

## YouTube におけるコンテンツモデレーションの選好分析

兼保圭介<sup>1</sup>  
高口鉄平<sup>2</sup>

### 要 旨

本研究は、直ちに違法とは言い切れないが一般的には好ましくないような情報に対するコンテンツモデレーションに焦点を当て、その在り方を利用者の選好の観点から分析したものである。具体的には、YouTube を分析対象とし、「年齢制限のあるコンテンツ」のうち「子どもの安全」「有害または危険なアクティビティ」「下品な言葉」の3つのテーマを取り上げ、利用者がどの程度表示・削除を望み、誰に削除権限を委ねたいと考えているのかを量的に検証した。

本研究では、はじめに政策動向と法的手当てを確認し、先行研究を整理したうえで、「コンテンツを見たいか(表示するべきか)、見たくないか(削除するべきか)」、「仮に削除するとしたら誰が判断をするべきか」、「(開放的な方針を望むものが)炎上加担者やフェイクニュースの拡大に寄与するものと特徴が一致していないか」という3つの疑問点を、明らかにすべきリサーチクエストとすることとした。

つぎに、Web アンケートをもとにリサーチクエストについて定量的に検証を行った。検証の結果、第一に、「(好ましくないような)コンテンツを見たいか(表示するべきか)、見たくないか(削除するべきか)」という判断は、大きく利用者の態度が二分する状況となっていることが明らかとなった。さらに、この態度の決定には利用者の「基本属性」、「YouTube上の経験や思想」が影響を与えていることが明らかとなった。

第二に、「仮に削除するとしたら誰が判断をするべきか」について、これも「このままYouTubeに判断を任せるべき」と「自分で判断するべき」で態度がおおよそ二分することが明らかとなった。さらに、この態度の決定にはニュースサイト閲覧やテレビ視聴といった編集型メディア接触が影響を与えていることが明らかとなった。

第三に、「(開放的な方針を望むものが)炎上加担者やフェイクニュースの拡大に寄与するものと特徴が一致していないか」については、いずれも先行研究等で炎上加担の典型要因とされた性別・年収などは一部しか重ならないことが明らかとなった。

本研究の結果を踏まえ、政策視点でコンテンツモデレーションを考える場合、利用者の意見形成を目指すべきと考えられる。基本属性や経験、思想などが異なる多様な意見を集めて、議論を進める事が適切であると考えられる。

**キーワード：** YouTube, コンテンツモデレーション, 統計解析, ユーザ調査

<sup>1</sup> 静岡大学創造科学技術大学院 博士課程  
<sup>2</sup> 静岡大学情報学部情報社会学科 教授

## 1. はじめに

インターネットには一部の地域を除き世界中にサービスを提供することが出来るという特徴がある。また、インターネットが世界の人口の 60%以上に普及している現在、インターネット上でサービスを提供する事業者（サービスプロバイダー）は比較的容易に世界中にサービスを提供することが出来るようになった（ITU 2022）。世界的に普及しているサービスの中には、デジタルプラットフォームと呼ばれるインターネット上のプラットフォームを提供する事業者が運営するものがある。例えば、動画コンテンツを配信するプラットフォームである YouTube や、投稿（テキスト）を配信するプラットフォームである X（旧 Twitter）などがこれに該当する。デジタルプラットフォームは、利用者の増加がその価値を高めるといふ、いわゆるネットワーク効果が発揮されるサービスであり、いくつかの主要なデジタルプラットフォームが世界のインターネットの市場を支配しており、これらが日本でも広く普及している。

デジタルプラットフォームでは、自社の基準によりプラットフォーム上のコンテンツの流通に一定の影響を与える「コンテンツモデレーション」という行為を行っている。コンテンツモデレーションという行為には決まった定義がみられないが、これを理解するために二つの例を参考とする。一つ目は EU の Digital Services Act(以下 DSA)である。DSA では、事業者がコンテンツモデレーションに関するレポートを公表するように定め、違法コンテンツが利用規約にそぐわないコンテンツの検出方法および対処内容を含めるよう求めている(Article15)。更に、対処を行った場合には理由を通知しなければならないと定め、通知に際して明示すべき情報としてアカウントの停止やコンテンツの降格、無効化、アクセス不能化、削除などを挙げている(Article17)。これらの条項は、一般的なコンテンツモデレーションを理解するのに役に立つ。二つ目に参考とするのは、The Santa Clara Principles2.0 である。The Santa Clara Principles2.0 は Google や Meta を含む主にアメリカの企業を中心としてまとめられたコンテンツモデレーションに関する原則として発表されている。この原則の中では、事業者はルールやポリシーをユーザが簡易にアクセスできる場所で公表すべきとし、その情報の中には、どのような種類のコンテンツに対してコンテンツの削除、ランクダウン、アカウントの停止などの行為を行うかを示すことを求めている(Foundational- 2. Understandable Rules and Policies)。このような原則も、デジタルプラットフォームが行うコンテンツモデレーションを理解するのに役立つ。コンテンツモデレーションの定義、具体的な対処内容は多様であり、いまだ曖昧な部分があるが、2例をもとに捉えると、コンテンツモデレーションとは、おおよそ、プラットフォーム上の問題があるコンテンツに対して、デジタルプラットフォームを運営する事業者が自らコントロールすることを指すといえる。

本稿では、このコンテンツモデレーションに焦点を当て、利用者の意識や選好を分析することを目的としている。前述のとおり、コンテンツモデレーションの定義は多様であるため、分析に当たってはその定義を明確にする必要がある。第 2 節で示すとおり、コンテンツモデレーションに関する研究は多くが法学的研究であり、利用者の視点で実証的に分析した研究は見当たらない。この点で、本研究は試行的、探索的な側面があることから、コンテンツモデレーションをあえて狭義にし、デジタルプラットフォームによるコンテンツの表示有無の決定と定義する。ただし、これには対象アカウントを削除し投稿されたコンテンツを

一気に非表示にする行為や、以降の投稿を受け付けられない行為も含まれる。このように定義することで、コンテンツモデレーションが及ぼす利用者への影響をより明確にできる。

この定義でコンテンツモデレーションを見た場合でも多くの検討がある。その背景には、多くの利用者がそれぞれ違う価値観でプラットフォーム上に投稿したり、コンテンツをアップロードしたり、またそれらを見たりするため、ある人にとっては適切であっても、別の人にとっては不快なコンテンツがそこに存在してしまうという問題がある。事業者は、利用者の何をどこまで許容するか、逆に何をどこまで規制するかについて、法による規律に加えて自主的に利用規約などの規律を設けて運営を行っている現状がある。一方で、デジタルプラットフォームの影響力の大きさから、事業者の定めるルールが適切か否かについて、世界中で法的な検討や政策的な検討がなされている。

次節で整理するように、我が国においても様々な検討が進んでいるが、いまだ不透明な部分が残されている。その原因として利用者の意識や選好を踏まえたコンテンツモデレーションの検討は少ない事が考えられる。利用者の観点からのコンテンツモデレーションにも様々な論点があるが、どのような基準が好ましいと思われるのかを元に検討することは今後の日本の様々な検討の政策を考えるにあたり、一つの出発点になりえる。本稿ではこのような視点に立ち、アンケート調査を実施し実証分析を行うとともに、分析結果を元に今後の政策の在り方を検討する。

