

デジタル空間における情報流通に関する現状と課題 参考資料

2023年11月7日

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
事務局

目次

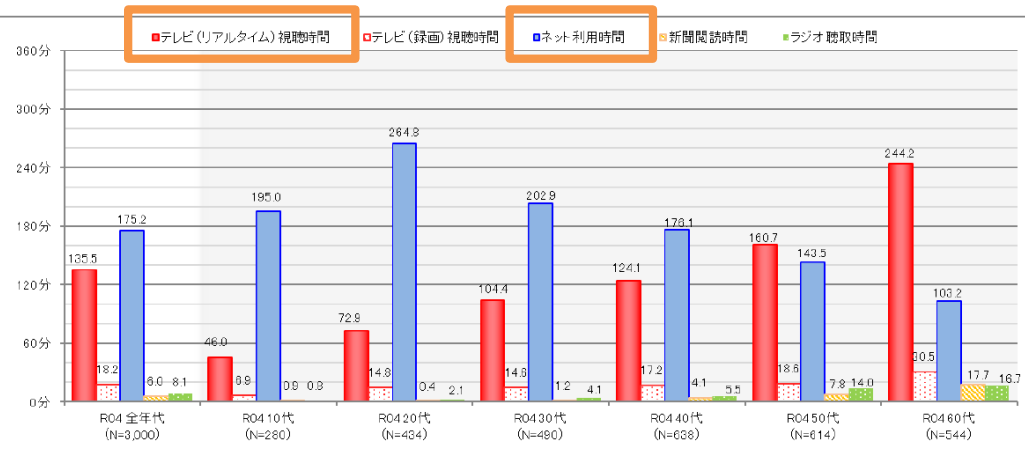
1. デジタル空間における情報流通の現状 ……P 2
2. 「プラットフォームサービスに関する研究会」における
インターネット上の偽・誤情報対策に関する主な議論 ……P16
3. プラットフォーム事業者等幅広い関係者による自主的な対応状況 ……P22
4. 総務省による主な取組状況（上記2. 以外の取組） ……P34
5. AI（特に生成AI）を巡る主な議論 ……P49
6. 諸外国の動向 ……P60
7. その他参考 ……P70

1. デジタル空間における情報流通の現状

インターネットやソーシャルメディアの国民への浸透

◆ 主なメディアの平均利用時間（平日 1日）

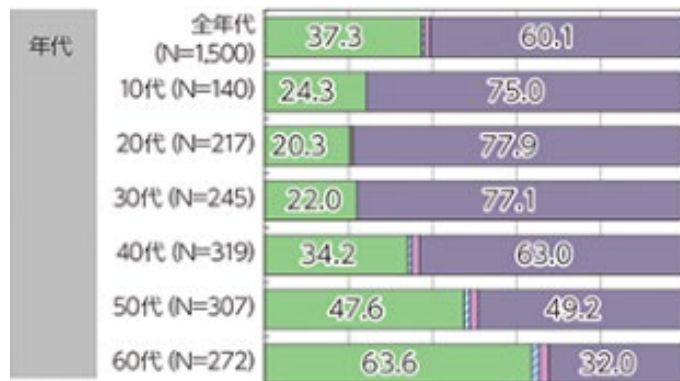
- 「50～60代」ではテレビ（リアルタイム）視聴が主流である一方、「全年代」「10代～40代」ではネットに費やす時間が最も長い。



【出典】令和4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（総務省）

◆ いち早く世の中のできごとや動きを知るために最も利用するメディア

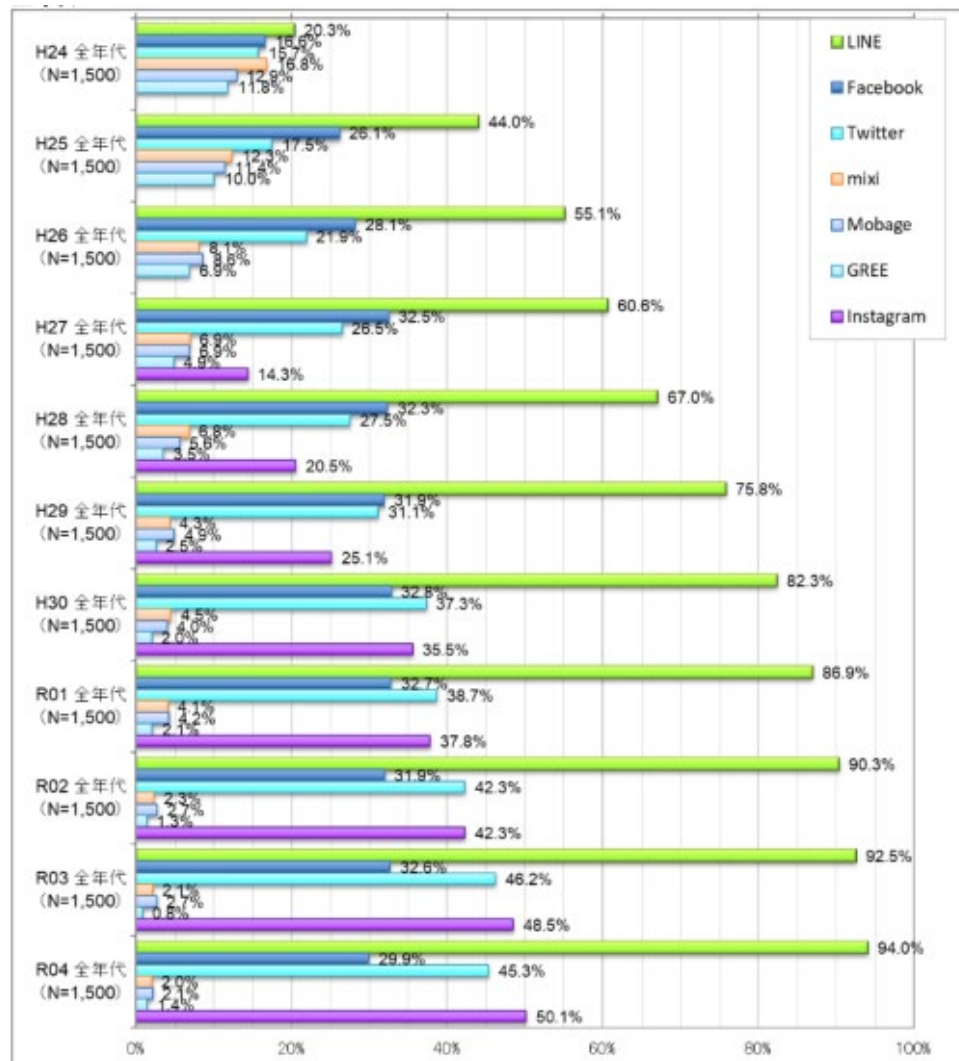
- 「全年代」「10代～50代」で最も利用するメディアはインターネット。
- 「60代」で最も利用するメディアはテレビ。



【出典】令和5年版情報通信白書（総務省）

◆ 主なソーシャルメディア系サービス／アプリ等の利用率

- LINE、Instagramは一貫して増加。LINEは90%超。SNSが国民の日常に深く浸透。



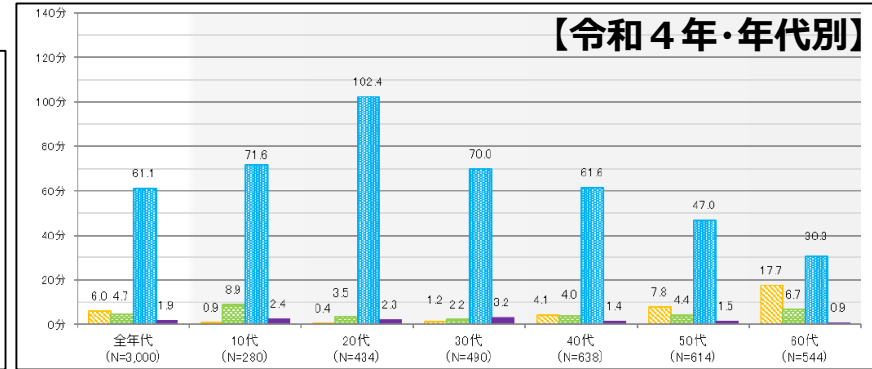
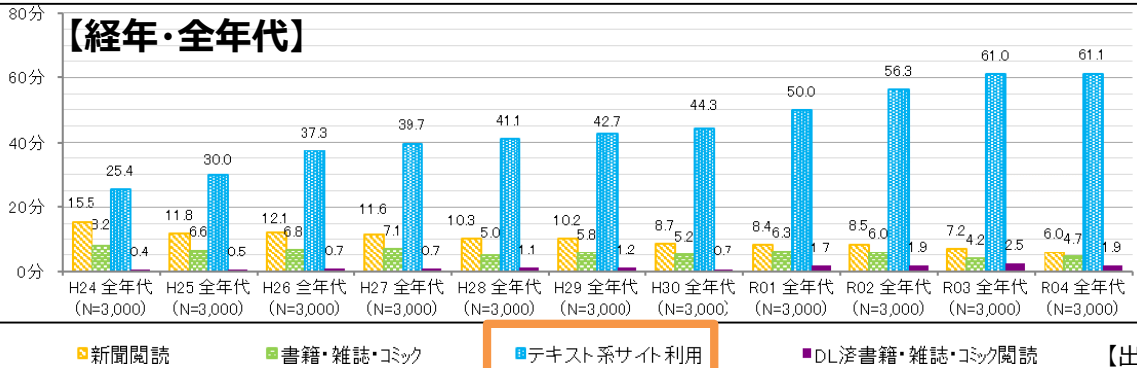
【出典】令和4年度情報通信メディアの利用時間と情報行動に関する調査報告書（総務省） 3

ソーシャルメディアを利用した情報入手の拡大

◆ テキスト系メディア※の平均利用時間（平日 1 日）

※ブログやウェブサイトに加え、ソーシャルメディアの利用も含む。

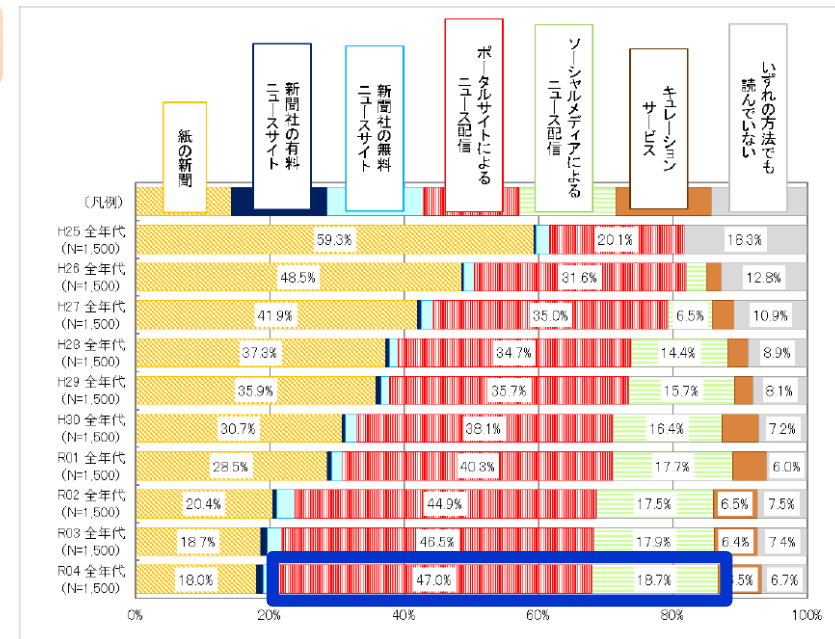
- ・テキスト系サイト利用に費やす時間が最も長い傾向が10年以上継続。次いで新聞閲読、書籍・雑誌・コミック。
- ・年代別に見ても、各年代でテキスト系サイト利用に費やす時間が最も長い。



【出典】令和4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（総務省）

◆ 最も利用しているテキスト系ニュースサービス

- ・「ポータルサイトによるニュース配信」と「ソーシャルメディアによるニュース配信」を合計した割合は一貫して増加する一方、「紙の新聞」の割合は一貫して減少。
- ・令和4年には「ポータルサイトによるニュース配信」と「ソーシャルメディアによるニュース配信」を合計した割合（65.7%）が「紙の新聞」の割合（18.0%）の3.5倍超。



