the German Basic Law (*Grundgesetz*, 'GG').<sup>65</sup> In that connection, the German criminal law allows for state prosecution of insult of a member of any group persecuted by the National Socialist regime,<sup>66</sup> while making holocaust denial and its justification an offence.<sup>67</sup> The German Federal Constitutional Court (FCC) embraced such laws as the 'general law' under Article 5(2) of the GG that could apply legitimately to restrict the dissemination of this type of hate speech.<sup>68</sup> Under this interpretive framework, holocaust denial constitutes a *false allegation of fact*, although sanction of hate speech with group target can be compromised where the speech is meant to criticise the government.<sup>69</sup>

By contrast, the Fourteenth Amendment to the US Constitution, which addresses equal protection, isn't and can never turn into an amendment to or containment of the First Amendment, so that US courts would not adopt a balancing act in its decisions. Instead, US courts abide by the bedrock principles of the First Amendment and refrain from restricting hateful speech merely for the reason that it contains discriminatory and derogatory messages. Following *Chaplinsky*,<sup>70</sup> a narrow category of speech was not treated as 'speech' for the purpose of the First Amendment. In *Rav*,<sup>71</sup> however, Scalia J, on behalf of the court, almost overthrew this rule by expressing the court's hostility toward any content-based regulations, i.e. there should be no content review of even fighting words. However, White J and colleagues suggested that states were free to ban any specific type of fighting words and that the court should consider whether the hate speech caused 'pressing public concerns'.<sup>72</sup> This line of reasoning was affirmed in *Virginia* when the court held it legitimate for states to ban a threatening message regardless of the content.<sup>73</sup> Such a logic indicative of an absolutist approach of the protection of free speech has continued to foment in digital context. In *Matal*, the Supreme Court struck down a

<sup>64</sup>ECtHR, *X v Germany* App no 9235/81 (ECHR, 14 July 1981).

<sup>65</sup>Ann Goldberg, 'Minority Rights, Honor, and Hate Speech Law in Post-Holocaust West Germany' (2021) 7(2) *Law, Culture and the Humanities* 224, at 227–28.

<sup>66</sup>German Criminal Code, s 194.

<sup>67</sup>*Ibid* s 130.

<sup>68</sup>BVerfG, *Auschwitz Lie* (13 April 1994), BVerfGE 90 241.

<sup>69</sup>BVerfG, *Soldiers Are Murderers* (10 October 1995), BVerfGE 93, 266.

<sup>70</sup>*Chaplinsky v New Hampshire* 315 US 568 (1942).

<sup>71</sup>RAV (n 21).

<sup>72</sup>*Ibid* 406–07.

<sup>73</sup>*Virginia* (n 21).

federal law prohibiting trademark names that disparage others, finding that the law itself constitutes ‘viewpoint discrimination’,<sup>74</sup> which makes it difficult even to pass the *Hudson* test.<sup>75</sup>

While hateful speech is not an exception to the protection of free speech under the First Amendment, its regulation does often pose an equalitarian dilemma. At best, the heterogeneous contours of regulation of online hate speech by European and US courts render the global regulatory framework a patchwork that calls for coordination and cooperation in legal, media and technological policies. This framework often requires redefining the roles of the state, private actors and media organisations.<sup>76</sup> Above all, multinational digital companies now play a decisive part across different jurisdictions, as the power of content review is relegated from public authorities to private content reviewers such as under the regulatory model of NetzDG. Clearly, the fear of dilated ‘private censorship’ is legitimate.<sup>77</sup> But it may also be true that such a ‘privatised process’ is channelled into the reasonable calculations of European and US jurisprudence.<sup>78</sup> Whichever way it is, such a new regulatory model is based on the premise of a largely free Internet in a democratic society, whereas this scenario leaves much technical leeway for non-liberal regimes to look to when designing their own regulatory system in an attempt to manipulate speech regulation.