## 2. コンテンツモデレーションに関する政府の動向

本節ではコンテンツモデレーションに関する議論について、法的側面と政策的側面から我が国の状況を整理する。

### 2. 1. 法的動向

近年の日本では、法的な手当てが行われてきた。「特定電気通信役務提供者の損害賠償責任の制限及び発信者情報の開示に関する法律」（プロバイダ責任制限法）が令和4年5月に改正され、法律の題名（略称）が「特定電気通信による情報の流通によって発生する権利侵害等への対処に関する法律」（情報流通プラットフォーム対処法）として改められるとともに、大規模プラットフォーム事業者に削除等の対応の迅速化と運用状況の透明化が義務付けられた。もっとも、同法の第二章損害賠償責任の制限に変更はなかった為、プロバイダが知りえた違法コンテンツをプロバイダ自ら削除しなければ削除義務違反による損害賠償責任を負う余地がある一方で、直ちに違法とは言えないコンテンツに対しては対処方法について明らかではない（渡辺・梅本・今村 2021）という問題は残る。

### 2. 2. 政策的動向

コンテンツモデレーションにかかる政府の検討として、2007年より総務省によって行われた「インターネット上の違法・有害情報への対応に関する検討会」がある。ここでは特に青少年の保護、児童ポルノを問題点として取り上げ規制を検討し、表現の自由の観点から規制を最小限にするべきとして、利用者によるフィルタリングの活用が提言された。また、2020年2月に総務省が行ったプラットフォームサービスに関する研究会から最終報告書が発表された。この報告書は、プラットフォームサービスが市場プレゼンスを拡大したことに

伴い発生した課題や懸念に対して論点整理や方向性を示している。フェイクニュースや偽情報の対策として事業者による対策を基本とし、事業者の取り組みの透明性が期待されている。また、この報告書では法による規律は自由の表現への萎縮効果や実効性の欠如及び恣意的運用への懸念があることから民間部門による自主的な取り組みを基本とすると位置づけている。その後、2023年から総務省が行った「デジタル空間における情報流通の健全性確保の在り方に関する検討会」でも検討が行われ、そのとりまとめでは、利用者の思想がフェイクニュース、フィルターバブルやエコーチェンバーなどの外的要因でゆがめられる恐れがあることの対処の検討を行った結果として、政府による直接的な規制を行わず事業者コンテンツモデレーションに関する活動の透明性を求めている方向が記されている。続く2024年からはデジタル空間における情報流通の諸課題への対処に関する検討会が発足され、デジタル広告や上流通の観点などから検討が行われている。このように、コンテンツモデレーションは多くの検討を必要とする複雑な分野であるといえよう。

### 3. 先行研究

前節で見たように、我が国では法的手当てや政府主導による検討が行われてきたが、コンテンツモデレーションの基準はいまだ明確になっていない。「プラットフォーム上において多くの不確かな情報や悪意のある情報が容易に流通・拡散することは、利用者が多様な情報をもとに物事を正確に理解して適切な判断を下すことを困難にし、結果として、利用者が安心・信頼してプラットフォームサービスを利用することができなくなる、また、利用者の知る権利が阻害されるのみならず、利用者に直接的な損害を与え得るなど、利用者にとって様々な不利益が生じるおそれがある」と総務省の報告書に示されるように、日本の政策的な検討には閲覧者の被害に着目した検討が多い。しかし、デジタルプラットフォームには様々な立場の利用者がいる為、閲覧者以外の様々な立場を踏まえた検討も必要となる。

本節ではデジタルプラットフォームの責任や、投稿者に着眼した先行研究を用いて政府動向を評価する。

#### 3. 1. デジタルプラットフォームの責任に関する先行研究

コンテンツモデレーションはサービス提供者であるデジタルプラットフォームによって行われているが、これが独占禁止法や表現の自由などに深くかかわるとして、賛否様々な意見がある。例えば、米国ではテクノロジー企業とソーシャルメディアが今より大きな責任を持つことを期待する声がある。現に多くのデジタルプラットフォームが存在する米国においてコンテンツモデレーションに大きく影響する通信品位法の改定案は複数回出されている。橋(2022)は、アメリカ大統領ドナルド・トランプの一連の情報配信やプラットフォームの対応を踏まえ、アメリカにおける通信品位法 230 条の分析を行っている。この分析を通じて、プロバイダに義務を課すことは競争法の問題が生じ、政府が関与することはプライバシー及び表現の自由に問題を生じさせる恐れがあるとして、透明性の確保を提案している。また、デジタルプラットフォーム上の情報の性質について、水谷(2022)はアメリカにおいてデジタルプラットフォームがコンテンツモデレートすることを禁止したフロリダ州法に関する分析を行っている。ソーシャルメディアが新聞などのメディアと同様に扱われるべきか、コモンキャリアとして扱われるべきかを挙げ「真実はその中間にある」とフロリダ

北部地区連邦地方裁判所が指摘したことを挙げ、プラットフォーム上に流通するコンテンツは、ニュース編集室からの生成物というよりも、工場で機械的に生成される製品に近いといえると主張している。このように様々な意見があるが、Jialun et al.(2022)は多くのコンテンツモデレーションの研究はある立場から見た課題の評価をするものが多く、本来トレードオフの関係にある事をより意識するべきと指摘する。

### 3. 2. 投稿に関する先行研究

デジタルプラットフォームと利用者の中でのコンテンツをアップロードする側の権利を検討するべきという検討がある。酒井(2024)は、著作権の面から EU の DSMCD やドイツの UrhDaG を参考にオーバーブロッキングのリスクに触れ、日本においても行き過ぎたブロッキングに対して、価値の衡量の面からアップロードする側の利益の救済を考慮すべきと指摘する。

多くのデジタルプラットフォームでは、アップロードされたコンテンツを他の利用者に転送し拡散することが出来る。この拡散に着目した検討として、山口(2015)はインターネット上の炎上の加担者の要因を分析している。この分析によると、性別、年齢、子どもの有無、年収、ラジオの視聴、ソーシャルメディアの利用等が炎上の加担に影響を与える事がみとめられた。一方で学歴やテレビの視聴時間は影響を与える要因とは認められなかった。また、山本(2022)はフェイクニュースの拡散に加担する要因として、情報リテラシーが関係するとしている。これらの先行研究は投稿者の加害者意識によらず、一定の特徴が炎上やフェイクニュースの拡大に寄与していることを明らかにしている。

これらの研究は、直ちに違法とは言えない情報へのコンテンツモデレーションを検討するにあたり有益である。すなわち、前節で見た通り、日本において明らかに違法である情報は事業者が知りえた時点で対処が必要と定められており、一定の法的手当てが行われている。しかし、直ちに違法とは判断しにくい情報については、事業者の自主努力の範疇といえる。直ちに違法とは言い切れない情報の一つとして、炎上コンテンツやフェイクニュースといったコンテンツがあるが、これらをどのように扱うべきかについては明らかではない。本稿では、このような直ちに違法とは言い切れない情報を有害コンテンツとして捉える。ただし、有害コンテンツが存在する事による権利侵害と表現の自由や情報の存在価値について法的整理を目的としていない。あくまで閲覧者と投稿者の両方を含めた利用者の意識や選好を把握した上での検討を目指す。

