

Title	Chapter 2-I : Constitutional governance in digital discussion spaces : governing the Behemoth
Sub Title	
Author	水谷, 瑛嗣郎(Mizutani, Eijiro)
Publisher	Keio University Global Research Institute
Publication year	2025
Jtitle	Monsterizing Platform Power and Law ; Volume 2. Platforms and Power: How to Tame the Monsters ,p.59- 66
JaLC DOI	10.14991/KO11003002.00000002-0059
Abstract	
Notes	Chapter 2 : The Behemoth as a communication infrastructure
Genre	Book
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO11003002-00000002-0059

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

Chapter 2 - I

CONSTITUTIONAL GOVERNANCE IN DIGITAL DISCUSSION SPACES

: Governing the Behemoth

Eijiro Mizutani^{*}

With the spread of digital technology and the Internet greatly expanding people's opportunities for communication, the contemporary digital environment of expression has brought a truly "free marketplace of ideas." However, the online environment that underpins our daily lives is forcing a paradigm shift that cannot be ignored in a liberalism-based democratic system. One aspect of this context is the "attention economy," which began with newspapers known as penny press and has now become an economic principle deeply rooted in the online environment. Communication there shifted from a "competition of ideas" to a "competition of stimuli."¹

For example, rational discussions in the speech marketplace consider it desirable to weed out the dissemination of false information or misinformation and avoid government intervention in the expression environment to the extent possible. However, in the current attention economy, the emphasis is placed on information that can more efficiently capture our attention, and as a result, false information and misinformation that is disseminated and distributed in a more sensational or subliminal manner have come to occupy a dominant position over factual news. Additionally, some digital platform companies ("DPF companies")—especially those that have become dominant in the global economy—have adapted to this environment by implementing various mechanisms to surveil users and collect, analyze, and use a lot of information to make users stay on their websites for a longer period of time (i.e., "stickiness"). At the same time, many of these companies are based in the U.S. and have grown enormously, enjoying the benefits of "First Amendment Lochnerism," which strongly rejects government intervention in activities of expression.²

As a result, the DPF companies, who have become more dominant in the economic market, have become

^{*} **Eijiro MIZUTANI:** Associate Professor of Constitutional Law and Media Law in Institute for Journalism, Media & Communication Studies, Keio University.

¹ Tatsuhiko Yamamoto, "Paradox of 'Individualism': A Constitutional Aphorism for an AI Society" (Kobundo, 2023), p. 212.

² See Eijiro Mizutani, "'Marketplace' of ideas and the State," Horitsu Jiho, Vol. 2, No. 9 (2020), 30-3. DPF companies have a wide variety of business forms and sizes, but this article focuses on search engines and social media (Google, Facebook, X, YouTube, etc.) that influence freedom of expression and the democratic system in Japan, which is of particular interest to us.

a force to be reckoned with in the speech market as well.

This article presents some considerations and perspectives on how to control the de facto power exerted by DPF companies in online spaces from the perspective of constitutional law, referencing recent trends in U.S. judicial precedents and theories.

1 DPF Companies and the Current Media Environment

(1) DPF companies' function in designing the information environment

First, what role do DPF companies play in the modern environment of expression? The author has long indicated that DPF companies have a “function to design the information environment” in online spaces. This refers to the function of providing many users with the “place” essential for sending and receiving information and then “designing, managing, and operating” that “place” itself. This does not apply only to the general public; politicians, corporations, and the media must also rely on the “places” provided by DPF companies to communicate more effectively in online spaces. At first glance, information appears to be freely transmitted and received in such places, but as is evident from the series of controversies regarding the freezing of former U.S. President Donald Trump’s social networking accounts, the users’ freedom is increasingly becoming nothing more than freedom “in the palms of the hands” of DPF companies.³ The more oligopolistic the “place,” the more powerfully it binds us.

(2) Three elements of DPF companies

This function of DPF companies described above consists of the following three elements working in an interrelated manner. The first is *architecture (or code)*. As Lawrence Lessig once indicated, the code, “the ‘built environment’ of social life in cyberspace,”⁴ constrains users’ behavior in online spaces and functions like the law. For example, X (Twitter) typically imposes a character limit per post (as of 2024, this limit can be removed by setting a premium feature). The “design” of such an environment not only constrains users from posting lengthy text, but it also encourages short update intervals and responsive communication.⁵

Second, the existence of *algorithms (AI)* is also indispensable. A typical example is a personalized system that prioritizes recommendations among social media posts.