## **An authoritarian approach: China’s regulation of online hate speech under the aegis of equality rights**

### ***China’s statist regulatory patchwork***

While China’s authoritarian constitutional framework may not provide an effective approach protective of free speech or equality rights, the flourishing use of an even incomplete Internet by a significant proportion of Chinese citizens does call for expedient regulation of online hate speech that exists in various forms of discrimination. Indeed, China’s constitution proscribes discrimination based on race, nationality, religion or gender in addition to the prescription of equal legal status of all Chinese citizens.<sup>79</sup> Moreover, comparable to the speech-regulatory framework in Europe, China’s constitution allows for restriction of speech in terms of national security, public morals and others’ rights.<sup>80</sup> Such constitutional mandates, however, are not to be construed as a balance between equality rights and speech rights

<sup>74</sup> *Matal* (n 22) 4.

<sup>75</sup> *Ibid* 23; see also *Central Hudson Gas & Elec Corp v Public Serv Comm’n of N Y* 447 US 557 (1980).

<sup>76</sup> Balkin, ‘Old School/New School Speech Regulation’ (n 5).

<sup>77</sup> Coe (n 63).

<sup>78</sup> Kohl (n 63).

<sup>79</sup> Constitution of the PRC 2018, arts 4, 33 and 36.

<sup>80</sup> *Ibid* arts 35, 41, 51, 53 and 54.

or as prioritising either when it comes to hate speech. Instead, they are conceptualised under a regulatory patchwork of hate speech in the digital environment, which China partly transplants from the Western legal system but steers in a different direction that hallows the ‘national security’ defined by the party-state.<sup>81</sup> Above all, China’s online hate speech regulation is geared to its cybersecurity laws, which specify the constitutional doctrine of national security enshrined in the National Security Act.<sup>82</sup> Under the Cybersecurity Act, an underlying legal source of this regime, such a statist approach takes the form of allowing the government to assume ‘cyber sovereignty’ over Internet governance whose overwhelming substance is the maintenance of the monopolistic rule of the authoritarian regime itself.<sup>83</sup>

This statist regulatory patchwork, under which state regulation is superior to the protection of speech, consists of three layers of rules. First, broadly defined racist and religious hate speech deemed a threat to China’s national security is regulated effectively through prior restraint and government censorship. This approach often results in suppression of both free speech and equality. Second, a harshly applied approach is the subsequent punishment of online hate speech that may trigger social unrest and jeopardise public morals. This approach may take the form of both administrative penalty and criminal offence. However, it presupposes a limited space for speech and involves less consideration of the protection of equality rights than of social stability. Finally, while China has no constitutional review system, a rights-based regulatory approach is taking shape, above all, through the development of intermediary liabilities of Internet service providers that host, cache and transmit hate speech. Whilst there is still rather limited latitude for free speech, such an approach represents that of an indirect tort which draws the attention of both legislators and courts, and might involve, ultimately, qualified considerations of balancing speech against equality. These three layers of rules are discussed in turn below.

### ***Administrative law as an instrument of prior restraint***

First, administrative regulations driven by state interests and underpinned by ‘demoralised pragmatism’ set low thresholds for banning online hate speech,<sup>84</sup> which go beyond the universally recognised principle of

<sup>81</sup>Hualing Fu, ‘China’s Imperatives for National Security Legislation’ in Cora Chan and Fiona de Londras (eds), *China’s National Security: Endangering Hong Kong’s Rule of Law?* (Hart 2020) 41, at 42–46.

<sup>82</sup>This includes e.g. the protection of the state power of the people’s democratic dictatorship and the socialist system, the integrity of state regime and sovereignty, national unity and territory, and the leadership of the CCP.; see The National Security Act of the PRC 2015, arts 1, 2 and 4.

<sup>83</sup>The Cybersecurity Act of the PRC 2016, art 1.

<sup>84</sup>‘Demoralised pragmatism’ was coined to depict and analyse the mentality of identifying ‘the superiority of Chinese self’ to Western ‘capitalist modernity’ in terms of an ‘allegedly pragmatic, rational and non-moralising approach to economic growth and social stability taken by the current authoritarian regime’; see Chenchen Zhang, ‘Right-Wing Populism with Chinese Characteristics? Identity, Otherness

proportionality. Above all, the Chinese government enacts a large number of administrative regulations in relation to online publication and information services, which share certain entrenched and overlapping norms. For instance, these regulations ban the publication of materials that ‘incite the national hatred or discrimination, undermine the solidarity of the nations, or infringe upon national customs and habits’, ‘propagate evil cults or superstition’ or ‘endanger public ethics or the fine cultural traditions of any nationalities’.<sup>85</sup> In practice, these terms are often hardly distinguishable from other terms prescribed in conjunction with them, such as ‘endanger the unity of the nation, sovereignty or territorial integrity’ or ‘endanger national security or damage the honour or benefits of the State’.<sup>86</sup>