### 3. 3. 残された課題とリサーチクエスチョン

前節で見たように、プラットフォーム(事業者)にどのような責任を持たせるべきかといった責任の観点での検討や、投稿者の視点でコンテンツの保護や拡散行動を検討したものがある。それでも我が国において政府検討会が続くようにコンテンツモデレーションは多くの論点を残している状況にある。この原因として、我が国のコンテンツモデレーションの実態を捉える為の利用者の選好に関する研究が少ない事が考えられる。

本稿では、利用者のコンテンツモデレーションに関する選好を、先行研究を踏まえて分析にするにあたり、着眼するべき疑問を3点設定する。第一に『直ちに違法とは判断できない情報に対して、コンテンツを見たいか(表示するべきか)、見たくないか(削除するべきか)』

という疑問がある。第二に、ステークホルダーを事業者と閲覧者の関係の中に限定すると、『仮に削除するとしたら誰が判断をするべきか』という疑問がある。第三に、開放的な方針を好む利用者、すなわち直ちに違法とは判断できない情報を消すべきではないと主張する利用者、あるいは消すとしても自分で判断したいと主張する利用者が、『炎上加担者やフェイクニュースの拡大に寄与するものと特徴が一致していないか』という疑問がある。三点目は有害コンテンツの発信者あるいは拡散者などの加担者がコンテンツモデレーションの基準を開放的にするように要望したとしても、有害コンテンツがより拡大する可能性があるため、社会的に受け入れる事は難しいと思われる。利用者の意識や選好の観点からもコンテンツモデレーションの基準についてより開放するべきと考える人々の特徴について踏まえる必要がある。以上3点を本稿におけるリサーチクエスション(以下RQ)とする。

次節ではこの問を解消するための方法を検討する。

### 3. 4. 検討の範囲と方法

直ちに違法とは判断しにくい情報の規制について、内容規制と内容中立規制による検討を行う。内容規制とはコンテンツの内容そのものに対する規制であり、内容中立規制は時間や場所に制約を持たせる考え方である。この際、内容規制と内容中立規制による二分論には、政府が表現規制を考えるとときの有効性について批判(木下 2020)や、戦略的な検討とともに限定的に利用すべきとの指摘(橋本 1992)もある。このような指摘を受け、本稿では主にコンテンツの内容規制に即して検討を進める。

また、前節で設定したRQにアプローチするために、現状の把握を量的側面から検討することにする。現状を量的側面から把握することで、先に行われた論理的検討がなされている先行研究を補完することも出来る。そこで、本稿では利用者の選好を量的に観測し、定量的な実証分析を行う手法で検討する。

### 3. 5. 検討対象のデジタルプラットフォーム

本節では検討対象とするデジタルプラットフォームを定める。RQの3つの疑問を明らかにするには、一つ目の条件として一定の網羅率の確保が必要である。その為には国民の多くが触れているサービスを調査する事が考えられる。もし、そのようなデジタルプラットフォームが無ければ、複数のデジタルプラットフォームを横断して調査する必要がある。二つ目の条件として、誰が判断しどのようなコンテンツが削除されているのかが明らかにされている必要がある。例えば、コンテンツモデレーションのガイドラインが公表されており、利用者が把握出来るデジタルプラットフォームが望ましい。

この2点を満たすデジタルプラットフォームとしてYouTubeがある。YouTubeはいわゆるネットワーク効果を働かせることで、急速にユーザを増やし幅広い年代に利用されている。YouTubeは投稿者にとっては日本にいながら世界中に情報発信することが出来る。また閲覧者にとっては世界中で発信された情報を日本にいながら閲覧することも出来る。このような投稿者と閲覧者の関係から表現の場および情報流通のプラットフォームとして、我が国において高い利用率を維持している。なお、日本で利用可能なメジャーな動画共有サ

ービスがいくつかあるが、その中で YouTube は最も多くの日本人に利用されている<sup>3</sup>。YouTube は特に 10 代から 40 代にかけて 90%を超える高い利用率となっている。また、利用者が多いということは多くの価値観を持つ人が利用することでもある。その為、YouTube 上のある人にとっては価値あるコンテンツであっても、別の人にとっては価値がないこともある。そればかりか、コンテンツによっては不快であると捉えられることがある。もちろん違法なコンテンツは対処が必要であるが、直ちに違法とは判断できないコンテンツにおいても、同様の事が起きる。そこで YouTube の提供・運営者である Google LLC は、利用者に何をどこまで許容するか、逆に何をどこまで規制するかについて、法による規律に加えて YouTube が自主的に利用規約などの規律をコミュニケーションガイドラインとして公開してコンテンツモデレーションを行っている。このようにコンテンツモデレーションの基準と判断者が公表されている点も検討対象としての条件に合致する。国民すべての選好を調査することは現実的でないことを踏まえると、上記理由によってデジタルプラットフォームの代表例として YouTube を調査することは十分に合理的と考えられる。

### 3. 6. 対象とするコンテンツ種類の検討

YouTube には前節でみたコミュニケーションガイドラインに加えて、年齢制限と制限付きモードという、年齢や環境によってより厳しい基準である「年齢制限のあるコンテンツ」で規制する機能がある。制限付きモードは制限をかける端末を管理者が設定することができるが、年齢制限機能はログインしたアカウント毎に年齢確認が行われる。また、年齢制限および制限付きモードの両方とも、コンテンツモデレーションを行うのは YouTube であって、利用者にコンテンツモデレーションをする機能は提供されていない。

利用者の視点で調査を行う際、利用者が実感しやすい基準で調査することが望ましい。通常のコンテンツモデレーションで削除されているコンテンツは、多くの利用者が閲覧できずコンテンツの存在をはじめから認識できていない事が想定される。一方、年齢制限機能の制限は YouTube にログインしていない環境で経験出来る。また、制限付きモードはその環境に入れば誰でも経験が出来る。すなわち、年齢制限機能と制限付きモードに適応されている基準は、利用者にとって通常のコンテンツモデレーションの基準の比較が実感しやすいと思われる。

なお、YouTube ヘルプによると、年齢制限機能に適応されている基準である「年齢制限のあるコンテンツ」には、子どもの安全、有害または危険なアクティビティ(規制されている薬物やドラッグなどを含む)、ヌードや性的なものを暗示するコンテンツ、暴力的で刺激の強いコンテンツ、下品な言葉の 5 種が挙げられている。また、同じく YouTube ヘルプによると制限付きモードで制限されるコンテンツとして、“年齢制限のあるコンテンツ”に加えて、更にいくつかの制限が挙げられているが、年齢制限機能と制限付きモードに共通して適応されている基準は、「年齢制限のあるコンテンツ」である。

---

<sup>3</sup> 例えば、総務省(2024),”令和 5 年度情報通信メディアの利用時間と情報行動に関する調査報告書”p.78-79 や Google による発表 <<https://www.thinkwithgoogle.com/intl/ja-jp/marketing-strategies/video/yt-trendsreport2022/>>

