

SNS事業者の規律モデルについて

2022/4/21 総務省プラットフォームサービス研究会

曾我部真裕（京都大学）

プロフィール

曾我部真裕（そがべまさひろ）

1974年生まれ、横浜市出身。京都大学大学院法学研究科教授（憲法・情報法）。聖光学院高等学校、京都大学法学部、同大学院法学研究科修士課程、博士課程（中退）、司法修習生（第54期）、京都大学大学院法学研究科講師、准教授を経て2013年から現職。放送倫理・番組向上機構（BPO）放送人権委員会委員長、情報法制研究所（JILIS）理事、情報法制学会運営委員など。『情報法概説（第2版）』（共著、弘文堂）、『憲法Ⅰ 総論・統治（第2版）』『憲法Ⅱ 人権（第2版）』（共著、日本評論社）など。

Twitter @masahirosogabe

E-mail sogabe@law.kyoto-u.ac.jp

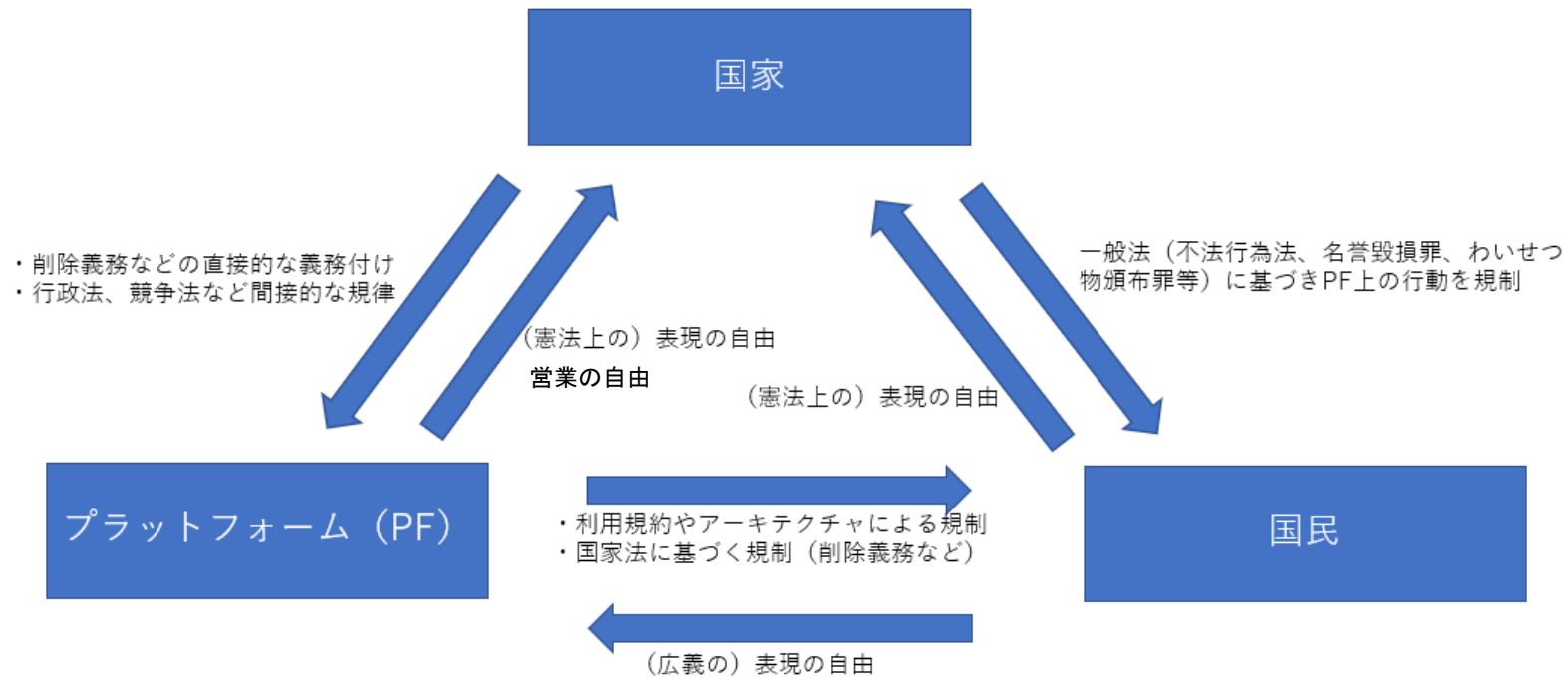
憲法・情報法の観点から見たSNS事業者の地位

- 成原慧「情報法からみたプラットフォームをめぐる法的課題（上）」法律時報93巻8号（2021年）
 - 情報流通の媒介者
 - データの集積者（アルゴリズムによる操作の問題）
 - アーキテクチャの設計者
- 情報流通の媒介者：現代における表現の自由の場
 - 社会的責任として、利用者の表現の自由の尊重が求められる（過剰削除を避ける要請）
- 危険責任としての責任
 - 情報流通の適正について一定の責任を負うのではないか。
- 基本権の主体としてのSNS事業者
 - 規制には合憲性の問題が生じ、また、法律の根拠を要する。

SNS事業者のあるべき規律をどのように構想するか

- 国家・事業者・個人の三面関係を踏まえること
 - この点については次のスライド参照。
- ミクロの観点とマクロの観点
 - マクロ的対策：誹謗中傷の総量を減らす対策（SNS事業者の取組、ユーザーのリテラシーなど）
 - ミクロ的対策：個別の被害者救済（最終的には訴訟に）
 - プロバイダ責任制限法は両者にかかわる（個々の救済に資すると同時に、削除対応を一般的に促す）。