A recent government regulation targeting online speech singles out contents ‘advocating terrorism or extremism, or instigating any terrorist or extremist activity’, ‘inciting ethnic hatred or discrimination to undermine ethnic solidarity’ and ‘detrimental to state religious policies, propagating heretical or superstitious ideas’ that must be filtered by information service providers (ISPs).<sup>87</sup> Where a user violates such government policies, that online platform is obliged to take measures such as ‘warning for rectification, restricting functions, suspending updating, and closing accounts’.<sup>88</sup> Meanwhile, the ISP must ‘eliminate illegal information and contents in a timely manner, keep relevant records, and report to the relevant competent authorities’.<sup>89</sup> Any ISP that fails to comply with these content review requests will face severe administrative penalties amounting to temporary bans from network information services or permanent industrial bans.<sup>90</sup>

Arguably, overburdening ISPs with strict censorship duties and administrative liabilities could result in overall bans of any politically sensitive contents by technological measures including marginal elements stirring up hatred or discrimination. Typically, ISPs undertake an opaque censorship process in which users are not allowed to participate.<sup>91</sup> Moreover, there are no accessible appeal mechanisms or legal remedies regarding the outcomes of censorship. This is observable in the regulation of speech ranging from racial and religious hatred to gender topics. Recently, the mobile social platform WeChat has shut down nearly 20 accounts run by

---

and Global Imaginaries in Debating World Politics Online’ (2020) 26(1) *European Journal of International Relations* 88, at 104.

<sup>85</sup>The State Council of the PRC, Regulation on the Administration of Publication (RAP), art 25(3), (4) and (9), 29 November 2020. For similar formulations, see Regulations on Broadcasting and Television (RBT), art 32(3) and (6), 29 November 2020.

<sup>86</sup>RAP, *ibid* art 25(1) and (2); RBT, *ibid* art 32(1) and (2).

<sup>87</sup>Cyberspace Administration of China, Provisions on Ecological Governance of Network Information Content, art 6(5)–(7), 1 March 2020.

<sup>88</sup>*Ibid* art 34.

<sup>89</sup>*Ibid*.

<sup>90</sup>*Ibid* art 39.

<sup>91</sup>Cybersecurity Act (n 83) arts 47–48.

campus LGBTQ clubs and non-profit organisations for suspicion of materials contravening mainstream political values, although they used to enjoy some privileges of equal protection sufficient to eschew stringent censorship.<sup>92</sup> That gap is now closed because, under the aforementioned legislation, public opinion authorities in China, which comprise both government propaganda departments and state media, are increasingly tying sexual minorities with ‘colour revolutions’ imported from overseas, as long as these groups trespass the red line of political speech.<sup>93</sup>

### ***Criminal law as an instrument of subsequent punishment***

Second, the statist approach of regulating online hate speech aims at restricting freedom of speech and religious freedom by virtue of criminal law. Whilst China’s constitution provides for the equality of ‘all nationalities’ and citizens’ equal legal status,<sup>94</sup> Articles 249 and 250 of the Criminal Code of the PRC build up the major framework that regulates the incitement of racial hatred and racial discrimination as well as the publication of racially discriminatory and insulting materials:

Article 249: Whoever provokes ethnic hatred or discrimination, if the case is serious, is to be sentenced to three years or fewer in prison, put under limited incarceration or surveillance, or deprived of their political rights. If the case is especially serious, they are to be sentenced to three to ten years in prison.