The third is *the content moderation mechanism*. DPF companies that provide social media have established policies (rules) to control user-generated content and enforce measures against violations. Although mechanisms vary from company to company, in most cases, other users or AI will “report” posts suspected of violating the rules (this is called “flagging”), and human moderators or AIs are responsible for judging.

3 Regarding the freezing of former President Trump’s SNS account, see Eijiro Mizutani, “Exploring the state of platforms from the freezing of the U.S. president’s account,” *Journalism* No. 71 (2021), pp. 54–61.

4 Lawrence Lessig (translated by Hiroo Yamagata), “CODE VERSION 2.0,” Shoeisha (2007), p. 171.

5 For example, Hamano Satoshi, “Ecosystem of architecture – How was the information environment designed,” Chikuma Bunko (2015), p. 212.

As discussed below, a key feature of the content-moderation process is the sheer volume of submissions to be reviewed, and hence, the implementation of AI-based automation (albeit not perfect).

2 Content Moderation in the U.S.

(1) “Censorship” by DPF companies?—Content moderation and the First Amendment

A sense of alarm is growing in the U.S. about the (de facto) power of DPF companies exercised in the environment of expression described above. The substance of that sense of caution, however, differs greatly between conservatives and liberals. The liberals’ issue is the DPF companies’ inadequate response to hate speech and disinformation. Conversely, some on the conservative and right-wing side have called content moderation by DPF companies “censorship.” They seem to attribute one case in particular—the freezing of President Trump’s social networking accounts—to DPF companies siding with the left and liberal side (although the truth is not clear).

Some have highlighted the opaqueness of the content moderation process. For example, a report by the UN Human Rights Council Special Rapporteur identified content moderation as the “platform law.”⁶ However, traditional American constitutional jurisprudence suggests that content moderation by DPF companies, which are private entities, is not a constitutional issue because the Constitution applies to state power, not private actors. Regardless of how much economic power DPF companies have, they cannot be bound by law on the basis of the Constitution. Rather, the First Amendment of the U.S. Constitution, which guarantees freedom of expression “from” the state, is a “weapon” for DPF companies. Even in the U.S., in some cases, DPF companies can be bound by the Constitution through court proceedings if extremely limited conditions called “state action” are met; however, for DPF companies, meeting these conditions is currently difficult.

Furthermore, in the U.S., Section 230 of the Communications Decency Act, which grants broad immunity to DPF companies to promote voluntary content moderation, may also be inconsistent with the regulation of DPF companies’ content moderation. Rather, the First Amendment operates as a protection against government regulation of DPF companies, who, in a sense, are granted the “freedom to moderate.”

(2) Series of lawsuits over preliminary injunctions against two state laws

Despite the conventional thinking described above, some interesting legislative developments have been observed in two U.S. states regarding content moderation by DPF companies.

One is a state law by the State of Florida called SB 7072. This includes prohibiting measures to freeze accounts of political candidates on social media platforms, a ban on post prioritization and shadowbanning of posts by candidates, and a ban on censorship of journalism companies, as well as the requirement for companies to be transparent and notify users and the imposition of considerable fines for non-

⁶ Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, UN Doc A/HRC/38/35, 18 June-6 July 2018, pp. 3.

compliance⁷. The other is HB 20 by the State of Texas. This law includes transparency and disclosure rules for businesses, in addition to prohibitions on “review” by social media platforms based on users’ “views.”⁸

These state laws have been cited as possible violations of the First Amendment (in addition, they may be inconsistent with Section 230 of the Communications Decency Act) because they restrict the freedom of moderation of private DPF companies, and NetChoice, an association of DPF companies, has filed a suit for a preliminary injunction against these state laws. The lawsuit began in federal district court, where an injunction was granted for alleged violations of both Florida and Texas First Amendment law but later split between the 11th Circuit Court of Appeals, which continued to grant the injunction, and the 5th Circuit Court of Appeals, which overturned the injunction, and the case will now go to the U.S. Supreme Court. As of this writing, oral arguments have been held, and a decision is expected in the summer of 2024.⁹

Then, what is the logic behind each of the court’s decisions that have been divided as described above? The decisions that have enjoined state laws (i.e., found them likely to violate the Constitution) have more or less equated moderation with the “editorial judgment” of newspapers and others protected by the First Amendment and cannot survive strict or intermediate review by courts to determine unconstitutionality. Conversely, the decisions that overturned the injunction (i.e., the state law was likely not a constitutional violation) did not view content moderation as similar to an editorial decision. Rather, as described below, DPF companies, unlike newspaper companies, are considered to function as common carriers that do not touch the information circulating in the market the way telecommunication companies do.

