

## NO LONGER PRIVATE: ON HUMAN RIGHTS AND THE PUBLIC FACET OF SOCIAL NETWORK SITES

Noa Mor\*

*Social network sites (“SNSs”) like Facebook and Twitter are gradually acquiring sweeping influence over the human rights of their users. In this Article, I offer an original framework for understanding SNSs’ impact and obligations in this regard, based on an interdisciplinary “zoom-out” examination of these platforms’ performance. I maintain that SNSs have taken the form of digital arenas in which human rights are crafted, designed, and allocated. The emergence of these arenas relies on two dimensions: (1) the broad spectrum of rights that are affected by SNSs, which consists of civil-political along with economic-social-cultural rights; and (2) the sophisticated control that SNSs have acquired over these diversified rights. Such control is obtained and cultivated through various practices that are incorporated in SNSs’ operation and include surveillance; profiling; censorship; practices relating to SNSs-authorities nexus; and finally, expulsion of individuals and communities. Together, these two dimensions create an overarching influence to shape human rights, thereby echoing the traditional task of the state and highlighting a salient public attribute of SNSs. Alongside the benefits afforded by these arenas, their emergence produces a troubling human rights regime for*

---

\* Research Fellow at the Center for Cyber Law and Policy at the University of Haifa (“CCLP”); Ph.D. Candidate, University of Haifa, Faculty of Law.

I thank Prof. Niva Elkin-Koren for her priceless guidance and support. I also thank Karni Chagal-Feferkorn, Sharon Haleva-Amir, Maayan Perel (Filmar), Amnon Reichman, as well as the participants of both the Internet Law Work-in-Progress 2018 conference, New York Law School and the Institute for Global Law and Policy 2018 conference, Harvard Law School, for their eye-opening comments. This work was supported by the CCLP in conjunction with the Israel National Cyber Directorate in the Prime Minister’s Office; the Center for the Rule of Law at the University of Haifa; and LINKS: The Israeli Center of Research Excellence.

† The Author is publishing a separate article in Hebrew with the *Haifa Law Review (Din Udvarim)* that shares similarities with this Article. However, the former focuses on the Israeli legal landscape and does not examine the application of public law norms in other legal regimes.

*their users, often characterized by unjust decision-making processes. The public character of SNSs justifies application of public law norms as proportionality, reasonableness, and reason-giving to them, thus urging these platforms to adopt a more accountable and user-oriented approach in all their diverse areas of performance. Such public law norms constitute a long-lasting and fundamental part of various jurisdictions and could provide a preliminary multinational or global standard for SNSs' conduct that might be further calibrated and adjusted by each legal regime.*

## I. INTRODUCTION

Starting in 2017, American actress Rose McGowan has been playing a leading role in resisting sexual abuse in Hollywood and in preparing the ground for the #MeToo movement. She used Twitter intensively in this regard and, inter alia, accused Harvey Weinstein of raping her.<sup>1</sup> However, on October 12, 2017, amid a vivid public discussion on these issues, Twitter temporarily and partially deactivated McGowan's account. In the notification sent to the actor, Twitter stated that McGowan's tweets violated its rules, but it failed to specify which rules and which tweets were at issue.<sup>2</sup> A Twitter representative later publicly explained that the suspension was induced by one of McGowan's tweets that included a private telephone number.<sup>3</sup> The suspension of McGowan's account drew fire to Twitter and stimulated a twenty-four-hour boycott of the platform led by #WomenBoycottTwitter.<sup>4</sup> While suspended, McGowan was prevented from using the platform to further speak out against sexual misconduct, contributing to the unfolding public discourse on this issue, supporting women who had experienced similar situations, and finding relief in social interactions. To these unfortunate outcomes, all

---

1. Sam Levin & Olivia Solon, *Rose McGowan Alleges Rape by Harvey Weinstein – and Amazon Ignored Claim*, GUARDIAN (Oct. 13, 2017), <https://www.theguardian.com/film/2017/oct/12/rose-mcgowan-amazon-studios-harvey-weinstein-ignore-rape>.

2. Nellie Bowles & Cara Buckley, *Rose McGowan's Twitter Account Locked After Posts About Weinstein*, N.Y. TIMES (Oct. 12, 2017), <https://mobile.nytimes.com/2017/10/12/arts/rose-mcgowan-twitter-weinstein.html>; Natasha Lomas, *Twitter Users Join 24hr Boycott to Protest Online Harassment*, TECHCRUNCH (Oct. 13, 2017), <https://techcrunch.com/2017/10/13/twitter-users-join-24-boycott-to-protest-online-harassment>.

3. Bowles & Buckley, *supra* note 2.

4. Lomas, *supra* note 2.

stimulated by Twitter's problematic decision-making processes, we shall return shortly.

SNSs—particularly Facebook and Twitter on which I focus here—have acquired immense popularity worldwide<sup>5</sup> and play a prominent role in diversified global cyber-related conflicts, among which are bullying and shaming, pornography, intellectual property, hate speech, and data security.<sup>6</sup> I suggest that these platforms have taken the form of digital arenas in which human rights are crafted, designed, and allocated. These arenas rely on two dimensions. The first dimension relates to the wide range of rights influenced by SNSs, which includes civil-political rights, as well as economic-social-cultural rights. SNSs' competence to affect such a broad spectrum of human rights rests on informational and social-communal key resources provided by the powerful social dynamics that thrive throughout these platforms.<sup>7</sup> Informational resources drive creation and dissemination processes of information pertaining to all areas of life, such as finance, health, politics, culture, religion, and employment. This content then plays an important part in users' ability to realize their rights in these various fields.<sup>8</sup> Social-communal resources, on the other hand, afford social virtues such as mutual help, social support, and collective action. These virtues encourage behaviors that have become very common in SNSs and include colleagues advising each other, students tutoring their peers, citizens organizing demonstrations, and workers striking.<sup>9</sup> The second dimension of these emerging digital arenas concerns the control SNSs have acquired over the foregoing broad spectrum of rights. This control is obtained and cultivated through diversified practices incorporated in the SNSs' operation. I focus on five major groups of such practices: (1) surveillance; (2) profiling; (3) censorship; (4) practices relating to SNSs-authorities nexus; and (5) expulsion of individual users and

---

5. As of June 30, 2018, Facebook had 2.23 billion monthly active users. Press Release, Facebook, Facebook Reports Second Quarter 2018 Results (June 25, 2018), <https://investor.fb.com/investor-news/press-release-details/2018/Facebook-Reports-Second-Quarter-2018-Results/default.aspx>. Twitter had 335 million monthly active users at that time. *Twitter Q2 2018 Earnings Report*, TWITTER (July 27, 2018), <https://twitter.com/i/moments/1022804623717875712>.

6. See Jacob Poushter, *Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies But Advanced Economies Still Have Higher Rates of Technology Use*, PEW RES. CTR., 21-23 (Feb. 22, 2016), [http://www.pewresearch.org/wp-content/uploads/sites/2/2016/02/pew\\_research\\_center\\_global\\_technology\\_report\\_final\\_february\\_22\\_\\_2016.pdf](http://www.pewresearch.org/wp-content/uploads/sites/2/2016/02/pew_research_center_global_technology_report_final_february_22__2016.pdf) (describing the global wide usage of SNSs while stressing their popularity in emerging and developing economies).

7. See *infra* Part II.

8. See *infra* Part II.A.

9. See *infra* Part II.B.

communities.<sup>10</sup> All these practices result in a sophisticated and effective control over various human rights of SNS users.

When put together, the two dimensions described create arenas in which the human rights of SNSs' users are being designed, crafted, and allocated. Along with its benefits, this phenomenon reflects a troubling human rights regime, in which users' human rights are often disproportionately, non-transparently, or inappropriately hindered.

The overarching influence on human rights, obtained throughout these new digital arenas, is rarely evident in other private entities, especially outside the cyber world. It loudly echoes the principal and traditional role of the state to manage rights, as reflected in the social covenant, and in the approach which considers the state as a fiduciary of the public.<sup>11</sup>

I suggest that this salient public attribute of SNSs justifies applying to them public law norms, such as proportionality, reasonableness, and reason-giving, originally designated to guide the state's discretion and prevent its abuse.<sup>12</sup> Such public norms constitute a long-lasting and fundamental part of various jurisdictions, though they might vary from one legal regime to another in application and scope.

Applying public law norms to SNSs may urge these platforms to take a more user-oriented approach and to carefully consider human rights in all facets of their performance, *inter alia*, when formulating policy, conducting different practices, and managing users' grievances.<sup>13</sup>

In Rose McGowan's case, applying public law norms such as reasonableness could have led to more desirable, level-headed decision-making, in which proper weight would have been given to McGowan's right to free speech, to her followers' and other users' right to access information, and to the public interest of eliminating sexual abuse. Twitter might also have been obliged to verify that it had applied its rules evenly, without discriminating against certain users and audiences. Proportionality might have indicated the platforms' duty to ensure that the disabling of McGowan's account was indeed fit for the intended purpose of such sanctioning and prompt it to prefer the least intrusive measure upon McGowan's and other users' human rights. Examples of such measures might be requiring McGowan to edit her tweet without disabling her account or using automated tools that warn users who

---

10. *See infra* Part III.

11. EVAN J. CRIDDLE & EVAN FOX-DECENT, FIDUCIARIES OF HUMANITY—HOW INTERNATIONAL LAW CONSTITUTES AUTHORITY 14 (2016); Evan Fox-Decent & Evan J. Criddle, *The Fiduciary Constitution of Human Rights*, 15 LEGAL THEORY 301, 309-10 (2009).

12. *See infra* Part V.

13. *See infra* Part VI.B.

include telephone numbers in their tweets. At the procedural level, applying public law norms could have compelled Twitter to promptly inform McGowan of the precise rules she was violating and the content that allegedly violated them. They might have also required Twitter to provide McGowan with appropriate grievance systems and appeal mechanisms. Such requirements could have included channels for McGowan to be heard and receive reasoned decisions and to access some form of the platform's (anonymized) database of cases in which users had been similarly sanctioned.

Enlisting the aid of public law norms in the context of SNSs could afford additional significant benefits. First, it might promote various desirable democratic values in relation to users' online activity, like fostering SNSs' accountability and transparency, and reinforcing the checks and balances exercised by the public and the courts pertaining to SNSs' conduct.<sup>14</sup> Applying public law norms is also expected to empower users, assist them in defining their expectations of these platforms, fortify their ability to demand answerability, and encourage them to participate and to take a more active role in protecting their rights in these important online avenues.<sup>15</sup> It might even elevate SNSs' legitimacy and users' trust in them, thus encouraging a more beneficial use thereof.<sup>16</sup> Second, thanks to the wide endorsement of public law norms in local and supranational legal regimes, they might serve as a preliminary legal standard that might obligate SNSs on a supranational or even global basis. Above this preliminary tier, each jurisdiction could calibrate and adjust its unique approach.<sup>17</sup>

Part II of this Article concerns the broad spectrum of rights that are influenced by SNSs,<sup>18</sup> and Part III deals with the various practices of control that are performed by SNSs and that might influence this wide range of rights.<sup>19</sup> Part IV discusses the role of the state in managing rights,<sup>20</sup> and Part V explores the public law norms that are applied to

---

14. See Martin Shapiro, *The Giving Reasons Requirement*, 1992 U. CHI. LEGAL F. 179, 180-81 (1992) (addressing reason-giving as providing such benefits); see also Jodi L. Short, *The Political Turn in American Administrative Law: Power, Rationality, and Reasons*, 61 DUKE L.J. 1811, 1820-21 (2012) (same).

15. Short, *supra* note 14, at 1821.

16. See Peter Cane, *Records, Reasons and Rationality in Judicial Control of Administrative Power: England, the US and Australia*, 48 ISR. L. REV. 309, 321 (2015) (relating to public norms as a means to achieve bureaucratic legitimacy).

17. *Id.*

18. See *infra* Part II.

19. See *infra* Part III.

20. See *infra* Part IV.

state authorities in different jurisdictions.<sup>21</sup> Finally, Part VI investigates and demonstrates the benefits of applying public law norms to SNSs.<sup>22</sup>

## II. SNSs' INFLUENCE ON A WIDE RANGE OF RIGHTS

The first dimension of the new arenas for crafting human rights, which SNSs have become, relates to the wide range of human rights that these platforms influence. This broad spectrum includes civil-political rights such as freedom of expression, freedom of thought and religion, freedom of assembly and demonstration, and the right to vote and to be elected. It also consists of economic-social-cultural rights such as the right to education, the right to culture, and various employees' rights.<sup>23</sup>

A crucial element of SNSs that afford their aptness to affect such a broad spectrum of rights lies in the powerful social dynamics that thrive throughout these platforms and that rests on the role of SNSs as a means to support relationships. SNSs are a particular category of "social media" that is "well suited to the kinds of interpersonal exchanges that serve to maintain and strengthen social bonds."<sup>24</sup> Distinguishing SNSs from other social media platforms, José van Dijck described them as sites that "primarily promote interpersonal contact, whether between individuals or groups" and that "forge personal, professional or geographical connections."<sup>25</sup> The majority of SNSs users do indeed use these platforms to support their existing offline bonds with family and friends.<sup>26</sup> This important role of SNSs yields vibrant social dynamics that propel the use of these platforms and produce two types of key resources: informational and social-communal. These make a significant contribution to the overarching influence that SNSs have gained over a range of their users' human rights, as will be explored below.

---

21. See *infra* Part V.

22. See *infra* Part VI.

23. See generally International Covenant on Civil and Political Rights, arts. 1& 2, Dec. 19, 1966, 999 U.N.T.S. 171; International Covenant on Economic, Social and Cultural Rights, arts. 1-41, Dec. 16, 1966, 993 U.N.T.S. 3; Philip Alston, *A Third Generation of Solidarity Rights: Progressive Development or Obfuscation of International Human Rights Law?*, 29 NETH. INT'L L. REV. 307 (1982).

24. Nicole B. Ellison & Jessica Vitak, *Social Network Site Affordances and Their Relationship to Social Capital Processes*, in THE HANDBOOK OF THE PSYCHOLOGY OF COMMUNICATION TECHNOLOGY 203, 207 (S. Shyam Sundar ed., 2015).

25. JOSÉ VAN DIJCK, THE CULTURE OF CONNECTIVITY: A CRITICAL HISTORY OF SOCIAL MEDIA 8 (2013).

26. Nicole B. Ellison et al., *Cultivating Social Resources on Social Network Sites: Facebook Relationship Maintenance Behaviors and Their Role in Social Capital Processes*, 19 J. COMPUTER-MEDIATED COMM. 855, 857 (2014); Ellison & Vitak, *supra* note 24, at 208; Weiwu Zhang et al., *The Revolution Will be Networked: The Influence of Social Networking Sites on Political Attitudes and Behavior*, 28 SOC. SCI. COMPUTER REV. 75, 79 (2010).

### A. Informational Resources

Relationships, offline or online, are based on the exchange of information<sup>27</sup> and are a critical component in creating and disseminating information.<sup>28</sup> Considering the importance of SNSs as a means to support relationships, it is no wonder that “[t]he desire to communicate and share content is a primary driver of SNS use.”<sup>29</sup> This close affinity between the exchange of information and relationships seems to underpin Danah Boyd’s and Nicole Ellison’s (2013) *revised* definition of an SNS, which depicts it as follows<sup>30</sup>:

A networked communication platform in which participants 1) have uniquely identifiable profiles that consist of user-supplied content, content provided by other users, and/or system-provided data; 2) can publicly articulate connections that can be viewed and traversed by others; and 3) can consume, produce, and/or interact with streams of user generated content provided by their connections on the site.<sup>31</sup>

Following their definition of an SNS as a “communication platform,”<sup>32</sup> the authors note the profile’s dynamic nature and the different informational sources that compose it, in comparison to the earlier static “self-descriptive” profile in their previous definition.<sup>33</sup> The updated definition also highlights the navigation and participation opportunities that SNSs currently provide and that are based on rich information streams, in which every piece of content is connected to another.<sup>34</sup>

27. See Andrea L. Kavanaugh et al., *Weak Ties in Networked Communities*, 21 INFO. SOC’Y 119, 119 (2005); Ellison & Vitak, *supra* note 24, at 218; Ellison et al., *supra* note 26, at 858.

28. Daniel Z. Levin & Rob Cross, *The Strength of Weak Ties You Can Trust: The Mediating Role of Trust in Effective Knowledge Transfer*, 50 MGMT. SCI. 1477, 1477 (2004).

29. Nicole B. Ellison & Danah M. Boyd, *Sociality Through Social Network Sites*, in THE OXFORD HANDBOOK OF INTERNET STUDIES 151, 159 (William H. Dutton ed., 2013).

30. For the *previous* definition given by the writers, see danah m. boyd & Nicole B. Ellison, *Social Network Sites: Definition, History, and Scholarship*, 13 J. COMPUTER-MEDIATED COMM. 210, 211 (2008). For the challenges in defining SNSs see Rosa Lemel, *A Framework for Developing a Taxonomy of Social Media*, 6 BUS. STUD. J. 67, 67 (2014) (exploring taxonomical difficulties); Jonathan A. Obar & Steve Wildman, *Social Media Definition and the Governance Challenge: An Introduction to the Special Issue*, 39 TELECOMMS. POL’Y 745, 745-46 (2015) (discussing terminological difficulties).

31. Ellison & Boyd, *supra* note 29, at 158 (emphasis omitted).

32. *Id.* at 158-60.

33. *Id.* at 153-55.

34. *Id.* at 155-58; see also Ellison & Vitak, *supra* note 24, at 207. For additional definitions of SNSs, see Andreas M. Kaplan & Michael Haenlein, *Users of the World, Unite! The Challenges and Opportunities of Social Media*, 53 BUS. HORIZONS 59, 63 (2010) (describing SNSs as “applications that enable users to connect by creating personal information profiles, inviting friends and colleagues to have access to those profiles, and sending e-mails and instant messages between each other.”) and OECD, PARTICIPATIVE WEB AND USER-CREATED CONTENT: WEB 2.0, WIKIS AND

The content exchanged via the myriad of interpersonal interactions occurring throughout SNSs creates an enormous trail of information that may cover countless topics. This outsized information may flow through different types of connections and sociability patterns. One prominent pattern relates to the maintenance and cultivation of relationships.<sup>35</sup> Users might deliver a wide range of messages, among which are mundane, casual communications, responses to requests and inquires of their friends, signals of social support (for instance, comments on a friend's post that conveys distress or woes), and signs of fondness or endorsement (for example, by tapping the "like" button or by tagging users).<sup>36</sup> These information-based gestures of attention<sup>37</sup> also raise their senders' expectations to receive similar tokens in return.<sup>38</sup> This reciprocity might exist in a relationship between two people, or in a more generalized public context, by encouraging individuals to "pay forward" different goods.<sup>39</sup> The publicity of many of these gestures creates "tie signs" visible to others and displays relational links among users<sup>40</sup> thereby further strengthening these ties.<sup>41</sup> All these behaviors drive the cyclic creation and dissemination of content among users and their various connections, be they close friends, peers, bosses, or others, consequently fostering mass production of information on politics, society, culture, religion, science, education, health, and more.