Article 250: Whoever is directly responsible for publishing materials that discriminate or insult minority nationalities, if the case is serious and results in grave consequences, is to be sentenced to three years or fewer in prison, or put under limited incarceration or surveillance.<sup>95</sup>

Notably, Article 249 encompasses any attempt of incitement of racial hatred or discrimination against all nationalities, including the majority Han nationality, through language or words or by any other means. By contrast, Article 250 is targeted at any attack on the customs, traditions and life habits of minorities in China (excluding the Han nationality itself but including any preferences and taboos in their material and spiritual lives). However, the components of both offences are vaguely defined, so the standards for criminalising what amounts to a ‘serious’ degree and an ‘especially serious’ degree of racist hate speech assume large uncertainties. In the case of *Pai*, for

<sup>92</sup>Vincent Ni and Helen Davidson, ‘Outrage over Shutdown of LGBTQ WeChat Accounts in China’ *The Guardian* (8 July 2021) <<https://www.theguardian.com/world/2021/jul/08/outrage-over-crackdown-on-lgbtq-wechat-accounts-in-china>> accessed 27 March 2022.

<sup>93</sup>See e.g. Lin Lan, ‘West’s Undisguised Intention of Using MeToo Movement to Create Antagonism in China’ *Global Times* (16 September 2021) <<https://www.globaltimes.cn/page/202109/1234461.shtml>> accessed 27 March 2022.

<sup>94</sup>Constitution of the PRC (n 79) arts 4 and 33.

<sup>95</sup>Criminal Law of the PRC 2020, arts 249 and 250.



instance, the defendant was charged with uploading ‘terrorist pictures’ on his personal blog whereas the defendant was unaware of the relationship between terrorism and the uploaded pictures of the flag of East Turkestan and the flag of Saudi Arabia.<sup>96</sup> Without explaining why this would elicit racial hatred and the specific harms that might be caused, the court adopted the prosecutors’ view that online dissemination of these pictures served to instigate racial hatred to a serious degree that merited sanctions.<sup>97</sup> Furthermore, these offences could be used to cover up government campaigns of cracking down on core political dissident speech. In the case of *Ma*, for example, the defendant was sanctioned for disseminating extreme religious thoughts through mobile phones and for uploading, caching and transmitting pictures relating to extreme religious thoughts in order to stir up racial conflicts – while the real purpose was to criticise the government’s religious policies.<sup>98</sup>

### ***Civil law as an instrument of marginal protection of speech and equality rights***

Finally, China has developed some basic instruments in its civil law to protect equality rights against hate speech, although this rights-based approach is far from constituting a balancing act and provides merely limited protection of either right. Above all, China’s new Civil Code defines the protection of ‘personality rights’ including the ‘inviolability and integrity of person’ and one’s ‘reputation’ and ‘honour’.<sup>99</sup> Additionally, it creates civil liabilities for tort of such rights:

Article 1183: Where any harm caused by a tort to a personal right or interest of a natural person inflicts a serious mental distress on the victim of the tort, the victim of the tort shall have the right to require compensation for the infliction of mental distress.<sup>100</sup>

In practice, such liabilities may be applicable to cases where hate speech concerning races across different regions, choice of religions, or gender equality causes serious harms to individuals, although harms of ‘mental distress’ are often measured by one’s materialistic loss, such as unemployment.<sup>101</sup> More importantly, the Civil Code has adopted the ‘notice and takedown’ regime in

<sup>96</sup>Beijing Chaoyang People’s Court, *Pai Ti Ri Ding Instigates Racial Hatred*, 2016 Verdict Jing 0105 Xing Chu No 871, 27 July 2016.

<sup>97</sup>Similar findings of online racial hate speech often come up in regions where racial conflicts abound; see e.g. Production and Construction Corps of Xinjiang, Mosuo Wanken District People’s Court, *Xian Fenyong Instigates Racial Hatred*, 2016 Verdict Bing 0802 Xing Chu No 33, 26 September 2016.

<sup>98</sup>Xinjiang Wusu People’s Court, *Ma Xiaolong Instigates Racial Hatred*, 2015 Verdict Wu Xing Chu Zi No 120, 5 December 2015.

<sup>99</sup>The Civil Code of the PRC 2021, art 990.

<sup>100</sup>*Ibid* art 1138.