【出典】令和4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（総務省）

「アテンションエコノミー」の広がり

- インターネット上では膨大な情報が流通する一方、人々の消費時間が有限である結果、事実に基づく正確な情報ばかりではなく、**過激なタイトル・内容や、憶測だけで作成された事実に基づかない情報等によって利用者を刺激し、より多くのアテンション（注目・関心）を集めて金銭的対価を得る「アテンションエコノミー」が拡大している**と指摘。

インターネットとソーシャルメディアにより、我々は絶えず情報を浴びせられると同時に、情報を発信する存在ともなった。そこでは、情報の消費と共有とが果てしなく繰り返され、情報に飢えることは、ない。

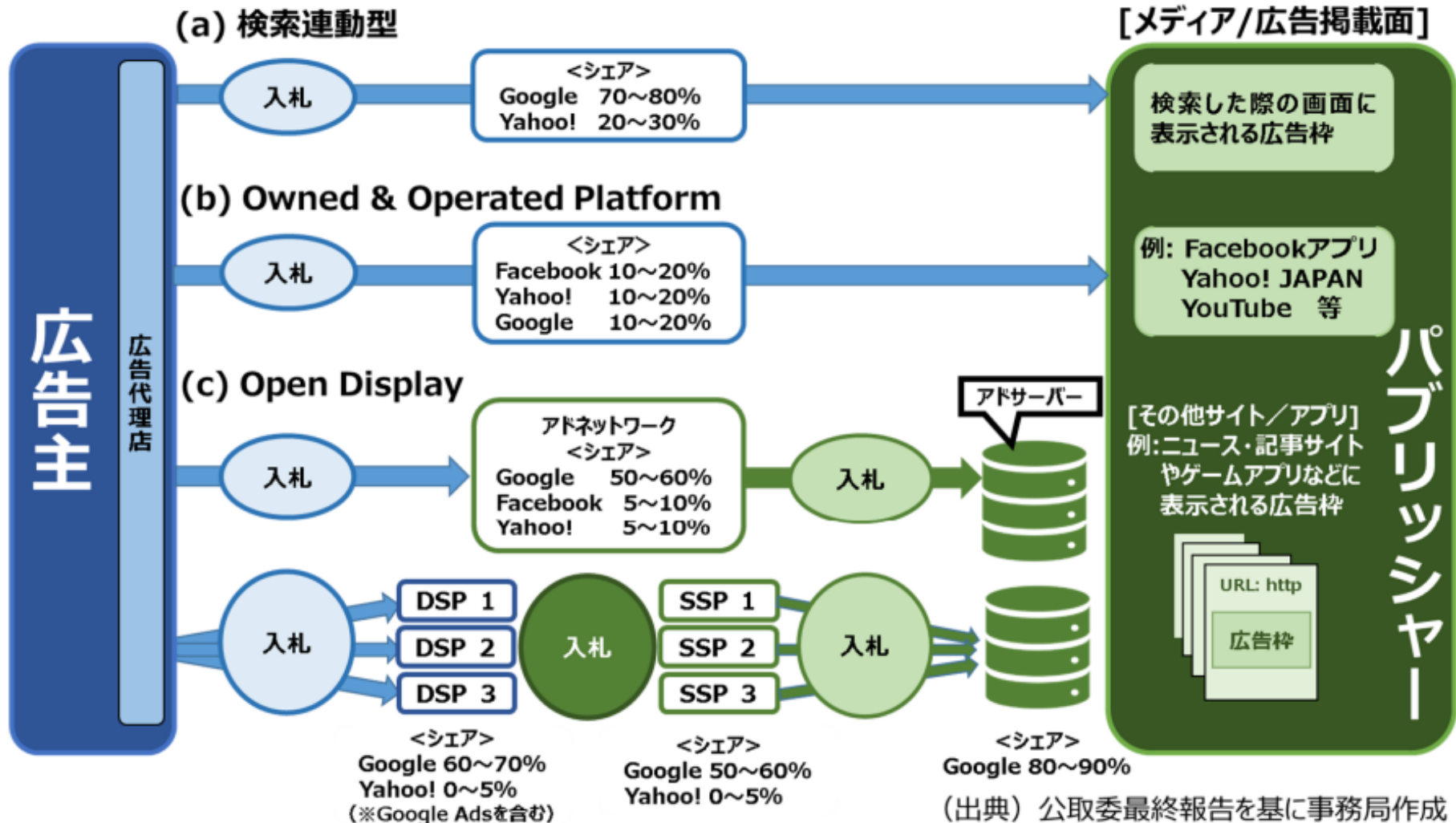
こうした情報過多の社会においては、供給される情報量に比して、我々が支払えるアテンションないし消費時間が希少となるため、それらが経済的価値を持って市場（アテンション・マーケット）で流通するようになる。**こうした経済モデルは、一般に「アテンション・エコノミー」と呼ばれる。**スマートフォンなどのモバイルデバイスが日々の生活に深く浸透したことで、我々の「身体」はますますアテンション・エコノミーに支配されるようになってくる。

心理学では、人間の思考モードを、①直観的で自動的な「システム1」というモードと、②「システム1」を補完し、熟慮を特徴とする「システム2」というモードに分けることが提案されている（二重過程理論）。**アテンション・エコノミーの世界では、前者、すなわち「システム1」を刺激することが重要だと言われる。「刺激」を与えることで得られたユーザーの脊髄反射的な反応が経済的な価値を有するからである**（こうした「反射」は、PV〔閲覧数〕や、ウェブページへの滞在時間や粘着性などにより評価され、経済的価値に換算される）。**その結果、事実に基づく正確な情報よりも、より多くの「刺激」を与えるコンテンツ（フェイクニュースを含む）の方が、数値化された客観的評価を受けやすく、広告収入などの金銭的メリットさえ享受しやすい状況**にある。

鳥海不二夫 = 山本龍彦 共同提言「健全な言論プラットフォームに向けてver2.0 – 情報的健康を、実装へ」（2023年5月）

デジタル広告市場の成長と課題

- デジタル広告費は年々増加。2020年には日本の広告費全体の約36%を占めるまでに成長。
- デジタル市場のマネタイズのためのインフラとして大きく成功したが故に、「アテンションエコノミー」の発生・拡大につながり、フィルターバブル、偽・誤情報の流通等、デジタル社会における新たな課題をもたらしていると指摘。



「ダークパターン」による消費者選択への影響の懸念

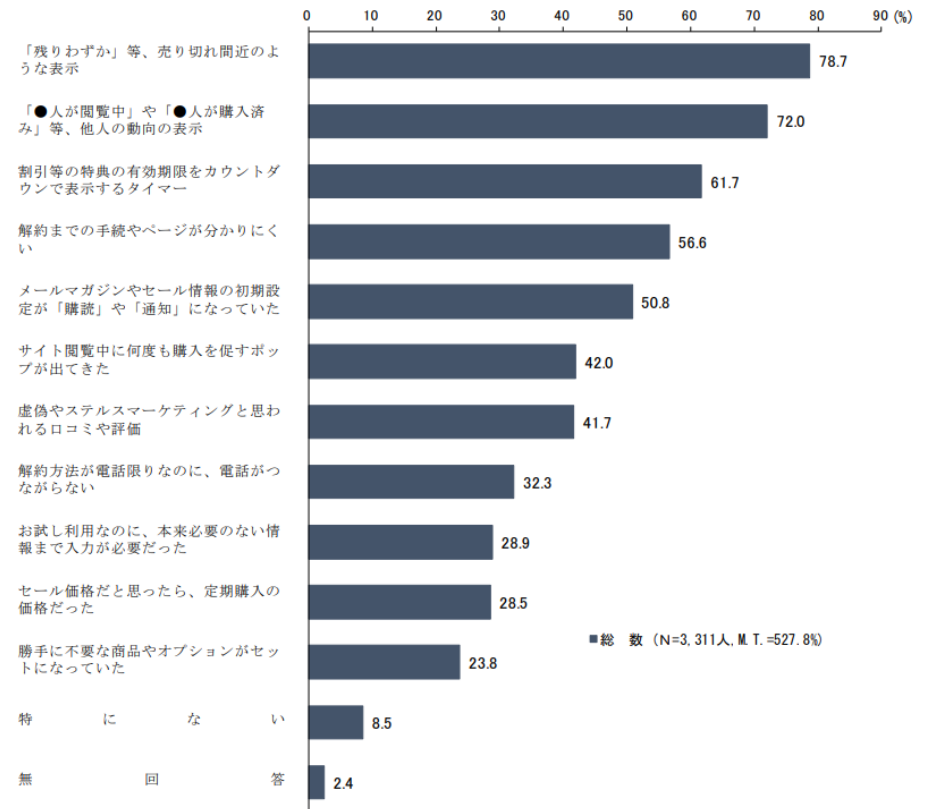
- 「ダークパターン」（オンラインユーザーインターフェース上で、消費者の自主性・意思決定・選択を損ない、又は弱めるように設計されたアーキテクチャを採用する商慣習）が、消費者に多大な被害を生じさせる可能性が指摘。

◆ ダークパターンの分類の例

| 分類 | 例 |
|---|--------------------------|
| 行為の強制 <i>Forced Action</i> | 望ましい範囲を超えた個人情報の開示の強制 |
| インターフェース干渉 <i>Interface interference</i> | 事業者にとって好都合な選択肢を視覚的に目立たせる |
| 執拗な繰り返し <i>Nagging</i> | 事業者にとって好都合な設定変更を繰り返し要請 |
| 妨害 <i>Obstruction</i> | サービスのキャンセルを困難にする |
| こっそり <i>Sneaking</i> | 取引の最終段階において料金を追加 |
| 社会的証明 <i>Social proof</i> | 他の消費者の購入行為を通知 |
| 緊急性 <i>Urgency</i> | オファー期限を知らせるカウントダウンタイマー |

【出典】 Dark Commercial Patterns, OECD Digital Economy Papers, No. 336 (2022)

◆ インターネットでの予約や購入で実際に目にしたり経験したもの



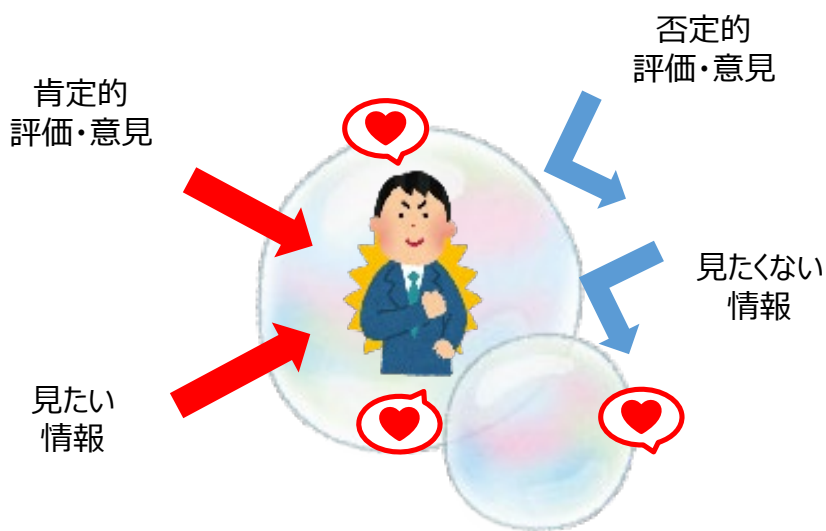
【出典】 令和4年度消費者意識基本調査（消費者庁）

アルゴリズム機能等をもたらす「フィルターバブル」と「エコーチェンバー」

- SNSの利用の進展に伴い、投稿・クリック履歴をもとに表示内容を決定するアルゴリズム等の影響により、「フィルターバブル」「エコーチェンバー」と呼ばれる現象が生じているとの指摘。
- これにより、人々は多様な情報を受信できず、**自らが「バブル」や「チェンバー」の中にいることすら気づかないまま、適切な判断を下すことが困難**となり、結果として**社会経済の混乱や民主主義への悪影響をもたらす可能性**が指摘。