- マクロの観点における3つのモデル
 - 自由放任モデル：特段の規律を設けず、一般法（民法、刑法など）の規律に委ねる。
 - 伝統的法規制モデル：（監視義務や）削除義務を課し、その違反には責任を問う。
 - 共同規制（規律された自主規制）モデル：情報流通の適正確保の義務を課し、透明性の確保等を通じ担保。
- 誹謗中傷問題以外との平仄：青少年保護、著作権保護、偽情報対策など。
 - アーキテクチャの設計者としての側面の考慮。

国家・プラットフォーム・国民の三面関係



(参照) 曾我部真裕ほか『情報法概説 (第2版)』(弘文堂、2019年) 34-35頁

S N S 事業者の規律：自由放任モデル

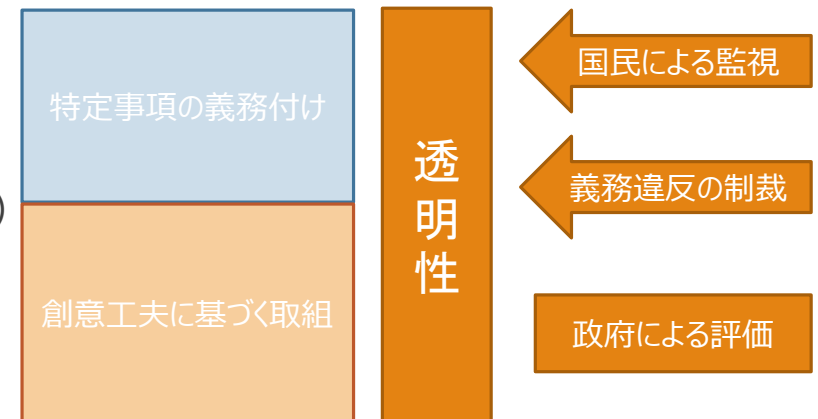
- SNS事業者に対する特段の規律を設けず、一般法（民法、刑法など）に委ねるモデル。
 - 日本では、プロバイダ責任制限法（2001年）制定以前の世界。
 - 同法制定後は純粋な自由放任モデルではないが、依然としてそれに近い状態であるが、その限界がこんにち問われている。
- ウェブ2.0による情報空間の変容と諸権利間の関係のリバランスの必要性
 - SNSの普及による情報拡散力の飛躍的向上：違法有害情報も同様。
 - アテンションエコノミー、選択的接触（コンテンツモデレーション）による感情の刺激：「炎上」の頻発、偽情報など。
 - 発信者に対する法執行の困難性。ゲートキーパーとしてのプラットフォーム規律の必要性。
- SNS事業者の責任
 - 表現の自由の場であるとともに、時には深刻な被害を生じ得る危険も伴う場の運営者としての責任があるのではないか。
- 表現の自由の変容
 - 「ソーシャルメディアの隆盛により、思想の自由市場に近い状態が現実化した結果、思想の自由市場論において素朴に前提とされてきたことが妥当しないことが露わになりつつあり、かつ、それによって弊害が生じつつある。これに対処し、思想の市場ないし情報空間を安定的に維持するためには、国家の介入が必要であるように思われ（…）」（拙稿「表現の自由(4)」法学教室492号（2021年）58頁）