<sup>101</sup>See e.g. Beijing Chaoyang District People’s Court, *Gao Yiming v Beijing Bide Chuangzhan Telecommunication Technology Corp Ltd*, 2008 Verdict Chao Min Chu Zi No 06688, 5 December 2015.

an overall manner by defining the ISPs' obligation to delete, block and disconnect upon prima facie evidence of tort as well as their joint liabilities for infringing contents in case of actual knowledge.<sup>102</sup> While this regime covers online hate speech in a way comparable to the German NetzDG, cases adjudicated so far dwell overwhelmingly in the area of copyright and personality rights such as reputation and privacy rights. This is because most online hate speech has been filtered or penalised as a result of applying the aforementioned two approaches by government authorities and courts. With regard to hate speech, there is no decisive role for ISPs to play beyond rigid compliance with government policies.

Out of fear of additional monetary liabilities, ISPs practise stringent censorship of spontaneous hate speech but leave the guided speech with hateful elements intact. In fact, a wide range of online hate speech is part of a systematic, decades-long statist propaganda project driven by the party-state apparatus and its zealous followers inculcated with blind nationalism for the purpose of criticising and denouncing a political target by a statist attack strategy reminiscent of the Cultural Revolution. Recent victims include Hong Kong pro-democracy protestors,<sup>103</sup> Wuhan novelist Fang Fang who reflected critically on the Covid-19 outbreak,<sup>104</sup> and *New Yorker* journalist Jiayang Fan who was attacked because of her overseas Chinese identity.<sup>105</sup> More recently, Xu, a China-born journalist and researcher working for an Australian think tank, became the target of a vehement online harassment campaign when official videos deprecated her work, numerous social media accounts ascribed the global Xinjiang boycott to her, and infuriated 'patriots' threatened to harass her parents in China.<sup>106</sup>

### ***Stigmatisation of 'White Left': laissez-faire in regulating hate speech?***

Whereas hate(ful) speech is often considered political speech, those hateful topics concerning race, religion and gender are not directly associated with the CCP's concern for regime survival. But this does not mean that hate speech is apolitical and irrelevant in China. On the contrary, the

<sup>102</sup>The Civil Code of the PRC 2021, arts 1194–97.

<sup>103</sup>Kiki Tianqi Zhao, 'China's Government Wants You to Think All Mainlanders View Hong Kong the Same Way. They Don't.' *ChinaFile* (27 August 2019) <<https://www.chinafile.com/reporting-opinion/view-point/chinas-government-wants-you-think-all-mainlanders-view-hong-kong-same>> accessed 27 March 2022.

<sup>104</sup>Fan Wang and Grace Tsoi, 'Fang Fang: Author Vilified for Wuhan Diary Speaks out a Year on' *BBC* (19 January 2021) <<https://www.bbc.co.uk/news/world-asia-54987675>> accessed 27 March 2022.

<sup>105</sup>Jiayang Fan, 'How My Mother and I Became Chinese Propaganda' *The New Yorker* (7 September 2020) <<https://www.newyorker.com/magazine/2020/09/14/how-my-mother-and-i-became-chinese-propaganda>> accessed 27 March 2022.

<sup>106</sup>Zeyi Yang, 'The Anatomy of a Chinese Online Hate Campaign' *Protocol* (9 April 2021) <<https://www.protocol.com/china/chinese-online-hate-campaigns>> accessed 27 March 2022.

regulation of hate speech is often complementary to that of subversive speech. This is palpable in the fact that the most prevalent hate speech in China is derived from, and supported by, government discourses that play up, galvanise and racialise the ‘national uniqueness’ of ‘Chineseness’.<sup>107</sup> In essence, the regulation or dis-regulation of such online hate speech is highly manipulable; this allows the one-party state to stir up a collective (un)consciousness of ‘racial patriotism’ which, chained to the mainstream ideology of nationalism, militates against any liberal-minded penchant in the mainstream parlance.<sup>108</sup>

Despite those aforementioned regulations of racial, religious, gender and regional discrimination, China’s government authorities, Internet platforms and fervent nationalistic netizens embrace collectively oriented hate(ful) speech against any target labelled as ‘anti-patriotic’ or ‘traitors of the motherland’. The stigmatisation of ‘White Left’ (Baizuo) in China’s Internet is a most eminent case that *prima facie* is part of the global racist discourse and at the same time betrays the politicisation of dis-regulated hate speech. A widely used definition of this term runs as follows:

