

総務省プラットフォームサービスに関する研究会：  
誹謗中傷等の違法・有害情報への対策に関するワー  
キンググループ（第5回）

2023年4月7日（金）17:00～

**「新たな統治者」の社会的責務とガバナンス**

関西大学社会学部メディア専攻 准教授  
水谷瑛嗣郎（ミズタニ エイジロウ）  
博士（法学）



# 1. 表現環境におけるDPF事業者の機能

➤DPF事業者が有している情報環境形成機能は、以下の3つが相互関連的に働くことで構成されている

## (1) アーキテクチャ

→例：Twitterの字数制限

## (2) アルゴリズム

→例：レコメンドシステム（パーソナライゼーション）、エンゲージメント強化システム（トレンドの仕組み）

## (3) モデレーション

→「インターネット企業が、ユーザー生成コンテンツが利用規約やその他の規則で明示された基準を満たしているかどうかを判断するプロセス」\*

➡違法有害情報の対策においてもこれらそれぞれの要素に目を配る必要がある。



## 2. 「新たな統治者」としてのDPF事業者

➤ DPF事業者は、自社が提供する場の管理者として、国家に“類似”する統治システムを有しつつある。

→ 「**オンライン言論の新たな統治者 (the New Governors)**」\*。

• 領域主権 (territorial sovereignty) から機能主権 (functional sovereignty) へ

→ DPF事業者はすでにルール形成、執行、紛争裁定といった機能を担い始めている。

= ユーザーは「**専制君主の臣民 (subjects of a despot)**」\*\*の地位へ。

• リヴァイアサンVSビヒモス\*\*\*

「国家とDPFとの関係について考えるうえで有用だと思われるのが、中世ヨーロッパ封建社会における国家的なるものと『中間集団』との関係である。……**特に重要なのはカトリック教会との類似性であろう**」。

\* See, Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 Harv. L. Rev. 1598 (2018) .

\*\* See, Frank Pasquale, *From territorial to functional sovereignty: the case of Amazon*, Open Democracy, 5 Jan 2018,

<<https://www.opendemocracy.net/en/digitaliberties/from-territorial-to-functional-sovereignty-case-of-amazon/>>; See also, Frank Pasquale, *Digital Capitalism—How to Tame The Platform Juggernauts*, Friedrich Ebert Stiftung, Jun 2018, <<https://library.fes.de/pdf-files/wiso/14444.pdf>>.

\*\*\* 山本龍彦「近代主権国家とデジタル・プラットフォーム——リヴァイアサン対ビヒモス——」山元一編『講座 立憲主義と憲法学〔第1巻〕』(信山社、2022年) 164頁。

### 3. コンテンツ・モデレーションの特質と限界

#### ➤ アルゴリズムック・コンテンツ・モデレーション\*

- 例えば、Metaは、1日当たり1億件のポリシー違反の執行措置を行っている\*\*。
- モデレーションの量が膨大で、人力だけでは限界がある。

#### ➤ 今ではモデレーションにおいて自動化は不可欠なツール

- 例えばテロリズムコンテンツのフラグ立てについて、Twitterはアカウント凍結の**93%**がスパム対策ツールによるもの\*\*\*であり、YouTubeに至っては削除動画の**98%**が機械学習アルゴリズムによるもの\*\*\*\*。
- またYahooニュースコメント欄で、**約7割**をAIが自動処理で削除している\*\*\*\*\*。

→ 必然的に生じるエラー（過少or過剰執行）。そして、**Metaの例でいえばわずか1%のエラーですら、一日100万件になる。**

\* See, Robert Gorwa, Reuben Binns & Christian Katzenbach, Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance, Big Data & Soc'y, Jan.-June 2020.

\*\* See, Over Sight Board, PAO-2021-02, <<https://oversightboard.com/decision/PAO-NR730OFI/>>.

\*\*\* See, Isobel Asher Hamilton, Facebook, YouTube, and Twitter could face fines if they fail to take down terrorist content within minutes, BussinessInsider, Aug 20, 2018, <<https://www.businessinsider.com/eu-could-fine-tech-firms-if-they-fail-to-remove-terror-content-2018-8>>.