S N S 事業者の規律：伝統的法規制モデル

- 個々の違法投稿につき削除義務を課し、その違反に制裁を加えるというモデル。
 - いわゆる「ルールベース」の規制。
 - 場合によっては監視義務も。
 - 民事法ではこのモデル：人格権に基づく削除請求とその違反に対する損害賠償請求（マイクロ対策＝個々の救済）。
- 執行リソースが大きい：行政機関による執行は疑問で、裁判所の関与を要するか。
- よくある指摘として判断の困難性と過剰削除のおそれ
 - ただし、要件論で対応可能か。プロ責法3条のような免責があれば過剰削除にはなりにくい。
- 監視義務を課すことはこれまでのコンセンサスに反する。
- 他の対策との平仄を視野に入れた場合、単調にすぎる。
 - 対策は削除だけではない。事業者の創意工夫の余地を阻害。

SNS事業者の規律：共同規制（規律された自主規制）モデル

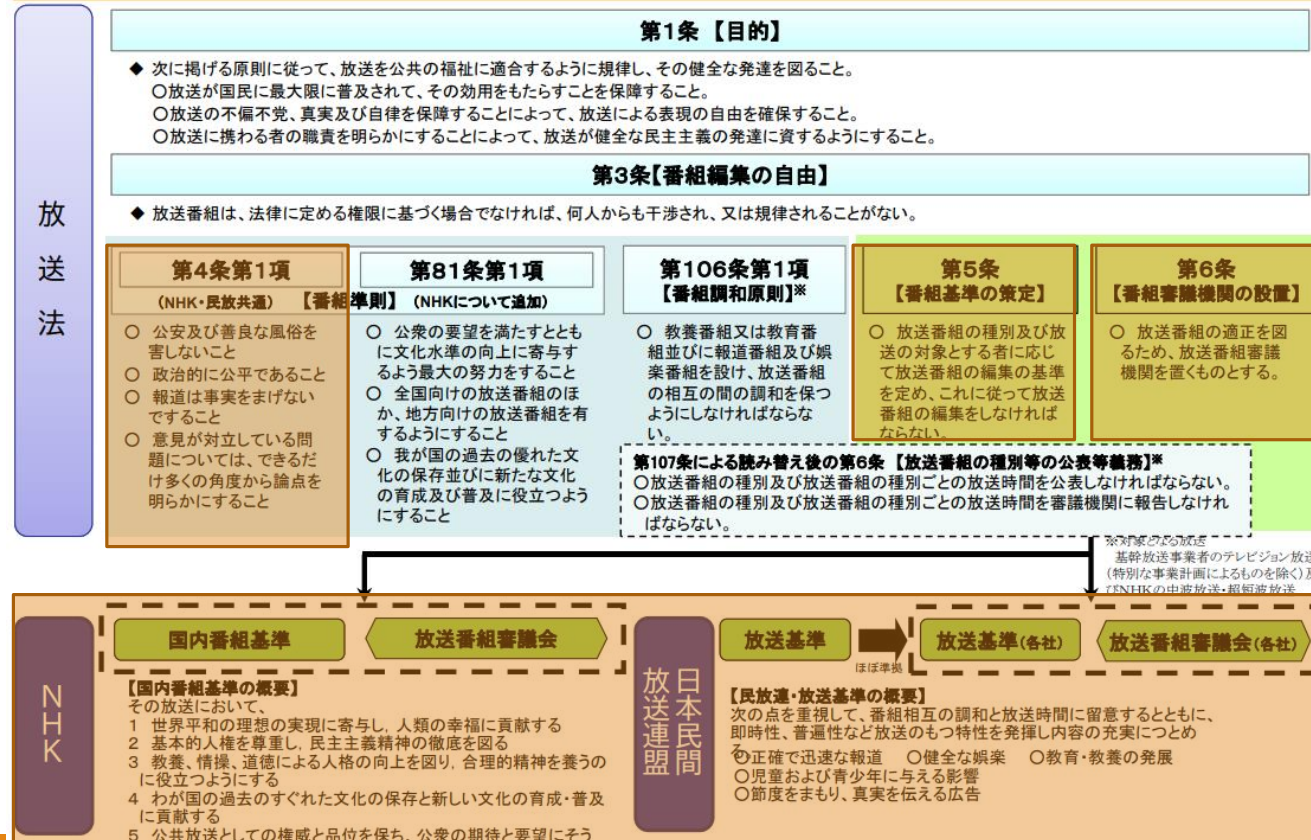
- 情報流通の適正確保の義務を課し、それを担保。
 - 透明性・説明責任の義務を課し、その違反に制裁。「ゴールベース」の規制。
 - より直接に体制整備（通報窓口、異議申し立て、監査など）の義務付け。
- 透明性・説明責任の趣旨：公開を通じた間接的な体制整備、情報流通の適正への取組の確保
 - 当該事業者がどの程度レピュテーションを重視するか、競合事業者の有無や状況などによって効果は異なる。
 - 特定DPF透明化法は政府が評価するプロセスを導入。
- 創意工夫の余地を認める。競争に持ち込めればなおよい。
- どこまで創意工夫に委ねるか：直接的な義務付けは不要か
 - 事業者に対する予測可能性や、規制効果のコントロールの観点から要検討。
 - 過剰削除のおそれ：異議申し立て制度の義務付けなど。
 - 過少削除のおそれ：直接的な義務付けよりは透明化により対処（民事もある）
- 他の対策との平仄：青少年保護や偽情報対策でも利用可能か。
 - アーキテクチャの設計者としての側面を考慮できる。
 - DSA法案など、近年の傾向とも合致
- 立法の必要性：SNS事業者の基本権を制約する側面がある。
 - 政府規制も透明化によってチェック可能となる。



