



2024年12月10日

情報空間のリスク研究会 「Foreign Malign Influence (FMI) 対策におけるデジタル・プラットフォームフォーマー (DPF) と政府の協力に関する論点」 実施報告

布施哲委員は「Foreign Malign Influence (FMI) 対策におけるデジタル・プラットフォームフォーマー (DPF) と政府の協力に関する論点」と題して報告を行った。

Foreign Malign Influence (FMI) とは、主に「偽情報を中心に、政府の意思決定や世論を一定方向に誘導しようとする有害な影響工作」を指す。前提として、海外でもまだ「偽情報」や「影響工作」の定義が明確に確立されていないことを指摘しておく。日本でも「国内発・国外発」といった定義が明確におらず、誹謗中傷や違法な情報も「偽情報・影響工作」に含む場合がある。今回は、ソーシャルネットワーキングサービス (SNS) 上の「外国勢力による偽情報」および「偽情報とは言えなくとも、有害なプロパガンダや社会的な影響が大きい情報の拡散」について論じる。デジタルプラットフォームフォーマー (DPF) は、Google (YouTube 含む)、Meta、X、LINE、Yahoo! を想定する。

■ FMI における DPF の重要性

偽情報を含む FMI 対策には DPF との連携が重要である。DPF は FMI が生じた際に攻撃源に関する情報収集能力を擁し、こうしたコメントやアカウントの削除等のコンテンツ・モデレーションが可能で、巨額な研究費をもとに FMI を検知する技術開発等をしているからである。サイバー空間の生殺与奪権を有すると言って過言ではない DPF と、政府はどのような協力関係が可能なのか。この問題意識に基づいて、日米それぞれの政策を比較する。

■ 偽情報に関する日米の DPF 規制

米国は、言論や表現の自由を手厚く保護する方針から、政府介入は最小限に留めて DPF の主体性に任せている (Soft Law)。EU は直接的に有害コンテンツ (ヘイトメッセージ、児童ポルノ、テロ、違法な画像等) の削除を義務付けていないが、政府が求めた場合は有害コンテンツに対する DPF の対処を説明するよう義務付けている。従わない場合は、最大で全世界売り上げの 6% 相当の罰金を課す (Hard Law)。日本は米国と同じく DPF の主体性に任せる Soft Law だが、議論はもっぱら有害コンテンツ (暴力、テロ、詐欺、児童ポルノ、等) と誹謗中傷をどう削除するかが中心で、肝心の外国からの偽情報や影響工作への対策は見られない。

■ 日米 DPF の FMI コンテンツ・モデレーションの比較

米国では、Meta が中国やロシア由来の偽情報やアカウントを削除した実績がある。グーグルと X もこれらの国々の偽情報を検知して削除しており、米国の DPF は感度、対応規模、速度ともケープビリティが高い。実質的に、偽情報対策を担っていると見られる。

日本では、Yahoo!がYahoo!ニュースのコメント欄における有害情報（暴力、詐欺、誹謗中傷、被害情報等）を削除した例があるが、自らが情報の削除判断の主体になるのは避けたい傾向だ。ファクトチェック機関や政府の判断が公表されれば依拠する姿勢に留まり、海外からの影響工作の実態把握や検知能力が不足しているという課題がある。

では、日本市場に参入している外資系 DPF に協力を求められるかと言うと、これも芳しくない。グローバル企業にとって、各国の事情や要望に合わせた個別対応は効率面で難しいのかもしれない。そこで、日本でも外資系 DPF の主体性に依存するだけでなく、Hard Law 的な規制を梃に彼らの協力を引き出す検討が始まっている。例えば、総務省の検討会における最近の議論では、「行政法規を所管する行政機関からの申出、申請のあったコンテンツ」と「権利侵害性や違法性はないが有害性や社会的影響の重大性が大きい偽情報」を対象に、外資系 DPF にコンテンツ・モデレーションを求めようとする提案が出ている。この提案は、日米デジタル貿易協定第4条の「安全保障の例外規定」に基づいて考えた場合、一層正当化されると見られる。第4条が「国家の存立や安全に関わる重大事案においては、DPF の免責を除外する」となっているからだ。