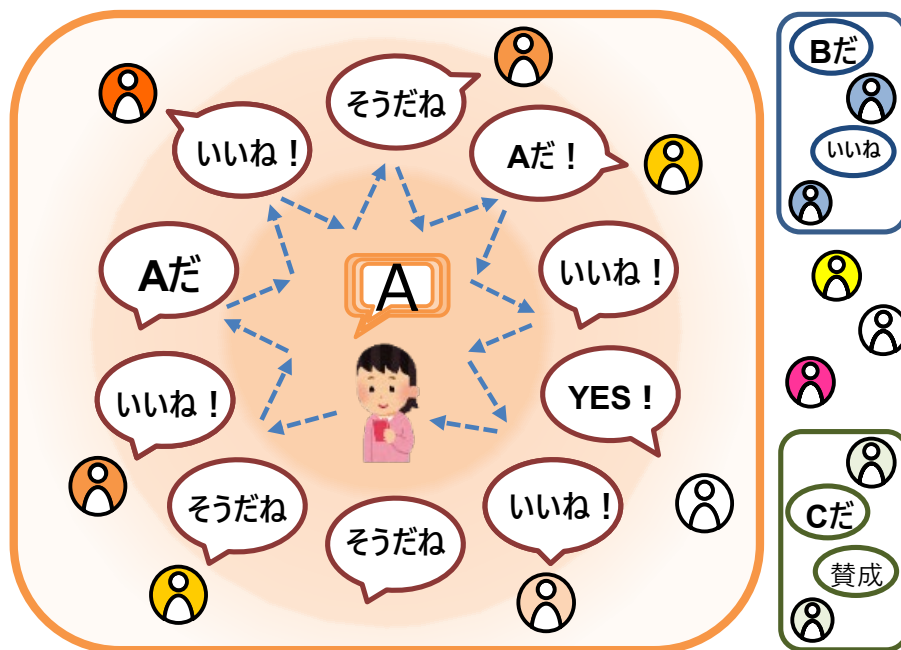
【フィルターバブル(Filter Bubble)】

アルゴリズムによって、利用者個人のクリック履歴に基づく情報（コンテンツのレコメンデーションや行動ターゲティング広告）が優先的に表示される（関心に合わない情報からは隔離される）結果、**自身の考え方や価値観に近い情報ばかりの「バブル（泡）」の中に包まれる状態**



【エコーチェンバー(Echo Chamber : こだまの部屋)】

ソーシャルメディア等において、自分と似た興味関心を持つユーザが集まる場でコミュニケーションする結果、**自分が発信した意見に似た意見が返ってくる(特定の意見や思想が増幅)状態**を、閉じた小部屋で音が反響する物理現象に例えたもの

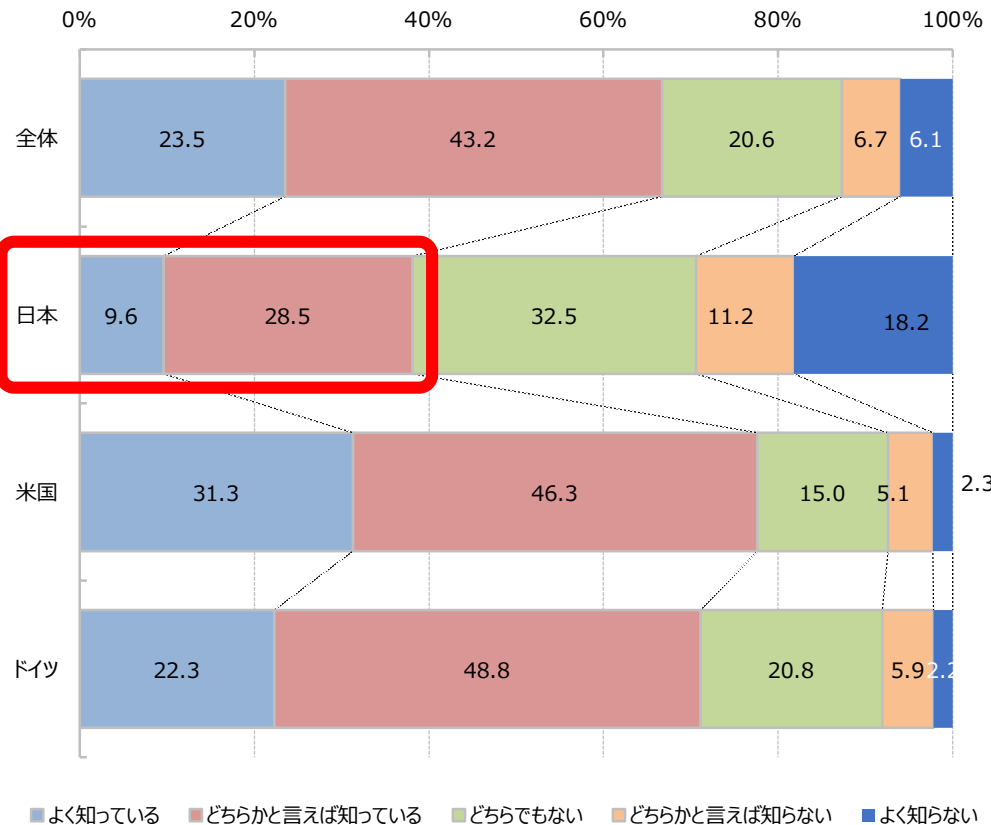


「自分に近い意見が表示されやすいこと」に関する認識調査

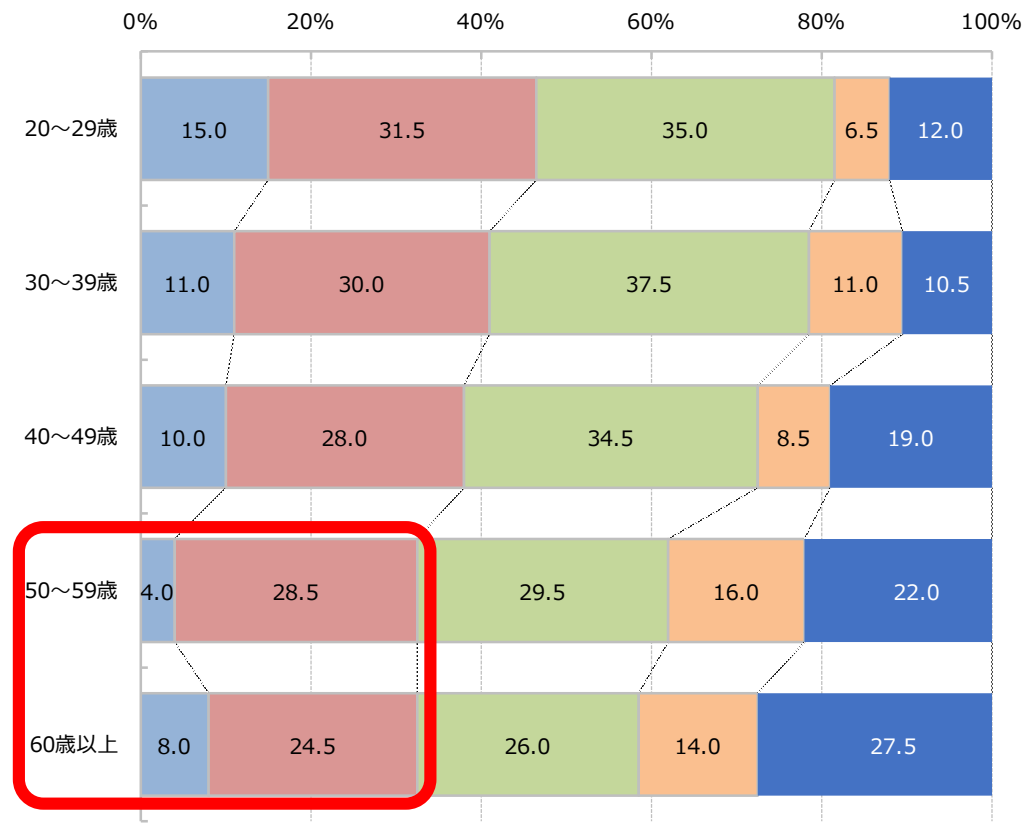
- SNS等では自分に近い意見や考え方等が表示されやすい傾向があることについて知っている（「よく知っている」と「どちらかと言えば知っている」の合計）と回答した割合は、欧米と比較すると低い。
- また、我が国について年代別に見ると、50歳代及び60歳以上では他の年齢層と比較すると低い。

【SNS等では自分に近い意見が表示されやすいことの認識】

（国際比較）



（日本：年代別）



フィルターバブル等に関する調査研究

- 令和5年度総務省調査研究（委託先：みずほリサーチ&テクノロジーズ(株)）において、**アルゴリズムによる選択的表示及び個人の情報に対する受動的態度の実態**の仮説を検討するため、①シミュレーション型検証と、並行して②大規模アンケート調査を実施予定。
- 手法の有効性の検討や結果の評価等を行うため、有識者4名から成る協議会を設置・運営。



【出典】
みずほリサーチ&テクノロジーズ作成資料

◆協議会構成員

| 構成員（敬称略） | 所属・役職 |
|------------|-------------------------------|
| 鳥海 不二夫（座長） | 東京大学大学院工学系研究科 システム創成学専攻 教授 |
| 笹原 和俊 | 東京工業大学環境・社会理工学院 准教授 |

| 構成員（敬称略） | 所属・役職 |
|----------|----------------------------------|
| 田中 優子 | 名古屋工業大学大学院工学研究科 基礎類 准教授 |
| 山口 真一 | 国際大学グローバル・コミュニケーション・ センター 准教授 |

インターネット上での偽・誤情報等の流通の顕在化

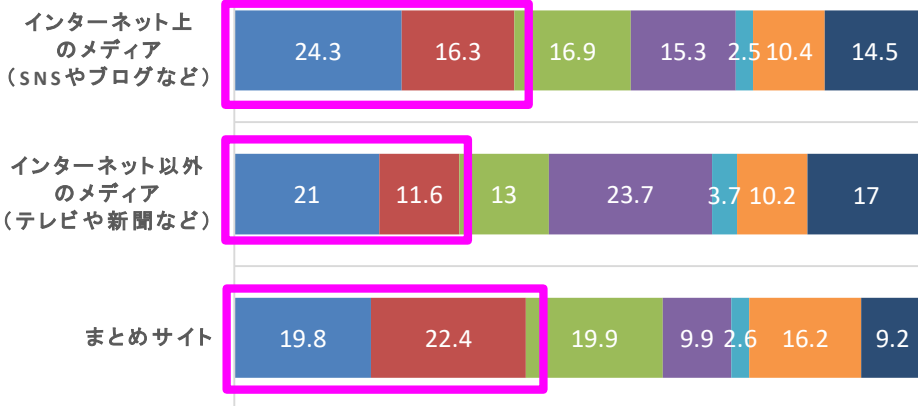
- 多くの利用者が情報を収集・閲覧するプラットフォームサービス等のインターネット上では、**違法・有害情報を含む偽・誤情報も流通するなどの問題が顕在化**。
- 一因として、**偽・誤情報は、SNS上において正しい情報よりも速く、より広く拡散する特性**があること等が指摘。

◆ メディアごとの偽・誤情報を見かける頻度

- ・ インターネット上のメディアやまとめサイトにおいては、40%強が週に1回以上接触。
- ・ インターネット以外のメディアにおいては、30%強が週に1回以上接触。

問) 直近の1ヶ月の間で、あなた自身が偽情報・誤情報※だと思う情報を次に示すメディアの中でどの程度見かけますか。
※ここでは、虚偽、または、誤解を招くと考えられる情報/ニュースを指します。

- 毎日、またはほぼ毎日
- 最低週1回
- 月に数回
- ほとんどない



【出典】令和4年度国内外における偽情報に関する意識調査 (総務省)

◆ 事実と誤情報の拡散ダイナミズム

- ・ 誤情報は事実よりも速く、深く、広範囲に拡散。
- ・ Twitter上で事実が1000人以上にツイートされるのは稀だが、誤情報はもっと多くの人にツイートされた。
- ・ 最初の投稿がリツイートされるまでの時間は誤情報の方が事実よりも20倍短く、1500人に届くまでにかかる時間も6倍短い。