News of any kind is similarly affected by these social mechanisms, therefore consumed, shared, and disseminated like any other content that

---

SOCIAL NETWORKING 16, 19 (2007) (suggesting an essentially similar definition). For a sociological approach, see Bernie Hogan & Barry Wellman, *The Relational Self-Portrait: Selfies Meet Social Networks*, in *SOCIETY AND THE INTERNET: HOW NETWORKS OF INFORMATION AND COMMUNICATION ARE CHANGING OUR LIVES* 53, 53-54, 61-63 (Mark Graham & William H. Dutton eds., 2014).

35. Victoria Schwanda Sosik & Natalya N. Bazarova, *Relational Maintenance on Social Network Sites: How Facebook Communication Predicts Relational Escalation*, 35 *COMPUTERS HUM. BEHAV.* 124, 124 (2014); Stephanie Tom Tong & Joseph B. Walther, *Relational Maintenance and CMC*, in *COMPUTER-MEDIATED COMMUNICATION IN PERSONAL RELATIONSHIPS* 98, 105-07 (Kevin B. Wright & Lynne M. Webb eds., 2011).

36. Ellison et al., *supra* note 26, at 858; Tong & Walther, *supra* note 35, at 106.

37. Tong & Walther, *supra* note 35, at 110-11. Such gestures might differ in the effort they entail: "Liking," for instance, requires a lower burden than writing a comment. The chosen gesture might, therefore, indicate the resources that one is willing to invest in cultivating a certain relationship. Ellison & Vitak, *supra* note 24, at 220; Sosik & Bazarova, *supra* note 35, at 125.

38. Ellison & Vitak, *supra* note 24, at 220.

39. *See id.* *See also* ROBERT D. PUTNAM, *BOWLING ALONE: THE COLLAPSE AND REVIVAL OF AMERICAN COMMUNITY* 134-47 (2000) (discussing the benefits of "generalized reciprocity").

40. Ellison & Vitak, *supra* note 24, at 219-21.

41. Tong & Walther, *supra* note 35, at 112-13.

flows through SNSs.<sup>42</sup> The importance of this is underscored by the large portion of users who acquire their news via SNSs.<sup>43</sup>

Another important factor with regard to availability and dissemination of information lies in users' "weak ties," which are the bonds that users share with people who are not part of their immediate milieu of friends and family.<sup>44</sup> SNSs provide users a platform for a simple and low-cost support in this type of ties, which might otherwise have faded away.<sup>45</sup> These "weak ties" contribute significantly to the flow of information, as they connect different social networks.<sup>46</sup> Thanks to this function, they provide users with diverse viewpoints and afford access to vast information and to new opportunities.<sup>47</sup>

The informational resources provided by SNSs are thus an integral part of the interpersonal interactions that flourish throughout these platforms and build on the creation, exchange, and dissemination of content within them. This information might arrive from various sources and relate to a variety of topics, touching on all areas of life. When such information is constantly delivered and aggregated, it becomes an important component in users' ability to exercise the wide range of their rights. Realizing the right to vote and to be elected, for instance, requires fair and reliable information on the various issues on the agenda and on

---

42. Ana Lucia Schmidt et al., *Anatomy of News Consumption on Facebook*, 114 PROC. NAT'L ACAD. SCI. 3035, 3035, 3038 (2017).

43. *Id.*

44. Kavanaugh et al., *supra* note 27, at 119-20. The authors explained that "[t]he strength of a tie is a combination of the amount of time, emotional intensity, intimacy (mutual confiding), and reciprocal services that characterize the tie." *Id.* at 120. Social ties also differ in the type of "social capital" they afford. Weak ties provide "bridging social capital" whereas strong ties afford "bonding social capital." *See* PUTNAM, *supra* note 39, at 22-23. The latter is associated with support and intimacy and with a higher level of trust, while the former is ascribed to informational resources. Kavanaugh et al., *supra* note 27, at 119-20; Ellison et al., *supra* note 26, at 856. "Social capital" could be defined as the "resources embedded in one's social networks, resources that can be accessed or mobilized through ties in the networks." Nan Lin, *A Network Theory of Social Capital*, in HANDBOOK OF SOCIAL CAPITAL 50, 51 (Dario Castiglione, Jan W. van Deth & Guglielmo Wolleb eds., 2008). For another definition that focuses on reciprocity and trust, see PUTNAM, *supra* note 39, at 19. Besides offline, these resources are also exchanged via online interpersonal interactions in general, and in SNSs in particular. Adam N. Joinson, 'Looking at', 'Looking up' or 'Keeping up with' People?: *Motives and Uses of Facebook*, in PROCEEDINGS OF ACM CHI 2008 CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS 1027, 1027-28 (2008).

45. Ellison & Boyd, *supra* note 29, at 159; Ellison & Vitak, *supra* note 24, at 212; Tong & Walther, *supra* note 35, at 110-11; *see also* Judith Donath, *Signals in Social Supernets*, 13 J. COMPUTER-MEDIATED COMM. 231, 235-36 (2007) ("SNSs can provide a richer social context for people one knows only superficially. Seeing who other people know and how they treat and are treated by others provides important cues for understanding them.").

46. Ellison & Vitak, *supra* note 24, at 209, 211; Kavanaugh et al., *supra* note 27, at 119-20 (explaining that "[w]ithout bridging ties to different groups, cliques lack the interpersonal ties that help spread information or innovation conveyed by mass media and other sources.").

47. Ellison & Vitak, *supra* note 24, at 211; Kavanaugh et al., *supra* note 27, at 120.

the candidates at hand. Likewise, realizing the right to work requires information pertaining to employment opportunities, professional innovation and updates, the nature of the relevant market, competitors, etc.

The importance and outcomes of the informational resources provided by SNSs are reflected in the U.S. Supreme Court's ruling in the famous *Packingham*<sup>48</sup> case, according to which, a statute that prohibited sex offenders from using SNSs was in violation of the First Amendment. The Court asserted that these platforms "are the principal sources for knowing current events, checking ads for employment, speaking and listening in the modern public square, and otherwise exploring the vast realms of human thought and knowledge."<sup>49</sup>

### B. Social-Communal Resources

The second type of key resources provided by the social dynamics generated in SNSs are social-communal resources. These resources focus on social and collective virtues, such as social support, mutual help, civic engagement, and collective action, that the use of SNSs affords.

These social-communal benefits rely on sentiments such as trust and belongingness which are often available throughout these spheres. SNSs encourage a sense of belonging for members of a certain group or community and stimulate the creation of a collective identity.<sup>50</sup> This holds for explicit communities, such as residents of a certain city, students of a specific faculty, professionals in a given branch, users who share the same political agendas or health concerns, etc. After all, "[t]oday's social network sites . . . are simply the latest, and arguably most complete, online social venues where virtual communities might form."<sup>51</sup>

Ties created in the settings of community and linkage among members are often also associated with trust.<sup>52</sup> Trust encourages users to

---

48. *Packingham v. North Carolina*, 137 S. Ct. 1730 (2017).

49. *Id.* at 1734. The Court also asserted that "[w]hile in the past there may have been difficulty in identifying the most important places (in a spatial sense) for the exchange of views, today the answer is clear. It is cyberspace—the 'vast democratic forums of the Internet' in general . . . and social media in particular." *Id.* at 1735.

50. Ellison & Boyd, *supra* note 29, at 162-65; Ellison & Vitak, *supra* note 24, at 218.

51. Malcolm R. Parks, *Social Network Sites as Virtual Communities*, in *A NETWORKED SELF: IDENTITY, COMMUNITY, AND CULTURE ON SOCIAL NETWORK SITES* 105, 117 (Zizi Papacharissi ed., 2010).

52. Zhang et al., *supra* note 26, at 77. Such trust might be "thick" (as often ascribed to strong ties) or "thin" (associated with weak ties). Kavanaugh et al., *supra* note 27, at 120. For the contribution of SNSs to the creation of trust among users, see Donath, *supra* note 45, at 235-36.

present their contacts with diverse queries and requests,<sup>53</sup> it betters the odds that such inquiries will be positively and duly answered,<sup>54</sup> and it elevates the value of the information provided to the users from their point of view.<sup>55</sup> The linkage among users creates accountability and motivation and fosters a more productive exchange of information.<sup>56</sup> Such processes in turn might further amplify trust among members, thus fueling and maintaining this cyclic process over and over again.<sup>57</sup> Trust also improves the chances that members of the community will work together toward certain common goals<sup>58</sup> and may promote active participation and civic and political engagement in various topics.<sup>59</sup>

The importance of such communities might further increase when their members belong to minorities and other vulnerable groups. For these audiences, “logging on to Facebook . . . is a way of building safe and supportive communities, sharing resources or planting the seeds of political change.”<sup>60</sup>

In addition, the weak ties discussed above regarding informational resources are also of value here, with respect to social-communal resources. Whereas “strong ties” are thought to fortify the community’s impregnability, weak ties are a source of participation and civic engagement.<sup>61</sup> People who are connected to others by weak ties tend to be connected to several social networks and are often characterized as having high levels of activism, participation, civic engagement, and community attachment.<sup>62</sup> SNSs that strongly support weak ties as noted earlier, enable these activists even to elevate their reach and influence among other people.<sup>63</sup>

---

53. Ellison & Vitak, *supra* note 24, at 209; Zhang et al., *supra* note 26, at 77.

54. Ellison & Vitak, *supra* note 24, at 222-23.

55. Leticia Bode et al., *A New Space for Political Behavior: Political Social Networking and its Democratic Consequences*, 19 J. COMPUTER-MEDIATED COMM. 414, 424-25 (2014); Homero Gil de Zúñiga et al., *Social Media Use for News and Individuals’ Social Capital, Civic Engagement and Political Participation*, 17 J. COMPUTER-MEDIATED COMM. 319, 331 (2012).

56. Ellison & Boyd, *supra* note 29, at 159.

57. Gil de Zúñiga et al., *supra* note 55, at 331; Kavanaugh et al., *supra* note 27, at 120.

58. Gil de Zúñiga et al., *supra* note 55, at 331; Zhang et al., *supra* note 26 at 76-77.

59. Bode et al., *supra* note 55, at 424; Thomas J. Johnson et al., *United We Stand?: Online Social Network Sites and Civic Engagement*, in *A NETWORKED SELF: IDENTITY, COMMUNITY, AND CULTURE ON SOCIAL NETWORK SITES* 185, 201 (Zizi Papacharissi ed., 2010).

60. Lil Miss Hot Mess, *Facebook’s ‘Real Name’ Policy Hurts Real People and Creates a New Digital Divide*, GUARDIAN (June 3, 2015, 7:30 AM), <https://www.theguardian.com/commentisfree/2015/jun/03/facebook-real-name-policy-hurts-people-creates-new-digital-divide>.

61. Johnson et al., *supra* note 59, at 189 (“Bridging social capital involves connecting more heterogeneous groups of people to bring about social and political change.”); Kavanaugh et al., *supra* note 27, at 120.

62. Kavanaugh et al., *supra* note 27, at 128.

63. *Id.* at 129.

Civic engagement is also encouraged through the establishment of discussions and deliberations on various topics that are an integral part of these platforms.<sup>64</sup> Information is just more likely to lead to civic engagement when delivered by interpersonal ties.<sup>65</sup>

The social-communal resources provided by SNSs therefore form the settings for sentiments such as trust and belongingness, which in turn, together with other factors, afford manifestations of social support, mutual help, collective action, and civic engagement. These sentiments and manifestations create the criteria for various social-communal behaviors that have become so common on SNSs, such as colleagues advising each other, students tutoring their peers, citizens organizing demonstrations, and workers striking.

Such behaviors, thus, afford realization of various civil-political rights as well as economic-social-cultural rights.<sup>66</sup>

By explaining how SNSs have become a source for the production, exchange, and dissemination of mass information on the one hand and for mutual help, social support, and civic participation on the other, the informational and social-communal resources shed light on the unique social mechanisms that propel SNSs' influence on a wide range of rights. This attribute of SNSs constitutes the first dimension of their emergence as arenas for crafting and allocating human rights. The second dimension, concerning the control SNSs have obtained over this very wide range of rights, is discussed next.

### III. PRACTICES OF CONTROL PERFORMED BY SNSs

The second dimension of the newly emerging digital arenas relates to the effective control that SNSs have acquired over the wide range of users' human rights, discussed in Part II.<sup>67</sup> This control regime is created and obtained through various practices that are incorporated in SNSs' operation and impact users' human rights, each in a different manner and volume. The following discussion focuses on five major categories of such practices: surveillance; profiling; censorship; practices relating to SNSs-authorities nexus; and finally, expulsion of individuals and

---

64. Johnson et al., *supra* note 59, at 202.

65. Kavanaugh et al., *supra* note 27, at 120. This holds also for weak ties. See Donath, *supra* note 45, at 236 ("One of the most valuable contributions of SNSs is their potential to add trust to weak ties. Trusted weak ties are very useful sources of information, combining the heterogeneity that such ties generally have with the believability that comes with trust.").

66. Informational and social-communal resources are not rigidly divided, but rather strongly connected. Social-communal resources, for example, both produce information and rely on it, as shown above. See *supra* Part II.B.

67. See *supra* Part II.

communities, as I elaborate in this Part. Together, all these practices result in intense and sophisticated control over various rights of SNS users.

SNSs' competence to influence such a broad spectrum of rights relies on their nature as online intermediaries, which provides them with considerable power to determine and affect the nature and scope of the informational and social-communal resources discussed earlier.<sup>68</sup> The control regime exerted by SNSs should also be understood against the backdrop of the increasing deployment of algorithmic and particularly AI tools throughout these platforms. AI-based technologies are integrated into various functions of SNSs, determining the flow of information within them, and "driving nearly all aspects of user experience," while making numerous decisions.<sup>69</sup> Among other ramifications, AI's prevalence in SNSs further magnifies and intensifies the control that SNSs exercise over their users' human rights.

### A. Surveillance

SNSs execute unprecedented, massive surveillance throughout their platforms, including collecting, storing, and processing information about their users.<sup>70</sup> Such surveillance aims to promote different objectives. Some lie in external demands, such as legal provisions that apply to them,<sup>71</sup> and some are fueled by self-driven initiatives.<sup>72</sup>

The surveillance conducted by SNSs is extremely extensive. It covers a myriad of aspects of users' personal, professional, and social lives, and pertains to their political agendas, their cultural-ethnic background, and their feelings and preferences in diverse areas.<sup>73</sup> Some

---

68. See generally OECD, THE ROLE OF INTERNET INTERMEDIARIES IN ADVANCING PUBLIC POLICY OBJECTIVES (2011); Derek E. Bambauer, *Middlemen*, 65 FLA. L. REV. F. 1 (2013). The current powers and attributes of SNSs and other intermediaries stem from various factors, among which is the vast-immunity regime provided by 47 U.S.C. § 230 (2012). See Niva Elkin-Koren, *After Twenty Years: Revisiting Copyright Liability of Online Intermediaries*, in THE EVOLUTION AND EQUILIBRIUM OF COPYRIGHT IN THE DIGITAL AGE 29, 29-30 (Susy Frankel & Daniel Gervais eds., 2014). See generally, Paul Ehrlich, *Communications Decency Act § 230*, 17 BERKELEY TECH. L.J. 401 (2002).

69. KIM HAZELWOOD ET AL., APPLIED MACHINE LEARNING AT FACEBOOK: A DATACENTER INFRASTRUCTURE PERSPECTIVE (2018), <https://research.fb.com/wp-content/uploads/2017/12/hpca-2018-facebook.pdf>.

70. As shall be discussed in this Part of the Article. See *infra* Part III.A.

71. For discussions about NetzDG and the Code of Conduct, see *infra* Part III.C.1, D.2.

72. For example, tailoring the commercial and non-commercial content that will be presented to users. See *infra* Part III.B.

73. See Yuval Karniel & Amit Lavie-Dinur, *Privacy in New Media in Israel: How Social Networks are Helping to Shape the Perception of Privacy in Israeli Society*, 10 J. INFO. COMM. & ETHICS SOC'Y 288, 290 (2012) (referring to four types of "privacy exposure" that exist in SNSs: (1) factual (personal) exposure (personal data); (2) visual exposure (physical representations); (3)

of the surveilled information is meant by users to be shared with others or with a limited audience (inter alia, in private messages).<sup>74</sup> However, a considerable volume of the information collected and used by SNSs may relate to data that users do not necessarily intend to share with anyone. Such information might lie in users' digital surfing patterns and include, for example, searches they have conducted, links they have "clicked" on,<sup>75</sup> duration of time spent on reading or watching a certain piece of content, and instances in which users turn on sound or switch to "full screen" mode when watching videos.<sup>76</sup> Moreover, the data collected might concern users' mental and psychological state and include indications that they themselves are not familiar with.<sup>77</sup>

SNSs' surveillance capabilities have advanced horizontally and vertically over time.<sup>78</sup> Horizontal growth lies in the variety of sources these platforms rely on, including acquired companies (in Facebook's case, Instagram and WhatsApp, among others), partnering platforms, and websites or mobile applications.<sup>79</sup> Vertical growth reflects the richness of the types of information that SNSs obtain, inter alia, by adding new tracking capabilities such as those afforded by "social plugins." These are "like buttons" inserted into third parties' websites, including sites that might contain private-sensitive information (such as

---

identity and emotions' exposure; (4) preferences' exposure).