表1 年齢制限のあるコンテンツ

制限カテゴリ	禁止内容
子どもの安全	<ul style="list-style-type: none"> <li>・爆発物の取り扱いや、怪我につながるチャレンジなど、未成年者が簡単に真似できる可能性のある危険なアクティビティに成人が参加している内容を含む動画</li> <li>・成人向けであるが、家族向けと間違えやすい動画</li> </ul>
有害または危険なアクティビティ (規制されている薬物やドラッグなどを含む)	<ul style="list-style-type: none"> <li>・フェイクであっても、あまりにリアルで視聴者には見分けがつかない有害ないたずら動画</li> <li>・大麻の販売所を宣伝する動画</li> </ul>
ヌードや性的なものを暗示するコンテンツ	<ul style="list-style-type: none"> <li>・性行為を誘う動画(挑発的なダンスや愛撫など)</li> <li>・登場人物が視聴者の性的興奮を引き起こすことを意図したポーズをとっている動画</li> <li>・登場人物が公共の場で一般的にふさわしくない服装(下着姿など)をしている動画</li> </ul>
暴力的で刺激の強いコンテンツ	<ul style="list-style-type: none"> <li>・交通事故の生存者のけがを写したコンテンツを含む動画</li> <li>・映画またはビデオゲームの中で最も刺激の強い暴力シーンのみが強調されているなど、暴力や残虐行為の画像を中心とする動画</li> </ul>
下品な言葉	<ul style="list-style-type: none"> <li>・タイトル、サムネイル、関連付けられたメタデータに非常に冒とく的な表現を含む動画</li> <li>・動画集や文脈を無視して抜き出したクリップなど、冒とく的な表現の使用に焦点を当てた動画</li> </ul>

出所：YouTube ヘルプから筆者転記

なお本研究では、実験の制限から3つのカテゴリに絞り調査を行うことにした。前節でみた有害コンテンツとして政府検討の対象となりやすい、フェイクニュースと誹謗中傷に着目し、「年齢制限のあるコンテンツ」のうちフェイクニュースと関連の強い「子どもの安全」、「有害または危険なアクティビティ」を選定し、誹謗中傷と関連の強い「下品な言葉」の3つのカテゴリを調査対象として選定した。

次節では、YouTubeの基準である「子どもの安全」、「有害または危険なアクティビティ」、「下品な言葉」を対象に、我が国の閲覧者の選好を量的側面から把握する。

## 4. 実験

### 4. 1. 調査の概要

本研究では、Web調査会社によるインターネットを通じたアンケート調査を実施した。調査にあたっては、分析のための調査(本調査)の前に、本研究の分析対象であるコンテンツについて説明等が理解されるかを確認するため、事前の調査(予備調査)を実施し、本調査の設計を行った。予備調査の概要はつぎのとおりである。

表2 予備調査の内容

調査期間	2024年9月17日 17:42～ 19:21
調査手法	クラウドソーシング
対象	YouTubeを使ったことのある人
回収数	203件

出所：筆者作成

予備調査では、回答者によって解釈の幅があることが回答に正確性を失わせると考え、対策として回答者には制限対象コンテンツを具体的にイメージして回答させる為に、想定しやすい具体例を提示した。また、明らかに違法であるコンテンツはデジタルプラットフォームに対処する義務があるため、あくまで直ちに違法と判断されない程度の表現行為を調査対象とすることとした。更に、カテゴリによって選好が違うことを想定し、3つのカテゴリそれぞれについて選好を確認する設問にした。これらを踏まえて設問の理解度を確認したが、本調査で想定している設問文の理解度はすべて97.5%以上であったことから、設問文は問題が無いと判断し、本調査に用いることとした。

予備調査を踏まえ、本調査を実施した。本調査の概要はつぎのとおりである。

表3 本調査の内容

調査期間	2025年2月21日～2025年2月27日
調査手法	オンラインアンケート
対象	YouTubeを使ったことのある18歳以上の男女
回収数	2,068件
回収方法	性・年代均等割り付け(性別：男女、年代：20代以下～60代以上)

出所：筆者作成

本調査では、調査項目として大きく「基本属性」、「日常の行動・知識」、「YouTube上の経験・思想」を取り上げた。「基本属性」では、性別、子どもの有無、年齢、年収を確認した。「日常の行動・知識」では、YouTube以外でフェイクニュースを見たことがあるか、フィルターバブルを知っているか、エコーチェンバーを知っているか、日常的に接しているメディアを確認した。「YouTube上の経験・思想」については閲覧頻度、投稿経験の有無、YouTubeでの不快な経験の有無、年齢制限機能および制限付きモードの認知の有無、不快なコンテンツの原因、違法ではないコンテンツに対する反応を確認した。

#### 4. 2. 回答内容の記述統計

本節ではアンケートで回収した回答内容について記述統計を整理する。

表4,表5は子どもの有無及び年収についての結果である。子どもがいると回答した人は42.2%であった。

表4 子どもの有無 (n=2,068)

分類	%
いる	42.2%
いない	57.8%

出所：筆者作成

表5 年収 (n=2,068)

選択肢	%
100万円未満	26.4%
100万円超～200万円以下	12.7%
200万円超～300万円以下	13.7%
300万円超～400万円以下	14.5%
400万円超～500万円以下	9.5%
500万円超～600万円以下	7.2%
600万円超～700万円以下	4.4%
700万円超～800万円以下	3.4%
800万円超～900万円以下	2.3%
900万円超～1,000万円以下	1.8%
1,000万円超～1,500万円以下	2.0%
1,500万円超～2,000万円以下	0.5%
2,000万円超～2,500万円以下	0.3%
2,500万円超～	1.4%

出所：筆者作成

表6はフェイクニュースの閲覧経験についての結果である。YouTubeによらずインターネット上でフェイクニュースを閲覧した経験があるかについては、見たことがあると回答した人は44%である。

表6 フェイクニュースの閲覧経験 (n=2,068)

選択肢	%
ある	44.0%
ない	56.0%

出所：筆者作成

表7はフィルターバブルおよびエコーチェンバーについての認知率についての結果である。フィルターバブルを聞いたことがあると答えた人が11.8%、意味も分かると答えた人が4.7%と合わせて16.4%いた。また、エコーチェンバーについては、聞いたことがあるが11.3%、意味もわかるが5.9%と合わせて17.2%いた。このうち、どちらも聞いたことがある、あるいは意味も分かると回答した人は10.8%だった。

表 7 フィルターバブルとエコーチェンバーの理解度 (n=2, 068)

フィルターバブルの理解度		エコーチェンバーの理解度	
選択肢	%	選択肢	%
意味もわかる	4.7%	意味もわかる	5.9%
聞いたことはあるが、 意味は知らない	11.8%	聞いたことはあるが、 意味は知らない	11.3%
知らない	83.6%	知らない	82.8%

出所：筆者作成

表 8 は日常的に接しているメディアについての結果である。日常的に TV と触れていると回答した人は 83.3%、ラジオが 30.4%、YouTube 以外の SNS は 81.3%、ニュースサイトは 16.0%であった。

表 8 日常的に接しているメディア (n=2, 068 複数選択)