Baizuo is used generally to describe those who ‘only care about topics such as immigration, minorities, LGBT and the environment’ and ‘have no sense of real problems in the real world’; they are hypocritical humanitarians who advocate for peace and equality only to ‘satisfy their own feeling of moral superiority’; they are ‘obsessed with political correctness’ to the extent that they ‘tolerate backwards Islamic values for the sake of multiculturalism’; they believe in the welfare state that ‘benefits only the idle and the free riders’; they are the ‘ignorant and arrogant westerners’ who ‘pity the rest of the world and think they are saviours’.<sup>109</sup>

In the past several years, hostility toward Baizuo on the Chinese Internet have escalated, and more obscene, slanderous and insulting words are used against any target that could be branded with this title. Indeed, Chinese online ‘patriots’ use the term ‘Baizuo’ to label both Merkel and Obama to attack the refugee and gender policies of Western countries,<sup>110</sup> regardless of the skin colour of the actual targets. Although a racial label itself, the term connotes more a politically ‘left’ identity than a racial one. To be sure, these debated elements of cultural and identity politics are rarely seen and only peripherally represented, if they exist at all, in Chinese society. The fact that these remote topics are scorned, belittled and disparaged widely on the Chinese Internet without being censored or

<sup>107</sup>Yinghong Cheng, *Discourses of Race and Rising China* (Palgrave Macmillan 2019) 295–303.

<sup>108</sup>*Ibid* 48–51 and 77–83.

<sup>109</sup>Chenchen Zhang, ‘The Curious Rise of the “White Left” as a Chinese Internet Insult’ *OpenDemocracy* (11 May 2017) <<https://www.opendemocracy.net/en/digitaliberties/curious-rise-of-white-left-as-chinese-internet-insult/>> accessed 27 March 2022.

<sup>110</sup>*Ibid*.

punished presents a stark contrast to the regulation of hate(ful) speech that could be channelled eventually to regime-related vulnerabilities such as unfair minority, religious and gender policies attributable to the authoritarian system.

As a highlight of its political nature, this hateful term can be applied *mutatis mutandis* to any Chinese who carry or exhibit similar hallmarks of such pro-liberalism ideology so as to limit their speech. For instance, many Chinese netizens conceived of Professor Rao, a famous neurobiologist in China who openly criticised the smearing use of Baizuo and Chinese netizens' support of Donald Trump, as an identifiable 'Baizuo' characteristic of their 'elitist sense of superiority, bigotry and ignorance'.<sup>111</sup> Seen from an ideological perspective, the ease of disseminating such an anti-Baizuo mentality among Chinese netizens is rooted in a widespread pre-Hobbesian philosophy of power politics and social Darwinist theory of 'survival of the fittest' under China's authoritarian regime.<sup>112</sup> Consequently, this hateful term acquires its domestic immediacy and influential implications in China in the wake of conflicts between established urban residents and mass immigrants from rural areas, ethnic and religious tensions between Han and non-Han minorities, and the confrontation between sexually conservative groups and LGBT or feminists.<sup>113</sup> By adopting the 'White' as a racial factor, such a debate on this prevailing hate(ful) term against alleged race traitors has gained a strong tinge of political discussion 'without an awareness of its racial nature'.<sup>114</sup>

Obviously, a guideline for a nationalist campaign encapsulating vague definitions of national security, cyber sovereignty and incitement to racial, religious and gender hatred and discrimination in China's legal framework is conducive to the ineptitude of the hate speech laws in regulating such a unique type of hate(ful) speech with rich and comprehensive externalities. The widespread use of the so-called 'Baizuo' on the Internet is evidence of a politicised speech campaign in China, which dwells in a *laissez-faire* state within China's existing regulatory patchwork. Notably, this campaign increasingly takes place under the tutelage of 'combatting Western hegemony',<sup>115</sup> and fits perfectly into China's statist approach of constructing the image of a rising power against that of a steadily falling Western civilisation. Oddly and ambivalently, the use of this term among many Chinese nationalists is coupled with their profound concern, anxiety and, perhaps, sympathy regarding the 'degeneration' of the White race itself, which is exemplified most saliently by the narrative 'blacks and Muslims will take

---

<sup>111</sup>*Ibid.*

<sup>112</sup>Zhang (n 84) 101–04.