それぞれの要素のバランスで多様な制度設計がありうる

放送法における「規律された自主規制」

- 番組準則を参照して番組基準を策定する義務。番審設置を義務付け、それを通じて説明責任を果たす。
 - 報告者は「日本モデル」と呼んでいる（拙稿「放送番組規律の『日本モデル』の形成と展開」（<http://hdl.handle.net/2433/173401>）。



【参考】 ドイツにおける「ネットワーク執行法」

- 2017年6月、ドイツでは難民の増加に伴う難民に対するSNS上のヘイトスピーチや偽情報等への対策として、2017年にネットワーク執行法が成立した。

対象事業者

- 利用者が任意のコンテンツを他の利用者と共有し、又は一般に公開可能なインターネット上のプラットフォームサービス（ソーシャルネットワーク）を営利目的で運営する事業者。
ただし、ドイツ国内の登録利用者数が200万人未満の事業者は、次の報告義務及び対応義務を負わない。

報告義務

- 年間100以上の苦情を受ける対象事業者は、違法コンテンツに係る苦情の処理について、半年ごとに当該期間を対象とする報告書を作成し、連邦官報及び自身のウェブサイト上で公表しなければならない。

申告のあった違法コンテンツへの対応義務

- 対象事業者は、違法コンテンツ申告のための手続窓口を設けた上、申告があった場合は、直ちに違法性を審査し、原則として以下の期間内に削除又はアクセスブロックをする義務を負う。
 - ① 明らかに違法なコンテンツ：申告を受けてから24時間以内
 - ② それ以外の違法コンテンツ：申告を受けてから7日以内
- 対象となる違法コンテンツは、ドイツ刑法の特定の犯罪に該当するものに限られる。

認定自主規制機関

- 対象事業者は、複数のソーシャルネットワーク提供事業者等が設立するなど、認定にあたり一定の要件を満たすことを要する自主規制機関に対し、コンテンツが違法か否かの判断を照会することができる。

過料

- 最大500万ユーロの過料（法人・団体には最大5,000万ユーロ【約65億円】の過料）
- なお、コンテンツが違法であるという事実に基づき、行政庁が過料を科そうとする場合には、コンテンツの違法性について、予め裁判所による先決的決定を得る必要がある。