分類	%
TV	83.3%
ラジオ	30.4%
YouTube以外のSNS	81.3%
ニュースサイト	16.0%
雑誌	26.7%
その他	29.7%

出所：筆者作成

表 9 は年齢制限機能および制限付きモードについての結果である。それぞれ知っていると回答したのは年齢制限機能が 44.4%、制限付きモードが 27.6%である。

表 9 年齢制限機能および制限付きモードの認知 (n=2, 068)

年齢制限機能の認知		制限付きモードの認知	
選択肢	%	選択肢	%
知っている	44.4%	知っている	27.6%
知らなかった	55.6%	知らなかった	72.4%

出所：筆者作成

表 10 は不快なコンテンツの原因についての結果である。違法ではないが有害であるコンテンツについて、YouTube と利用者のどちらに原因があると思うかを問う趣旨で、コンテンツをデジタルプラットフォームに乗せる人と YouTube というメディアそのものを比べた

時、YouTube のほうに原因があると思うかを 5 段階で回答させた。結果は、どちらとも言えないが 46.9%となり最も多かった。次いで、まったくそう思わない、どちらかといえばそう思わないと回答したコンテンツを上げる人が原因と考える人が 39.5%となり、どちらかといえばそう思う、強くそう思うと回答した YouTube に原因があると考える人は 13.6%に留まった。

表 10 不快なコンテンツの原因 (n=2, 068)

選択肢	%
まったくそう思わない	11.9%
どちらかといえばそう思わない	27.6%
どちらともいえない	46.9%
どちらかといえばそう思う	9.4%
強くそう思う	4.2%

出所：筆者作成

表 11 は違法ではないコンテンツに対する反応についての結果である。違法ではないが有害とされるコンテンツのうち、子どもの安全、有害で危険なアクティビティ、下品な言葉それぞれのガイドラインで規制されるようなコンテンツについて、表示すべきか削除すべきかを聞いた。回答は児童の安全に関するコンテンツは「表示すべき」が 45.4%、有害で危険な行為に関するコンテンツは「表示すべき」が 31.7%、下品な言葉に関するコンテンツは「表示すべき」が 41.3%となった。

表 11 違法ではないコンテンツに対する反応 (n=2, 068)

選択肢	児童の安全	有害で危険な アクティビティ	下品な言葉
表示すべき	45.4%	31.7%	41.3%
削除すべき	54.6%	68.3%	58.7%

出所：筆者作成

この結果が示すとおり、削除を希望する割合はガイドラインで規制されるコンテンツの種類によってバラつきが見られる。最も削除希望が多い有害で危険なアクティビティでも 7 割に届かない。また、児童の安全や下品なコンテンツでは 6 割未満である。やや乱暴に言えば、どの項目も意見が割れたと評価できる。よって、ガイドラインの示す基準は半数を少し超える程度の利用者の希望は満たしているが、大多数の価値観とあっているとは言い難い状況にあるといえる。

#### 4. 3. コンテンツモデレーションの強弱に関する推計

本節では、第一の RQ である『直ちに違法とは判断できない情報に対して、コンテンツを見たいか(表示すべきか)、見たくないか(削除すべきか)』を検証するための推計を行う。

同時に、第三の RQ である『炎上加担者やフェイクニュースの拡大に寄与するものと特徴が一致していないか』についても検証する。

第一の RQ の検討の前に、まず第三の RQ の比較条件について確認する。

第三の RQ の尺度として先行研究によって明らかになっている炎上加担の要因として、山口(2015)から、性別、年齢、子供、個人年収、ラジオ、ソーシャルメディアを、山本(2022)から情報リテラシーとの相関関係を確認する。まず、山口(2015)の性別、年齢、子供、個人年収、ラジオ、ソーシャルメディアはそれぞれ、本研究における性別、年齢、子どもの有無、年収、ラジオ視聴、SNS 接触と比較を行う。山本(2022)の情報リテラシーは本研究におけるリテラシーと比較を行う。

なお、これらの変数については、本研究の特徴を活かすために、必ずしも計測方法を一致させていない点に注意が必要である。違いについてはまず、山口(2015)において、年齢は年齢そのものを扱っているが、本研究では世代による違いを明らかにするために、10 歳毎の段階で扱う。子供は子どもとの同居の有無によって計測されているが、その後の検討において子持ちの親であることが原因であるとの分析がなされている。よって本研究では子どもの有無を要因として扱う。また、ソーシャルメディアについては、平日におけるソーシャルメディアの利用時間を計測しているが、本研究ではソーシャルメディアの中で炎上問題の対象となる SNS の利用時間を曜日に関わらず 0 か 1 以上かの観点で扱う。これは、アンケートによる利用時間の把握の限界を踏まえたものであり、数値の信ぴょう性による揺らぎを排除するための処置である。ただし、今回の調査対象は全員 YouTube を利用していることから、YouTube 以外の SNS の利用時間を対象とし、SNS 接触とする。ラジオについても平日の視聴時間を計測しているが、本研究ではソーシャルメディアと同様の理由で 0 か 1 以上かの観点で扱う。次に、山本(2022)において情報リテラシーとは、「未知の専門的な概念について、最も信頼でき、検証済みで、簡潔で包括的な説明はどこに記載されていると思いますか」などの 4 つの間に対する正解数を得点として扱っている。本研究では鳥海・山本(2023)において、情報の真偽を判定する際の前提となるフィルターバブルやエコーチェンバーの認知率に着目し、これを情報リテラシーとした。情報リテラシーの項目について、観測範囲が違うことを踏まえて以降の検討では参考値として扱う。

第三の RQ の比較条件を以上のように整理したうえで第一の RQ の検討を始める。第一の RQ について、前節で示したとおり、「基本属性」、「日常の行動・知識」、「YouTube 上の経験・思想」の回答は多様であり、これらについての回答(態度)が「違法ではないコンテンツに対する反応」に影響を与えている可能性がある。そこで、推計では次のようなモデルを考える。また、推計に用いた具体的な変数は表 12 のとおりである。

「違法ではないコンテンツに対する反応」=f(「基本属性」, 「日常の行動・知識」, 「YouTube 上の経験・思想の回答」)