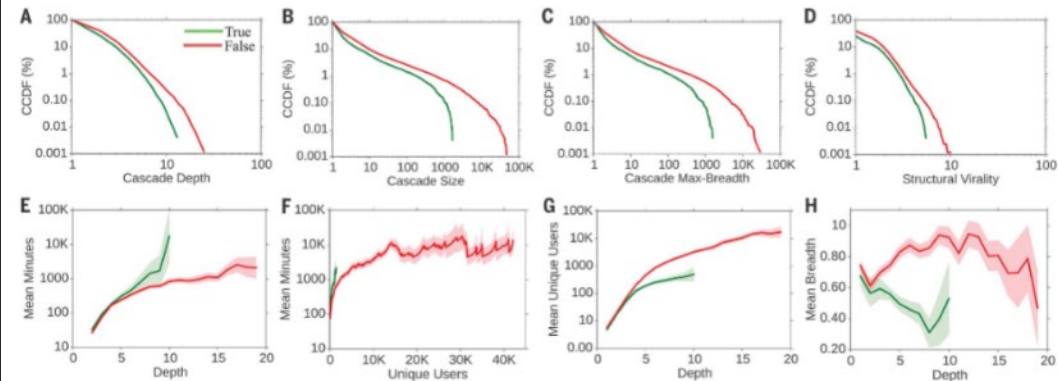


Fig. 2. Complementary cumulative distribution functions (CCDFs) of true and false rumor cascades. (A) Depth. (B) Size. (C) Maximum breadth. (D) Structural virality. (E and F) The number of minutes it takes for true and false rumor cascades to reach any (E) depth and (F) number of unique Twitter users. (G) The number of unique Twitter

users reached at every depth and (H) the mean breadth of true and false rumor cascades at every depth. In (H), plot is lognormal. Standard errors were clustered at the rumor level (i.e., cascades belonging to the same rumor were clustered together; see supplementary materials for additional details).

【出典】Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151.

インターネットにおける偽・誤情報の拡散の事例

| 年 | エリア | 概要 |
|------|-------|---|
| 2011 | 日本 | 東日本大震災直後、「今回の地震は人工地震である」、「被災地で外国人による犯罪が横行している」等の偽・誤情報（流言飛語）が拡散。 |
| 2016 | 英国 | EUからの離脱を問う国民投票に際し、「英国はEUに毎週 3 億5000万ポンドを支払っている」という偽・誤情報が拡散。 |
| 2016 | 米国 | 大統領選挙に際し、各候補者について「ローマ法王が支持を表明した」、「あるピザ店を拠点とする児童売春組織に関与している」等の偽・誤情報が拡散。後者の拡散は、噂されたピザ店に実際に銃を持って押し入り、発砲する人物まで現れる事態につながった（いわゆる「ピザゲート事件」）。 |
| 2016 | ドイツ | 移民によるテロ事件の発生後、事件と無関係の難民とメルケル首相（当時）との写真が利用され、同首相がテロリストと関係があったかのような偽・誤情報が拡散。 |
| 2017 | フランス | 大統領選挙に際し、「マクロン候補（当時）が租税回避地にペーパーカンパニーや銀行口座を保有している」という偽・誤情報が拡散。 |
| 2019 | 日本 | 常磐自動車道であり運転や傷害行為に及んだ人物の同乗者であるとして、無関係の人物が特定され、誹謗中傷を受けた。 |
| 2020 | 世界 | 新型コロナウイルス感染症の流行に際し、「第 5 世代移動通信システム（5G）が流行に寄与している」、「ワクチンにマイクロチップが含まれている」等の偽・誤情報や、医学的な根拠のない感染予防効果を標榜する健康商品等のインターネット広告が拡散。 |
| 2022 | 世界 | ロシアによるウクライナへの軍事侵攻を巡り、特定の国・個人・企業がいずれかの陣営を支持又は非難しているかのように見せかけた偽画像・偽情報等が拡散。 |
| 2022 | フィリピン | 大統領選挙に際し、候補者に不利な事実の存在を否定するなどの内容の偽・誤情報が拡散。 |
| 2023 | 日本 | X（旧Twitter）に、財務官の氏名や肩書、写真等を使用して本人のアカウントと偽ったなりすましアカウントが出現、550人ほどがフォロー（後にアカウント凍結）。 |
| 2023 | 日本 | 東京電力福島第一原子力発電所からのALPS処理水の海洋放出に際し、無関係の画像等を使用して「魚が大量死している」、「奇形の魚が発見された」等とする偽・誤情報が拡散。 |
| 2023 | 世界 | イスラエルとパレスチナ武装勢力ハマスとの衝突を巡り、無関係の画像等を使用して現地の状況を実際よりも軽く見せたり、逆に深刻に見せたりする偽・誤情報が拡散。 |

東日本大震災に係るインターネット上の流言飛語への適切な対応に関する要請

報道資料

平成23年4月6日

東日本大震災に係るインターネット上の流言飛語への適切な対応に関する電気通信事業者関係 団体に対する要請

総務省は、本日、電気通信事業者関係団体に対し、東日本大震災に係るインターネット上の流言飛語について、各団体所属の電気通信事業者等が表現の自由に配慮しつつ適切に対応するよう、周知及び必要な措置を講じることを要請しました。

本日、「被災地等における安全・安心の確保対策ワーキングチーム」において、「被災地等における安全・安心の確保対策」が決定されました。

同対策においては、東日本大震災後、地震等に関する不確かな情報等、国民の不安をいたずらにあおる流言飛語が、電子掲示板への書き込み等により流布している状況に鑑み、インターネット上の流言飛語について関係省庁が連携し、サイト管理者等に対して、法令や公序良俗に反する情報の自主的な削除を含め、適切な対応をとることを要請し、正確な情報が利用者に提供されるよう努めることとされています。

同対策を踏まえ、総務省では、社団法人電気通信事業者協会、社団法人テレコムサービス協会、社団法人日本インターネットプロバイダー協会及び社団法人日本ケーブルテレビ連盟に対して、東日本大震災に係るインターネット上の流言飛語について、各団体所属の電気通信事業者等が表現の自由に配慮しつつ適切に対応するよう、周知及び必要な措置を講じることを要請しました。

別紙：[東日本大震災に係るインターネット上の流言飛語への適切な対応に関する要請](#)

別紙

平成23年4月6日
総基消第145号

社団法人 電気通信事業者協会
社団法人 テレコムサービス協会
社団法人 日本インターネットプロバイダー協会
社団法人 日本ケーブルテレビ連盟
御中

総務省総合通信基盤局長

東日本大震災に係るインターネット上の流言飛語への 適切な対応に関する要請

平素より、情報通信行政に対し、格別の御高配を賜り、厚く御礼申し上げます。

さて、本日、「被災地等における安全・安心の確保対策ワーキングチーム」(※)において、「被災地等における安全・安心の確保対策」(別紙参照)が決定されました。

東日本大震災後、地震等に関する不確かな情報等、国民の不安をいたずらにあおる流言飛語が、電子掲示板への書き込み等により流布しており、被災地等における混乱を助長することが懸念されます。

つきましては、インターネット上の地震等に関連する情報であって法令や公序良俗に反すると判断するものを自主的に削除することを含め、貴団体所属の電気通信事業者等に、表現の自由にも配慮しつつ、「インターネット上の違法な情報への対応に関するガイドライン」や約款に基づき、適切な対応をおとりいただくよう御周知いただくとともに、貴団体においても必要な措置を講じてくださいますようお願い申し上げます。

※ 平成23年3月31日、関係省庁が緊密に連携し、被災地等における安全・安心の確保に係る総合的な対策を検討・推進することを目的に設置(議長：内閣官房副長官補(内政))。

以上

「コグニティブセキュリティ」の重要性に関する指摘

- 科学技術振興機構（JST）の研究開発戦略センター（CRDS）による「研究開発の俯瞰報告書」（2023年5月）において、その前提として必要な「科学と社会」の視座として、「**人の認知（コグニティブ）を守るコグニティブセキュリティの研究も望まれている**」旨が記載。
- また、7つの俯瞰区分の1つ「セキュリティ・トラスト」における研究開発領域「人・社会とセキュリティ」にて、「これまでの研究開発の流れとトレンド」として、**プラットフォームにおけるコンテンツ・モデレーション等**が、そして、「新展開・技術トピックス」として、**ファクトチェック等**が、**Misinformation・Disinformationの対策技術**として紹介。

| 技術トレンド | 社会・経済の動向 | 俯瞰 | 重点的に取り組むべき研究開発領域 | 推進シナリオ | ビジョン |
|---|---|---|--|--|---|
| 社会的要請との整合 研究開発活動や科学技術そのものに対する社会的要請の高まり。データに関するプライバシーの考慮やAI技術へのトラスト担保など。 あらゆるもののスマート化・自律化 機器のスマート化が進み、大量のデータの収集と解析が可能になった。ビッグデータと機械学習を組み合わせたサービスが多数生み出された。 あらゆるもののデジタル化・コネクティッド化 無線化・大容量化・グローバル化。ウェブ、スマートフォン、IoT、クラウドなど、社会基盤のデジタル化とコネクティッド化。 | 世界 ロシアのウクライナ侵襲。新型感染症。経済のブロック化。格差、貧困、食料備在化。温暖化、自然災害リスク。産業・労働構造の変化。 日本 DX推進。サプライチェーンリスクの顕在化。少子高齢化。経済低成長。 | デジタル安全保障に対する総合知による取り組み スマート化・自律化の根本である知能の原理探究 サステナブル社会のためのICT基盤 | ① デジタル社会におけるトラスト形成 ② コグニティブセキュリティ ③ データ共有 ④ 知能モデルの解明・探求/身体性に宿る知能 ⑤ 人間中心インタラクション ⑥ バイオハイブリッドロボット ⑦ 最適化 ⑧ 社会課題解決に向けたメタバースデザイン ⑨ ネットワークのスマート化 ⑩ 社会デジタルツイン ⑪ 社会システムを支えるAIアーキテクチャー | 技術 社会 基盤 社会 基盤 産業 基盤 技術 産業 技術 社会 技術 産業 技術 社会 基盤 技術 産業 基盤 技術 基盤 | 社会課題解決と人間中心社会の実現 経済発展と社会問題解決を両立し、誰もが快適で活力に満ちた質の高い生活を送れる社会の実現。ITは人間の判断や決定を補助する道具として働く。 データ駆動型・知識集約型の価値創造 知識・情報・データベース化と統合活用を実現するプラットフォームやAIにより、データ駆動型・知識集約型の価値創造とDXが加速される。 サイバー世界とフィジカル世界の高度な融合 IoTやCPSが社会生活を支える基盤となる。オープンなサービスプラットフォームなどが実現し、多くの産業が効率化・省エネルギー化する。 |