(3) Are DPF companies “common carriers”?

The reason for this split in judgment appears to be related to the movement to position DPF companies as “common carriers.” To understand the existence of common carriers, understanding the classic American classification of what legal responsibilities should be borne by the intermediaries of information is necessary. First, businesses that conduct their own editorial activities, such as newspapers and magazines, are called “publishers.” While they have freedom of expression and editorial freedom under the First Amendment, they are just as liable as the originators themselves for the circulation of illegal expressions (such as defamatory expressions), even if they are merely publishing someone’s discourse. Conversely, businesses such as bookstores and libraries are called “distributors” and have the freedom to distribute content, but they are liable only if they know that illegal expression is circulating. Finally, common carriers are also called “conduits,” meaning that they are required to distribute the expression from A to B exactly as it is expressed. Typically,

7 The Florida Senate Website, SB 7072: Social Media Platforms, (<https://www.flsenate.gov/Session/Bill/2021/7072>) For the SB7072 bill and the federal district court injunction, see Eijiro Mizutani, “Content Moderation on Social Media Platforms and ‘Freedom of Expression,’” Media Communication: Bulletin of the Institute of Media and Communication, Keio University, No. 72, pp. 27–40.

8 Texas Legislative Online Website, HB20, (<https://capitol.texas.gov/BillLookup/BillStages.aspx?LegSess=872&Bill=HB20>)

9 Amy Howe, Social media content moderation laws come before Supreme Court, Scotusblog (Feb. 23, 2024), (<https://www.scotusblog.com/2024/02/social-media-content-moderation-laws-come-before-supreme-court/>). The Supreme Courts of Florida and Texas have already ruled on July 1, 2024, to throw out the injunction lawsuits that dealt with the content moderation laws in those two states. See, Amy Howe, Court sends social media moderation cases back to lower courts, Scotusblog (Jul. 1, 2024), (<https://www.scotusblog.com/2024/07/court-sends-social-media-moderation-cases-back-to-lower-courts/>).

telecommunications carriers, such as telephone companies, do not have the same editorial freedom as publishers and do not intervene in the expression circulating, but they are also not liable for it.

In fact, in the ruling overturning the injunction against the state law restricting content moderation discussed earlier, the court found that the state was regulating DPF companies to treat them as common carriers. Justice Clarence Thomas, who believes that DPF companies should be classified as common carriers, is also on the U.S. Supreme Court, where the issue will be decided. He claimed the similarity of social media to common carriers or public accommodations in another court case.¹⁰ Some scholars have also proposed amending Article 230 of the Communications Decency Act to allow DPF companies to choose between common carrier status or distributor status.¹¹ However, in my opinion, the “DPF companies = common carriers” argument seems to strongly reflect the unique circumstances of the United States. In the U.S., First Amendment rights are so highly valued that many government regulations (especially content regulations) are declared unconstitutional; freedom of expression has recently become such a powerful de-regulatory tool that it has become known as “First Amendment Lochnerism,” and such arguments may be a means of bypassing the mighty “wall” erected by the First Amendment.

3 Governance of Design of Online Discourse Space

(1) The significance of “The New Governors”—DPF companies as intermediary organizations

A legal scholar in the United States, Kate Klonick, has examined the position of DPF companies from a different perspective than the discussion of the legal liability of information intermediaries. Through interviews and document research with DPF companies, Klonick argues that they have a centralized organization, detailed rules and training, and a content moderation mechanism with a three-tier escalation structure from Tier 1 to Tier 3, and are subject to external influences when updating policies and rules; this positions them as “the New Governors of online speech.”¹²

In Japan, studies have been conducted from a similar perspective. Constitutional scholar Tatsuhiko Yamamoto indicates that when the relationship between the state and DPF companies is considered, the “the relationship between the state and ‘intermediate groups’ in medieval European feudal societies” is particularly important, and notes the similarities between DPF companies and “the Catholic Church.”¹³ DPF companies, like the “Church” in the Middle Ages, “have their own law (*ius canonicum*) of the ‘code-algorithm,’” (2) “have their own sanctions such as freezing (banning) accounts,” and (3) “have their own