74. Monika Bickert & Brian Fishman, *Hard Questions: Are We Winning the War on Terrorism Online?*, FACEBOOK NEWSROOM (Nov. 28, 2017), <https://newsroom.fb.com/news/2017/11/hard-questions-are-we-winning-the-war-on-terrorism-online>. See Facebook's answer to a users' question regarding the nature of the surveillance conducted by the platform to battle terror: "Our policies apply across the site, including posts, comments, messages, pages and events." *Id.*

75. Ira S. Rubinstein, Ronald D. Lee & Paul M. Schwartz, *Data Mining and Internet Profiling: Emerging Regulatory and Technological Approaches*, 75 U. CHI. L. REV. 261, 270-73 (2008); see also *Facebook's Data Pool*, EUR. VERSUS FACEBOOK, [http://europe-v-facebook.org/EN/Data\\_Pool/data\\_pool.html#Connections](http://europe-v-facebook.org/EN/Data_Pool/data_pool.html#Connections) (last visited Feb. 3, 2019) (documenting the categories of personal data that are being collected by Facebook, as reflected in the review provided to users who request it).

76. Vibhi Kant & Jie Xu, *Taking into Account Live Video When Ranking Feed*, FACEBOOK NEWSROOM (Mar. 1, 2016), <https://newsroom.fb.com/news/2016/03/news-feed-fyi-taking-into-account-live-video-when-ranking-feed>; Meihong Wang & Yue Zhuo, *Taking into Account More Actions on Videos*, FACEBOOK NEWSROOM (June 29, 2015), <https://newsroom.fb.com/news/2015/06/news-feed-fyi-taking-into-account-more-actions-on-videos>.

77. See discussion *infra* Part III.B.

78. BRENDAN VAN ALSENOY ET AL., FROM SOCIAL MEDIA SERVICE TO ADVERTISING NETWORK: A CRITICAL ANALYSIS OF FACEBOOK'S REVISED POLICIES AND TERMS 10 (2015), <https://www.law.kuleuven.be/citip/en/news/item/facebook-revised-policies-and-terms-v1-3.pdf> (addressing Facebook).

79. In the context of Twitter, see @KyleB, *Introducing Partner Audiences*, TWITTER (Mar. 5, 2015), [https://blog.twitter.com/marketing/en\\_us/a/2015/introducing-partner-audiences.html](https://blog.twitter.com/marketing/en_us/a/2015/introducing-partner-audiences.html).

health and governmental sites), and that enable cookies-based surveillance of users' (and non-users') browsing activity.<sup>80</sup>

The embrace of algorithmic and AI mechanisms by SNSs considerably heightens the surveillance conducted by them and further tightens the control it affords. These AI-based technologies allow outsized monitoring and analyzing routines through various tools including matching photos and videos, extracting meaning from texts, and providing facial recognition.<sup>81</sup>

But how does SNS surveillance impact the wide range of rights addressed earlier? Put differently, how does the said surveillance regime determine the availability of the informational and social-communal resources that are provided to SNSs' users?

Surveillance, as explained by Foucault, is a method of control induced from a coerced visibility and that provides those who exercise its disciplinary power. The subjects' visibility, he explained, "assures the hold of the power that is exercised over them. It is the fact of being constantly seen, of being able always to be seen, that maintains the disciplined individual in his subjection."<sup>82</sup> Such "coerced visibility" and self-discipline also occur in the context of SNSs' massive surveillance and might lead to a "chilling effect" caused by "the normalizing of behavior" induced by SNSs' surveillance.<sup>83</sup> Surveillance might deter users from interacting with certain information, individuals, and groups, thereby preventing them from enjoying the full scope of both informational and social-communal resources necessary for proper realization of their various human rights. This restraint might concern users' financial or health condition, their political opinions, the social or religious dilemmas they face, and many other issues they might feel reluctant to share, sometimes in the very nascent phases of obtaining knowledge and forming beliefs and opinions.<sup>84</sup>

---

80. VAN ALSENOY ET AL., *supra* note 78, at 89-90; *see also* Dia Kayyali, *New Report Shows European Data Protection Authorities are Taking Facebook's Questionable Terms of Service Seriously*, ELECTRONIC FRONTIER FOUND. (Feb. 26, 2015), <https://www EFF.ORG/deeplinks/2015/02/new-report-shows-european-data-protection-authorities-are-taking-facebooks> (addressing Alsenoy et al.'s report).

81. Bickert & Fishman, *supra* note 74.

82. MICHEL FOUCAULT, *DISCIPLINE AND PUNISH: THE BIRTH OF THE PRISON* 187 (Alan Sheridan trans., Vintage Books 2d ed. 1995) (1977).

83. Ben Marder et al., *The Extended 'Chilling' Effect of Facebook: The Cold Reality of Ubiquitous Social Networking*, 60 *COMPUTERS HUM. BEHAV.* 582, 583 (2016).

84. *See* Elkin-Koren, *supra* note 68, at 43. For an empirical demonstration of the troubling ramifications of online "chilling effect," *see* Jonathon W. Penney, *Chilling Effects: Online Surveillance and Wikipedia Use*, 31 *BERKELEY TECH. L.J.* 117, 125-29 (2016) (showing that traffic to certain Wikipedia articles has decreased in the aftermath of NSA/PRISM surveillance programs revelations).

Moreover, with the passage of time, if many users hold back from creating or sharing certain information or interacting with certain people, the overall informational and social-communal picture depicted might be distorted and incomplete. This could undermine all users' ability to properly realize their rights within SNSs' realms.<sup>85</sup> Such an outcome might be worsened by certain algorithmic features, among them the preference that SNSs might give to already popular content when displaying information to users.<sup>86</sup>

To highlight my point about the chilling effect that might impede a broad range of rights online, I would like to go back to Foucault, who argued that alongside the visibility of the surveeye, "disciplinary power" itself "is exercised through its invisibility."<sup>87</sup> This holds in the SNSs context as well. SNSs conduct surveillance in the dark, with abundant vagueness, which might aggravate the challenges created to human rights.

### B. Profiling

Another control practice performed by SNSs is profiling. In general, profiling, or "social sorting," relates to processes of classifying, categorizing, and clustering users according to various criteria, to serve for assessment, measurement, comparison, and judgment of users.<sup>88</sup> This practice is a salient manifestation of control achieved by surveillance<sup>89</sup> and could be understood as the depiction of each individual as a case. Foucault perceived profiling as conduct that "at one and the same time constitutes an object for a branch of knowledge and a hold for a branch of power."<sup>90</sup> In the SNSs context, profiling might be

any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse [sic] or predict aspects

---

85. *E.g.*, Complaint for Declaratory and Injunctive Relief at 23-24, Knight First Amendment Inst. v. Trump, 302 F. Supp. 3d 541 (2018) (No. 1:17-cv-05205), 2017 WL 2952634 (arguing that the blockage of users from Presidents' Twitter account "violates the Knight Institute's First Amendment rights by distorting the expressive forum in which the Knight Institute and other non-blocked users participate").

86. Stuart Dredge, *How Does Facebook Decide What to Show in my News Feed?*, GUARDIAN (June 30, 2014, 6:49 AM), <https://www.theguardian.com/technology/2014/jun/30/facebook-news-feed-filters-emotion-study>.

87. FOUCAULT, *supra* note 82, at 187.

88. David Lyon, *Surveillance as Social Sorting: Computer Codes and Mobile Bodies*, in SURVEILLANCE AS SOCIAL SORTING: PRIVACY, RISK, AND DIGITAL DISCRIMINATION 13-15 (David Lyon ed., 2003); Rubinstein, Lee & Schwartz, *supra* note 75, at 271-73.

89. FOUCAULT, *supra* note 82, at 187.

90. *Id.* at 191.

concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location or movements.<sup>91</sup>

SNSs use profiling for various purposes, a prominent one lies in targeted advertisements. These advertisements are the bedrock of SNSs' business model, leveraging their profits and subsidizing their operation<sup>92</sup> and are the factor that guides many aspects of these platforms' conduct.<sup>93</sup> The content subject to advertisement through SNSs is by no means limited to products and services in their traditional sense, but include diversified promoted content such as posts of individuals, groups, non-profit organizations, political ads, etc.<sup>94</sup>

Profiling processes may also pertain to SNSs' efforts to personalize the general, non-commercial content to be displayed and available to users, including the content in their feed and new friends and pages suggested to them. This kind of profiling serves to foster users' engagement with content and to augment time spent using these platforms.<sup>95</sup>

Alongside the benefits provided by profiling practices, including the ability to match users with information that could be of interest to them,<sup>96</sup> they might endanger users' human rights in various ways.

First, the aspiration to tailor the information shown to users according to their interests might result in the famous "Filter Bubbles."<sup>97</sup> These depict situations wherein content (and in the SNSs context also

91. Regulation (EU) 2016/679 On the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), art. 4(4), 2016 O.J. (L 119) 1, 1-88.

92. FED. TRADE COMM'N, ONLINE PROFILING: A REPORT TO CONGRESS 9-10 (2000).

93. *Id.* Twitter, for instance, suggests various targeting options, based, among other factors, on language, interests, an account's followers, keywords used by users (when searching or tweeting), and offline consuming patterns. *Ad Targeting: Connect with a Relevant Audience*, TWITTER BUS., <https://business.twitter.com/en/targeting.html> (last visited Feb. 3, 2019); see also OECD, THE ECONOMIC AND SOCIAL ROLE OF INTERNET INTERMEDIARIES 16-22 (2010) (discussing online intermediaries' ads-driven business model and "two-sided markets").

94. See discussion on political ads *infra* text accompanying notes 113-14.

95. Adam Mosseri, *Building a Better News Feed for You*, FACEBOOK NEWSROOM (June 29, 2016), <https://newsroom.fb.com/news/2016/06/building-a-better-news-feed-for-you>. Mosseri explained that:

Our aim is to deliver the types of stories we've gotten feedback that an individual person most wants to see. We do this not only because we believe it's the right thing but also because it's good for our business. When people see content they are interested in, they are more likely to spend time on News Feed and enjoy their experience.

*Id.*

96. FED. TRADE COMM'N, *supra* note 92, at 8-10.

97. ELI PARISER, THE FILTER BUBBLE: HOW THE NEW PERSONALIZED WEB IS CHANGING WHAT AND HOW WE THINK 5, 9 (2011).

individuals and groups) which aligns with users' opinions and preferences—is amplified at the expense of content (and people) representing different viewpoints. Different opinions and content are thus weeded out, creating a distorted and incomplete informational and social-communal picture that will be presented to users.<sup>98</sup> This is a rickety foundation for forming opinions, establishing relationships, and making decisions in various areas of life. It affects a person's ability to realize human rights, which depends on a fair informational structure in these fields. It might even hinder users' ability to properly grasp reality and take personal, civic, or political action when needed.<sup>99</sup> In addition, preference given to like-minded content, individuals, and groups cultivates “likeminded cliques” that might have adverse effects on social polarization, stereotyping, and trust of those who are different from them.<sup>100</sup> Such outcomes, besides their troubling ramifications with regard to securing equal rights and equal opportunities for different audiences, might also foster misinformation and the spread of fake news and false rumors.<sup>101</sup> This in turn impedes users' reliance on informational and social-communal resources in SNSs as a means to realize their human rights.

Second, the profiles created by SNSs may determine the scope of opportunities and resources provided to users in various life domains. This might occur as individuals and groups are “quietly” being disfavored and their social mobility being unfairly hindered, sometimes with circumvention of regulations that might prohibit such practices.<sup>102</sup> Profiling processes might, for instance, influence the premiums charged by insurance companies (according to health risks identified in one's account),<sup>103</sup> affect a user's likelihood of receiving a desirable loan offer, or determine whether a housing advertisement will be available to her.<sup>104</sup> Profiles might also influence users' potential for employment<sup>105</sup> or their

---

98. *Id.* at 9-10.

99. Engin Bozdag & Jeroen van den Hoven, *Breaking the Filter Bubble: Democracy and Design*, 17 ETHICS INFO. TECH. 249, 249-52 (2015).

100. Nicholas DiFonzo, *Rumour Research Can Douse Digital Wildfires*, 493 NATURE 135, 135 (2013); Schmidt et al., *supra* note 42, at 3035.

101. DiFonzo, *supra* note 100, at 135; Schmidt et al., *supra* note 42, at 3038.

102. Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 99-101 (2014).

103. FED. TRADE COMM'N, *supra* note 92, at 13.

104. Crawford & Schultz, *supra* note 102, at 99-101.

105. This is underscored in light of the efforts done by SNSs such as Facebook to establish themselves as avenues for matching users and employees. Among other measures, employers may post job openings and target the desired users. *E.g.*, Alex Himel, *Helping People Find Jobs and Local Businesses Hire*, FACEBOOK NEWSROOM (Feb. 28, 2018), <https://newsroom.fb.com/news/2018/02/jobs/>; *see also* Josh Constine, *Facebook Rolls Out Job Posts to Become the Blue-Collar*

admission to certain educational institutions.<sup>106</sup> These kinds of risks were recently demonstrated when Facebook enabled its advertisers to exclude various ethnic groups from the audience exposed to the ads<sup>107</sup> or to target anti-Semitic users.<sup>108</sup> Discrimination or unfair allocation of opportunities based on classification of SNSs' users might be very elusive and hard to detect by users themselves, Non-Governmental Organizations ("NGOs"), or enforcement authorities.<sup>109</sup> As put by Michael Fertik: "If you live on the wrong side of the digital tracks, you won't even see a credit offer from leading lending institutions, and you won't realize that loans are available to help you with your current personal or professional priorities."<sup>110</sup>

Third, the profiles created might build on extremely sensitive information, regarding users' personality, preferences, and connections; identify their vulnerabilities and facilitate their abuse by these platforms or by other stakeholders.<sup>111</sup> These profiles might depict users' social, emotional, or psychological state, while hindering privacy and autonomy and causing additional losses. It was reported that Facebook informed advertisers of its ability to monitor posts and photos in real time to identify situations in which teenagers felt "defeated," "overwhelmed," "anxious," "useless," and a "failure."<sup>112</sup> Another area in which such profiles can be abused is the growing industry of targeted political ads. These allow political figures and groups to adjust the propaganda they place on SNSs according to users' personality attributes. In the recent Cambridge Analytica scandal, for instance, users' psychological profiles were designed to support targeted political advertising. Such profiles

---

*LinkedIn*, TECHCRUNCH (Feb. 28, 2018), <https://techcrunch.com/2018/02/28/facebook-job-posts>.

106. For the interest that educational institutions might find in users' accounts, see Darian Somers, *Do Colleges Look at Your Social Media Accounts?*, U.S. NEWS (Feb. 10, 2017, 8:00 AM), <https://www.usnews.com/education/best-colleges/articles/2017-02-10/colleges-really-are-looking-at-your-social-media-accounts>.

107. Julia Angwin & Terry Parris Jr., *Facebook Lets Advertisers Exclude Users by Race*, PROPUBLICA (Oct. 28, 2016, 1:00 PM), <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>.

108. Julia Angwin, Madeleine Varner & Ariana Tobin, *Facebook Enabled Advertisers to Reach 'Jew Haters'*, PROPUBLICA (Sept. 14, 2017, 4:00 PM), <https://www.propublica.org/article/facebook-enabled-advertisers-to-reach-jew-haters>.

109. Crawford & Schultz, *supra* note 102, at 99-101.

110. Michael Fertik, *The Rich See a Different Internet Than the Poor*, SCI. AM. (Feb. 1, 2013, 4:00 PM), <https://www.scientificamerican.com/article/rich-see-different-internet-than-the-poor>.

111. Karniel & Lavie-Dinur, *supra* note 73, at 291; Rubinstein, Lee & Schwartz, *supra* note 75, at 271-73.

112. Sam Levin, *Facebook Told Advertisers it Can Identify Teens Feeling 'Insecure' and 'Worthless'*, GUARDIAN (May 1, 2017, 3:01 PM), <https://www.theguardian.com/technology/2017/may/01/facebook-advertising-data-insecure-teens>; *see also* Karniel & Lavie-Dinur, *supra* note 73, at 291.

relied on five users' traits (known as the "Big Five" or "OCEAN"): openness, conscientiousness, extroversion, agreeableness, and neuroticism.<sup>113</sup> This information was then combined with additional data gathered on users.<sup>114</sup>

Fourth, profiling may further aggravate the aforementioned "chilling effect" induced by surveillance.<sup>115</sup> SNSs might expose users mainly to information akin to their previous use patterns. This could be a constant reminder of SNSs' extensive surveillance and thus cultivate (or worsen) users' self-constraint and reluctance to interact with certain content or people.

### C. Censorship

The third SNS practice of control is censorship. This practice is sophisticatedly and intensely deployed throughout SNSs and increasingly relies on automated filtering mechanisms.<sup>116</sup> SNSs' censorship shapes the informational and social-communal arsenal available to users and thus affects many of their human rights. It is executed in three different time-frames: ex-ante, real-time, and ex-post. All three share similarities in their influence on human rights, but each one of them bears unique attributes which should be taken into account.<sup>117</sup> While it is certainly essential and desirable to remove certain content, it should be performed cautiously and level-headedly. The way this censorship is currently performed in SNSs in all three time-frames, however, poses risks to human rights, as discussed next.

#### 1. Ex-Post Censorship

Ex-post censoring in SNSs is the removal of pre-existing content, whether reactively, based on users' reports, or proactively, following

---

113. *Opinion 3/2018: Opinion on Online Manipulation and Personal Data*, EUR. DATA PROTECTION SUPERVISOR 8 (Mar. 2018), [https://edps.europa.eu/sites/edp/files/publication/18-03-19\\_online\\_manipulation\\_en.pdf](https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf).

114. *See id.* at 8-9. To the role SNSs play in political targeted campaigns, see Robert Booth, *Facebook Employs Ex-Political Aides to Help Campaigns Target Voters*, GUARDIAN (May 8, 2017, 2:06 PM), <https://www.theguardian.com/technology/2017/may/08/facebook-political-aides-campaigns-target-voters>.

115. *See infra* Part III.A.

116. *Community Standards Enforcement Preliminary Report*, FACEBOOK TRANSPARENCY (Nov. 2018), <https://transparency.facebook.com/community-standards-enforcement> ("How much content do we take action on?" in "Violence and Graphic Content").

117. Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1635-47 (2018) (addressing ex-ante content moderation, ex-post proactive manual content moderation, and ex-post reactive manual content moderation, as performed by Facebook, Twitter, and YouTube).

self-initiated and algorithm-based searches for hazardous content.<sup>118</sup> This is a constant and common practice whereby SNSs might endanger the informational and the social-communal resources considered above and thus affect a wide range of their users' human rights.<sup>119</sup>

Silencing content in SNSs also means hindering one's ability to engage with individuals and groups that are connected to this content, as activists and people who share the same views or struggle with the same woes or dilemmas. Indeed, SNSs have long assumed the important role of articulating the limits of public discourse and determining the normative boundaries of "right" and "wrong." Consider the case of nakedness in cultural-historical creations: in 2016, Facebook removed a Norwegian author's post that explored some photos that "changed the history of warfare."<sup>120</sup> The reason for that sanction was the Pulitzer Prize-winning "napalm girl" photo, which documented a naked child running in horror in the aftermath of napalm bombing during the Vietnam war.<sup>121</sup> This sparked a huge protest against Facebook, in which Norway's Prime Minister, Erna Solberg, uploaded the photo to her Facebook account, resulting in the deletion of her post as well. Eventually Facebook backtracked, commenting that the value of permitting the photo outweighed protecting the community by removing it.<sup>122</sup> Yet the lesson was apparently not learned. More recently, it was reported that Facebook censored a photo of a monumental prehistoric Venus of Willendorf statue depicting a naked woman and addressed it as "dangerously pornographic."<sup>123</sup> Another case of censoring art concerns a photo of the landmark sixteenth-century statue of the sea god Neptune in Bologna, Italy.<sup>124</sup> SNSs might also censor body images that they consider "undesirable" by prohibiting ads containing them.<sup>125</sup> Another

---

118. *Id.* at 1638-47.

119. Maayan Perel & Niva Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, 19 STAN. TECH. L. REV. 473, 491 (2016).

120. Alice Ross & Julia Carrie Wong, *Facebook Deletes Norwegian PM's Post as 'Napalm Girl' Row Escalates*, GUARDIAN (Sept. 9, 2016, 8:29 AM), <https://www.theguardian.com/technology/2016/sep/09/facebook-deletes-norway-pms-post-napalm-girl-post-row>.

121. *Id.*

122. *Id.*

123. Dagmar Breitenbach, *Facebook Apologizes for Censoring Prehistoric Figurine 'Venus of Willendorf'*, DW (Mar. 1, 2018), <http://www.dw.com/en/facebook-apologizes-for-censoring-prehistoric-figurine-venus-of-willendorf/a-42780200>.

124. *Censorship Continues on Online Platforms*, FREEMUSE (Feb. 21, 2018) <https://freemuse.org/news/censorship-continues-online-platforms>. In both cases, however, Facebook eventually regretted the removal of the photos. *Id.*; see Breitenbach, *supra* note 123.

125. See Will Heilpern, *Facebook Apologizes After Banning an 'Undesirable' Ad Featuring Plus-Size Model Tess Holliday*, BUS. INSIDER (May 24, 2016, 6:21 AM), <http://www.businessinsider.com/facebook-sorry-for-ban-on-fat-model-2016-5> (describing an instance in which the platform did not allow an Australian feminist group's ad that used a plus-size

field to which SNSs pay significant attention is the censoring of hate speech and terror-related content, as shall be elaborated below.<sup>126</sup>

SNSs often enjoy vast discretion in determining their censoring policy. They operate in the absence of clear legal boundaries set by nation-states or supranational institutions to guide them.<sup>127</sup> When more specific legal arrangements apply, such as the German NetzDG or the Code of Conduct, the emphasis is often on the speed and volume of the performed censorship, rather than on thoughtfully guiding and circumscribing it.<sup>128</sup> SNSs are granted such discretion even though their censorship suffers from severe difficulties, in substance and procedure alike. One prominent challenge, which might shed light on some of the examples discussed so far in this Part of the Article, is SNSs' difficulty identifying the context of a given content. This might hold for human moderation,<sup>129</sup> but it is more heavily stressed in the context of algorithmic decision-making, or hybrid processes.<sup>130</sup> Poor extraction of meaning may lead to the censoring of utterly legitimate (even desirable)

---

model to promote an event they organized); Stephen Matthews, *Furious Mother Hits Out at Facebook After a Photo of Her Toddler is Banned - Because Her Rare Skin Condition Looks 'Undesirable' and Would Make Others 'Feel Bad About Themselves'*, DAILY MAIL (Oct. 3, 2017, 6:23 AM), <http://www.dailymail.co.uk/health/article-4943816/Furious-mother-hits-Facebook.html> (discussing the banning the photo of a toddler displayed by her mother to raise awareness of her rare skin disease).

126. See *infra* Part III.D.2.

127. See *infra* Part III.D.2.

128. The *Act to Improve Enforcement of The Law in Social Networks* ("NetzDG") entered into force on October 10, 2017 and requires SNSs to swiftly evaluate and remove or restrict access to content reported as illegal, as prescribed by selected provisions of the German Criminal Code. Except for evaluation processes of the reported content, the Act requires SNSs to store the banned content for ten weeks. Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken [NetzDG] [Act to Improve Enforcement of the Law in Social Networks], Oct. 1, 2017, NETZWERKDURCHSETZUNGSGESETZ VOM 1 at § 3(2)(4) (Ger.), [https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG\\_engl.pdf?\\_\\_blob=publicationFile&v=2](https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG_engl.pdf?__blob=publicationFile&v=2); see *Germany: The Act to Improve Enforcement of the Law in Social Networks*, ARTICLE 19 (2017). For elaboration about the Code of Conduct, see *infra* Part III.D.2.

Additional legal provisions and sources might relate to SNSs' obligation to remove certain content that breaches upon criminal law, such as child pornography. This kind of content is not included in the immunity provided to online intermediaries under the CDA-230. See *supra* note 68 and accompanying text.

129. Monika Bickert, *At Facebook We Get Things Wrong – But We Take Our Safety Role Seriously*, GUARDIAN (May 22, 2017, 6:59 PM), <https://www.theguardian.com/commentisfree/2017/may/22/facebook-get-things-wrong-but-safety-role-seriously>.

130. Algorithmic instruments are very dominant throughout content moderation processes, even when human factors are involved. Such processes might be used to evaluate the enormous number of users' reports of hazardous content, to recognize overlapping reports, and to deliver the reported content to the most appropriate human moderator. Kristina Cooke, *Facebook Developing Artificial Intelligence to Flag Offensive Live Videos*, REUTERS (Dec. 1, 2016, 11:05 AM), <https://www.reuters.com/article/us-facebook-ai-video/facebook-developing-artificial-intelligence-to-flag-offensive-live-videos-idUSKBN13Q52M>.

content such as satire and parody mistakenly identified as hate speech<sup>131</sup> or to the banning of public-domain content, wrongly regarded as IP-protected material.<sup>132</sup> In addition, SNSs' censorship lacks transparency and due process. It is difficult to comprehend the actual scope and nature of this censorship regime because, like surveillance, much of it is concealed and not reported.<sup>133</sup> Even the precise standards and rules that guide SNSs' censorship policy are sometimes opaque and only partly communicated to the users, making it difficult for the latter to properly grasp the actual boundaries of the permitted behavior on these platforms and adapt their actions accordingly.<sup>134</sup> Moreover, some SNSs, like Facebook, do not always allow users to appeal (via the platforms' internal grievances system) when censored.<sup>135</sup>

Furthermore, it is important to keep in mind that censorship in SNSs is not all about removing content or restricting access to it. We should be mindful of the much more elusive and sophisticated algorithm-based means of censoring afforded by down-prioritizing content shown to users in the information stream, search results, or other parts of these sites.<sup>136</sup>

---

131. Troy Wolverton, *AI is Great at Recognizing Nipples, Mark Zuckerberg Says*, BUS. INSIDER (Apr. 25, 2018, 6:47 PM), <http://www.businessinsider.com/ai-can-identify-nipples-but-not-hate-speech-mark-zuckerberg-says-2018-4>; see *infra* Part III.E.2 (discussing the suspension of the German magazine *Titanic*'s Facebook account).

132. Current AI mechanisms might restrict content that does not constitute a breach of copyright or that is safeguarded under the "fair use" clause, thus limiting the capacity and richness of the public domain and hindering the delicate balance between the interests of rightsholders and the public, as reflected in various legal regimes' IP regulation. See Perel & Elkin-Koren, *supra* note 119, at 486-87.

133. *But see Community Standards Enforcement Preliminary Report*, *supra* note 116. In this report, Facebook reveals, for the first time, preliminary data regarding action taken against content it deemed violating some of its community standards. As this report is a step in the right direction, it is only partial and leaves many open questions. First, it does not cover all categories of content which the company takes action against for violating its community standards. Second, the data provided does not distinguish between the different responses triggered by such violations, but jointly includes instances of content removal, placing warning screens, and disabling accounts. Third, the data given is apparently global and is not available on a country or regional level. *Id.*

134. Klonick, *supra* note 117, at 1639. Nevertheless, last year, Facebook published its internal enforcement guidelines. Monika Bickert, *Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process*, FACEBOOK NEWSROOM (Apr. 24, 2018), <https://newsroom.fb.com/news/2018/04/comprehensive-community-standards>.

135. Klonick, *supra* note 117, at 1648; *How to Appeal*, ONLINECENSORSHIP, <https://onlinecensorship.org/resources/how-to-appeal> (last visited Feb. 3, 2019). It should be noted that during 2018, Facebook launched an appeal mechanism for content that it had removed due to "nudity, sexual activity, hate speech, or graphic violence." Bickert, *supra* note 134. The appeals are not available at this stage, in other occasions of content removal, and while providing a human-second review, it seems that they do not enable the users to specify their claims, nor do they establish a commitment to provide the users with reasons for the decision reached by Facebook in the appeal procedure. *Id.*

136. *E.g.*, Mark Zuckerberg, FACEBOOK (Jan. 11, 2018), <https://www.facebook.com/>

## 2. Ex-Ante Censorship

Ex-ante censorship is increasingly deployed throughout SNSs to preemptively control the content that users upload.<sup>137</sup> This type of censoring is mainly performed by algorithmic mechanisms, without the active involvement of human moderators.<sup>138</sup>

In the IP context, for instance, in addition to applying a regular Notice and Takedown procedure, as legally required in the U.S. and other jurisdictions,<sup>139</sup> Facebook is “going beyond notice-and-takedown” as explained by the company.<sup>140</sup> It uses Audible Magic, a third-party technology that enables rightsholders to digitally sample their content and prevent matching videos from being uploaded.<sup>141</sup> SNSs might also use ex-ante censoring against other types of information, such as hate speech, terror-related content,<sup>142</sup> child pornography,<sup>143</sup> and spam.<sup>144</sup> It is likely to become more and more pervasive and encroach on additional areas over time.

In addition to many shared challenges with ex-post censoring, this method of excluding speech prevents content from appearing in advance, thereby reflecting the most severe and dangerous harm to free speech.<sup>145</sup>

Preventing content from being published in advance also denies all social interactions that might have occurred in relation to it and hampers the creation or growth of communities which this kind of content might

---

zuck/posts/10104413015393571 (announcing the content-prioritizing shift embraced by the platform: “I’m changing the goal I give our product teams from focusing on helping you find relevant content to helping you have more meaningful social interactions. . . . The first changes you’ll see will be in News Feed, where you can expect to see more from your friends, family and groups. . . . [Y]ou’ll see less public content like posts from businesses, brands, and media”).

137. See Klonick, *supra* note 117, at 1636. This kind of censorship is carried out “between ‘upload’ and publication.” *Id.*

138. *Id.*

139. The American procedure is enshrined in the Digital Millennium Copyright Act. See 17 U.S.C. § 1201 (2012).

140. *Intellectual Property*, FACEBOOK TRANSPARENCY, [https://transparency.facebook.com/intellectual\\_property](https://transparency.facebook.com/intellectual_property) (last visited Feb. 3, 2019).

141. *Reporting Copyright Infringements: What Tools Does Facebook Provide to Help Me Protect My Intellectual Property in My Videos?*, FACEBOOK, [https://www.facebook.com/help/400287850027717?helpref=faq\\_content%20\(note%3A%20maybe%20not%20requires%3F](https://www.facebook.com/help/400287850027717?helpref=faq_content%20(note%3A%20maybe%20not%20requires%3F) (last visited Feb. 3, 2019).

142. Press Release, Eur. Comm’n, Fighting Terrorism Online: Internet Forum Pushes for Automatic Detection of Terrorist Propaganda (Dec. 6, 2017), [http://europa.eu/rapid/press-release\\_IP-17-5105\\_en.html](http://europa.eu/rapid/press-release_IP-17-5105_en.html). In this press release, Dimitris Avramopoulos stated: “We are not only able to detect it but also remove it quickly, in some cases before it even sees the light of day.” *Id.*; see also *id.* (Monika Bickert’s statement).

143. Klonick, *supra* note 117, at 1636-37.

144. *Id.* at 1637.

145. For a discussion on the concealed nature of ex-ante filtering, see Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2318 (2014).

concern. Ex-ante censorship could also limit public oversight of the removal of the content and the formation and expression of opinions relating to the content's legitimacy. Finally, this type of concealed censorship might be vulnerable to errors and to exploitation attempts by different stakeholders who might have an interest in silencing certain content.<sup>146</sup> This is, by no means, to say that this kind of censoring is not required in certain instances, but to highlight that it should be done with much deliberation.

### 3. Real-Time Censorship

SNSs might also censor in "real-time" by restricting access to live streams or content that document ongoing events. In addition to the virtues it shares with other forms of expression via SNSs, real-time documenting has further important advantages. It is an effective and affordable instrument that empowers users to point out injustice, arbitrariness, or discrimination by officials and other actors, to prove such misconduct in legal proceedings, and to stimulate public debate about the filmed events.<sup>147</sup> In the context of SNSs, documenting real-time events might produce social support, mental or practical help, or other social benefits that might become available to the documenting users in the course of the events.

In general, SNSs enable users to take live videos and see great importance in encouraging this type of communication. Facebook, for instance, even promotes live videos in its News Feed algorithm, thereby increasing the likelihood that this content will be viewed and interacted with.<sup>148</sup> This emphasizes the importance of identifying the challenges that filtering such videos might pose. Facebook declared that it is testing ways to identify and block content during live streaming using AI tools.<sup>149</sup> Until that effort materializes, that platform (and other SNSs) might stop live streaming by other measures that could be harsher for users, for example, by disabling their accounts.<sup>150</sup>

Apart from the desire to silence certain types of speech, SNSs' real-time censoring methods sometimes aim to prevent the unfolding of

---

146. Perel & Elkin-Koren, *supra* note 119, at 489-91.

147. See Klonick, *supra* note 117, at 1600-01.

148. Vibhi Kant & Jie Xu, *Taking into Account Live Video When Ranking Feed*, FACEBOOK NEWSROOM (Mar. 1, 2016), <https://newsroom.fb.com/news/2016/03/news-feed-fyi-taking-into-account-live-video-when-ranking-feed>.

149. Cooke, *supra* note 130. Blocking live streaming has long been used by YouTube. See Geeta Dayal, *The Algorithmic Copyright Cops: Streaming Video's Robotic Overlords*, WIRED (June 9, 2012, 6:00 AM), <https://www.wired.com/2012/09/streaming-videos-robotic-overlords-algorithmic-copyright-cops>.

150. See *infra* notes 152-53 and accompanying text (discussing Korryn Gaines's example).

interpersonal interactions that occur and that might influence live developments.<sup>151</sup> An example is the tragic case of Korryn Gaines. In 2016, the police confronted Gaines, when at home with her five-year-old child. Amid that set off, Gaines posted videos that documented the event in her Facebook and Instagram accounts. The police requested Facebook to deactivate the accounts, aiming, *inter alia*, to stop the comments Gaines was getting from other users which, according to the police, discouraged her from quietly surrendering.<sup>152</sup> Facebook accepted the police's request. After the deactivation of the accounts, and during the confrontation, Gaines was shot and killed. Her child was wounded.<sup>153</sup>

In light of SNSs' effort to promote live streams and the virtues of such videos in empowering users, real-time censoring might continue while challenging human rights and limiting the information and social-communal resources available to them.

#### *D. Practices Relating to the Nexus Between SNSs and the Authorities*

Additional practices of SNSs that influence a wide range of users' human rights stem from the nexus between these platforms and the authorities (national or supranational) or representatives of such authorities. This nexus has several prominent manifestations, including information-sharing processes, restricting access to content, and public-servants'/public-institutions' use of SNSs' accounts. These manifestations sometimes portray a multi-layered and multi-faceted relation between these two sides, characterized by push-back and push-forward dynamics as shall be now discussed.

##### 1. Information-Sharing Processes

An important element of the nexus between SNSs and the authorities lies in information-sharing processes, in which SNSs provide governments with information about their users. Because of the extensive and rich private information SNSs accumulate, they are becoming of increased interest to enforcement authorities.<sup>154</sup> Thus

---

151. Ian Sherr, *How Facebook Censors Your Posts (FAQ)*, CNET (Sept. 9, 2016, 9:58 AM), <https://www.cnet.com/news/how-zuckerberg-facebook-censors-korryn-gaines-philando-castile-dallas-police-your-posts-faq>.

152. Mathew Ingram, *Facebook Shuts Down Live Stream of Shooting at Police Request*, FORTUNE (Aug. 5, 2016), <http://fortune.com/2016/08/05/facebook-video-police>.

153. *Id.* Gaines's family filed a civil suit in the aftermath of the event and was awarded compensation. See Pamela Wood & Alison Knezevich, *Jury Awards More than \$37M to Family of Korryn Gaines in Civil Case Against Baltimore County*, BALT. SUN (Feb. 16, 2018, 4:45 PM), <http://www.baltimoresun.com/news/maryland/baltimore-county/bs-md-co-korryn-gaines-civil-deliberate-20180216-story.html>.