施行日

- 2017年10月1日

- ドイツ国内の登録者数が200万人以上のソーシャルネットワーキングサービスは、ネットワーク執行法に基づき、透明性レポートを半年に1回公開する義務がある。

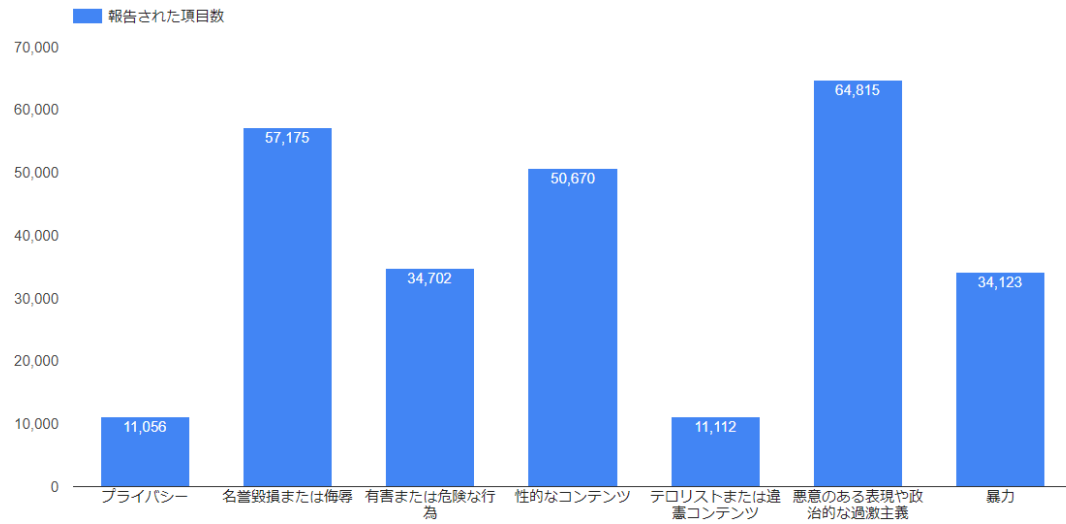
Google + ^{*1}	YouTube ^{*2}	Facebook ^{*3} /Instagram ^{*4}	Twitter ^{*5}
<特徴>			
2018年下半期の違反報告数は約2,800件。(2019年4月でサービス廃止。)	Twitterに次いで違反コンテンツ報告件数と削除された件数が多い。2020年上半期の違反報告数は約39万件。	Facebookは、2018年上半期(2018年7月提出)のレポートで、ユーザーからの違反報告数が過小報告されていたため、科料が科された。今後、FSMとの合意の元、監査が入る。2020年上半期の違反報告数は合わせて6,300件。	NetzDGの違反コンテンツ報告件数、削除された件数が最も多い。削除にあたっての判断を外部委託している件数も他の企業と比して多い。2020年上半期の違反報告数は約76万件。
<違法な投稿を防止のための取り組み>			
自動検出と人による報告を併用。画像や映像それぞれに付けられるハッシング(Digital Fingerprint)機能を使用し、違反と判断されたコンテンツの再アップロードを検知する。既に違反と判断されたコンテンツをもとに、2017年6月から機械学習を基盤とした機械での違反コンテンツのスクリーニングを導入。違反のコンテンツが明らかに存在する時に非常に有効。コミュニティガイドラインの作成。	自動検出と人による報告を併用。機械学習による画像の自動検出のほか、児童ポルノの場合はシステム上で、何人が特定のアカウントをブロックしているのか、急に特定のアカウントが子供のアカウントと連絡を取ろうとしているか、などの情報も確認する。コミュニティガイドラインの作成。	自動検出と人による報告を併用。機械学習による自動検出独自のスクリーニングツールを整備。Global internet forum to counter terrorismのデータベースなどを利用している。コミュニティガイドラインの作成。	
<報告への対応に関する社内体制>			
コミュニティガイドラインで違反するものが第一段階で削除され、削除されなかったものに対してNetzDGの基準との照らし合わせが行われている。			
11人体制の他言語対応なNetzDGチームがドイツに在籍している。	66人体制の他言語対応なNetzDGチームがドイツに在籍している。	19人で構成される2部体制のチームが世界3拠点に置かれている。	50名以上の他言語対応のチームが存在している。