10 See, *Biden v. Knight First Amendment Institute at Columbia Univ*, 141 S. Ct. 1220 (2021) (THOMAS, J, concurring).

11 Eugene Volokh, *Treating Social Media Platforms Like Common Carriers?*, 1 J. FREE Speech L. 377, 454 (2021).

12 Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1603 (2018)

13 Tatsuhiko Yamamoto, “Modern States and Digital Platforms: Leviathan versus Behemoth,” in Hajime Yamamoto (ed.), *Lectures on Constitutionalism and Constitutional Law, Volume 1: Basic Theories of Constitutional Law*, Shinzansha (2022), p. 164.

‘taxation’ system, such as data collection.”¹⁴ He then contrasts the DPF companies with the state, which is a leviathan, and sees them as a “behemoth,” advocating “building a strategic cooperative relationship” between the state and the DPF companies based on a “concordat.”¹⁵

By considering some of these behemoth DPF companies as having functions comparable to those of the state, we may find it necessary to examine the governance of such companies from a constitutional perspective. However, the DPF companies, as “the new governors,” are not purely the state itself; this is because they are still fundamentally private companies pursuing profit in the economic marketplace. First, economic incentive seems to be one of the factors behind DPF companies’ voluntary content curation, including content moderation. Under the economic principle of the Attention Economy, DPF companies want users to stay on their platform as long as possible. In this regard, the spread of consecutive lawful but harmful information in the space provided by the company may cause users to leave.¹⁶ Additionally, considering that the advertising revenue system is a major support for the attention economy, to maintain the advertiser’s brand safety, hate speech should not spread excessively widely on platforms.

Thus, if the state is an actor that realizes “public welfare” that cannot be reduced to activities in the economic market alone, DPF companies must be considered as a different actor, one whose operating principle is the pursuit of economic profit. Therefore, in principle, we should not (or would find it difficult to) directly apply “freedom of expression” as a user’s “human right” to discipline DPF companies in the same way as it is for the state.

However, in light of these differences with the state’s function, seeking a certain method of “governance” for the “power” that DPF companies can exercise in their own space to realize the constitutional value of maintaining and developing the democratic system that supports “freedom of expression” has room for consideration. The main question is how to design governance in a form that is (constitutionally) appropriate.

(2) Limitations of rights-based governance

In considering the constitutional governance of these DPF companies, understanding how they manage user content in their own venues (i.e., how content moderation works) is important. A court case involving a state law that attempts to limit content moderation by DPF companies has presented the idea of positioning the same as “editorial judgment” by newspapers. The world of content moderation, however, seems to be more like a factory floor process than a newspaper editorial one. And it seems to treat content more public health model than how it would be done in courtrooms, where each expression is carefully discussed and handled. For example, many DPF companies prohibit problematic content such as hate speech in their rules, but their efforts are made from the perspective of how to reduce the risk of users coming into contact with

¹⁴ Yamamoto, *supra* note (13), pp. 165-168.

¹⁵ Yamamoto, *supra* note (13), p. 177.

¹⁶ Shoshana Zuboff (trans. Kyoko Nonaka), *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Toyo Keizai Shinposha (2021), points out that since the threat to DPF companies is the intervention of public authorities and the mass departure of users, content management is only a “defensive measure” against such points.

problematic content (i.e., how to reduce the impressions).

Additionally, every day, a large amount of content is posted on social media operated by DPF companies, a huge part of which must be vetted. To deal with such quantitative problems, relying on human moderators is not a solution, and as a result, many DPF companies rely on AI-based automatic detection and deletion mechanisms. Such aggressive adoption of AI inevitably gives rise to errors (false positives/false negatives), thus making content moderation a probabilistic world. Therefore, in the governance of such a worldview, a systemic perspective that allows for and takes into account the errors that inevitably occur, rather than one based on individual rights, is essential.¹⁷

(3) Direction of constitutional governance for DPF companies

In light of the limitations of the approach based on the subjective theory of rights emphasized in traditional constitutional law (i.e., reconciling the interests of individual expression with the disadvantages caused by it), taking a more macro or systemic perspective and considering governance methods for DPF companies, who manage users' content with the above characteristics, from the viewpoint of maintaining the process of public opinion formation (the foundation of the democratic system behind freedom of expression) seems necessary. As a perspective on such constitutional governance, we would like to offer two directions.