表 12 推計に使用した変数

	本分析上の項目	調査での質問文	回答形式、項目	変数の作成
個人属性	性別	あなたの性別を選んでください	「男、女、答えたくない」の3択	男性を1、女性を0とする2値変数
	子どもの有無	あなたにお子様がいる場合、お子様の年齢層で当てはまる選択肢をすべて選択してください。	「子どもはいない」を含む子どもの就学状況の6択	「子どもはいない」を1、それ以外を0とする2値変数
	二十代、三十代、五十代、六十代	あなたの年齢を選んでください	「10代、20代、30代、40代、50代、60代」の6択	10代か20代を二十代、30代を三十代、50代を五十代、60代を六十代でそれぞれ1とする各年代の2値変数
	年収	あなたの年収を教えてください	100万円未満から1000万円まで100万円単位、1000万円以上2500万円までは500万円単位、2500万円以上の14段階選択肢	「100万円未満」を50、「2500万円以上」を2500とし、それ以外は各選択肢の中央値
日常の行動・知識	フェイクニュース経験	あなたはインターネットで「フェイクニュース」を見たことはありますか？(YouTubeに限らず)	「ある、ない」の2択	「ある」を1、「なし」を0とする2値変数
	リテラシー	あなたは「フィルターバブル」という言葉を知っていますか？	「意味もわかる、聞いたことがあるが意味は知らない、知らない」の3択	フィルターバブルの理解度が「聞いたことがあるが意味は知らない」以上で且つエコーチェンバーの理解度が「聞いたことがあるが意味は知らない」を1、それ以外を0とする2値変数
		あなたは「エコーチェンバー」という言葉を知っていますか？	「意味もわかる、聞いたことがあるが意味は知らない、知らない」の3択	
	SNS接触	あなたはYouTube以外のインターネットメディアを毎日どれくらい使いますか？	「YouTube以外のSNS」の利用時間(分)	1分以上を1、0を0とする2値変数
	ニュースサイト接触		「ニュースサイト」の利用時間(分)	1分以上を1、0を0とする2値変数
	TV視聴	インターネット以外で日常的に触れるメディアを選んでください	「新聞、TV、ラジオ、雑誌、その他」の5択	TVの選択者を1、非選択者を0とする2値変数
ラジオ視聴			ラジオの選択者を1、非選択者を0とする2値変数	
YouTube上の経験や思想	閲覧頻度	あなたのYouTubeの閲覧頻度として最も近いものを選んでください。(1回の視聴時間は問いません)	「毎日見る～1か月に1回未満」の5択	「毎日見る」を1、それ以外を0とする2値変数
	投稿経験	あなたはYouTubeに 動画(ショート等を含む) 投稿したことはありますか？	「ある、ない」の2択	どちらかの質問で「ある」の選択者を1、どちらも「ない」の選択者を0とする2値変数
		あなたはYouTubeにコメント投稿したことはありますか？	「ある、ない」の2択	
	不快経験	あなたはYouTubeで不快な思いをしたことがありますか？あてはまるもの全てを選んでください。	「特にない」を含む閲覧箇所の5択	「特にない」を0、それ以外を1とする2値変数
	年齢制限機能認知	あなたはYouTubeの「年齢制限」の機能を知っていましたか？	「知っている、知らなかった」の2択	「知っている」を1、「知らなかった」を0とする2値変数
責任主体	あなたはコンテンツを上げる人とYouTubeを比べた時、YouTubeに原因があると思いますか？	「まったくそう思わない～強くそう思う」の5択	「どちらかというと思う」か「強くそう思う」を1、それ以外を0とする2値変数	

出所：筆者作成

モデルの推計にあたっては、被説明変数に当たる「違法ではないコンテンツに対する反応」が「表示するべき」(0)か「削除するべき」(1)の2値であることから、2項ロジスティック回帰による推計を採用する。結果は表13のとおりである。

表 13 違法ではないコンテンツに対する反応の回帰分析の結果

変数	子どもの安全		有害で危険なアクティビティ		下品な言葉	
	係数	P値	係数	P値	係数	P値
切片	0.330	0.132	0.549	0.020 *	0.585	0.009 **
性別	0.400	0.000 ***	0.656	0.000 ***	0.530	0.000 ***
子どもの有無	-0.259	0.012 *	-0.253	0.024 *	-0.336	0.001 **
二十代	0.191	0.214	0.220	0.174	0.095	0.543
三十代	0.134	0.357	0.336	0.032 *	0.111	0.455
五十代	0.105	0.469	0.233	0.135	0.078	0.596
六十代	0.236	0.115	0.354	0.031 *	0.361	0.021 *
年収	0.000	0.695	0.000	0.948	0.000	0.785
閲覧頻度	-0.411	0.000 ***	-0.404	0.000 ***	-0.394	0.000 ***
不快経験	0.206	0.048 *	0.345	0.002 **	0.065	0.539
投稿経験	-0.124	0.307	-0.310	0.014 *	-0.049	0.690
年齢制限機能認知	-0.336	0.001 **	-0.303	0.006 **	-0.414	0.000 ***
責任主体	0.288	0.035 *	0.516	0.001 **	0.484	0.001 ***
フェイクニュース経験	-0.232	0.028 *	-0.160	0.159	-0.252	0.019 *
リテラシー	-0.274	0.076	-0.449	0.004 **	-0.159	0.304
SNS接触	-0.204	0.118	-0.133	0.351	-0.244	0.072
ニュースサイト接触	0.215	0.113	0.054	0.716	0.085	0.540
TV視聴	0.099	0.435	0.309	0.020 *	0.203	0.116
ラジオ視聴	0.010	0.920	0.026	0.817	-0.046	0.663

\*\*\*は 0.1%、\*\*は 1%、\*は 5%水準で有意であることを示す

出所：筆者作成

表 13 の推計結果について 5%の有意水準で有意になった変数を確認する。児童の安全についてみると、性別、不快経験、責任主体がプラスに有意、子どもの有無、閲覧頻度、年齢制限機能認知、フェイクニュース経験がマイナスに有意となった。有害で危険なアクティビティでは、性別、三十代、六十代、不快経験、責任主体、TV 視聴がプラスに有意、子どもの有無、閲覧頻度、投稿経験、年齢制限機能認知、リテラシーがマイナスに有意となった。下品な言葉では、性別、六十代、責任主体がプラスに有意、子どもの有無、閲覧頻度、年齢制限機能認知、フェイクニュース経験がマイナスに有意となった。

この結果から、テーマ(コンテンツの種類)によらない特徴がみられた。まず、閲覧頻度や不快経験、年齢制限機能認知、責任主体など、YouTube 上の経験や思想の多くが有意になっていることがあげられる。基本属性の中では性別(プラス)と子どもの有無(マイナス)が共通して有意となっている。また、一方で SNS 接触、ニュースサイト接触、ラジオ視聴といった、他のメディアの視聴の多くは有意になっていないこともあげられる。これらの事から、YouTube 上のコンテンツを毎日見るヘビーユーザは、出来るだけ多くのコンテンツを閲覧する事を希望しており、違法でないコンテンツは一旦表示するべきと回答する傾向があると解釈できる。また、年齢制限機能でフィルターをかけられる事を知っているユーザは、そのような機能でより強い規制が掛けられるのであれば、通常は違法でないコンテンツを表示するべきと回答する傾向があると解釈できる。他方、違法コンテンツが YouTube 上に存在する事は YouTube が原因と考えるユーザは削除を求める傾向にあると言える。

テーマごとの傾向もみられる。有害で危険な行為と下品な言葉使いでは六十代がプラス

に有意となったが児童の安全では有意にならなかった。児童の安全に関する規制は世代に対するバイアスがなく、閲覧頻度などの YouTube 上の経験や思想の影響を受けていると思われる。一方、有害で危険な行為と下品な言葉使いでは、世代によってコンテンツの印象が変わっていることがうかがえる。有害で危険な行為に対しては TV 視聴(プラス)とリテラシー(マイナス)が有意になっている。有害で危険な行為に含まれるフェイクニュースに対して、TV と接触の多い層は信頼のおける情報に触れたいと考える傾向にあり、リテラシーの高い層は自ら防衛できると考えて表示するべきと考える傾向にあると解釈できる。炎上加担との特徴の比較では、性別および子どもの有無が有意となったが、年収やラジオの視聴、SNS 利用は有意にならなかった。先行研究によって明らかになっている炎上加担の要因である性別、年齢、子供、個人年収、ラジオ、ソーシャルメディア、情報リテラシーといった特徴のうち、性別は当てはまるものの、子どもの有無に関してはむしろ真逆の反応が示され、その他の要因は当てはまらない結果となった。よって、炎上加担の要因と有害コンテンツの削除を希望するものは完全に一致しているとは言い難い。