\*\*\*\* See, Susan Wojcicki, Expanding our work against abuse of our platform, YouTube Official Blog, Dec 05, 2017, <<https://blog.youtube/news-and-events/expanding-our-work-against-abuse-of-our/>>.

\*\*\*\*\* Yahoo! JAPAN 「メディア透明性レポート（2021年度版）」 <<https://about.yahoo.co.jp/common/transparencyreport/>>を参照。

### 3. コンテンツ・モデレーションの特質と限界

➤ DPF事業者はなぜこうした統治システムを構築・運用するのか？

- 「粘着性」（＝より長く、繰り返しPFを利用してもらう）が重視されるアテンション・エコノミーのもとでは、より多くのユーザーに「**安心・安全**」な空間を提供する必要がある（**トラスト&セーフティ**）。
- 広告主に対するブランド維持（**ブランドセーフティ**）。

→ **経済的必要性**

➤ モデレーションはDPF事業者にとって商品であるコンテンツの「品質管理」であり、ポリシーはそのための「品質決定指針」として機能（≠職人的＝生産ライン的）。

- **保険数理的(すなわち確率論的)世界\*であり、「公衆衛生」的モデル\*\***
- 公的機関の介入とユーザーの大量離脱が脅威となるため、その「防御策」としてコンテンツ管理を行う\*\*\*

➤ DPF事業者＝「新たな統治者」≠政府

➔ 違法情報対策に関しては、「公共の福祉」の主たる担い手である政府による一定程度の関与も必要。

\* See, Evelyn Douek, *Governing Online Speech: From 'Posts-As-Trumps' to Proportionality and Probability*, 121 Colum. L. Rev. 759, 797(2021).

\*\* See, Jonathan Zittrain, *Three Eras of Digital Governance* (Sept. 23, 2019), <<https://www.ssrn.com/abstract=3458435>>.

\*\*\* ショシヤナ・ズポフ（野中香方子訳）『監視資本主義』（東洋経済新報社、2021年）参照。



## 4. DPF事業者の責務—法的責任から社会的責務へ

- ▶ これまではDPF事業者も情報媒介者の側面を有することから、違法情報の流通に対して、民事上の個別事案に対する法的責任を調整する施策が適用されてきた（日・プロ責法3条、米・CDA230条）\*。
- Google (YouTube) については、CDA230条の免責をめぐる裁判が連邦最高裁で係争中（**Gonzalez v. Google LLC**）。
  - ⇒ 確率論的世界ゆえに必然的にエラーが生じうるため、過度な法的責任負担はオーバーブロッキングにつながる。
- DPF事業者 = コモンキャリア論
  - ⇒ DPF事業者を「導管」に徹しさせることは、モデレーションを封じることになる。
- ▶ しかしここまで見てきた通り、DPF事業者は単なる媒介者を超えて、表現環境における特異な地位を占めることから、**違法情報に対する法的責任論から（公法上の義務を含む）社会的責務論へのシフトしていく必要がある。**
- EUのDSAは、「媒介者の免責の制度は維持しつつ、答責性・透明性に係る一定の行政上の義務を課すという近年の立法の流れをくむものと考えられる」との指摘\*



\* 例えば、See, Amy Howe, "Not, like, the nine greatest experts on the internet": Justices seem leery of broad ruling on Section 230, SCOTUSblog (Feb. 21, 2023, 4:31 PM), <https://www.scotusblog.com/2023/02/not-like-the-nine-greatest-experts-on-the-internet-justices-seem-leery-of-broad-ruling-on-section-230/>