The first direction is to increase the transparency of DPF companies' content curation process and introduce a due process perspective, asking whether the design of the process is legitimate or not. These methods are also included in the Digital Services Act recently enacted in the EU. Moreover, a set of private principles developed by scholars called the Santa Clara Principles 2.0, which have been endorsed by several DPF companies, also call for transparency in content moderation, user notification, and an appeal mechanism.¹⁸ Furthermore, the existence of the Oversight Board, an external independent body funded by Meta, is interesting to note as a voluntary initiative of the DPF companies. The Oversight Board is primarily a mechanism for receiving appeals against content moderation on Meta's platform, which may not be sufficient in terms of guaranteeing individual due process and can externally review the moderation process after the fact, point out policy design flaws, and encourage improvements.

The second direction is a more proactive governance model, bearing in mind the limitations to discussing individual errors (over-regulation or under-regulation) in the social media world, based on schemes to verify, correct, and remedy them after the fact. Such a proactive governance model would make use of methods such as continuous transparency at the policy-making level of DPF companies, the establishment of an independent audit system for risk assessment of content distribution and curation, and conducting quality assessments of content moderation.¹⁹ What DPF companies can do on a voluntary and continuous basis has limits; thus, a system of "collaborative governance" in which efforts are made jointly by the public and

17 See, Evelyn Douek, *Governing Online Speech: From 'Posts-As-Trumps' to Proportionality and Probability*, 121 COLUM. L. REV. 759, 791 (2021).

18 The Santa Clara Principles. On Transparency and Accountability in Content Moderation, (<https://santaclaraprinciples.org>).

19 See, Evelyn Douek, *Content Moderation as Systems Thinking*, 136 HARV. L. REV. 526 (2022).

private sectors may be worth considering.²⁰

Conclusion

As described above, when considering governance for DPF companies, whichever of the above two directions is sought, classifying companies according to their size and function rather than treating them as a single group is necessary. The EU Digital Services Act classifies DPF companies by scale, including such “Very large online platforms,” and establishes tiered regulations. In terms of functional differences, the Supreme Court of Japan has recently distinguished between the functions of the search engine Google (“plays an important role as a foundation for information distribution on the Internet”)²¹ and the social networking service X (“provides its users with a place to disseminate information and a means to obtain the information they need from among the postings, among other things”)²² in its ruling on the removal of criminal record information.

Additionally, understanding the “governance structure” and “ecosystem” of DPF companies and then discussing how to control them from the perspective of constitutional values is important. Although this point was not examined in depth in this article owing to space limitations, the trend of “digital constitutionalism” being developed mainly by European constitutional scholars is noteworthy.²³ Developing and extending the constitutionalist experiment of control over the power exercised by the “new governors” of the online environment of expression seems to be the key to the future of Internet governance. Examining the remaining issues in this report is the author’s obligation going forward.

* Note: this article is a substantially revised version of my article, “Constitutional Governance Perspectives on Social Media Platforms,” *Law and Computers*, Vol. 41 (2023), p. 33 and following, with substantial additions and revisions.

20 See, Margot E. Kaminski, *Binary Governance: Lessons from the GDPR’s Approach to Algorithmic Accountability*, 92 S. CAL. L. REV. 1529 (2019); also noteworthy is a similar approach advocated in Japan, the agile governance mechanism. For more on agile governance, see the Ministry of Economy, Trade and Industry’s new study group for a new governance model “Government Innovation Ver. 2: Toward the Design and Implementation of Agile Governance” (2021). See also Hiroki Habuka, Tatsuhiko Inatani, and others, “[Roundtable Discussion] Considering the Ideal Form of Governance Innovation,” *Japan Institute of Law and Information Systems* No. 10 (2021), p. 3 and following.

21 Supreme Court, 3 January 2016, *Minshu*, Vol. 71, No. 1, p. 66.

22 Supreme Court 24 June 24 2022, *Minshu*, Vol. 76, No. 5, p. 170.

23 For more on digital constitutionalism, see Kento Yamamoto, “Digital Constitutionalism and Constitutional Law,” *JILIS* No. 13 (2023), pp. 55-71.