154. Niva Elkin-Koren & Eldar Haber, *Governance by Proxy: Cyber Challenges to Civil*

government's surveillance, "whether targeted or programmatic, for law enforcement or foreign intelligence," now "relies on the cooperation of a small number of technology companies,"<sup>155</sup> among which are SNSs. Alan Rozenshtein terms these companies, which stand between governments and private data, "surveillance intermediaries."<sup>156</sup>

Such information could be shared with governments in keeping with established legal procedures. Yet it may also be supplied as part of the informal and non-transparent linkage between these parties in a somewhat "regulatory twilight zone"<sup>157</sup> that lacks an adequate legal source of authority.<sup>158</sup> This sort of cooperation might be conducted by governments to bypass constitutional barriers they face when searching through users' electronic data.<sup>159</sup> A limited window into the scope and nature of these processes of personal data sharing (and into additional manifestations of the link between SNSs and governments), is provided by Transparency Reports voluntarily published by some SNSs.<sup>160</sup> Facebook's reports, for instance, reflect a general increase in the number of U.S. government requests over the years.<sup>161</sup> In the first half of 2017, 32,716 total requests for users' data were submitted to Facebook by the U.S. government, addressing 52,280 users/accounts.<sup>162</sup> The rate of compliance (provision of some information) by Facebook was eighty-five percent,<sup>163</sup> the highest rate of acceptance reported as to that time.<sup>164</sup> Often, users themselves are not even aware that some of their personal data has been delivered to governments. More than half of the requests Facebook received from U.S. law enforcement bodies during the first

---

*Liberties*, 82 BROOK. L. REV. 105, 125 (2016); see Michael D. Birnhack & Niva Elkin-Koren, *The Invisible Handshake: The Reemergence of the State in the Digital Environment*, 8 VA. J.L. & TECH. 1, 55-56 (2003).

155. Alan Z. Rozenshtein, *Surveillance Intermediaries*, 70 STAN. L. REV. 99, 112 (2018).

156. *Id.* at 105.

157. Elkin-Koren & Haber, *supra* note 154, at 107, 126 ("Put simply, Internet companies hand out information without any legal obligation to do so.").

158. *Id.* at 125-26.

159. *Id.* at 107.

160. See FACEBOOK TRANSPARENCY, *Transparency Report: Government Requests for User Data*, <https://transparency.facebook.com/government-data-requests> (last visited Feb. 3, 2019) (including data concerning Instagram, Messenger, and WhatsApp); TWITTER, *Transparency Report: Information Requests*, <https://transparency.twitter.com/en/information-requests.html> (last visited Feb. 3, 2019).

161. FACEBOOK TRANSPARENCY, *supra* note 160.

162. *Id.*

163. *Id.*

164. *Id.* Over the years documented in the Transparency Reports in that time frame, the rate of Facebook's compliance to U.S. authorities' requests varied and ranged between seventy-nine percent to eighty-five percent. *Id.*

half of 2017 contained a non-disclosure order that prohibited the platform from notifying the relevant users.<sup>165</sup>

However, SNSs' role in overall government surveillance is more complex. Besides allowing and facilitating government surveillance, they might also restrain it, *inter alia*, by using various legal tools (such as "proceduralism" and "litigiousness") and technical techniques (such as end-to-end encryption).<sup>166</sup> This underlines the large discretion that SNSs enjoy in relation to government requests for private information.<sup>167</sup>

Sharing users' private information with governmental authorities, especially in the vague and nontransparent manner that characterizes it, might harm users' privacy and worsen the chilling effect produced by SNSs, thereby suppressing users' liberty to engage with people and information through these platforms.<sup>168</sup> The said cooperation might also lead to interrogation and arrests devoid of legal authorization, thus harming additional human rights.<sup>169</sup> At the same time, as "surveillance intermediaries," SNSs affect governments' ability to tackle various challenges, such as security and safety, and thereby influence other fundamental rights such as the right to life.<sup>170</sup>

## 2. Restricting Access to Content

Another area of cooperation concerns states' and supranational authorities' requests that SNSs restrict access to content.<sup>171</sup> If a particular piece of content is illegal under a certain jurisdiction, but does not

---

165. Chris Sonderby, *Reinforcing Our Commitment to Transparency*, FACEBOOK NEWSROOM (Dec. 18, 2017), <https://newsroom.fb.com/news/2017/12/reinforcing-our-commitment-to-transparency> (explaining that "[f]ifty-seven percent of the data requests we received from law enforcement in the U.S. contained a non-disclosure order that prohibited us from notifying the user, up from 50% in our last report").

166. Rozenshtein, *supra* note 155, at 122-49.

167. *Cooperation or Resistance? The Role of Tech Companies in Government Surveillance*, 131 HARV. L. REV. 1722, 1722-23 (2018).

168. Elkin-Koren & Haber, *supra* note 154, at 125, 133.

169. Daniel J. Solove, *Digital Dossiers and the Dissipation of Fourth Amendment Privacy*, 75 S. CAL. L. REV. 1083, 1085 (2002); Rubinstein, Lee & Schwartz, *supra* note 75, at 263.

170. *Cooperation or Resistance? The Role of Tech Companies in Government Surveillance*, *supra* note 167, at 1723 (SNSs have the power to determine "at least in part, the government's access to information about our personal relationships, professional engagements, travel patterns, financial circumstances, and much more. They also impact the government's ability to prevent terrorist attacks, solve murders, and locate missing children. In short, companies such as Facebook, Google, and Twitter are now responsible for decisions that have major consequences for our privacy, on the one hand, and our safety, on the other.").

171. Like information-sharing processes, this manifestation of linkage to state authorities is partly documented in SNSs' Transparency Reports. See FACEBOOK TRANSPARENCY, *Transparency Report: Content Restrictions Based on Local Law*, <https://transparency.facebook.com/content-restrictions> (last visited Feb. 3, 2019); TWITTER, *Transparency Report: Removal Requests*, <https://transparency.twitter.com/en/removal-requests.html> (last visited Feb. 3, 2019).

violate SNSs' policy, the latter might restrict access to content only in the geographical territory at issue.<sup>172</sup> In some cases, these incidents might indicate government abuse (backed by local court orders or other legal sources) of this cooperation channel, to silence criticism and impede public oversight. One recent example is a case in which Instagram (owned by Facebook) removed content relating to a prominent Russian opposition activist who claimed corruption by a Russian official after being urged to do so by the Russian internet censor.<sup>173</sup> It was reported that the demand was accompanied by an implicit threat to block the platform and prevent its operation in Russia.<sup>174</sup> Another example is the restriction of a video in which the King of Thailand was documented in a German shopping mall wearing minimal clothing (a “crop top”) and accompanied by a young woman.<sup>175</sup>

The said cooperation between SNSs and the authorities, could, however, be carried out as part of a wider censoring partnership—like that established by the European Union (“EU”) and the prominent internet platforms—that aims to encourage prompt removal of hate and terror-related speech. An important manifestation of this partnership is reflected in the Code of Conduct on countering illegal hate speech online (“Code of Conduct”), signed, *inter alia*, by Facebook and Twitter (which are also involved in the EU Internet Forum) in 2016.<sup>176</sup> The Code of

---

172. FACEBOOK TRANSPARENCY, *supra* note 171; *see also Thailand Backs Down on Threat to Ban Facebook*, TECHCRUNCH (May 16, 2017), <https://techcrunch.com/2017/05/16/thailand-backs-down-on-threat-to-ban-facebook/> (“When governments believe that something on the Internet violates their laws, they may contact companies like Facebook and ask us to restrict access to that content. When we receive such a request, it is scrutinized to determine if the specified content does indeed violate local laws. If we determine that it does, then we make it unavailable in the relevant country or territory and notify people who try to access it why it is restricted.”).

173. *Instagram Submits to Russia Censor’s Demands*, BBC NEWS (Feb. 15, 2018), <http://www.bbc.com/news/technology-43070555>.

174. Rhett Jones, *Instagram Bows to Russian Censor’s Pressure, Will YouTube Be Next?*, GIZMODO (Feb. 18, 2016, 11:35 AM), <https://gizmodo.com/instagram-bows-to-russian-censors-pressure-will-youtub-1823066754>. This demand, it was reported, was also represented to YouTube, but the latter refused to comply. *Id.*

175. David Gilbert, *Censored by Facebook*, VICE NEWS (May 8, 2017), [https://news.vice.com/en\\_us/article/434m4m/facebook-is-blocking-thai-users-from-seeing-video-of-their-king-in-a-crop-top](https://news.vice.com/en_us/article/434m4m/facebook-is-blocking-thai-users-from-seeing-video-of-their-king-in-a-crop-top).

176. EUROPEAN COMMISSION, JUSTICE AND CONSUMERS, CODE OF CONDUCT ON COUNTERING ILLEGAL HATE SPEECH ONLINE, [https://ec.europa.eu/newsroom/just/document.cfm?doc\\_id=42985](https://ec.europa.eu/newsroom/just/document.cfm?doc_id=42985) (last visited Feb. 3, 2019) [hereinafter CODE OF CONDUCT]. It is also signed by Microsoft and YouTube. *Id.*

The EU Internet Forum, established in 2015, plays an important role in promoting such partnership. Some of its members are EU Interior Ministers, the Europol, and prominent internet platforms. Its goal is “to reach a joint, voluntary approach based on a public-private partnership to detect and address harmful material online.” Press Release, Eur. Comm’n, EU Internet Forum: Bringing Together Governments, Europol and Technology Companies to Counter Terrorist Content and Hate Speech Online (Dec. 3, 2015), [http://europa.eu/rapid/press-release\\_IP-15-6243\\_en.htm](http://europa.eu/rapid/press-release_IP-15-6243_en.htm).

Conduct concludes that when these platforms receive a report of allegedly illegal content, they will assess its legality “against their rules and community guidelines”<sup>177</sup> which imparts a somewhat flexible meaning to the concept of legality. The Code of Conduct further enshrines additional mechanisms to foster the cooperation between the platforms and the European state authorities.<sup>178</sup> It also encourages internet platforms to cooperate with respect to hate speech removal.<sup>179</sup> Such cooperation is achieved, inter alia, by the Global Internet Forum to Counter Terrorism established in 2017.<sup>180</sup> Through their collaboration, the platforms cultivate a Shared Industry Hash Database of what they perceive as illegal content, which facilitates preemptive censoring of such content across these various platforms.<sup>181</sup> The internet platforms have asserted that illegal content is removed independently according to each platform’s policy, but it seems more likely that in cases in which censorship was carried out by one platform (even if unjustifiably), the others might be encouraged to conceive the given content as illegal and follow suit.<sup>182</sup> In these cases, cooperation among these prominent internet platforms might lead to a vast denial of legitimate content.

However, whereas the EU demonstrates tolerance regarding how content’s lawfulness is determined, it is more rigid about the pace and volume of the content removed.<sup>183</sup> A proper understanding of the cooperation reflected in the Code of Conduct also requires taking into

---

177. CODE OF CONDUCT, *supra* note 176, at 2. The platforms are also required to take into account certain national laws “where necessary.” *Id.*

178. *Id.* (“The IT companies to provide information on the procedures for submitting notices, with a view to improving the speed and effectiveness of communication between the Member State authorities and the IT Companies, in particular on notifications and on disabling access to or removal of illegal hate speech online. The information is to be channeled through the national contact points designated by the IT companies and the Member States respectively. This would also enable Member States, and in particular their law enforcement agencies, to further familiarize themselves with the methods to recognise and notify the companies of illegal hate speech online.”).

179. *See, e.g., id.* at 3 (“The IT Companies to intensify cooperation between themselves and other platforms and social media companies to enhance best practice sharing.”).

180. *Facebook, Microsoft, Twitter and YouTube Announce Formation of the Global Internet Forum to Counter Terrorism*, FACEBOOK NEWSROOM (June 26, 2017), <https://newsroom.fb.com/news/2017/06/global-internet-forum-to-counter-terrorism>; *Partnering to Help Curb Spread of Online Terrorist Content*, FACEBOOK NEWSROOM (Dec. 5, 2016), <https://newsroom.fb.com/news/2016/12/partnering-to-help-curb-spread-of-online-terrorist-content>.

181. *Facebook, Microsoft, Twitter and YouTube Announce Formation of the Global Internet Forum to Counter Terrorism*, *supra* note 180; *Partnering to Help Curb the Spread of Terrorist Content Online*, GOOGLE (Dec 5, 2016), <https://blog.google/topics/google-europe/partnering-help-curb-spread-terrorist-content-online>.

182. *See Partnering to Help Curb the Spread of Terrorist Content Online*, *supra* note 181.

183. For the importance ascribed to these factors see *Code of Conduct on Countering Online Hate Speech – Results of Evaluation Show Important Progress*, EUR. COMMISSION: JUST. & CONSUMERS (June 1, 2017), [http://ec.europa.eu/newsroom/just/item-detail.cfm?item\\_id=71674](http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=71674).

account that this agreement was “voluntarily” embraced against the backdrop of explicit threats made by the European Commission to otherwise regulate these areas and to impose fines when necessary.<sup>184</sup>

Nevertheless, SNSs do not always conduct censorship on behalf of governments from a weakened or coerced position and may even initiate it. An important example is the reported efforts invested by Facebook to develop a censorship mechanism that might pave the platform’s way to resume operations in China.<sup>185</sup> The link between SNSs and governments in censorship is thus not monolithic, but rather a multi-layered and multi-faceted one. It consists of a variety of interests and motivations involved in the formulation of information and social-communal resources made available to users online and that influence many of their human rights.

### 3. Public Use of SNSs’ Accounts

Another important affinity between SNSs and governments lies in the use made by public servants and public institutions of SNSs. Over the years, such use has constantly risen, and the mutual dependence of SNSs and public actors on each other has accordingly increased. SNSs are now ubiquitously used by heads of states, ministers, state governors, members of Congress, heads of municipalities and cities, and many other public servants and state officials. They are also embraced by various state and public institutions such as governments, courts, ministerial and governmental offices, public-services institutions, etc.<sup>186</sup> Many of these actors might use SNSs to communicate with the public, receive feedback and inform the public with various updates. SNSs might even be used as “a tool of governance,” as recognized by the District Court for the Eastern District of Virginia.<sup>187</sup> The most prominent example is probably the frequent use of Twitter by the sitting U.S. President, Donald Trump.<sup>188</sup> This, of course, is of priceless value to Twitter. Among other

---

184. Jacob Kastrenakes, *EU Says It’ll Pass Online Hate Speech Laws if Facebook, Google, and Others Don’t Crack Down*, VERGE (Sept. 28, 2017, 3:27 PM), <https://www.theverge.com/2017/9/28/16380526/eu-hate-speech-laws-google-facebook-twitter>.

185. Samuel Gibbs, *Facebook Developed Secret Software to Censor User Posts in China, Report Says*, GUARDIAN (Nov. 23, 2016, 10:08 AM), <https://www.theguardian.com/technology/2016/nov/23/facebook-secret-software-censor-user-posts-china>.

186. See *Packingham v. North Carolina*, 137 S. Ct. 1730, 1735 (2017) (“[O]n Twitter, users can petition their elected representatives and otherwise engage with them in a direct manner. Indeed, Governors in all 50 States and almost every Member of Congress have set up accounts for this purpose.”).

187. *Davison v. Loudoun Cty. Bd. of Supervisors*, 267 F.3d 702, 713 (E.D. Va. 2017). With respect to such blockage, see also *infra* notes 191-95 and accompanying text.

188. See Donald Trump (@realDonaldTrump), TWITTER, <https://twitter.com/realDonaldTrump> (last visited Feb. 3, 2019).

benefits, it has increased the number of consumers of news via this platform.<sup>189</sup> Peter Greenberger, Twitter's global news content partnerships director, said in this regard:

Twitter has never been more influential or more relevant. Every single day, the story of the day, that leads on the newspaper, that leads online, that leads on television, is very often based on what the President tweeted or how people responded to that tweet . . . and so whatever happens, whether you support the administration or whether you disagree with the administration, Twitter is the place to go to find out what's coming next, and how the opposition is responding to what the administration, for instance, is tweeting about.<sup>190</sup>

Besides the unprecedented opportunity for direct communication with once much less accessible servants and institutions, this development might impede users' human rights in several ways, of which I now discuss two.

First, the communication between public servants and the public is mediated by SNSs, which exert significant influence on its nature and features. This communication depends on the architecture of SNSs, their policies, and algorithmic mechanisms that determine the manner and scope in which their pages will be displayed (or suggested) to users. One example comes from recent suits brought to U.S. courts about public servants' accounts being blocked to certain users. A prominent case was a lawsuit filed by the Knight First Amendment Institute at Columbia University and additional plaintiffs against U.S. President Trump (and others).<sup>191</sup> The plaintiffs asserted that because of their criticism of the President, they were blocked from "viewing the President's tweets, from replying to the tweets, from viewing the discussions associated with the tweets, and from participating in those discussions."<sup>192</sup> Such blockage, they argued, denied their taking part in a "public forum" and violated the First Amendment.<sup>193</sup> In a ruling rendered in May 2018, the court accepted these arguments.<sup>194</sup> The plaintiff did not address SNSs' part in

---

189. Elisa Shearer & Jeffrey Gottfried, *News Use Across Social Media Platforms 2017*, PEW RES. CTR. 4 (2017) ("Since 2013, at least half of Twitter users have reported getting news on the site, but in 2017, with a president who frequently makes announcements on the platform, that share has increased to about three-quarters (74%), up 15 percentage points from last year.").

190. Robert Andrews, *Twitter's Greenberger on News Platform for Advertisers*, BEET.TV at 2:33–2:50; 3:03–3:20 (Apr. 16, 2017), <https://www.beet.tv/2017/04/174atwittergreenberger.html>.

191. Complaint for Declaratory and Injunctive Relief, *supra* note 85.

192. *Id.* at 2-3.

193. *Id.* at 3.

194. Knight First Amendment Inst. v. Trump, 302 F. Supp. 3d 541, 580 (S.D.N.Y. 2018); *see also* Davison v. Loudoun Cty. Bd. of Supervisors, 267 F. Supp. 3d 702, 718 (E.D. Va. 2017) (holding that the Chair of the Loudoun County, Virginia, Board of Supervisors violated the First

the subject matter, but it is important to keep in mind that the entire constitutional debate at issue stems from a technical affordance of SNSs that allows public servants (like any other users) to block content in a split second. However, a “who’s the boss” recall was given by the assigned judge shortly before delivering the court’s final decision, suggesting that if Trump uses the “muting” option provided by Twitter, rather than “blocking” the plaintiffs, it might settle the dispute.<sup>195</sup>

Second, the influence operating between public servants’ extensive use of SNSs and SNSs’ standards, architecture, and performance also works in the other direction, while determining the characteristics and nature of these platforms. SNSs might, for instance, accommodate their policies to enjoy the commercial benefits that public servants and public organizations afford them as their users. An illustration of this is Twitter’s recent update of its policy after being criticized for not removing President Trump’s tweets concerning North Korea, seemingly in violation of the platform’s policy.<sup>196</sup> Its updated policy states that when determining whether a tweet violates the company’s rules, it may also take into account its “newsworthiness” and whether it is of public interest.<sup>197</sup> While such revision of Twitter’s rules might be positive, it underscores the enormous effect that public representatives have in designing the settings in which the other users operate. In certain instances, this power might be embodied in informal editions of SNSs policies. It was reported, for example, that Twitter removed Trump-related results from certain adverse-oriented search results.<sup>198</sup> This might indicate SNSs’ willingness to adapt their rules to accommodate their interests, in order to cultivate and secure public-actors use of their

---

Amendment when she blocked a user from her Facebook page). *But cf.* Davison v. Rose, No. 1:16cv0540, slip op. at 19 (E.D. Va. July 28, 2017) (providing that “[i]t is not clear as a legal matter whether the Facebook pages at issue in this litigation can be said to constitute either type of public forum. In any event, it cannot be said that such a First Amendment right was a ‘clearly established’ right, ‘of which a reasonable person would have [sic] known.’”).