\*\* 神足 祐太郎「EUの違法有害コンテンツ対策と意見の多様性：デジタルサービス法パッケージの経緯と概要を通じて」情報法制研究 10巻（2021年）64頁。

## 4. DPF事業者の責務—法的責任から社会的責務へ

### ➤ DPF事業者の違法情報流通に対する社会的責務

- (1) ネットワーク効果により、多数のユーザーをロックインしているため、我々は実生活においてその「場」に依存している（**寡占・依存**）
- (2) ユーザーを「統治」する仕組みを有している（**統治システム**）
- (3) アテンション・エコノミーのもとで違法情報を発信・拡散を（結果的にはあるが）誘発するアーキテクチャ・アルゴリズムが形成されている（**誘発的デザイン**）

→ 違法情報に関しては、その対象が（ある程度）明確化されており、また個人の人格的利益侵害という危害性（harm）に関して表現の自由との間の調整も（一応は）済んでいる  
（侮辱、名誉毀損、プライバシー侵害が典型）

→ DPF事業者はその特異な地位に鑑み、違法情報流通の低減に対する社会的責務を負い、それに伴い公法上の義務も一定程度果たす必要がある。

⇨ 他方、有害情報に関しては、対象をまず確定する必要があり、民主政社会全般に対して具体的にどのようなリスクを孕むのかについて、継続的に検証を行う必要もある。

### ➤ 検索エンジン事業者とSNS事業者の責務の違い

検索エンジン：「インターネット上の情報流通の基盤としての大きな役割」\*

SNS：「その利用者に対し、情報発信の場やツイートの中から必要な情報を入手する手段を提供するなどしている」\*\*



# 5. デュー・プロセスと協調的ガバナンスの可能性

➤ 社会的責務の実装としての二つの方向性\*

## ① デュー・プロセスおよび権利保障ベース

- ユーザー個々に対するアカウントビリティを果たすための透明化と適切な告知、さらに不服申立ての仕組み
- Metaの監督委員会のような事後的な独立した個別審査機関

→ ルールの課題発見と是正勧告（トランプ元大統領のアカウント凍結、クロスチェックプログラム） ⇔ Twitterの「恩赦」の流れ（非常に「きまぐれ」的）

But…モデレーションが保険数理的世界であることを踏まえ、事後的な手続き保障のみならず、より「**プロアクティブ**」な制度設計が必要\*。

## ② 協調的ガバナンス（あるいは共同統治）ベース

- 例えば、継続的な透明性報告書の公表、影響評価、外部監査システムなど

→ DPF事業者の統治において生じうる潜在的リスクを事前に洗い出し、アルゴリズムの設計段階等での問題点を予防・改善すること等を促進し、ひいてはその統治に対する社会的受容性を担保することを目指す。

➡ この二つの方向性を相互互換的に捉え、両立を検討すべき。



\* 以下については、次の論稿を参考にしている。See, Margot E. Kaminski, Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability, 92 S. CAL. L. REV. 1529 (2019); See also, Evelyn Douek, Content Moderation as Systems Thinking, 136 Harv. L. Rev. 526 (2022).; Ari Waldman, Shifting the Content Moderation Paradigm, 2022 Jotwell: J. Things We Like 1 (2022).



## 6. 「検閲代理人」化の防止—政府の透明性とアカウントビリティ

➤ただし、こうした政府による関与を行っていく場合には、「間接規制」及びDPFが「検閲代理人」化するリスクを考慮しなければならない。

→政府側にも相応の透明性・アカウントビリティが求められる (See, Santa Clara Principles 2.0\*).

+ 政府関与の仕組みの濫用リスクを軽減するための制度設計

- 民主政システムにとってもっとも重要な世論形成の機能（世論による政府・政治家に関連する批判）が損なわれないように、違法情報の削除義務化や行政庁によるDPF事業者への削除強制、さらに特定アカウントの凍結強制を行う場合など、DPFへの強制力を伴った措置には、（仮に判例法理の「検閲」に該当しないとしても）例えば第三者機関の関与など、別途でガバナンスの仕組みを設ける必要である\*\*。

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

\*\* 例えば、川崎市のヘイトスピーチ対策（差別のない人権尊重のまちづくり）条例においては、「拡散防止措置」等を市長が講じる際に、「差別防止対策等審査会」に諮問を行う仕組み（18条）となっている。



ご清聴ありがとうございました。