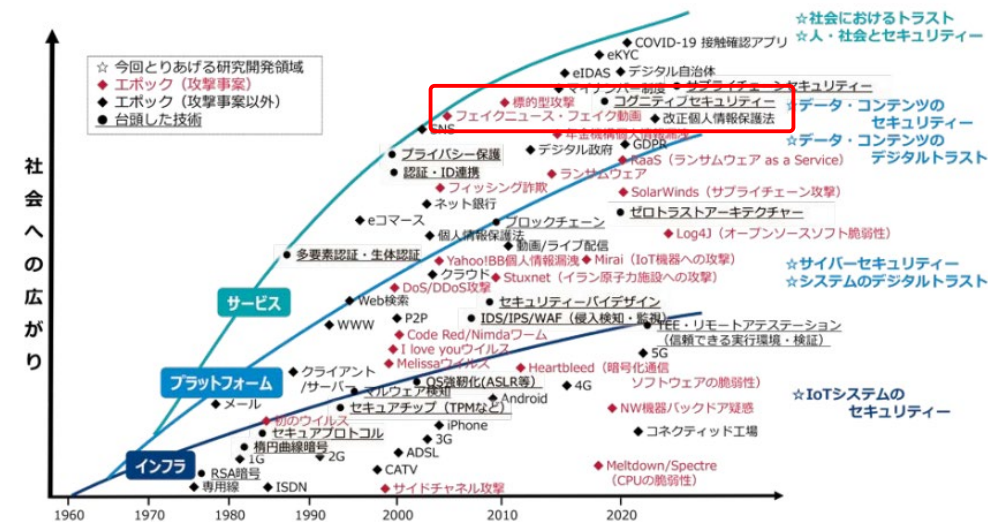


図2-4-1 セキュリティ・トラストの俯瞰図（時系列）

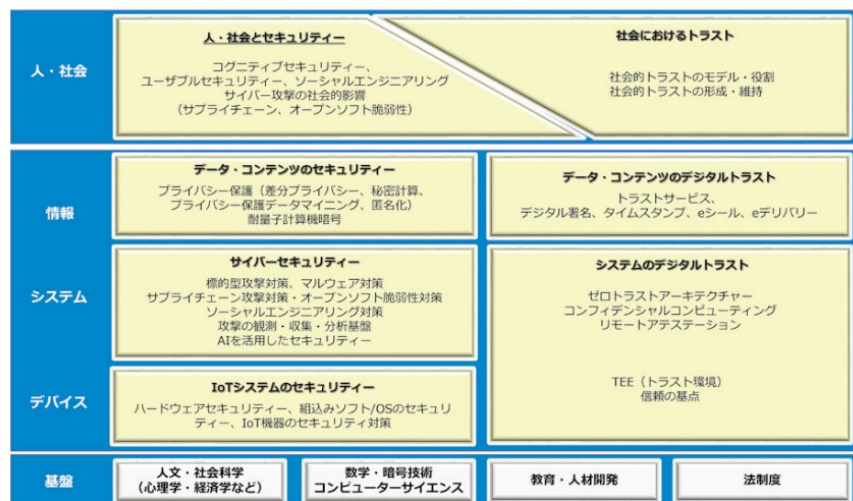


図2-4-2 セキュリティ・トラストの俯瞰図（構造）

日本心理学会第87回大会における偽・誤情報関連シンポジウム

- 2023年9月15日から17日まで神戸市で開かれた日本心理学会第87回大会において、公募シンポジウム「誤情報への挑戦 心理プロセスの理解にもとづいた介入の可能性」が開催。
- 誤情報に対する訂正情報の効果と限界、誤情報の拡散に対する心理的介入の類型、事前介入（プレバンキング）の有効性等について報告・議論。

【シンポジウム概要】

- 日時：2023年9月15日（金）9:00～11:00
- 話題提供者：
 - **犬塚 美輪** 東京学芸大学 教育心理学講座学校心理学分野 准教授 ※企画代表者
「誤情報からの表象構築と訂正」
 - **田中 優子** 名古屋工業大学 大学院工学研究科 准教授
「誤情報訂正：その効果と限界」
 - **藤本 和則** 近畿大学 経営学部経営学科 教授
「心理的介入の類型化について」
- 指定討論者：
 - **菊池 聡** 信州大学 人文学部人文学科文化情報論・社会学 教授
 - **眞嶋 良全** 北星学園大学 社会福祉学部心理学科 教授

2. 「プラットフォームサービスに関する研究会」における インターネット上の偽・誤情報対策に関する主な議論

「プラットフォームサービスに関する研究会」におけるこれまでの検討①

- 「プラットフォームサービスに関する研究会」では、オンライン上のフェイクニュースや偽情報への対応について、**第1回会合（2018年10月）から検討**。
- 2019年4月、同研究会の中間報告書の公表後、同研究会において、同年6～7月に**プラットフォーム事業者等からのヒアリングを実施し、2020年2月に最終報告書が公表**。
- 上記最終報告書において、我が国におけるフェイクニュースや偽情報への対応の在り方として、**まずはプラットフォーム事業者を始めとする民間部門における関係者による自主的な取組を基本とした対策を進めていくことが適当**と提言。

【最終報告書（2020年2月7日）】（抄）

第6節 我が国におけるフェイクニュースや偽情報への対応の在り方

1. 自主的スキームの尊重

偽情報への対応の在り方の基本的な方向性としては、法律による規制、自主的取組とその実効性を支える法的規律を組み合わせた共同規制的な枠組みの構築、業界に対して行動規範の策定を求める等による自主的な対応、個社ごとの自主的な対応など、様々なレベルの対応が考えられる。

この点、第3節における記載のとおり、選挙への影響等により早くから問題が顕在化していた米国や欧州においては、プラットフォーム事業者による自主的な対応を中心に取組が進められてきたところである。一方、一部の国々、具体的には、ドイツ・フランス・シンガポール等の国々では、偽情報の削除義務等を定めた法律が制定されているが、表現の自由への萎縮効果への懸念や、偽情報の該当性判断の困難性による実効性の欠如及び恣意的運用への懸念など、多くの課題や批判の声もある。

したがって、**我が国における偽情報への対応の在り方の基本的な方向性としては、まずはプラットフォーム事業者を始めとする民間部門における関係者による自主的な取組を基本とした対策を進めていくことが適当**である。

政府は、これらの民間による自主的な取組を尊重し、その取組状況を注視していくことが適当である。特に、プラットフォーム事業者による情報の削除等の対応など、個別のコンテンツの内容判断に関わるものについては、表現の自由の確保などの観点から、政府の介入は極めて慎重であるべきである。

他方で、**仮に民間による自主的スキームが達成されない場合、あるいは、将来的に偽情報の拡散等の問題に対して効果がないと認められる場合には、透明性・アカウントビリティの確保方策に関して、プラットフォーム事業者に対する行動規範の策定や対応状況の報告・公表など、行政からの一定の関与も視野に入れて検討を行うことが適当**である。

「プラットフォームサービスに関する研究会」におけるこれまでの検討②

- 「プラットフォームサービスに関する研究会」では、2020年2月の最終報告書の公表後、同研究会において、2021年3～5月にプラットフォーム事業者等からのヒアリングを通じて、**モニタリング・検証評価を実施し**、2021年9月、以下の**モニタリングの結果を含む中間とりまとめが公表**。

- **全体的な傾向として、プラットフォーム事業者の偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は限定的であった。**
- 他方、多様なステークホルダーによる協力関係の構築、ファクトチェック推進、ICT リテラシー向上に関しては、まだ十分とは言えないものの、我が国においても取組が進められつつある。

プラットフォームサービス事業者におけるフェイクニュースや偽情報への対策状況ヒアリング 質問項目

- | | |
|--|---|
| <p>1. 前提(サービス概要)</p> <ul style="list-style-type: none">① 対象サービス名・サービス分類② 加入者数、月間アクティブユーザー数又は書き込み数 <p>2. 「我が国における実態の把握」関係</p> <ul style="list-style-type: none">① 偽情報等の発生・拡散状況を把握できる体制 分析・調査の有無② 日本における偽情報等の発生・拡散状況、結果公表③ 研究者への情報提供、利用条件④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等)⑤ 不正な申告や削除要請への対策の方法・仕組み、対応件数 <p>3. 「多様なステークホルダーによる協力関係の構築」関係</p> <ul style="list-style-type: none">① 産学官民の協力関係の構築② 具体的な役割③ 議論を踏まえた取組 <p>4. 「プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保」関係</p> <p>(1) 偽情報等に関するポリシー</p> <ul style="list-style-type: none">① ①禁止行為 (ii)削除等の対応 <p>(2) 削除等の対応</p> <ul style="list-style-type: none">① 偽情報等に関する申告や削除要請の件数② (i)偽情報等に関する申告や削除要請に対する削除件数、(ii)アカウントの停止数③ 偽情報等に関する主体的な削除件数(AIを用いた自動検知機能の活用等)④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等)⑤ 削除以外の取組 (i)警告表示 (ii)表示順位の低下 (iii)その他の取組内容⑥ 不正な申告や削除要請への対策の方法・仕組み、対応件数 <p>(3) 削除要請や苦情に関する受付態勢・プロセス</p> <ul style="list-style-type: none">① 一般ユーザーからの申告・削除要請への受付窓口・受付態勢、対応プロセス② 対応決定時における通知の内容、理由の記載の程度③ 一般ユーザーからの申告や削除要請に対応する部署・チームの規模・人数④ ③の部署・チームに関する日本国内の拠点の有無、日本における責任者の有無⑤ 削除等への苦情や問い合わせに対する苦情受付態勢及び苦情処理プロセス <p>(4) 透明性・アカウントビリティの確保</p> <ul style="list-style-type: none">① コンテンツモデレーションのアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照② 透明性レポート 日本語で閲覧可能か③ 取組の効果分析 | <p>5. 「利用者情報を活用した情報配信への対応」関係</p> <ul style="list-style-type: none">① 広告表示先の制限② 広告出稿制限③ ターゲティング技術の適用に関する規定④ 広告のアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照⑤ 出稿者の情報や資金源の公開⑥ 広告とコンテンツの分離⑦ その他の透明性・アカウントビリティ確保方策 ユーザーへのツール提供 <p>6. 「ファクトチェックの推進」関係</p> <ul style="list-style-type: none">① ファクトチェック結果の表示 具体的な仕組み・基準② ファクトチェックを容易にするツールの開発及び提供③ ファクトチェックを実施する人材の育成④ ファクトチェック機関との連携 <p>7. 「ICTリテラシー向上の推進」関係</p> <ul style="list-style-type: none">① 普及啓発の取組・投資② 他のステークホルダーとの連携・協力・投資 <p>8. 「研究開発の推進」関係</p> <ul style="list-style-type: none">① AI技術に関する研究開発② 「ディープフェイク」対策の研究開発 <p>9. 「情報発信者側における信頼性確保方策の検討」関係</p> <ul style="list-style-type: none">① 信頼性の高い情報の表示② ニュースの選別・編集に関する透明性・アカウントビリティ確保方策③ メディアとの連携体制構築 具体的検討・取組④ 情報源のトレーサビリティ確保、なりすまし防止・認証 <p>10. その他</p> <ul style="list-style-type: none">① 意見・補足 |
|--|---|

「プラットフォームサービスに関する研究会」によるモニタリング結果（2021年）

◆偽情報対策に関するプラットフォーム事業者における透明性・アカウントビリティ確保状況の主な評価項目の比較

| 項目 | |
|--|---|
| 我が国における実態の把握(2. 関係) | ・すべての事業者について、我が国における偽情報の実態把握及び結果の分析・公開は未実施 |
| 多様なステークホルダーによる協力関係の構築(3. 関係) | ・「Disinformation対策フォーラム」や「Innovation Nippon」により一定の進捗がみられる |
| プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保(4. 関係) | <ul style="list-style-type: none"> ・ヤフー(CGMサービス)は偽情報を直接禁止するポリシーが設けられておらず、対応件数等も非公開 ・LINEは、虚偽の情報の流布を包括的に禁止している。対応件数は一部公開 ・Googleはコロナ関連偽情報に関してポリシーがあり、我が国における削除件数等について一部公開 ・Facebook・Twitterは、削除ポリシーの整備や削除結果の公表等がグローバルでは進められているが、我が国における取組状況やその効果については非公開 |
| 利用者情報を活用した情報配信への対応(5. 関係) | <ul style="list-style-type: none"> ・いずれの事業者も、偽情報に関する広告や政治広告について、広告配信先の制限や、広告内容に関する何らかの制限を規定するなど、一定の対応は行われている ・LINE及びTwitterでは政治広告禁止。Googleでは日本において選挙広告は禁止 ・政治広告についてのターゲティング技術の適用については明確になっていない |
| ファクトチェックの推進(6. 関係) | <ul style="list-style-type: none"> ・ヤフーやgoogleなど一部のプラットフォーム事業者において、プラットフォーム上の情報へのファクトチェック結果の紐付けや、資金提供等の取組が開始されている ・Facebookでは、グローバルな取組は進められているが、我が国において国際ファクトチェックネットワーク(IFCN)に加盟しているファクトチェッカーがいないため、ファクトチェッカーとの連携は行われていない ・LINE、Twitterでは、取組は未実施(LINEは台湾では取組を実施) |
| 情報発信者側における信頼性確保方策の検討(9. 関係) | <ul style="list-style-type: none"> ・「Disinformation対策フォーラム」において、既存メディアや有識者との情報共有や協議が進展 ・いずれの事業者も、新型コロナウイルス感染症関係の政府(内閣官房や厚労省等)の情報をサービス内で優先表示させる仕組み等を積極的に実施しているが、認知度は低い ・Googleでは、ジャーナリズムの支援策を実施 |
| ICTリテラシー向上の推進(7. 関係) | <ul style="list-style-type: none"> ・Facebook・Google・LINEでは、我が国において偽情報の問題に対応したリテラシー教育に関する取組を実施 ・Twitterの取組は、偽情報対策に資する内容となっているか不明。ヤフーは今後実施予定 |
| 研究開発の推進(8. 関係) | <ul style="list-style-type: none"> ・Facebook、Google、Twitterではグローバルな取組としてディープフェイク対策のための研究開発が行われている ・ヤフー、LINEではディープフェイク対策の研究開発は行われていない |