195. *Trump Told Mute Twitter Critics, Not Block Them, by New York Judge*, BBC NEWS (Mar. 9, 2018), <http://www.bbc.com/news/world-us-canada-43344837>.

196. Bill Chappell, *‘Declaration Of War’ Means North Korea Can Shoot Down U.S. Bombers, Minister Says*, NPR (Sept. 25, 2017, 12:53 PM), <https://www.npr.org/sections/thetwo-way/2017/09/25/553475174/declaration-of-war-means-north-korea-can-shoot-down-u-s-bombers-minister-says>.

197. Chloe Watson, *Twitter Says Trump’s Threat to North Korea was ‘Newsworthy’ and Will Not be Taken Down*, GUARDIAN (Sept. 26, 2017, 6:05 AM), <https://www.theguardian.com/technology/2017/sep/26/twitter-trump-threat-north-korea-newsworthy>. For the revised rules, see *Our Approach to Policy Development and Enforcement Philosophy*, TWITTER: HELP CTR., <https://help.twitter.com/en/rules-and-policies/enforcement-philosophy> (last visited Feb. 3, 2019).

198. Jeff Parsons, *Twitter Removes Donald Trump’s Account From Search Results for ‘Asshole’, ‘Bigot’ and ‘Racist’*, MIRROR (Jan. 26, 2017, 1:57 PM), <https://www.mirror.co.uk/tech/twitter-removes-donald-trumps-account-9696657>.

platforms, even when such adaptation does not fall in line with their other users' rights.

### *E. Expulsion of Individuals and Communities*

The fifth and last practice of control exerted by SNSs relates to their expulsion of users, whether individuals or communities, achieved by disabling accounts, groups, and pages.

#### 1. Expulsion of Individuals

The disabling of accounts might vary in duration and scope. It might be permanent or temporary. It could be partial (allowing users to consume content but not to produce it) or absolute (preventing users from logging in to their accounts).<sup>199</sup>

One example of a temporary and partial blockage of an account is reflected in the disabling of Rose McGowan's account, discussed above.<sup>200</sup> Another example is the 2017 suspension of Guo Wengui's Facebook account after he pointed out corruption by officials in China last year. Facebook later claimed that the deactivation was caused by an error and restored the account.<sup>201</sup> A few months later though, Facebook blocked Wengui's account once again. Facebook admitted receiving a complaint from a representative of the Chinese government, but insisted that the blockage stemmed from Wengui's violation of its rules which prohibited the publishing of private information concerning other

---

199. *E.g.*, *Disabled Accounts*, FACEBOOK: HELP CTR., <https://www.facebook.com/help/185747581553788> (last visited Feb. 3, 2019); *Our Range of Enforcement Options*, TWITTER: HELP CTR., <https://help.twitter.com/en/rules-and-policies/twitter-rules> (last visited Feb. 3, 2019).

200. *See supra* Part I.

201. Paul Mozur, *Facebook Briefly Suspends Account of Outspoken Chinese Billionaire*, N.Y. TIMES, Apr. 22, 2017, at B2, <https://www.nytimes.com/2017/04/21/technology/guo-wengui-china-facebook-account-suspended.html>. Facebook further argued that the reason for the suspension is hard to determine and that publishing it may allow the manipulation of the system. *Id.* For an example of a temporary suspension of a Pulitzer-winning reporter's Facebook account, reportedly due to a post in which he argued for corruption in the government of Malta, see Julie Carrie Wong, *Facebook Blocks Pulitzer-Winning Reporter Over Malta Government Exposé*, GUARDIAN (May 19, 2017, 4:00 PM), <https://www.theguardian.com/world/2017/may/19/facebook-blocks-malta-journalist-joseph-muscat-panama-papers>.

individuals.<sup>202</sup> SNSs' users have been expelled in various other cases involving nudity,<sup>203</sup> hate speech,<sup>204</sup> and the use of fake names.<sup>205</sup>

Expulsion of users from SNSs, implemented by the disabling of their accounts, abruptly denies them the entire inventory of informational and social-communal resources afforded by these platforms and exerts significant adverse effects on their various human rights. It deprives them of consuming, sharing, and distributing information or interacting with other users, groups, or pages. This practice (at least in its absolute form) therefore imposes the most extreme and severe limitations on users' wide range of civil-political and economic-social-cultural rights. It might prevent users from accessing public servants' updates, chatting with a public health institution, advancing professionally, benefiting from discussions in various education-related groups, etc. Expulsion also raises specific questions about users' ownership of assets aggregated in their accounts, such as "likes," friends, posts, and photos.

Despite the significant danger to human rights involved in the suspension of accounts, the criteria guiding this practice suffers from an extreme lack of transparency and predictability. Users cannot know what precise behaviors will lead to the disabling of their accounts<sup>206</sup> and

---

202. Heather Timmons, *Facebook Blocked Dissident Guo Wengui After the Chinese Government Complained*, QUARTZ (Nov. 1, 2017), <https://qz.com/1117822/facebook-blocked-dissident-guo-wengui-after-the-chinese-government-complained>. The reliability of Wengui reports is unclear, however. See John Ruwitch, *Facebook Pulls Page, Limits Posting for Exiled Chinese Tycoon Guo*, REUTERS (Oct. 2, 2017, 4:21 AM), <https://www.reuters.com/article/us-china-facebook-tycoon/facebook-pulls-page-limits-posting-for-exiled-chinese-tycoon-guo-idUSKCN1C70R0>.

203. In an occasion that had reached legal proceedings, Facebook deactivated a French educator's accounts, allegedly for posting a photo of the iconic painting "L'Origine du monde" ("The Origin of the World"), painted in 1866 by Gustave Courbet. See Naomi Rea, *A French Court Rules That Facebook Was Wrong to Censor Gustave Courbet's Provocative 'Origin of the World'*, ARTNET NEWS (Mar. 5, 2018), <https://news.artnet.com/art-world/facebook-censorship-gustave-courbet-1245458>.

204. In Israel, Facebook accounts of many journalists and media-figures have been disabled during 2016, apparently for using words that the platform deemed offensive. In a letter sent to Facebook Israel CEO, they explained that in certain cases, the disabling of their accounts was attached to posts they had written years before. They further stated that the platform has grown to be a significant part of their lives and work and that they wished to proceed using it without the constant fear of unexplained restrictions. See Anat Bein-Lobovitch, *Content Creators: Facebook Silences our Voice and Tramples Freedom of Expression*, GLOBES (Nov. 13, 2016, 2:29 PM), <http://www.globes.co.il/news/article.aspx?did=1001160720>. For a different example, see Alex Hern, *Twitter Suspends American Far-Right Activists' Accounts*, GUARDIAN (Nov. 16, 2016), <https://www.theguardian.com/technology/2016/nov/16/twitter-suspends-american-far-right-activists-accounts>.

205. Lil Miss Hot Mess, *supra* note 60.

206. For instance, Facebook's vague policy, according to which disabling accounts might take place, inter alia, when "[p]osting content that doesn't follow the Facebook Terms" or "[c]ontinuing

sometimes are not provided with an explanation in tandem with such disabling.<sup>207</sup> It might also be difficult for users to evaluate the instances where they will be given a warning and a tentative partial suspension of their accounts and to predict the nature and duration of these partial “timeouts.”<sup>208</sup> These challenges are further stressed in light of users’ difficulties contacting SNSs representatives to convey their grievances, as well as the poor automated appeal mechanisms offered by these platforms.<sup>209</sup>

## 2. Expulsion of Communities

Users’ expulsion might also apply to groups and communities and be realized by shutting down groups and pages that connect users with shared interests or other commonalities. The communities and groups that operate throughout SNSs may form important settings for the exchange of information and participation in various areas, such as politics, health, parenthood, education, hobbies, entertainment, residence in a certain city or country, profession, etc.

Like the expulsion of individuals, this kind of expulsion inhibits both the informational and the social-communal resources afforded by SNSs. However, often these resources address specific or confined topics (and people). Certain communities may represent unpopular, controversial, and even rejected viewpoints or gather minority individuals and provide them with a safe sphere to consume, produce, and deliberate on their beliefs, concerns, and thoughts.<sup>210</sup> Shutting groups down puts an end to these important benefits. For example, Facebook shut down the account of the satirical German magazine *Titanic* which was followed by 100,000 users, apparently for parodying racist tweets written by a politician from the anti-immigrant Alternative for Germany (“AfD”) Party.<sup>211</sup> A few months later, Facebook shut down a group of University of Chicago students, in which members had been able to anonymously consult with each other or raise issues relating to

---

behavior that’s not allowed on Facebook by violating our Community Standards.” *Disabled Accounts*, *supra* note 199.

207. See *supra* Part I (discussing Rose McGowen’s case).

208. Facebook, for example, does not commit to notify users before disabling their accounts, nor does it explain the different disabling possibilities. *Disabled Accounts*, *supra* note 199.

209. For users’ difficulties to reach Facebook after deactivation of their accounts, see *My Account is Suspended, (Name Not Authentic) How Can I Contact Facebook?*, FACEBOOK: HELP COMMUNITY, <https://www.facebook.com/help/community/question/?id=10205954842851666> (last visited Feb. 3, 2019).

210. Lil Miss Hot Mess, *supra* note 60.

211. Ivana Kottasová, *Is Germany’s New Hate Speech Law Killing Press Freedom?*, CNN MEDIA (Jan. 4, 2018, 12:53 PM), <https://money.cnn.com/2018/01/04/media/twitter-satire-free-speech-germany/index.html>.

their studies and campus life. The stoppage was evoked by one student's post that mocked French people.<sup>212</sup> Controversial and extremist groups can also be subject to community expulsion. Following the violent Charlottesville events of "Unite the Right" rally, both Instagram and Facebook shut down several right-wing extremist groups.<sup>213</sup> Closing down communities could also be effectuated by the blockage of accounts of individuals who attract or lead certain groups. An example is the temporary suspension of Black Lives Matter activist Shaun King, who had about 800,000 followers<sup>214</sup> after he had posted a screenshot of a racist email he received.<sup>215</sup> Likewise, suspension of the account of Rose McGowan, a leader of the resistance to sexual misconduct in Hollywood followed by hundreds of thousands of people, may also be seen as a community expulsion.<sup>216</sup>

This Part explored five different SNSs practices, that influence their users' various human rights, mainly through determining the availability of informational and social-communal resources. This, of course, is not a definitive list. Additional practices might relate, for instance, to manipulation of users' choices and cooperation with other internet platforms. Each of the five explored practices influences a wide range of users' human rights, and together, they generate a highly effective control regime exerted over these diverse rights. The outcome is digital, sophisticated arenas in which human rights are crafted, designed, and allocated. This echoes something familiar: management of rights, as conducted by states. This is the theme of the next Part.

#### IV. THE STATE AND THE TASK OF MANAGING HUMAN RIGHTS

The task of managing human rights is traditionally associated with and tackled by the state. The state determines a great deal of our human rights. It decides when and how we may express our thoughts and are exposed to other peoples' thoughts, protest and associate with others,

---

212. Lindsay McKenzie, *Facebook Says Au Revoir to Student Group*, INSIDE HIGHER ED (Apr. 24, 2018), <https://www.insidehighered.com/news/2018/04/24facebook-bids-adieu-anonymous-student-group-over-anti-french-post>.

213. Keith Collins, *A Running List of Websites and Apps That Have Banned, Blocked, Deleted, and Otherwise Dropped White Supremacists*, QUARTZ (Aug. 16, 2017), <https://qz.com/1055141/what-websites-and-apps-have-banned-neo-nazis-and-white-supremacists>.

214. Presently, this account has over 1.9 million followers. Shaun King, FACEBOOK, <https://www.facebook.com/shaunking> (last visited Feb. 3, 2019).

215. Sam Levin, *Facebook Temporarily Blocks Black Lives Matter Activist After He Posts Racist Email*, GUARDIAN (Sept. 12, 2016, 5:26 PM), <https://www.theguardian.com/technology/2016/sep/12/facebook-blocks-shaun-king-black-lives-matter>. Facebook explained that the account's suspension was a mistake. *Id.*

216. See *supra* Parts I, III.E.1.

choose our occupation and do business, walk about freely, follow our cultural and religious customs, acquire education, and receive health services.<sup>217</sup>

Such understanding of states' influence over human rights (and the obligation it entails) has ancient roots, well situated in the scholarship of prominent Western-political thinkers, and encapsulated in the notion of the social covenant and in the approach that considers the state as a fiduciary of the public.<sup>218</sup>

John Locke, for instance, conducted a profound examination of the fiduciary role of the state in his *Second Treatise of Government*,<sup>219</sup> where he located the objective of the state in the value of human rights.<sup>220</sup> He asserted that the social order established by people aims to promote their lives, liberties, and estate, which together he named "property."<sup>221</sup> Locke explained that in the "state of nature," the people were indeed their own masters, but their rights were under a constant threat of invasion by others. The state therefore was established to better safeguard these rights.<sup>222</sup> Later on, Jean-Jacques Rousseau wrote the famous *The Social Contract*,<sup>223</sup> in which he set forth his perception of the desirable political community. He weighed up the benefits and the losses resulting from the establishment of the state and enthusiastically stressed how advantageous this exchange was for individuals. People lost their natural liberty and their unlimited right to do as they pleased, he explained, but they won civil liberties and ownership of their possessions.<sup>224</sup> This self-choice to subject themselves to the state afforded them their current liberties.<sup>225</sup> Later scholars, such as Thomas Paine, likewise pointed to a theoretical tradeoff as justification for the creation of political societies.<sup>226</sup> According to Paine, people established

217. This is not to underestimate limitations introduced to states' sovereignty by international institutions and other factors. *See infra* Part VI.A.

218. This approach can be tracked back to Plato, Aristotle, and Cicero. *See* CRIDDLE & FOX-DECENT, *supra* note 11, at 13. For a fascinating description of the unfoldment of the state's fiduciary concept (coined "the Public Trust Doctrine"), see Robert G. Natelson, *The Constitution and the Public Trust*, 52 BUFF. L. REV. 1077 *passim* (2004).

219. JOHN LOCKE, *SECOND TREATISE OF GOVERNMENT* (Thomas P. Peardon ed., 1979) (1690).

220. DIETER GRIMM, *SOVEREIGNTY: THE ORIGIN AND FUTURE OF A POLITICAL AND LEGAL CONCEPT* 30 (Belinda Cooper trans., 2015); *see also* Natelson, *supra* note 218, at 1115-18 (addressing Locke's approach).

221. LOCKE, *supra* note 219, at 70-71.

222. *Id.*

223. JEAN-JACQUES ROUSSEAU, *THE SOCIAL CONTRACT: OR PRINCIPLES OF POLITICAL RIGHT* (H.J. Tozer trans., 1998) (1762).

224. *Id.* at 19.

225. *Id.*

226. Paul F. Boller, *Thomas Paine and Natural Rights: A Reconsideration*, 52 SOC. SCI. 67, 68 (1977).

governments and entrusted them with the power to protect their rights. They exchanged certain natural rights for civil rights that protected their security and property while retaining their natural rights to freedom of thought, conscience, and speech.<sup>227</sup> This social covenant and the theoretical tradeoff it entails, thus constitute the perception of the state as a fiduciary of the public. It explains the origin of the immense power of the state to manage and allocate human rights and highlights states' obligation to promote and protect such rights.<sup>228</sup>

The philosophical perception of the state as a fiduciary was later internalized and embraced by various jurisdictions and founding documents, such as the American Declaration of Independence and enjoyed "real-world legal implications."<sup>229</sup> Its importance is well reflected in the words of Evan Fox-Decent and Evan J. Criddle:

[H]uman rights emanate from a fiduciary relationship between the state and persons subject to its powers. Fiduciary relations denote the subjection of one party to the ongoing administrative power of another . . . . The fiduciary principle authorizes states to exercise power on behalf of their people, but subject to strict limitations . . . .<sup>230</sup>

This fiduciary role, therefore, authorizes states to establish legal order and govern people, "to define their several rights, and redress their several wrongs,"<sup>231</sup> but only as long as this falls in line with their obligation to promote human rights. Some scholars even linked this obligation of the state to its sovereignty. Criddle and Fox-Decent argued that "because this authorization is constrained and constituted by a duty to respect, protect, and implement human rights, state sovereignty is likewise constrained and constituted by human rights,"<sup>232</sup> and when a state fails to uphold its obligation to respect human rights, it weakens the sources of its sovereignty.<sup>233</sup> Ann Peters explained that state sovereignty

---

227. *Id.*

228. CRIDDLE & FOX-DECENT, *supra* note 11, at 14; *see also* Natelson, *supra* note 218, at 1137-38. The notion of the state as a fiduciary, counter to fiduciaries' duties in private law (that apply to professionals as lawyers and doctors), is not a legal requirement, but rather a theoretical concept. Jack Balkin suggested applying duties of such private-sector fiduciaries to online service providers with relation to private information they obtain. Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C. DAVIS L. REV. 1183, 1205 (2016).