<sup>113</sup>Cheng (n 107) 285–86.

<sup>114</sup>*Ibid* 288.

<sup>115</sup>Zhang (n 84) 90–94.

over Europe' due to the extremely low birth rate of the White race and the 'extremely humanitarian' mindset in the Western society toward those immigrants.<sup>116</sup> Thus, such a campaign of hateful speech can generate extra-territorial influence over Western politics.<sup>117</sup>

## Conclusion

The post-war hate speech law around the world is characterised by a transition from the classical liberty-focused discourse to a balance-oriented approach in Europe and, alternatively, deeper engagement of the First Amendment with equality in race, religion and gender in the US. This imbroglio of free speech and equality rights continues to dominate the digital scenario with regard to most hate(ful) speech despite the technological, political and legal challenges pertaining to diversified structures of various ISPs and their social media platforms. While their regulatory approaches are taking shape with different legitimate stances, the disparity between the EU and the US in dealing with online hate speech by equality rights does provide a multitude of instruments and flexibilities for authoritarian regimes including China to emulate, albeit with different purposes.

In post-socialist China, the regulation of online hate speech under the aegis of protecting equality rights is largely characterised by a guideline of nationalist campaign and demoralised pragmatism that informs a patchwork of administrative law, criminal law and civil law. Administrative law serves as an online valve of prior restraint guided by a vague concept of national security and cyber sovereignty. Criminal law acts as subsequent punishment to deter any anti-state and anti-social online speech. Civil law provides merely marginal protection of speech and equality rights by imposing stringent liabilities on ISPs and content providers. All these legal boundaries bear the hallmark of the equal protection of different nationalities, races, religions, regions and genders.

A deeper look into a unique online hate(ful) term in China (Baizuo) and its immunity from such regulation reveals that China's online hate speech law represents a statist approach that deviates from the paradigm protective of equality rights in liberal democracies. Whereas the law imposes restrictions on online speech by virtue of mandates of equal protection, the statist campaign of hate(ful) speech often trumps the latter and serves to fend off open criticism of government policies and public discussion of topics that potentially contravene the mainstream political ideologies.

<sup>116</sup>Cheng (n 107) 289.

<sup>117</sup>For example, this Chinese slang was used to protest against the US Democratic Party's educational policies that many perceived as racist and against Asian students. Tucker Carlson, 'Even the Chinese Know America Won't Survive with "Woke" Liberals in Charge' *Fox News* (20 March 2021) <<https://www.foxnews.com/opinion/tucker-china-america-white-liberalism-biden>> accessed 27 March 2022.

Moreover, such a legal regime encourages the promotion of its extraterritoriality. Ultimately, the regulation and dis-regulation of such statist hate(full) speech runs counter to both European law and US law, which aim to reconcile potential conflicts between liberal interests and equalitarian interests.

## Acknowledgements

The author thanks Sophie Turenne and Oliver Butler for organising the BACL workshop and their insightful comments on an earlier draft. Thanks also go to the Yale ISP writing workshop for thoughtful discussions of the draft. The usual disclaimer applies.

## Disclosure statement

No potential conflict of interest was reported by the author.

## Notes on contributor

**Ge Chen** is Assistant Professor in Global Media and Information Law at Durham Law School. His research interests are media and information law and their constitutional and rule-of-law aspects in international and comparative perspectives, with a focus on China. He was a Visiting Academic of the Programme in Comparative Media Law and Policy at the University of Oxford. He was a Postdoctoral Research Associate at, and remains an Associate of, the Centre for Intellectual Property and Information Law at the University of Cambridge. He was a Resident Fellow, and remains an Affiliated Fellow, of the Information Society Project at Yale Law School. He was a research associate at Mercator Institute for China Studies, the largest China-related think tank of the European Union. He is the author of *Copyright and International Negotiations: An Engine of Free Expression in China?* (Cambridge University Press 2017), a research monograph featured in *Harvard Law Review*.