「プラットフォームサービスに関する研究会」におけるこれまでの検討③

- 「プラットフォームサービスに関する研究会」では、2021年9月の中間とりまとめの公表後、同研究会において、2022年3月にプラットフォーム事業者等からのヒアリングを通じて、**モニタリング・検証評価を実施し**、2022年8月、以下の**モニタリングの結果を含む第二次とりまとめが公表**。

- **全体的な傾向**として、プラットフォーム事業者の偽情報への対応及び**透明性・アカウントビリティ確保の取組の進捗は未だ限定的**であった。
- 他方、多様なステークホルダーによる協力関係の構築、特定のトピックに関する偽情報や誤解を招く情報の流布に関するポリシーの設定、ファクトチェック推進、ICTリテラシー向上に関しては、まだ十分とは言えないものの、我が国においても取組が進められつつある。

| プラットフォームサービス事業者における偽情報等への対策状況ヒアリング（2022年） 質問項目 | |
|--|--|
| 1. 前提(サービス概要) | 5. 「利用者情報を活用した情報配信への対応」関係 |
| ① 対象サービス名・サービス分類 | ① 広告表示先の制限 |
| ② 加入者数、月間アクティブユーザー数又は書き込み数 | ② 広告出稿制限 |
| 2. 「我が国における実態の把握」関係 | ③ ターゲティング技術の適用に関する規定 |
| ① 偽情報等の発生・拡散状況を把握できる体制・分析・調査の有無 | ④ 広告のアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照 |
| ② グローバル・日本における偽情報等の発生・拡散状況、結果の分析・公表 | ⑤ 出稿者の情報や資金源の公開 |
| ③ 研究者への情報提供、利用条件 | ⑥ 広告とコンテンツの分離 |
| ④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等) | ⑦ その他の透明性・アカウントビリティ確保方策 ユーザーへのツール提供 |
| ⑤ 不正な申告や削除要請への対策の方法・仕組み、対応件数 | |
| 3. 「多様なステークホルダーによる協力関係の構築」関係 | 6. 「ファクトチェックの推進」関係 |
| ① Disinformation対策フォーラムなどの産学官民の協力関係の構築 | ① ファクトチェック結果の表示 具体的な仕組み・基準 |
| ② 多様なステークホルダーによる協力関係の構築に向けての具体的な役割 | ② ファクトチェックを容易にするツールの開発及び提供 |
| ③ 議論を踏まえた取組 | ③ ファクトチェックを実施する人材の育成 |
| 4. 「プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保」関係 | ④ 日本におけるファクトチェック推進に向けた取組 |
| (1) 偽情報等に関するポリシー | ⑤ ファクトチェック機関との連携 |
| ① 禁止行為 (ii) 削除等の対応 | |
| ② ポリシー等の見直し状況及び外部レビューの有無とそのタイミング | 7. 「ICTリテラシー向上の推進」関係 |
| (2) 削除要請や苦情に関する受付態勢・プロセス | ① 普及啓発の取組・投資 |
| ① 一般ユーザーからの申告・削除要請への受付窓口・受付態勢、対応プロセス | ② 他のステークホルダーとの連携・協力・投資 |
| ② 対応決定時における通知の内容、理由の記載の程度 | |
| ③ 一般ユーザーからの申告や削除要請に対応する部署・チームの規模・人数 | 8. 「研究開発の推進」関係 |
| ④ ③の部署・チームに関する日本国内の拠点の有無、日本における責任者の有無 | ① AI技術に関する研究開発 |
| ⑤ 削除等への苦情や問い合わせに対する苦情受付態勢及び苦情処理プロセス | ② 「ディープフェイク」対策の研究開発 |
| (3) 削除等の対応 | 9. 「情報発信者側における信頼性確保方策の検討」関係 |
| ① 偽情報等に関する申告や削除要請の件数 | ① 信頼性の高い情報の表示 |
| ② (i) 偽情報等に関する申告や削除要請に対する削除件数、(ii) アカウントの停止数 | ② ニュースの選別・編集に関する透明性・アカウントビリティ確保方策 |
| ③ 偽情報等に関する主体的な削除件数(AIを用いた自動検知機能の活用等) | ③ メディアとの連携体制構築 具体的検討・取組 |
| ④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等) | ④ 情報源のトレーサビリティ確保、なりすまし防止・認証 |
| ⑤ 削除以外の取組 (i) 警告表示 (ii) 表示順位の低下 (iii) その他の取組内容 | |
| ⑥ 不正な申告や削除要請への対策の方法・仕組み、対応件数 | 10. その他 |
| (4) 透明性・アカウントビリティの確保 | ① 意見・補足 |
| ① コンテンツモデレーションのアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照 | |
| ② 透明性レポート 日本語で閲覧可能か | |
| ③ 取組の効果分析 | |

「プラットフォームサービスに関する研究会」によるモニタリング結果（2022年）

◆ 偽情報対策に関するプラットフォーム事業者における透明性・アカウントビリティ確保状況の主な評価項目の比較

| | |
|--|---|
| 我が国における実態の把握(2. 関係) | ・すべての事業者について、我が国における偽情報の実態把握及び結果の分析・公開は未実施 |
| 多様なステークホルダーによる協力関係の構築(3. 関係) | ・「Disinformation対策フォーラム」などにより一定の進捗がみられる |
| プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保(4. 関係) | <ul style="list-style-type: none"> ・ヤフー(ニュース及び知恵袋)は偽情報を直接禁止するポリシーを設定。対応件数等は一部公開 ・LINEは、虚偽の情報の流布を包括的に禁止している。対応件数は一部公開 ・Googleはコロナ関連偽情報に関してポリシーがあり、我が国における削除件数等について一部公開 ・Meta及びTwitterは、削除ポリシーの整備や削除結果の公表等がグローバルでは進められており、我が国における削除件数等について、ヒアリングの事後に一部公開 |
| 利用者情報を活用した情報配信への対応(5. 関係) | <ul style="list-style-type: none"> ・いずれの事業者も、偽情報に関する広告や政治広告について、広告配信先の制限や、広告内容に関する何らかの制限を規定するなど、一定の対応を実施 ・LINE及びTwitterでは政治広告禁止。Googleでは日本において選挙広告は禁止。政治広告についてのターゲティング技術の適用については、ヤフー及びGoogleにおいて、政治広告に関する行動ターゲティングを制限 |
| ファクトチェックの推進(6. 関係) | <ul style="list-style-type: none"> ・ヤフーやGoogle、LINEなど一部のプラットフォーム事業者において、ファクトチェック結果のまとめの掲出やプラットフォーム上の情報へのファクトチェック結果の紐付け、資金提供等の取組が開始されている ・Metaでは、グローバルな取組は進められているが、我が国において国際ファクトチェックネットワーク(IFCN)に加盟しているファクトチェッカーがいないため、ファクトチェッカーとの連携は行われていない ・Twitterでは、特定の投稿にユーザが文脈を追加できるBirdwatch機能を試行予定。 |
| 情報発信者側における信頼性確保方策の検討(9. 関係) | <ul style="list-style-type: none"> ・「Disinformation対策フォーラム」において、既存メディアや有識者との情報共有や協議が進展 ・いずれの事業者も、新型コロナウイルス感染症やウクライナ侵攻関係の情報について、信頼性のあるニュースメディアやファクトチェック機関、政府(内閣官房や厚労省等)の情報をサービス内で優先表示を実施 ・Googleでは、ジャーナリズムの支援策を実施 |
| ICTリテラシー向上の推進(7. 関係) | <ul style="list-style-type: none"> ・すべての事業者において、偽情報の問題に対応したリテラシー教育に関する取組を実施 ・ヤフーでは、不確かな情報に惑わされないための「Yahoo!ニュース健診」を提供。 |
| 研究開発の推進(8. 関係) | <ul style="list-style-type: none"> ・Google、Meta、Twitterではグローバルな取組としてディープフェイク対策のための研究開発が行われている。ヤフー、LINEではディープフェイク対策の研究開発は行われていない |

3. プラットフォーム事業者等幅広い関係者による自主的な対応状況

Disinformation対策フォーラムの設立（2020年6月）

- 「プラットフォームサービスに関する研究会」最終報告書（2020年2月）を踏まえ、偽情報対策について**多様なステークホルダーによる協力関係構築を図り、対話の枠組み**を設けるべく、「**Disinformation対策フォーラム**」が**民間主導で設立**。

目的

・偽情報流通の実態を正確に把握し、その対応について多面的に検討すべく、産学官民の多様なステークホルダーによる協力関係の構築を図り、対話の枠組みを設ける

基本方針

- 多様：各社方針の尊重による多元性確保、情報の多様性促進
- 対話：各社取組みの公表等による信頼性・説明性の向上、リテラシー向上施策の推進等
- 連携：既存の違法有害情報対策との接続性、研究開発やファクトチェック推進のための協働

構成員等

| | | | |
|----------|---|------|----------------|
| 穴戸常寿（座長） | 東京大学大学院法学政治学研究科 教授 | 藤代裕之 | 法政大学社会学部 教授 |
| 沢田登志子 | 一般社団法人ECネットワーク 理事 | 安野智子 | 中央大学 文学部 教授 |
| 瀬尾傑 | スマートニュース メディア研究所 所長 | 山口真一 | 国際大学GLOCOM 准教授 |
| 西田亮介 | 東京工業大学リベラルアーツ研究教育院 准教授 | | |
| （事業者） | Facebook Japan株式会社、Google合同会社、Twitter Japan株式会社、ヤフー株式会社 | | |
| （オブザーバ） | 総務省、消費者庁、日本新聞協会、日本放送協会、日本民間放送連盟 | | |
| （事務局） | 一般社団法人セーフターインターネット協会 | | |

スケジュール

| | | | |
|------------------|-------------|-----------------|------------------|
| 2020年6月 | 設立会合開催 | 2021年7月 | ワクチンデマ対策シンポジウム開催 |
| 2020年10月～2021年3月 | 第2回～第6回会合開催 | 2021年9月～2022年1月 | 第7回会合～第10回会合開催 |
| 2021年3月 | 中間とりまとめ公表 | 2022年3月 | 報告書公表 |

ワクチンデマ対策シンポジウムの開催（2021年7月）

- 国内におけるコロナワクチンの接種開始に伴い、ワクチンの効果や副作用に関する情報が大量に飛び交う中で、インターネットを通じて正しい情報を基に適切な判断を行える環境醸成が重要との考えの下、2021年7月、（一社）セーファーインターネット協会（SIA）が「**ワクチンデマ対策シンポジウム**」を開催。
- ワクチンデマの現状と分析、インターネット事業者が行っているデマ防止の対策について紹介。

主催

（一社）セーファーインターネット協会（SIA）

後援

総務省・厚生労働省

日時

2021年7月15日（木）13時～15時

内容

※肩書は当時

- 1. 来賓ご挨拶**
河野 太郎 内閣府特命担当大臣
- 2. ワクチンデマの具体的事例と影響**
古田 大輔 氏（ジャーナリスト／メディアコラボ代表）
- 3. ワクチンデマの分析（最新の研究成果）**
山口 真一 氏（国際大学GLOCOM准教授）
- 4. 適切な接種判断のために必要な情報や考え方**
忽那 賢志 氏
（大阪大学大学院医学系研究科・医学部感染制御学講座）
- 5. インターネット事業者のワクチン情報に関する取組紹介**
Facebook Japan株式会社、Google合同会社、
LINE株式会社、Twitter Japan株式会社、ヤフー株式会社
- 6. 閉会の挨拶**
吉田 奨 氏（SIA専務理事／ヤフー株式会社政策企画本部長）