229. Natelson, *supra* note 218, at 1136-37.

230. Fox-Decent & Criddle, *supra* note 11, at 301-02.

231. WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND § 48 (S. Sweet, Chancery Lane, 19th ed. 1836).

232. Fox-Decent & Criddle, *supra* note 11, at 315.

233. CRIDDLE & FOX-DECENT, *supra* note 11, at 95. *But cf.*, GRIMM, *supra* note 220, at 125-26.

is not only limited by human rights, but also that its qualifications lay in its humanity and in respecting “human rights, interests, and needs.”<sup>234</sup>

The vulnerability of human rights and public interests<sup>235</sup> to the powerful state lies in the bedrock of public law and is the justification for its emergence as a separate branch of law.<sup>236</sup> Contrary to private law, “which regulated the affairs of subjects as between themselves,” public law is called upon to regulate “the affairs of subjects vis-à-vis public authorities.”<sup>237</sup> The notion of the influence that the state has on human rights and human needs became incorporated into fundamental public law themes<sup>238</sup> and foremost, the Rule of Law. This principle, which emerged as a result of the establishment of the modern state,<sup>239</sup> can bear different meanings and interpretations.<sup>240</sup> At its core, it requires that “every act which affects the legal rights, duties or liberties of person, must be shown to have a legal strictly pedigree.”<sup>241</sup> This authorizing legal basis should constrain the discretionary power of the state and prevent it from being abused by governments.<sup>242</sup> Some ascribe to the Rule of Law a richer interpretation that consists of additional requirements, such as fairness.<sup>243</sup> This tension between the state’s power and human rights also propelled modern constitutional development. Given the wide powers that governments enjoy in the modern era, liberty could only be ensured by confining these powers. Accordingly, the principal interest of modern constitutions is “to maintain a balance

234. Anne Peters, *Humanity as the A and Ω of Sovereignty*, 20 EUR. J. INT’L L. 513, 514 (2009) (explaining that “[s]tate sovereignty is not only—as in the meanwhile canonical view—limited by human rights, but is from the outset determined and qualified by humanity, and has a legal value only to the extent that it respects human rights, interests, and needs. It has thus been humanized.”). *But see* Emily Kidd White, *Humanity as the A and Ω of Sovereignty: Four Replies to Anne Peters*, 20 EUR. J. INT’L L. 545, 546-47 (2009).

235. It could be justifiably argued that even public interests aim, eventually, to safeguard human rights. As put by William Blackstone: “[T]he public good is in nothing more essentially interested, than in the protection of every individual’s private rights.” BLACKSTONE, *supra* note 231, at § 139. For the complex and changing relationship between individuals’ rights and public interests, see Morton J. Horwitz, *The History of the Public/Private Distinction*, 130 U. PA. L. REV. 1423, 1428 (1982); see also Elisabeth Zoller, *Public Law as the Law of the Res Publica*, 32 SUFFOLK TRANSNAT’L L. REV. 93, 95-97 (2008).

236. See H.W.R. WADE & C. F. FORSYTH, *ADMINISTRATIVE LAW* 295-96 (11th ed., 2014).

237. *O’Reilly v. Mackman* [1983] (C.A.) 2 A.C. 237, 255 (Lord Denning, J.).

238. WADE & FORSYTH, *supra* note 236, at 240 (“Statutory power conferred for public purposes is conferred as it was upon trust, not absolutely—that is to say, it can validly be used only in the right and proper way . . .”).

239. MARTIN LOUGHLIN, *FOUNDATIONS OF PUBLIC LAW* 314 (2010).

240. WADE & FORSYTH, *supra* note 236, at 15. It may also vary depending on the jurisdiction at issue. See LOUGHLIN, *supra* note 239, at 314.

241. WADE & FORSYTH, *supra* note 236, at 15.

242. *Id.* at 15-16.

243. *Id.* at 17-18.

between the conferral of powers on government and the preservation of the liberties of the individual.”<sup>244</sup> The principle of “Separation of Powers” (the “trias politica”) also stems from the same aspiration to restrict the power of the state and to promote liberties.<sup>245</sup>

The task of designing and managing rights is, as shown above, tightly and inherently attached to the state and constitutes one of its ancient and principal roles. Thus, the emergence of SNSs as new arenas for crafting and allocating human rights—where a wide range of civil-political and economic-social-cultural rights are subject to a sophisticated and effective control regime—echoes this very role of the state and reflects a salient public attribute of these platforms. This public character, I suggest, justifies application of public law norms to SNSs, thereby urging them to take a more user-oriented approach.

Before exploring the benefits of applying public law norms to SNSs, I examine some of these norms as they occur in various jurisdictions.

## V. PUBLIC LAW NORMS

Over time, guiding norms and principles have developed to control the extortionate power of the state, to confine it and guide it so as to better safeguard human rights.<sup>246</sup> These principles may be often associated with fairness and natural law,<sup>247</sup> and include, among other standards, reasonableness, proportionality, good-faith, reason-giving, and due process.<sup>248</sup>

These Public Law norms play a fundamental role in various jurisdictions. However, they may differ, in scope and application, depending on the jurisdiction at issue, and the historical, legal, and social heritage thereof. In general, while every jurisdiction has “public law,” namely its own set of instruments and disciplines that sustains the management of the “res publica,” states might differ in the “density and the thickness” of such disciplines.<sup>249</sup> However, as will be demonstrated

---

244. LOUGHLIN, *supra* note 239, at 312; *see also* WADE & FORSYTH, *supra* note 236, at 17-18.

245. Ana J. Samuel, *The Design of Montesquieu's The Spirit of the Laws: The Triumph of Freedom over Determinism*, 103 AM. POL. SCI. REV. 305, 309-10 (2009); *see also* Denis Baranger, *The Apparition of Sovereignty*, in SOVEREIGNTY IN FRAGMENTS: THE PAST, PRESENT AND FUTURE OF A CONTESTED CONCEPT 47, 61-62 (Hent Kalmo & Quentin Skinner eds., 2010).

246. WADE & FORSYTH, *supra* note 236, at 295-96.

247. Robert Thomas, *Reason-Giving in English and European Community Administrative Law*, 3 EUR. PUB. L. 213, 214 (1997).

248. *See infra* Part V.A-B.

249. ÉLISABETH ZOLLER, INTRODUCTION TO PUBLIC LAW: A COMPARATIVE STUDY 17 (2008). For instance, in some countries (such as France) there are separate courts to adjudicate public law cases, whereas in other places (like the U.S. and England), “public law litigation” is conducted

below, there are also many commonalities between different jurisprudences regarding the way they perceive and apply these public law norms. This commonality is related to a vibrant and important discussion on the globalization of constitutional law, which is “becoming a shared enterprise that transcends the borders of the nation-state” and relies on shared templates.<sup>250</sup> I shall now dedicate a thumbnail discussion to some prominent public law principles and their implications in certain jurisdictions. I start by discussing proportionality and reasonableness as representatives of substantive fairness and then address reason-giving, which reflects themes of procedural fairness. This brief sketch may highlight how fundamental these principles are in various legal regimes and clarify the aims they strive to accomplish.

### A. Proportionality and Reasonableness

Proportionality, one of the most important doctrines in constitutional adjudication,<sup>251</sup> is invoked when rights are limited by an administrative action and essentially investigates the relationship between the objective of this action and the means to realize it.<sup>252</sup> The birthplace of the proportionality test is Germany,<sup>253</sup> where the courts identified its basis in the Rule of Law and in the very essence of fundamental rights. This unwritten principle was regarded by German courts as constitutional,<sup>254</sup> and as an instrument that “carries the main burden of fundamental rights protection in Germany.”<sup>255</sup> From Germany, the proportionality test spread to other European countries and later to additional jurisdictions outside Europe, such as Canada, New Zealand, and Israel.<sup>256</sup> It came to “dominate the dockets of constitutional and supreme courts around the world,”<sup>257</sup> and, as some assert, to

---

within the regular court system. *Id.* at 1-2.

250. See Moshe Cohen-Eliya & Iddo Porat, *American Balancing and German Proportionality: The Historical Origins*, 8 INT'L J. CONST. L. 263, 263-64 (2010).

251. Moshe Cohen-Eliya & Iddo Porat, *The Hidden Foreign Law Debate in Heller: The Proportionality Approach in American Constitutional Law*, 46 SAN DIEGO L. REV. 367, 369 (2009).

252. Aharon Barak, *Proportional Effect: The Israeli Experience*, 57 U. TORONTO L.J. 369, 371 (2007).

253. The German Federal Constitutional Court has been using the proportionality test since the 1950s to review laws impeding human rights or decisions that apply such laws. However, this test was developed before the German Constitution was established, during the late nineteenth century, in Prussia. See Cohen-Eliya & Porat, *supra* note 250, at 271-72; Dieter Grimm, *Proportionality in Canadian and German Constitutional Jurisprudence*, 57 U. TORONTO L.J. 383, 384-85 (2007).

254. Grimm, *supra* note 253, at 385-86.

255. *Id.* at 386.

256. Grimm, *supra* note 253, at 384; Alec Stone Sweet & Jud Mathews, *Proportionality Balancing and Global Constitutionalism*, 47 COLUM. J. TRANSNAT'L L. 72, 73-74 (2008).

257. Sweet & Mathews, *supra* note 256, at 73-74.

reflect “a global constitutional standard.”<sup>258</sup> The basic structure of the proportionality test is shared by many of these jurisprudences. The court starts by assessing the legitimacy of the objective. Then it applies a threefold proportionality test: first, the chosen means need to be adequate to furthering its objective (“suitability”); second, the means must be that which is least intrusive upon the individual’s rights (“necessity”); and third, the court performs a cost-benefit analysis, and checks whether a proper balance between the burden on the individuals’ rights and the gain afforded has been struck (“proportionality in the strict sense”).<sup>259</sup> Notwithstanding the common structure of the principle, its application<sup>260</sup> (including the interpretation provided to its three subtests) may vary depending on the given legal regime.<sup>261</sup>

The proportionality principle is also considerably used on a supranational level, inter alia, by the courts of the European community, the European Convention on Human Rights (“ECHR”), and the World Trade Organization (“WTO”).<sup>262</sup>

Another fundamental and frequently used doctrine in administrative and constitutional law, primarily in various common law countries, is reasonableness.<sup>263</sup> Reasonableness, like proportionality, could be understood as an instrument that aims to prevent governmental authorities from abusing their power<sup>264</sup> and reflects a stance that rejects the notion of absolute power.<sup>265</sup> In the U.K. this principle is mainly associated with “*Wednesbury* unreasonableness,” which addresses

---

258. *Id.* at 111-37; see also DAVID BEATTY, *THE ULTIMATE RULE OF LAW* 159-88 (Oxford Univ. Press ed. 2004) (referring to proportionality as the “ultimate rule of law”); David S. Law, *Generic Constitutional Law*, 89 MINN. L. REV. 652, 693-98 (2005).

259. Cohen-Eliya & Porat, *supra* note 250, at 267; Grimm, *supra* note 253, at 387-88.

260. E. THOMAS SULLIVAN & RICHARD S. FRASE, *PROPORTIONALITY PRINCIPLES IN AMERICAN LAW CONTROLLING EXCESSIVE GOVERNMENT ACTIONS* 5 (2009) (explaining that the application of the proportionality test in every jurisdiction might vary in accordance with the extent to which an individual’s rights are being recognized in each legal regime. The difference in this regard often concern social, economic, and cultural rights because “civil and political rights generate more uniform government acceptance and protection.” Welfare states, thus, use this principle more extensively).

261. *E.g.*, Grimm, *supra* note 253, at 387-95 (explaining that German courts attribute a rather narrow interpretation to the required adequacy of the objective, by only demanding a purpose that is not prohibited by the constitution, whereas the Canadian courts apply a stricter approach (though German courts may take this factor into consideration later, when addressing the third subtest)). For the differences between Germany and Canada in applying the proportionality test, see Grimm, *supra* note 253, at 387-95. For the Israeli application of proportionality in comparison to other jurisdictions, see generally Barak, *supra* note 252.

262. Sweet & Mathews, *supra* note 256, at 138-59; see also GORDON ANTHONY, *UK PUBLIC LAW AND EUROPEAN LAW: THE DYNAMICS OF LEGAL INTEGRATION* 110-16 (2002).

263. WADE & FORSYTH, *supra* note 236, at 294.

264. *Id.* at 295.

265. *Id.* at 294.

absurd decisions that no reasonable authority could have reached,<sup>266</sup> but also relates to a wide category of “irrelevant considerations” and mistakes.<sup>267</sup> Alongside the differences between reasonableness and proportionality,<sup>268</sup> it seems that these doctrines “cover a great deal of common ground,” and that a “disproportionate administrative act could also often be perceived as unreasonable.”<sup>269</sup> Moreover, the proportionality principle is infiltrating U.K. law by the ECHR and the European Court of Justice.<sup>270</sup> Against this background, many advocate for the use of proportionality as a stand-alone principle also in the U.K.<sup>271</sup>

Another approach which constitutes an exception to the wide endorsement of proportionality<sup>272</sup> is the U.S. jurisdiction, which essentially relies on three tiers of scrutiny tests when evaluating challenged laws or acts.<sup>273</sup> The tiers differ according to the nature of the purpose of the act/law at issue (legitimate; important; compelling governmental interest) and the linkage between that act/law and its purpose (rational/reasonable; substantially related; or necessary/narrowly tailored/less restrictive alternative).<sup>274</sup> The type of scrutiny applied is determined through a categorical classification of the law/act at issue. Strict scrutiny, the highest level of scrutiny, will be applied, for instance, when designated fundamental rights are threatened

---

266. *Id.* at 302. This approach was established in *Associated Provincial Picture Houses Ltd. v. Wednesbury Corp.* [1948] 1 K. B. 223 at 229.

267. WADE & FORSYTH, *supra* note 236, at 302-03.

268. *Id.* at 316-318. Wade and Forsyth explained that these doctrines also reflect a difference in the nature of judicial review. In proportionality lies a preliminary assumption of a harm to a protected right, and the administration is required to justify it, whereas in reasonableness, the assumption is that the act was legitimate, and the claimer needs to prove otherwise. *Id.* at 317-18.

269. *Id.* at 315-16.

270. *Id.* at 305-06.

271. *Regina (Alconbury Devs. Ltd.) v. Sec’y of State for the Env’t* [2001] UKHL 23, [2001] 2 All ER 929 (Lord Slynn’s opinion, para. 51) (“[T]he time has come to recognise that this principle is part of English administrative law, not only when judges are dealing with Community acts but also when they are dealing with acts subject to domestic law. Trying to keep the *Wednesbury* principle and proportionality in separate compartments seems to me to be unnecessary and confusing.”).

272. Vicki C. Jackson, *Constitutional Law in an Age of Proportionality*, 124 YALE L.J. 3094, 3096 (2015); *see also* Cohen-Eliya & Porat, *supra* note 250, at 276 (shedding light on the differences between the origins of these two approaches and explaining that “[w]hile in Germany the idea of proportionality was the way in which the protection of rights was introduced into a system that provided limited formal protection of rights, in the United States balancing was applied to address the opposite problem. In the U.S. there was strong textual support for the protection of rights but little textual basis for limiting rights. Balancing, therefore, was an important interpretative tool for the prevention of absolutism in the protection of rights.”).

273. Shapiro, *supra* note 14, at 189-93.

274. Richard H. Fallon, Jr., *Strict Judicial Scrutiny*, 54 UCLA L. REV. 1267, 1273-75, 1328-29 (2007).

or when individuals are classified on the basis of race.<sup>275</sup> Notwithstanding the uniqueness of the American approach, American constitutional law clearly shares significant conceptual similarities with both “proportionality”<sup>276</sup> and “reasonableness.”<sup>277</sup> Regarding reasonableness, Martin Shapiro further argued that “[t]he First Amendment really reads ‘Congress shall make no law abridging the freedom of speech that the Supreme Court finds on balance to be unreasonable.’”<sup>278</sup> Moreover, before the formation of the three scrutiny tests in the twentieth century, American courts generally evaluated the government’s regulation against a single reasonableness standard.<sup>279</sup> In recent years, calls have been made to return to that principle, or embrace a different general standard of judicial review, such as proportionality.<sup>280</sup>

### B. Reason-Giving

Reason-giving is a requirement applied to state authorities that obliges them to back their decisions with reason.<sup>281</sup> An important benefit achieved by this principle is the quality of the decision-making process, because “[a] decisionmaker required to give reasons will be more likely to weigh pros and cons carefully before reaching a decision than will a decisionmaker able to proceed by simple fiat.”<sup>282</sup> In the U.S. this duty is enshrined in the Administrative Procedure Act (“APA”)<sup>283</sup> and has been unfolded in adjudication over the years.<sup>284</sup> Courts have interpreted the “arbitrary and capricious” test under the APA<sup>285</sup> as a requirement to

---

275. Tara Leigh Grove, *Tiers of Scrutiny in a Hierarchical Judiciary*, 14 GEO. J.L. & PUB. POL’Y 475, 486 (2016).

276. *E.g.*, Jackson, *supra* note 272, at 3094 (“[S]ome areas of U.S. constitutional law embrace proportionality as a principle, as in Eighth Amendment case law, or contain other elements of the structured ‘proportionality review’ widely used in foreign constitutional jurisprudence, including the inquiry into ‘narrow tailoring’ or ‘less restrictive alternatives’ found in U.S. strict scrutiny.”).

277. *E.g.*, Shapiro, *supra* note 14, at 190-91 (addressing the rational/reasonable linkage required in the American approach).

278. *Id.* at 191.

279. Grove, *supra* note 275, at 488.

280. *Id.* at 487-89; *see also* Cohen-Eliya & Porat, *supra* note 250, at 284-86; Jackson, *supra* note 272, at 3193-96; Shapiro, *supra* note 14, at 189-93.

281. Short, *supra* note 14, at 1813.

282. Shapiro, *supra* note 14, at 191.

283. 5 U.S.C. §§ 500–596 (2012).

284. Sidney A. Shapiro & Richard E. Levy, *Heightened Scrutiny of the Fourth Branch: Separation of Powers and the Requirement of Adequate Reasons for Agency Decisions*, 1987 DUKE L.J. 387, 405-07 (1987) (explaining that starting in the 1970s, American courts have imposed a stricter level of scrutiny over administrative authorities and incorporated a substantive component to this originally-procedural requirement).