#### 4. 4. コンテンツモデレーションの主体に関する推計

本節では、第二の RQ である『仮に削除するとしたら誰が判断をするべきか』を検証する。同時に、第三の RQ である『炎上加担者やフェイクニュースの拡大に寄与するものと特徴が一致していないか』についても検証する。

表 11 でみた「違法ではないコンテンツに対する反応」で、「削除するべき」を選んだ人に対して削除判断の主体者を確認するために、YouTube に判断を任せるか、自分で判断するべきかを聞いた。表 14 は削除判断の主体者についての結果である。表 14 でみた「児童の安全」では、「削除するべき」を選んだものが 1,130 あったが、そのうち 54.1%が「自分で判断するべき」と回答した。「有害で危険なアクティビティ」では、「削除するべき」を選んだものが 1,412 あったが、そのうち 49.0%が「自分で判断するべき」と回答した。「下品な言葉」では、「削除するべき」を選んだものが 1,214 あったが、そのうち 51.7%が「自分で判断するべき」と回答した。

表 14 削除判断の主体者の結果

選択肢	児童の安全 (n=1,130)	有害で危険な アクティビティ (n=1,412)	下品な言葉 (n=1,214)
このままYouTubeに判断を任せるべき	45.9%	51.0%	48.3%
自分で判断するべき	54.1%	49.0%	51.7%

出所：筆者作成

結果が示すとおり、誰が判断するべきかという質問に対しては概ね半々に意見が割れた。つまり、事業者が判断している現在の状態は半数程度が納得しているものの、残りの半数の価値観とは合っていないと言える。

更に、「基本属性」、「日常の行動・知識」、「YouTube 上の経験・思想」の影響を確認し、同時に炎上加担者との関係を確認するために推計を行う。

推計では次のようなモデルを考える。また、推計に用いた具体的な変数は表 12 のとおりである。

「削除判断の主体者」=f(「基本属性」, 「日常の行動・知識」, 「YouTube 上の経験・思想の回答」)

モデルの推計にあたっては、被説明変数に当たる「削除判断の主体者」が「このまま YouTube に判断を任せるべき」(0)か「自分で判断するべき」(1)の 2 値であることから、2 項ロジスティック回帰による推計を採用する。この際、使った項目は先の表 15 と同じものを使った。結果は表 15 のとおりである。

表 15 削除判断の主体者に対する 2 項ロジスティックの結果

変数	子どもの安全		有害で危険なアクティビティ		下品な言葉	
	係数	P値	係数	P値	係数	P値
切片	1.061	0.000 ***	1.006	0.000 ***	0.723	0.011 *
性別	-0.196	0.147	-0.235	0.051	-0.204	0.115
子どもの有無	0.006	0.967	0.039	0.746	0.036	0.780
二十代	0.200	0.357	0.182	0.354	0.154	0.466
三十代	-0.008	0.966	0.080	0.655	0.235	0.225
五十代	-0.205	0.288	0.086	0.622	0.249	0.181
六十代	-0.030	0.874	0.134	0.443	0.061	0.740
年収	0.000	0.201	0.000	0.084	0.000	0.005 **
閲覧頻度	-0.048	0.723	-0.054	0.653	-0.130	0.311
不快経験	0.078	0.582	0.219	0.084	0.047	0.731
投稿経験	-0.297	0.086	-0.149	0.338	-0.129	0.439
年齢制限機能認知	-0.201	0.156	-0.308	0.015 *	-0.336	0.014 *
責任主体	0.177	0.307	0.137	0.374	0.184	0.263
フェイクニュース経験	-0.176	0.212	-0.103	0.412	0.167	0.228
リテラシー	0.026	0.908	0.100	0.617	-0.121	0.569
SNS接触	-0.021	0.899	-0.029	0.845	0.186	0.247
ニュースサイト接触	-0.437	0.019 *	-0.630	0.000 ***	-0.623	0.000 ***
TV視聴	-0.373	0.037 *	-0.591	0.000 ***	-0.442	0.012 *
ラジオ視聴	0.043	0.756	0.100	0.414	0.009	0.945

\*\*\*は 0.1%、\*\*は 1%、\*は 5%水準で有意であることを示す

出所：筆者作成

表 15 の推計結果について 5%の有意水準で有意になった変数を確認する。

子どもの安全ではニュースサイト接触と TV 視聴がマイナスに有意となった。有害で危険なアクティビティでは、年齢制限機能認知とニュースサイト接触、TV 視聴がマイナスに有意となった。下品な言葉では、年収がプラスに有意となり、年齢制限機能認知、ニュースサイト接触、TV 視聴がマイナスに有意となった。

テーマによる違いがないものとして、ニュースサイト接触と TV 視聴が挙げられる。ニュースサイトも TV も編集者がコンテンツを編集しているという点で、利用者は受動的に情報

を扱うことが出来る。このようなメディアとの接触経験がある事で、SNS にも同様の期待をかける傾向がうかがえる。なお、リテラシーはいずれも有意とならなかった。次に、テーマによる違いとして、下品な言葉では、年収が有意となっている点があげられる。誹謗中傷につながる下品な言葉では、年収が影響している可能性を示唆している。

これらの結果から炎上加担者の特徴である性別、年齢、子供、個人年収、ラジオ、ソーシャルメディア、情報リテラシーとの共通性は、下品な言葉に限り年収が該当するもののほとんど共通性が見られない。

また、この分析では年代や性別、子どもの有無による違いは見られなかった。むしろ、他のメディアとの接触に有意な要素が見られたことが特徴として挙げられる。

## 5. おわりに

本研究では、直ちに違法とは言い切れないが一般的には好ましくないような情報に対するコンテンツモデレーションに焦点を当て、その在り方を利用者の選好の観点から分析した。具体的には、YouTube を分析対象とし、「年齢制限のあるコンテンツ」のうち「子どもの安全」「有害または危険なアクティビティ」「下品な言葉」の3つのテーマを取り上げ、利用者がどの程度表示・削除を望み、誰に削除権限を委ねたいと考えているのかを量的に検証した。

検証では、はじめにデジタルプラットフォームが行うコンテンツモデレーションに対する、政府で行われている検討や法的な研究を確認した。次に先行研究を踏まえて、『直ちに違法とは判断できない情報に対して、コンテンツを見たいか(表示するべきか)、見たくないか(削除するべきか)』、『仮に削除するとしたら誰が判断をするべきか』、『炎上加担者やフェイクニュースの拡大に寄与するものと特徴が一致していないか』の3つのRQを設定した。これらのRQに量的側面からアプローチするために、アンケート調査を実施し、2つの推計による分析を行った。