Disinformation対策フォーラム報告書の策定・公表（2022年3月）

- Disinformation対策フォーラムにおける2022年1月までの計10回の会合やシンポジウムにおける議論を踏まえ、民間事業者をはじめとする関係者の具体的な取組を検討・実施するに当たり踏まえるべき前提、検討する際のポイント、取組を実践する際の留意点、中長期的に偽情報・誤情報対策を進めていく上での課題や論点についてとりまとめ、同フォーラムより、2022年3月に報告書を公表。

第1章 「Disinformation」の定義、議論の範囲及び踏まえるべき前提

- Disinformationの定義：「あらゆる形態における虚偽の、不正確な、または誤解を招くような情報で、設計・表示・宣伝される等を通して、公共に危害が与えられた、又は、与える可能性が高いもの」
⇒発信者の悪意や騙す意図の有無は問題とせず、「偽情報・誤情報」として記述。
- 議論の範囲：インターネット上のSNS等で個人ユーザが発信する「デマ」の類を対象とする。
- 偽情報・誤情報対策の目的：情報の誤りそのものではなく、誤った情報の拡散、当該情報を信じた者の行動変容、結果として生じる個人・社会・経済への被害を問題とし、当該被害を抑制・防止・回復することを目的とする。
⇒民主主義や政策過程に悪影響を与え得るものから民主主義や自由かつ健全な言論空間を守ることも目的に含まれる。
- 表現の自由への配慮 ○ 偽情報・誤情報の多様性 ○ 対策の多様性 ○ 実態把握の重要性
- プラットフォーム事業者への社会的期待 ○ インターネット上のサービスにおけるアーキテクチャ上の工夫

第2章 ファクトチェックの取組

- プラットフォーム事業者の取組の現状：公的機関の情報やファクトチェック結果の拡散、ファクトチェック機関等との連携 等
- 更なる取組の必要性と具体化に向けた留意点：ファクトチェックを行う目的の明確化と表現の自由の尊重 等
- 取組の具体化及び実施の方向性：ファクトチェックの総合的な実施、対象分野・案件の選定 等

第3章 リテラシー向上の取組

- 取組具体化に向けた留意点：リテラシーの中身、個人の自由な理解と立場の尊重、リテラシーの多様性 等
- 取組の具体化及び実施の方向性：一般的と分野別コンテンツ、自身や周囲が誤る可能性の自認 等
- 今後に向けた課題：情報リテラシー指標の開発と持続的な実態把握、量的な拡充

第4章 その他の論点：偽情報・誤情報生成のエコシステムの把握、属性ごとの集団の分析

偽情報対策に係る取組集の策定・公表（2023年3月）

- 2023年3月、「プラットフォームサービスに関する研究会」において、各ステークホルダーによる自主的な対応をまとめた、関係者間で参照しやすい取組集を公表。

| 事業者 | 課題 | 取組 | 効果 |
|----------------------|---|---|---|
| ヤフー株式会社 | 正確な情報の迅速な伝達は行っていたが、ファクトチェックに特化した記事の配信は少ない | <ul style="list-style-type: none"> Yahoo!ニュースでは、公共性の高い情報やデマを打ち消す情報を最も目立つ場所に掲載 Yahoo!トップページでは、生命財産に関わる重大事項について、メディアから提供を受けたコンテンツや情報収集した特設サイトに誘導 など | <ul style="list-style-type: none"> ユーザ動線の各所にフェイクニュース対策のコンテンツや偽情報打ち消し記事等を配置することにより、多数のユーザに注意喚起を含めた情報を提供。 信用できる情報の届け方のパターンを開発 |
| LINE株式会社 | 「LINEオープンチャットが有害情報の温床となっている」かのような論調 | <ul style="list-style-type: none"> オープンチャット安心・安全ガイドラインに基づく削除等の強化、ユーザーへの啓発 画面を開いた瞬間に出るポップアップにより注意喚起 | <ul style="list-style-type: none"> 明らかなデマやフェイクニュースをテーマとしたオープンチャットが検索結果に出現しなくなった 「オープンチャットが有害情報の温床となっている」という論調が減少 |
| Google LLC | 社会における適切な対策を検討するために、日本の実態調査が必要 | <ul style="list-style-type: none"> 3年間で15,000人以上を対象に、偽情報・誤情報に関する人々の行動を調査分析 対象とした偽・誤情報は広範囲で、国内で広く拡散されたコロナワクチンデマも含む | <ul style="list-style-type: none"> 日本における偽・誤情報の拡散範囲や、騙されやすい人の特徴が明らかになった 総務省の有識者会議や啓発教材で共有・引用 |
| Meta Platforms, Inc. | 戦略的目標のために公共の議論を操作したり誤らせたりすることを目的とした組織的な「影響工作・組織的偽装行為」 | <ul style="list-style-type: none"> 政治的広告・投稿の透明性を高めるツールや製品の導入 アカウントの保全性と実名の使用に係るポリシーに基づき毎日数百万の偽アカウントを削除 | |
| セーファーインターネット協会 (SIA) | 「Disinformation 対策フォーラム」において「ファクトチェック機関の創設」の提言 | <ul style="list-style-type: none"> 「日本ファクトチェックセンター (JFC)」の創設 | <ul style="list-style-type: none"> 2022年12月時点で40件の記事を配信、217万Page View |
| ファクトチェックイニシアチブ (FIJ) | 国内におけるファクトチェック記事の絶対量が少ない | <ul style="list-style-type: none"> 疑義言説の自動検知システム「Fact Checking Console」 疑義言説データベース「Claim Monitor」を提供 | <ul style="list-style-type: none"> ファクトチェック記事の量的拡充に寄与 主要な利用団体が実施したファクトチェックの約7割がこのシステムを利用 |

ファクトチェックの推進

- 「**ファクトチェック**」とは、**公開された言説のうち、客観的に検証可能な事実について言及した事項に限定して真実性・正確性を検証し、その結果を公表**する営み。 [NPO法人ファクトチェック・イニシアティブ (FIJ)「ファクトチェック・ガイドライン」より]
- 現在はインターネットを經由して、国外からも真偽不確かなものも含め様々な情報が瞬時に国内へ到達。**オンライン上の情報に対してファクトチェックを推進する必要性**が高まっていると指摘。

疑義言説（画像一体型）

ドローンで撮影された
静岡県の水害



静岡県の水害直後にTwitterで拡散

ファクトチェック記事

「ドローンで撮影された静岡県の水害」は虚偽 AIで偽画像を作成

× 虚偽

リトマス

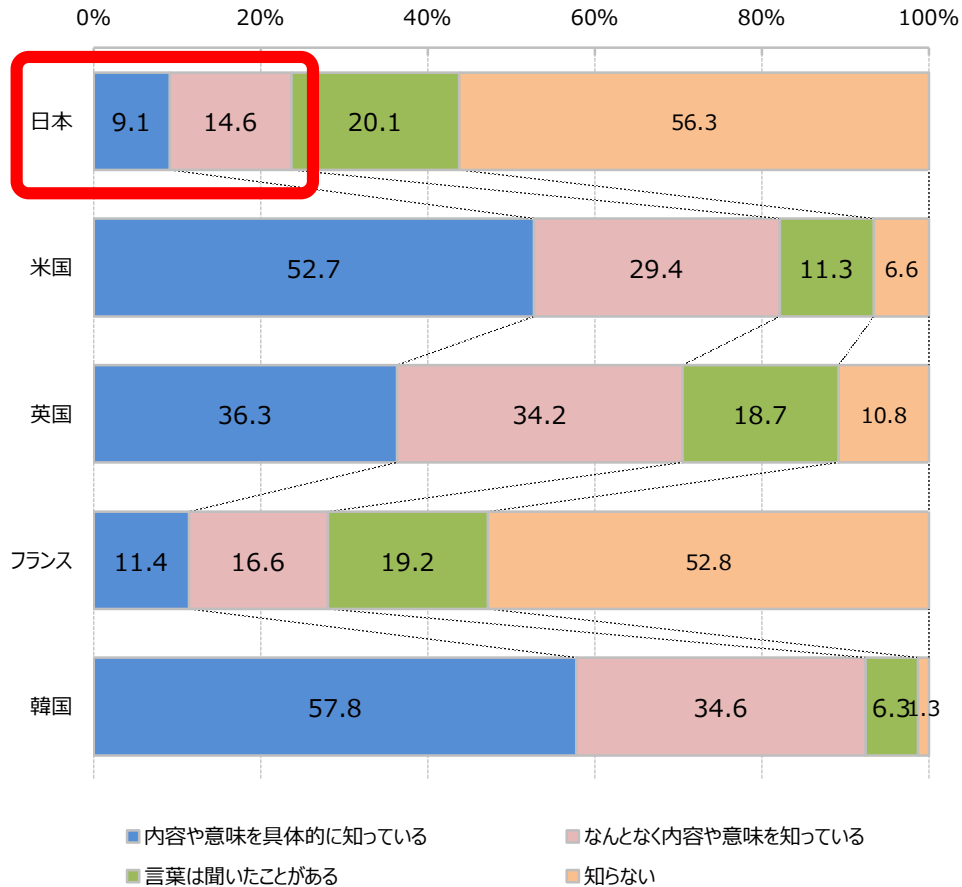
ファクトチェックを実施し
記事化して配信

ファクトチェックの認知度

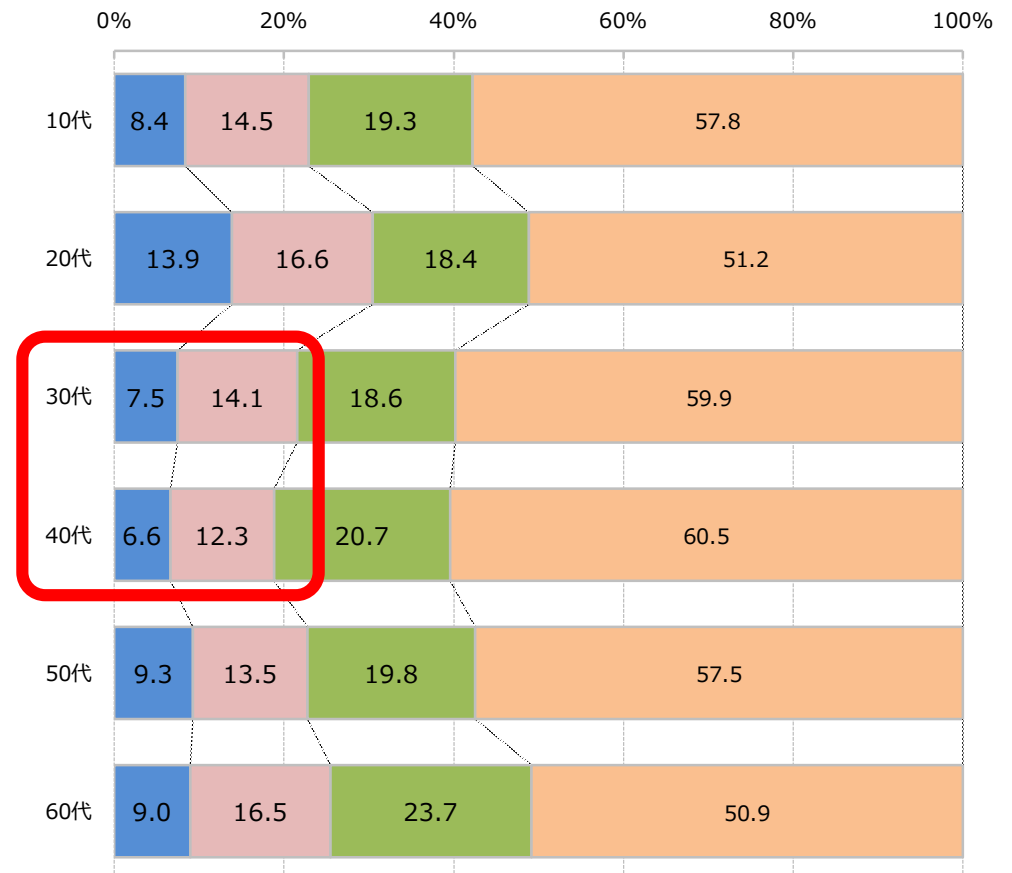
- 我が国における「ファクトチェック」という用語の認知度は、他国と比較すると低い。
- 我が国につき、年代別に見ると、20代と60代では他の年齢層と比較すると高いのに対し、30代及び40代では低い。