285. 5 U.S.C. § 706(2)(A). This test constitutes the standard of review for informal rulemaking. In general, the APA relates to two forms of administrative rulemaking: formal and

investigate whether the agency's decision "was based on a consideration of the relevant factors and whether there has been a clear error of judgment"<sup>286</sup> and demonstrated a "rational connection between the facts found and the choice made."<sup>287</sup> In the 1983 *State Farm* case, the U.S. Supreme Court further articulated the criteria for reason-giving, and provided:

Normally, an agency rule would be "arbitrary and capricious" if the agency has relied on factors which Congress has not intended it to consider, entirely failed to consider an important aspect of the problem, offered an explanation for its decision that runs counter to the evidence before the agency, or is so implausible that it could not be ascribed to a difference in view or the product of agency expertise.<sup>288</sup>

In fact: "[T]he APA reason requirement has grown into a substantive demand that the agency give complete and even synoptically correct reasons."<sup>289</sup> Moreover, the APA requirements to give public notice and to allow the public's comments in the administrative rulemaking process ("Notice and Comment"), incorporate additional values, such as participation, public oversight, and transparency.<sup>290</sup>

The reason-giving requirement exists in many other jurisprudences. In England, for instance, administrators are under no general obligation of reason-giving,<sup>291</sup> but are required to follow specific common law-based obligations located in natural justice, especially in the Fair Hearing rule.<sup>292</sup> When such specific obligations apply, they might require "adequate" reasons that are clear and explanatory, and address

---

informal. Formal rulemaking (5 U.S.C. §§ 556-557) requires "rather rigid court-like procedures, with a complete record culminating in findings of fact and law supporting the rule announced." Informal rulemaking (5 U.S.C. § 553) is more commonly used and imposes fewer procedural requirements. It demands "that the agency give notice of its intention to make a rule, then accept comments by the public, and ultimately publish its rule accompanied by a concise and general statement of its basis and purpose . . ." Shapiro, *supra* note 14, at 185. For further discussion regarding the "arbitrary and capricious" test, see Shapiro & Levy, *supra* note 284, at 410-11.

286. *Motor Vehicle Mfrs. Ass'n of the U.S. v. State Farm Mut. Auto. Ins.*, 463 U.S. 29, 43 (1983); *Citizens to Preserve Overton Park v. Volpe*, 401 U.S. 402, 416 (1971); Shapiro & Levy, *supra* note 284, at 410-11.

287. *Motor Vehicle Mfrs.*, 463 U.S. at 43; *Burlington Truck Lines v. United States*, 371 U.S. 156, 168 (1962).

288. *Motor Vehicle Mfrs.*, 463 U.S. at 43; Short, *supra* note 14, at 1818-19; see also Patrick M. Garry, *Judicial Review and the "Hard Look" Doctrine*, 7 NEV. L.J. 151, 159 (2006).

289. Shapiro, *supra* note 14, at 198.

290. *Id.* at 204-05; see *supra* note 285.

291. Cane, *supra* note 16, at 313-14; Thomas, *supra* note 247, at 214.

292. Cane, *supra* note 16, at 313; Thomas, *supra* note 247, at 214-16; see *Lloyd v. McMahon* [1987] AC 625, 1135-36 [702 H] [PNR], *R v. Univ. Funding Council ex-parte Inst. of Dental Surgery* [1993] EWHC Admin 5 (QBD) (1993).

all substantial points at issue.<sup>293</sup> Reason-giving is apt to promote both substantive and procedural fairness.<sup>294</sup> Regarding the latter, in the famous *Doody* case, the court held that “a person who may be adversely affected” by a decision, should be able to argue his claims. In order to accomplish this, such person should be informed of the gist of the case and know “what factors may weigh against his interests.”<sup>295</sup> Nevertheless, it seems that the role of reason-giving in English law is less central than in U.S. law.<sup>296</sup> One reason for this divergence is that the former “deal[s] primarily with the making of individual decisions by administrative officials, [whereas] the account of U.S. law focuses on administrative rule-making.”<sup>297</sup> Like proportionality, the duty to give reasons is also recognized by multinational institutions.<sup>298</sup> European Community courts, for example, have noted the importance of reason-giving to secure effective judicial review and to facilitate protection of the affected individuals.<sup>299</sup>

## VI. BENEFITS OF APPLYING PUBLIC LAW NORMS TO SNSs

### A. *Why Public Law?*

In the Parts above, I explored the emergence of SNSs as arenas for crafting and allocating human rights and argued that this feature loudly echoes the principal and ancient role of the state. This, I suggest, highlights a salient public attribute of SNSs that justifies the application of public law norms to them. Such norms, I showed, play a vital role in various jurisprudences as a means to confine the immense power of the state with regard to managing human rights and to promote substantive and procedural fairness.

The enlistment of public law norms in the context of SNSs is required, *inter alia*, because private law was not designed with far-reaching influence on human rights in mind and is currently ill-equipped

---

293. Cane, *supra* note 16, at 314; *see* South Bucks Dist. Council v. Porter [2004] 1 WLR 1953, para. 24.

294. Cane, *supra* note 16, at 313-14.

295. Regina v. Sec’y of State for the Home Dep’t, *ex-parte* Doody: HL 25 Jun 1993 (Lord Mustill), at 106.

296. Shapiro, *supra* note 14, at 184-85.

297. Cane, *supra* note 16, at 310.

298. Vlad Perju, *Reason and Authority in the European Court of Justice*, 42 VA. J. INT’L L. 308, 316-19 (2007); Shapiro, *supra* note 14, at 197.

299. *E.g.*, Case 222/86 Unectef v. Heylens [1987] ECR 4097, para. 15; Perju, *supra* note 298, at 316-17.

to tackle the considerable challenges SNSs present.<sup>300</sup> Private law consists of particular and circumscribed regulation that often fails to apply to new issues arising in the cyber realm in general and in SNSs in particular. In the absence of legal coverage of these “vacuum” areas, SNSs, which, at the end of the day are commercial corporations aiming to maximize their profits, do not have sufficient legal incentives to promote and secure human rights. In addition, enactment processes might last a considerable amount of time and not always catch on to the dynamic developments that thrive in SNSs. Public law norms, on the other hand, offer a broad blanket of standards capable of coping with the various unfolding challenges that arise in these online platforms. Such standards may offer holistic guidance to SNSs’ discretion in relation to all areas of conduct, whether covered by a particular regulation or not.<sup>301</sup>

Applying public law norms to SNSs aligns with broader processes in which the traditional distinction between public and private law was challenged. These processes included erosion of states’ sovereignty and growing power of international institutions,<sup>302</sup> privatization and outsourcing of government services,<sup>303</sup> the growth of powerful private bodies, and the widening notion that corporations are gaining “coercive power that had formerly been reserved to governments.”<sup>304</sup> The blurring between private and public may intensify in the cyber environment. This is reflected in recent American adjudication that closely linked SNSs’ services with the realization of the constitutional right to free speech<sup>305</sup> and perceived SNSs accounts of public servants as “Public Forums.”<sup>306</sup> It is also conveyed in the U.S. District Court for the District of Columbia’s recent acknowledgement that “[r]egulation of the Internet presents serious line-drawing problems that the public/private distinction in physical space does not.”<sup>307</sup> In light of these developments, calls

---

300. See *supra* Part IV.

301. However, some scholars argue that public norms are inherent to private law, thereby weakening the justification for recruiting public law as a means to realizing these norms in the private sphere. See Amnon Reichman, *Property Rights, Public Policy and the Limits of the Legal Power to Discriminate*, in HUMAN RIGHTS IN PRIVATE LAWS 245, 248 (Dan Friedmann & Daphne Barak-Erez eds., 2001) (presenting four approaches which challenge the notion that property rights allow discrimination in access to commercial property).

302. GRIMM, *supra* note 220, at 106-07.

303. PAUL R. VERKUIL, *OUTSOURCING SOVEREIGNTY: WHY PRIVATIZATION OF GOVERNMENT FUNCTION THREATENS DEMOCRACY AND WHAT WE CAN DO ABOUT IT* 23-25 (2007).

304. Horwitz, *supra* note 235, at 1428.

305. *E.g.*, *Packingham v. North Carolina*, 137 S. Ct. 1730, 1735-36 (2016); *Davison v. Loudoun Cty. Bd. of Supervisors*, 267 F. Supp. 3d 702, 718 (E.D. Va. 2017).

306. *E.g.*, *Knight First Amendment Inst. v. Trump*, 302 F. Supp. 3d 541, 549 (S.D.N.Y. 2018).

307. *Sandvig v. Sessions*, 315 F. Supp. 3d 1, 12 (D.D.C. 2018).

have been made to apply public law norms to certain online platforms as well.<sup>308</sup>

In the next two final Subparts, I wish to further explore how public law norms could be of value when applied to SNSs. My intention is by no means to assert that public norms ought to be applied to SNSs exactly as they apply to governments or that they should apply to every area of their operation.<sup>309</sup> Instead, I suggest that the essence of these principles aligns with these platforms' extensive influence over human rights and provides instruments to handle the challenges they present. First, I will discuss and illustrate how applying public law norms to SNSs could promote user-oriented decision-making-processes. Second, I examine additional benefits provided by the introduction of public law norms to the SNSs environment.

### *B. Obtaining User-Oriented Decision-Making-Processes*

If applied to SNSs, public law norms might encourage level-headed and user-oriented decision-making for all areas of SNSs' operation. It might, inter alia, urge SNSs to deliberately consider human rights when formulating policy, carrying out different practices, sanctioning users, and handling grievances.

I also believe that applying public law norms to SNSs will reap most benefits if applied with the promotion of the informational and social-communal key resources we explored above in mind.<sup>310</sup> Embracing proportionality or reasonableness as binding principles in the context of SNSs might stimulate thoughtful deliberation regarding harm to human rights and could encourage more moderate conduct from the very earliest stages of decision-making processes.

Proportionality, for instance, could offer a structured doctrine applicable in various areas, to guide SNSs when handling the numerous challenges to human rights that arise throughout their platforms and to encourage them to look for and to favor less offensive options. Thus, in accordance with the components of the proportionality test we explored

---

308. E.g., Ira Steven Nathenson, *Super-Intermediaries, Code, Human Rights*, 8 INTERCULTURAL HUM. RTS. L. REV. 149-56 (2013) (calling for the application of "Digital Due Process" on online intermediaries); see also K. Sabeel Rahman, *The New Utilities: Private Power, Social Infrastructure, and the Revival of the Public Utility Concept*, 39 CARDOZO L. REV. 1621, 1668-70 (2018) (arguing that online platforms should be perceived as public utilities). To this fascinating and lively discourse, I wish to add SNSs' role in crafting and allocating a broad spectrum of human rights as a salient public attribute of these platforms.

309. For instance, my arguments concerning application of public law norms to SNSs do not relate to their relationships with third parties, such as their employees or suppliers (to the extent that such relationships do not bear a substantive effect on users' human rights).

310. See *supra* Part II.

earlier,<sup>311</sup> if, for example, an SNS intends to deploy a new feature of automated censorship, then, *inter alia*, it ought to be “the least restrictive” one in terms of hampering human rights (and in the SNSs context—to minimally impair the availability of informational and communal resources). This might require, for example, enhancing and calibrating the automatic censorship’s technical capabilities and tightening human oversight thereof to prevent over-censoring or errors. In certain cases, “the least restrictive” measure might be achieved by “warning screens” or restricting the age of users who may watch harsh content, as is in fact done by some SNSs in certain cases.<sup>312</sup> This test might also suggest, for instance, that *ex-ante* censorship, considering its adverse ramifications for human rights,<sup>313</sup> should be considerably restricted and limited. In addition, the benefits introduced by the implemented censorship measure must outweigh the costs and burdens on human rights. When examining this requirement, it is important to ascribe the proper weight to the rights and values encapsulated in art depicting nudity,<sup>314</sup> fair use of IP-protected content,<sup>315</sup> satiric political posts,<sup>316</sup> and the exposure of corruption<sup>317</sup> or sexual abuse.<sup>318</sup>

Application of reasonableness could sometimes lead to the same results in the context of these censoring challenges. Insisting on deploying a wide censoring method that excessively harms human rights; not taking into account all relevant considerations; or considering discriminatory criteria, could be found unreasonable.<sup>319</sup> Likewise, American strict scrutiny might suggest that hindering First Amendment rights would require “narrow tailoring” or preferring “less restrictive alternatives.”<sup>320</sup>

Requiring SNSs to give reasons is also expected to serve the human rights of their users. This seemingly procedural duty may significantly influence the quality of the decisions taken by SNSs pertaining to censoring or any other field. It might apply when SNSs intend to internalize changes to their policy and practices, for instance. In such cases the reasons could become available to the public through various

---

311. *See supra* Part V.A.

312. *See supra* Part V.A.; *see also* *Community Standards Enforcement Preliminary Report*, *supra* note 116.

313. *See supra* Part III.C.

314. *See supra* Part III.C.1.

315. *See supra* Part III.C.1–2.

316. *See supra* Part III.C.1, E.2.

317. *See supra* Part III.D.2, E.1.

318. *See supra* Part I.

319. *See supra* Part V.A.

320. *See supra* Part V.A.

channels, including public proactive disclosure (possibly periodic) or disclosure by demand (brought by users or other actors, such as NGOs). As shown earlier, this principle might further indicate the duty to encourage users' participation in important decision-making, by directly communicating with them or with bodies that advocate for their rights.<sup>321</sup> In light of the benefits presented by reason-giving and the openness it entails, Facebook's approach: "We don't always share the details of our policies, because we don't want to encourage people to find workarounds"<sup>322</sup> is not persuasive. This requirement could also oblige SNSs to inform users whom they have sanctioned (for instance, by removing their posts or disabling their accounts/pages) of the particular reason that triggered such step. The logic and values embodied in the principle of reason-giving might also indicate SNSs' need to enhance users' grievance mechanisms and communication channels with SNS representatives, and perhaps to establish a support center, reachable via telephone or chat. It might also suggest that platforms should consider more positively requests to comment on reported incidents of misconduct on their part, brought by actors like journalists, activists, etc.

Many other public law norms and obligations that are ascribed to states, such as good faith and due process,<sup>323</sup> can also contribute considerably to attaining more responsible, open, and moderate decisions by SNSs, and to better safeguard human rights. Some of these norms could, for instance, require SNSs to form appropriate appeal procedures, which *inter alia* make room for users to be heard.<sup>324</sup> Proper appealing mechanisms might also require that sanctioned users be permitted access to some form of the platforms' (anonymized) database of cases in which other users were likewise sanctioned.

Public law norms might also suggest that when SNSs take the drastic measure of deactivating accounts (after ruling out less restrictive measures), they should make the property aggregated therein (list of friends, messages, photos, videos, etc.) available to the sanctioned users, and perhaps even facilitate mobility of such property to different SNSs. They might also be of great value for many information security issues, by imposing a duty to design and realize an effective capability to resist misuse of data by various actors. Moreover, public law norms might indicate SNSs' obligation to inform users immediately after a data

---

321. *See supra* Part V.B.

322. Bickert, *supra* note 129.

323. *E.g.*, Nathenson, *supra* note 308.

324. *See supra* note 135 and accompanying text (discussing appeal mechanisms concerning censored content); *supra* note 209 and accompanying text (discussing appeal mechanisms in general).

misconduct event has occurred, to properly and openly investigate it, and to take all measures to minimize the losses caused. Finally, applying public law norms might commit SNSs to consider their own responsibility with regard to the troubling routine in which public servants block users from their pages. This might lead, for instance, to a more user-driven architecture, which will require public servants to specify the reasons for the blockage, and will include transparent data with respect to such blockage instances.<sup>325</sup>

### C. Additional Benefits

Applying public law norms to SNSs carries significant additional advantages.

First, it could promote various democratic values. It might foster accountability and transparency on SNSs' part and reinforce the checks and balances exercised by public and state authorities in securing human rights online.<sup>326</sup> It might also reflect respect for users and their rights, assist users in crystallizing their legitimate expectations of these platforms, and encourage their engagement in these important avenues.<sup>327</sup> Finally, it might elevate the legitimacy of SNSs and users' trust in them, thereby encouraging a more meaningful and beneficial use from the users' point of view.<sup>328</sup>

Second, the application of public norms to SNSs could suit various jurisdictions and be realized in accordance with their specific political, social, and legal culture. These norms might, of course, vary from one legal regime to another, but they constitute a fundamental part of many of them. They could be applied through existing or designated legal tools and through different channels as legislation or adjudication. Relevant legal tools in this regard are, for instance, the American "state action" doctrine<sup>329</sup> and the designated provision included in the U.K. Human Rights Act.<sup>330</sup>

---

325. See *supra* Part III.D.3.

326. See Shapiro, *supra* note 14, at 179-80; Short, *supra* note 14, at 1820-21.

327. See Cane, *supra* note 16, at 313.

328. Short, *supra* note 14, at 1823.

329. Courts interpreted the doctrine as such that applies constitutional duties to private bodies if a "public function" or certain types of linkage to the state is demonstrated. See, e.g., *Marsh v. Alabama*, 326 U.S. 501, 506, 509 (1946).

330. Human Rights Act 1998, c.42 (Eng.) § 6(3)(b). This provision considers private bodies performing "functions of a public nature" as "public authorities" that must abide by requirements set in the ECHR.

Third, thanks to the wide endorsement of public norms in local and supranational legal regimes,<sup>331</sup> they might serve as a multinational, or even global, preliminary legal standard for SNSs' required performance. The absence of a shared legal framework to guide SNSs' conduct carries substantial drawbacks. Some concern the difficulties in properly evaluating the risks and benefits introduced by these platforms, formulating a unified stance on their obligations; reaching legal certainty; and encouraging users to participate and take action in promoting their rights in these important realms. Above this preliminary legal tier, each jurisdiction could calibrate and nuance its unique approach, according to its particular needs and legal-economic heritage. This point should be understood against the backdrop of the lively discussion on global constitutionalism, which was demonstrated in passing above.<sup>332</sup>

## VII. CONCLUSION

SNSs have taken the form of digital arenas in which human rights are crafted, designed, and allocated. The emergence of these arenas relies on two dimensions: (1) the broad spectrum of rights affected by SNSs, which consists of civil-political and economic-social-cultural rights; and (2) the sophisticated control obtained by SNSs over these diversified rights. Such control is maintained through various practices incorporated into SNSs' operation, including surveillance, profiling, censorship, practices related to SNS-authorities nexus, and expulsion of individuals and communities. Together, these dimensions create overarching influence in shaping and allotting human rights that loudly echoes the principal and traditional task of the state in this regard. This salient public attribute of SNSs justifies the application of public law norms to them, thereby urging these platforms to take a more user-oriented approach in all various areas of performance. Application of public norms to SNSs affords additional advantages. These norms might promote accountability and users' participation, suit to various legal regimes, and even form a multinational or global preliminary shared legal standard, which could be later calibrated according to each jurisdiction's legal heritage and nature.

---

331. *See supra* Part V.

332. *See supra* Part V.