- ネットワーク執行法に基づく透明性レポートでは、期間の間での違反報告数や削除した件数などが記載されている。主要事業者が公表している対応件数は以下のとおり。

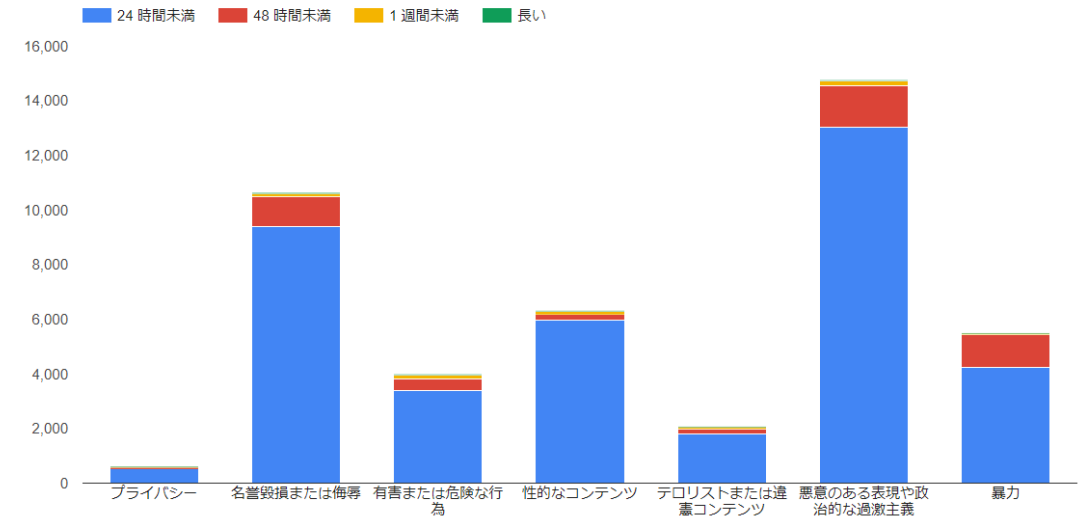
	Google+ ^{*1}	YouTube ^{*2}	Facebook ^{*3}	Instagram ^{*4}	Twitter ^{*5}	
違反報告の総件数 (直近1年間)	5,604	666,302	7,379	2,390	1,609,242	
違反報告のうち、 利用者からの報告(内数)	5,511	455,856	5,950	1,907	1,505,484	
違反報告のうち、 苦情処理機関からの報告(内数)	93	210,446	1,429	483	103,758	
直近1年間削除した件数 (報告の総件数に対する削除率)	2,779 (50%)	162,721 (24%)	3,351 (45%)	1,288 (54%)	259,473 (16%)	
削除 件数 の内 訳	24時間以内に削除したもの (削除総数における割合)	2,638 (95%)	150,015 (92%)	4,762 (90%)	1,365 (74%)	226,588 (87%)
	48時間以内に削除したもの (削除総数における割合)	86 (3%)	4,388 (3%)	273 (5%)	265 (14%)	28,507 (11%)
	1週間以内に削除したもの (削除総数における割合)	40 (1%)	5,385 (3%)	220 (4%)	172 (9%)	4,004 (2%)
	1週間より後に削除したもの (削除総数における割合)	15 (1%)	2,933 (2%)	40 (1%)	44 (2%)	298 (0.1%)
データ参照期間	2018年1月～12月	2019年7月～2020年6月	2019年7月～2020年6月	2019年7月～2020年6月	2019年7月～2020年6月	

Googleの報告書(2021年7月～12月分)

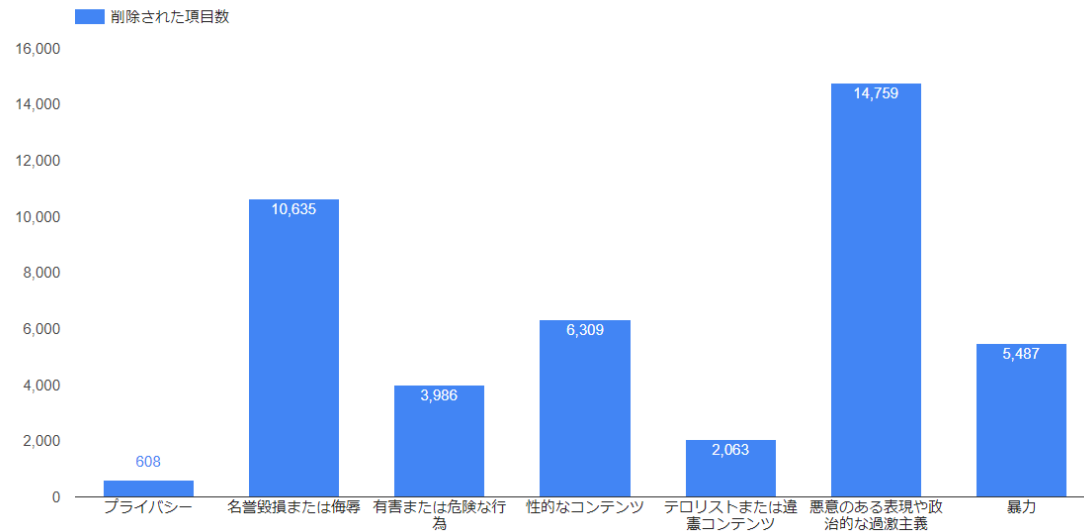
報告された項目数 (申告理由別)