■直近の米国の変化と取り組み

米国の施策は日本の一つのモデルになるので、米国の取り組みを概観する。米国政府は「DPF は出版社ではない。ユーザーの意見を流通させているだけ」というロジックから DPF に広い裁量を与えて自主性を優先してきたが、近年、民主党を中心に相次いで DPF に規制を求める法案が出された。共和党が「X のトランプアカウントの削除に象徴されるよう、規制は政府の検閲だ」と強く批判したので、この議論は党派対立を巻き込んで展開された。

現在、米国は政府も FMI に関与する体制をとっており、主要アクターはサイバーセキュリティや重要インフラの防護を担う国土安全保障省 (CISA) と連邦捜査局 (FBI) としている。CISA と FBI が、SNS 空間上の有害な情報を検知して DPF に削除を依頼する。しかし、ここでもやはり「政府は情報を提供するが、対応の全ては DPF の独立的な判断である」との姿勢は崩していない。FBI は公式見解として「DPF の対応がどうであれ、FBI は対抗措置を取らない」と発表している。では、この構図において「米国政府は、どこでどんな方法で顕著な兆候を把握するのか」「DPF は自身の運用指針と照らして対応を判断するわけで、このスピード感でどうなのか」等、いくつか明らかにしたい点がある。日本政府が必要な能力を備えていくに当たって、必要な情報だと思うからである。

米国の情報機関のネット監視手法、能力、体制等の情報は非公開なのでベールに包まれているが、機密解除文書から、ある程度読み説くことはできる。例えば、2022 年の米国中間選挙に関する事案を見ると、明らかに政府が、SNS 上のオープンソースの情報を見ているだけでは把握できないトレンドを握っていた可能性が示唆されている。国家安全保障局 (NSA) はメタ情報に留まらず、個人の投稿や画像、動画などのレベルまで把握していると見られる。更に、NSA は海底ケーブルの通信傍受の必要性についても指摘している。果たして通信傍受は今後の対策なのか、既に運用されているのか実態は不明だが、日本にとってこれも論点になってくると思われる。明らかな公開情報として、米国会計検査院による調査レポートがある。これは、米国の政府機関の能力として、(1) SNS のコ

ンテンツ分析（外国による偽情報ポストの探知、キーワード、位置情報等）、（2）ネットワーク分析（SNS 上のメッセージやユーザー情報の追跡、拡散動向とインパクト）、（3）自然言語処理（機械学習による偽情報のコンテンツの探知）、（4）合成コンテンツの探知（ディープフェイク、AI の生成による動画、音声、テキストの探知と解析）がある、と示している。こうした文書から、米国政府はオープンソース情報、プラットフォームからの情報、通信傍受情報の総合力を持って、FMI に対応していると推測される。

■日本の論点

以上を踏まえて、日本政府の FMI 対策における DPF の協力における論点は、（1）行為主体者の識別、特定、兆候の検知、意図の推測する能力、（2）偽情報の無力化を目的とした自動化された DPF によるモデレーション、および（3）事前に対象をリスト化する等、DPF に対するモデレーション要請手法、の3点だと考える。一方で、言論と表現の自由の保護は極めて重要である。「これだけは国家として譲れない」という有害性の判断基準や場面（例：国政選挙、憲法改正議論、台湾有事等）を明確にして、国民・DPF との相互理解を促す議論が必要だと考える。

報告を踏まえて、委員との質疑応答では「日本の DPF に対する評価」、「有害情報に関するコンテンツ・モデレーション規制と他のデジタルプラットフォーム規制の相乗効果」、「日本におけるセキュリティ適用の可否」、「DPF による偽情報識別・特定能力及び政府による規制や削除要請のあり方」、そして「米国政府がどのように DPF との協力体制を構築したか、日本との協力を仰げるか」といった質問とコメントが交わされた。

（了）