分析結果から、『直ちに違法とは判断できない情報に対して、コンテンツを見たいか(表示するべきか)、見たくないか(削除するべきか)』については、概ね半分に意見が割れていることが明らかとなった。さらに、基本属性やYouTube上の経験や思想が利用者の態度に影響を与えていた事が明らかになった。

また、『仮に削除するとしたら誰が判断をするべきか』についても、YouTubeが判断すべきという意見と自分で判断すべきという意見がほぼ半分に割れている実態が明らかとなった。加えて、これらの意見を持つ要因について、普段のニュースとの接し方が影響を与えている事が明らかになった。

『炎上加担者やフェイクニュースの拡大に寄与するものと特徴が一致していないか』については、コンテンツを消すべきではない、あるいは消すとしても自分で判断したいと主張するものと、一部共通の項目で有意な項目はあったものの、特徴が一致しているとは言い難いという結果となった。

本研究から得られた知見を踏まえると、つぎのような政策的方向性、また提言を示すことができる。

違法ではないコンテンツを消すか消さないかというような単純な2項対立で考えると、コンセンサスを得られている部分は少なく、現状では大勢を納得させるような解を持つこ

とは極めて難しいといえる。違法ではないが好ましくないコンテンツを一律で削除する、あるいは削除しないことのみを推進するような施策ではなく、テーマや利用者の基本属性、経験などを踏まえた柔軟な施策にすることで意見の対立を生じさせにくくなると考えられる。例えば、単純な炎上やフェイクニュースの対策として、より広いコンテンツを YouTube が削除するような対策は、拡大加担者と拡大防止策が一致していない為、利用者のコンセンサスは得られにくいと考えられる。

好ましい対策としては次のように考えられる。本研究でみたように違法ではないコンテンツに対する反応は、子どもの安全という分野に関して世代に影響を受けず、むしろ YouTube 上の経験や思想に影響を受けることから、年代よりも人々の経験を考慮した対策を行う施策が好まれると考えられる。また、利用者の経験に注目すれば、閲覧頻度の高い人は何らかの情報を探している可能性が高く、より多くの情報を求めがちであるため、開放的な方針が好まれると考えられる。逆に不快な経験をした人は、より安全なコンテンツだけを求める傾向があり、より保守的な政策を好むと考えられる。また、削除判断の主体に関しては、メディアの接触に影響を受けている。メディア側の編集が強く働く TV やニュースサイトなどとの接触が高い人は、YouTube に責任主体を求める傾向がある。このように、テーマ、削除主体、利用者の属性や経験など、細かく検討範囲を区切ったうえで、丁寧に検討する事が好まれる政策を生み出す可能性が高い。

本研究で得られた知見を総合的に見た点からの示唆として、基本属性に限らず、経験や思想などの観点を踏まえた様々な意見が必要となる。そうした背景を理解するためにも、マルチステークホルダーによる議論を進める事が適切であると思われる。また、それらの議論の透明性を高め、利用者に対して一定のコンセンサスを取ることから始めるべきではないか。

最後に本研究の分析上の課題を挙げる。本研究では、国民の多くが利用している YouTube を代表例として扱ったが、他のデジタルプラットフォームについてもその影響を検討する余地は残されている。また、特定時期の観測データであり、意識の変化については対応ができていない。経年変化について今後継続的な調査を行うことで明らかにすることが出来るだろう。

## 参考文献

1. 木下智史(2021),“表現内容規制・表現内容中立規制二分論の現在”, 立命館法学 393/394 号, p.255-277
2. 木下昌彦(2021),“デジタル・メディア・プラットフォームの憲法理”, 情報法制研究 2021 年 9 巻, p.16-33
3. 酒井麻千子(2024),“著作権侵害に対するコンテンツ・モデレーションのあり方”, 情報通信政策研究 7 巻 2 号, p.29-45
4. 総務省(2020),“プラットフォームサービスに関する研究会 最終報告書”
5. 総務省(2024),“令和 5 年度情報通信メディアの利用時間と情報行動に関する調査報告書”
6. 総務省(2014),“プラットフォームサービスに関する研究会 第三次とりまとめ”, p.45
7. 橘雄介(2022),“米国通信品位法 230 条の動向とプロバイダ責任のあり方への示唆”, 情報通信学会誌 39 巻 4 号, p.119-126

8. 鳥海不二夫・山本龍彦(2023),“共同提言『健全な言論プラットフォームに向けて—デジタル・ダイエット宣言 ver.2.0』”, KGRJ Working Papers
9. 橋本基弘(1992),“表現内容規制・内容中立的規制二元論”, 高知女子大学紀要 40 巻
10. 水谷瑛嗣郎(2022),“ソーシャルメディア・プラットフォームのコンテンツ・モデレーションと「表現の自由」”,メディア・コミュニケーション: 慶應義塾大学メディア・コミュニケーション研究所紀要 No72, p.27-40
11. 三菱総合研究所(2022),“インターネット上の誹謗中傷情報の流通実態に関するアンケート調査結果”
12. 山口真一(2015),“実証分析による炎上の実態と炎上加担者属性の検証”, 情報通信学会誌 33 巻 2 号, p.53-65
13. 山本明(2022),“誤った情報の拡散とメディア・リテラシー”, メディア・コミュニケーション: 慶應義塾大学メディア・コミュニケーション研究所紀要 No72,p.117-131
14. 渡邊涼介・梅本大祐・今村敏(2021),“デジタルプラットフォームの法律問題と実務” 青林書院
15. Democratic Party(2020),“2020 Democratic Party Platform” p.80
16. ITU(2022), “The State of Broadband 2022: Accelerating broadband for new realities”, p.2
17. Jialun Aaron Jiang, Peipei Nie, Jed R. Brubaker, Casey Fiesler(2023)“A Trade-off-centered Framework of Content Moderation”, ACM Transactions on Computer-Human Interaction, Volume 30

<Web サイト等>

18. European Commission(2022), “The Digital Services Act package”, <<https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>> (2025.05.04 確認)
19. Google(2023),”YouTube をマーケティングのヒントに — トrendが生まれ、マルチフォーマットなどで生活に定着進む”,<<https://www.thinkwithgoogle.com/intl/ja-jp/marketing-strategies/video/yt-trendsreport2022/>> (2025.05.04 確認)
20. Google 検索ヘルプ, <<https://support.google.com/websearch/answer/7315336>> (2025.05.04 確認)
21. The Santa Clara Principles(2021),”The Santa Clara Principles 2.0”, <<https://santaclaraprinciples.org/>> (2025.05.04 確認)
22. YouTube ヘルプ,<<https://support.google.com/youtube/answer/2802167>> (2024.11.16 確認)
23. YouTube ヘルプ,<<https://support.google.com/youtube/answer/7354993?sjid=17272901086248713581-AP>>(2024.11.16 確認)

〔付記〕

本研究は、JST-RISTEX-JPMJRS23L1「可視化によるトラスト形成: パーソナライズされたデジタル情報空間のリテラシー教育」および静岡大学『高柳記念未来技術創造基金支援事業』『AI 時代のデータ利活用に資する制度研究拠点形成事業』による支援を受けたものである。

(掲載決定日: 令和7年11月28日)