所要時間 (申し立て理由別)



申し立ての理由に基づき削除されたアイテム



2021年7月～2021年12月

理由	24時間未満	48時間未満	1週間未満	長い
プライバシー	514	57	21	16
名誉毀損または侮辱	9,394	1,109	101	31
有害または危険な行為	3,393	404	165	24
性的なコンテンツ	5,969	205	124	11
テロリストまたは違憲コンテンツ	1,805	183	67	8
悪意のある表現や政治的な過激主義	13,044	1,508	191	16
暴力	4,249	1,197	35	6

Metaの報告書(2021年7月~12月分)

Table 2. NetzDG Complaints by German Criminal Code Provision(s) cited

Criminal Code Provision	Reporter Type		
	Complaints from Complaint Bodies	Complaints from Other Individuals	Total
Dissemination of propaganda material of unconstitutional organizations (§ 86)	952	19,332	20,284
Using symbols of unconstitutional organizations (§ 86a)	954	12,419	13,373
Preparation of a serious violent offense endangering the state (§ 89a)	716	9,991	10,707
Encouraging the commission of a serious violent offence endangering the state (§ 91)	675	9,048	9,723
Treasonous forgery (§ 100a)	666	9,152	9,818
Public incitement to crime (§ 111)	772	15,264	16,036
Breach of the public peace by threatening to commit offenses (§ 126)	754	12,778	13,532
Forming criminal or terrorist organizations (§§ 129 - 129b)	713	9,356	10,069
Incitement to hatred (§ 130)	1,125	32,260	33,385
Dissemination of depictions of violence (§ 131)	974	14,424	15,398
Rewarding and approving of offenses (§ 140)	717	13,074	13,791
Defamation of religions, religious and ideological associations (§ 166)	1,071	22,365	23,436
Distribution, acquisition, and possession of child pornography (§ 184b in conjunction with § 184d)	724	9,144	9,868
Insult (§ 185)	1,580	45,241	46,821
Defamation (§ 186)	1,357	35,099	36,456
Intentional defamation (§ 187)	1,246	29,905	31,151

Table 3. Number of Complaints Resulting In Deletion/Blocking

Criminal Code Provision	Reporter Type		
	Complaints from Complaints Bodies	Complaints from Other Individuals	Total
Dissemination of propaganda material of unconstitutional organizations (§ 86)	60	2,050	2,110
Using symbols of unconstitutional organizations (§ 86a)	50	1,651	1,701
Preparation of a serious violent offense endangering the state (§ 89a)	39	691	730
Encouraging the commission of a serious violent offence endangering the state (§ 91)	32	603	635
Treasonous forgery (§ 100a)	30	552	582
Public incitement to crime (§ 111)	45	2,127	2,172
Breach of the public peace by threatening to commit offenses (§ 126)	46	1,133	1,179
Forming criminal or terrorist organizations (§§ 129 - 129b)	32	545	577
Incitement to hatred (§ 130)	82	3,554	3,636
Dissemination of depictions of violence (§ 131)	54	1,462	1,516
Rewarding and approving of offenses (§ 140)	35	1,300	1,335
Defamation of religions, religious and ideological associations (§ 166)	88	2,420	2,508
Distribution, acquisition, and possession of child pornography (§ 184b in conjunction with § 184d)	42	1,207	1,249
Insult (§ 185)	186	8,194	8,380
Defamation (§ 186)	123	3,715	3,838
Intentional defamation (§ 187)	97	2,739	2,836

Fin

ご清聴ありがとうございました。