【ファクトチェックの認知度】

(国際比較)



(日本：年代別)



日本ファクトチェックセンター（JFC）

- 総務省「プラットフォームサービスに関する研究会」の提言を受け、産学官民の多様なステークホルダーによる協力関係の構築を図り、対話の枠組みを設けることを目的として、「Disinformation対策フォーラム」が設置。
- 同フォーラムでの議論を経て、2022年10月1日、（一社）セーファーインターネット協会に、新たなファクトチェック団体として「日本ファクトチェックセンター（JFC : Japan Fact-check Center）」が設立。

設立団体の概要

➤ 体制：

運営委員長：曾我部 真裕 京都大学教授

監査委員長：穴戸 常寿 東京大学大学院法学政治学研究科教授

編集長：古田 大輔 ジャーナリスト/NPO法人ファクトチェック・イニシアティブ理事

➤ 業務：ファクトチェックの実施及びリテラシー教育の推進等の幅広い偽情報・誤情報対策を推進

➤ 運営資金（発足時）：Google.org、Yahoo



JFC ロゴ

ファクトチェック実績

➤ 2022年10月～2023年4月：記事89本

これまでの経緯

2020年2月：プラットフォームサービスに関する研究会最終報告書の公表

2020年6月：Disinformation対策フォーラムの設立

2022年3月：Disinformation対策フォーラム報告書の公表

2022年10月：日本ファクトチェックセンター（JFC）の設立

2023年5月：ICFN（国際ファクトチェックネットワーク）に加盟

JFC ファクトチェック判定基準

| | | |
|---|------|---|
| ✓ | 正確 | 誤りが無く、重要な要素が欠けていない。 |
| | ほぼ正確 | 一部に誤りを含んでいるが、重要な部分を含む大部分は正しく、十分に正確な言説である。 |
| ✓ | 根拠不明 | 根拠がないか不十分であり、事実の検証ができない。 |
| ✗ | 不正確 | 一部は正しいが、重要な部分に誤りや又は欠落がある。またはミスリード。 |
| | 誤り | 誤りである。又は重要な要素が大きくかけている。 |

国際ファクトチェックネットワーク (IFCN)

- 世界中のファクトチェッカー及び事実に基づく情報の擁護者のコミュニティを結集させる目的で、米国において、非営利組織としてPoynter Institute (ポインター学院) が2015年に設立。
- ファクトチェック団体が遵守すべき**行動規範 (Code of Principles)** を規定、趣旨に賛同するファクトチェック団体に対して審査の上で**認証を付与**。

【IFCN行動規範 (Code of Principles)】

1. **非党派性及び公平性** (A commitment to Non-partnership and Fairness)
2. **情報源の透明性** (A commitment to Standards and Transparency of Sources)
3. **財源及び組織の透明性** (A commitment to Transparency of Funding & Organization)
4. **方法論の透明性** (A commitment to Standards and Transparency of Methodology)
5. **明確かつ誠実な訂正** (A commitment to an Open & Honest Corrections Policy)

※ 認証取得 (加盟) 団体数 (2023年11月3日時点) : 108
うち日本を拠点とする団体数 : 3 (InFact、JFC、リトマス)

国際ファクトチェックデー (毎年4月2日)

- すべての人々に、政治やジャーナリズム、そして日常生活において事実の大切さを呼びかける日として、IFCNが2017年に制定。
- この日の前後に、世界各地でファクトチェックに関連するワークショップなど様々な行事が開催。

偽・誤情報検知等を目的とする研究開発①

- 諸外国（一部日本含む）における偽・誤情報検知等を目的に研究開発が行われたICTツール例について、「コンテンツの検証」、「リソースの信頼性検証」、「フェイクニュース検出」、「コンテンツの検証作業支援」等の目的毎に分類。
- 日本においては、インターネットの「信頼」には早期から取り組んでおり、TrustedWeb構想や、最近では、情動的健康的の共同宣言、Originator Profile技術の取組が行われている。

| No. | ツール名 | 目的 | 内容 |
|-----|---------------------------------------|--|---|
| 1 | SYNTHETIQ VISION (日本) | <ul style="list-style-type: none"> • コンテンツ検証（動画、画像） | <ul style="list-style-type: none"> • AIにより生成されたフェイク顔映像を自動判定するプログラムを開発。 • 国立情報学研究所シンセティックメディア国際研究センター長の越前功氏と副センター長の山岸順一氏のグループが開発。サイバーエージェントが採用した。 |
| 2 | Microsoft Video Authenticator | <ul style="list-style-type: none"> • コンテンツ検証（動画、画像） | <ul style="list-style-type: none"> • ディープフェイク動画、画像を検出する技術。 • リアルタイムで動画の信頼性が表示される。ディープフェイク部分を赤枠で示される。 • マイクロソフトが開発。 |
| 3 | Reality Defender | <ul style="list-style-type: none"> • コンテンツ検証（動画、画像） | <ul style="list-style-type: none"> • ディープフェイクやGenerative AIで生成されたメディアを検出することができる。 • NATO、米国防総省、米国土安全保障省などの政府機関や、米国やアジアの放送メディア等が真偽検証用に利用。 • Reality Defenderが開発。 |
| 4 | SBIs(Self-Blended Images) (日本) | <ul style="list-style-type: none"> • コンテンツ検証（動画、画像） | <ul style="list-style-type: none"> • ディープフェイク検出AIを開発。 • 東京大学情報理工学系研究科電子情報学 准教授山崎俊彦氏と院生による研究。 |
| 5 | FakeCatcher | <ul style="list-style-type: none"> • コンテンツ検証（動画） | <ul style="list-style-type: none"> • 人間の生物学的信号とデータを使用してディープフェイクを96%の精度で識別および分類するツールを開発。 • インテルとピンガムトン大学のグラフィックス&イメージコンピューティング研究所の共同研究。 |
| 6 | deepware scanner | <ul style="list-style-type: none"> • コンテンツ検証（動画） | <ul style="list-style-type: none"> • YouTube, Facebook, Twitter上の動画の信頼性をリアルタイムで表示する。 • ディープフェイク動画を検知するとツール内のメーターが赤を示す。 • Deepwareが開発。PCのウェブブラウザ、Androidのアプリで利用できる。 |
| 7 | TinEye | <ul style="list-style-type: none"> • コンテンツ検証（画像） | <ul style="list-style-type: none"> • リバースイメージ検索ツール。「Idée Inc. (カナダ)」が提供する。 |
| 8 | The News Provenance Project | <ul style="list-style-type: none"> • コンテンツ検証（画像） | <ul style="list-style-type: none"> • 偽・誤情報を防止するために、ブロックチェーン技術を利用してニュース画像の信頼性を確認できる基盤を構築する。IBMのHyper Ledger Fabric上に構築される。 • New York Timesが2019年に計画を公表した。 |

偽・誤情報検知等を目的とする研究開発②

| No. | ツール名 | 目的 | 内容 |
|-----|---|--|--|
| 9 | Sphere | <ul style="list-style-type: none"> コンテンツ検証（テキスト） | <ul style="list-style-type: none"> 記事の真正性を評価する偽情報検出AIツール。開発にあたってはWikipedia上のデータを使用した。 Metaが開発。 |
| 10 | WeVerify/InVID | <ul style="list-style-type: none"> コンテンツ検証（テキスト、動画、画像） リソースの信頼性検証（作成者） | <ul style="list-style-type: none"> EUのHorizon 2020の資金拠出を受けて開発された動画・画像検証ツール。 Ontotext ADがプロジェクト代表機関。 機械学習技術を用いて、ソーシャルメディア上のテキスト、画像、動画を分析しコンテンツの出典を明らかにし、また、大量データから偽情報を作成・宣伝するアカウントを判別する。 検証をおこなった偽情報の画像や動画を蓄積した分散型データベースを作成して公開する（ブロックチェーンを使用）。 |
| 11 | NewsGuard | <ul style="list-style-type: none"> リソースの信頼性検証 | <ul style="list-style-type: none"> ニュースメディアの信頼性評価を行う。AIではなく人間が評価する点が特徴。 |
| 12 | Disinformation Index (GDI) | <ul style="list-style-type: none"> リソースの信頼性検証 | <ul style="list-style-type: none"> 特定のメディアに関する偽情報の可能性を評価、国単位での評価も実施。 |
| 13 | TruthNest | <ul style="list-style-type: none"> リソースの信頼性検証（作成者） | <ul style="list-style-type: none"> Twitter上の情報を対象として、リソースの信頼性検証を行うための支援ツール。 アカウントの投稿傾向を分析して、Botや偽アカウントの検出を行うことができる。 |
| 14 | SocialTruth Project | <ul style="list-style-type: none"> リソースの信頼性検証（コンテンツ、作成者） フェイクニュース検出 | <ul style="list-style-type: none"> AI、ブロックチェーンを活用し、①SNS上に流通するコンテンツや発信者（著者）の信頼性検証、②偽情報の増加検出を行う。 画像、ビデオ、テキストが対象。 |
| 15 | Grover | <ul style="list-style-type: none"> フェイクニュース検出 | <ul style="list-style-type: none"> 英語の文章を対象に人間かAIが書いたものか否かを92%以上の精度で判別できる（2019年の公表時点において）。 アレン人工知能研究所の研究。 |
| 16 | SummarizeBot API 「Fake News Detection」 | <ul style="list-style-type: none"> フェイクニュース検出 | <ul style="list-style-type: none"> 文章の要約サービス。フェイクニュース検出も可能。URL先のニュース記事文章をAIが分析し、事実/偽の確率を表示する。 API形式で提供。対象文章のURLを入力すると、判定結果が表示される。 |

偽・誤情報検知等を目的とする研究開発③

| No. | ツール名 | 目的 | 内容 |
|-----|--|--|---|
| 17 | ファクトチェック支援システム (日本) | <ul style="list-style-type: none"> コンテンツの検証作業支援 | <ul style="list-style-type: none"> 疑似言説収集システム (FCC) + 疑似言説データベース (Claim Monitor) + ファクトチェックナビから構成される。 認定NPO法人ファクトチェック・イニシアティブ (FIJ) が提供する。 11メディア・団体 (新聞社・放送局4社含む)、2つの教育機関が利用。 |
| 18 | Truly Media | <ul style="list-style-type: none"> コンテンツの検証作業支援 コンテンツ検証 (画像) | <ul style="list-style-type: none"> Twitter、Facebook、YouTube等上から集めたソーシャルメディア上のコンテンツの検証を支援するためのツール。 画像分析アルゴリズムにより画像の加工状況を分析する。 |
| 19 | ClaimBuster | <ul style="list-style-type: none"> コンテンツの検証作業支援 フェイクニュース検出 | <ul style="list-style-type: none"> Webベースの自動化されたライブファクトチェックツール。事実と虚偽の情報を特定する。 テキサス大学アーリントン校が開発。 |
| 20 | Full Fact AI 「Automated Fact Checking」 | <ul style="list-style-type: none"> コンテンツの検証作業支援 コンテンツ検証 (テキスト) | <ul style="list-style-type: none"> 偽誤情報の研究機関である「FULLFACT」がGoogleの言語モデル「BERT」を利用したAI分析ツールを開発し、実運用している。有料ライセンスで外部への提供もおこなう。 BERTは膨大なテキスト化された文章データの中からクレーム部分を検出して分類化 (原因と結果、数量など) する。 |
| 20 | FactChat | <ul style="list-style-type: none"> ライブファクトチェック結果共有 | <ul style="list-style-type: none"> 2020年米大統領選期間中、メッセージングアプリの「WhatsApp」上に、米国内の10のファクトチェッカーがFactChatに、ファクトチェック結果を英語とスペイン語で登録公開した。 Poynter InstituteのIFCN (米国) が開発・提供した。 |
| 21 | フェイクアラート (日本) | <ul style="list-style-type: none"> フェイクニュースへのアラート | <ul style="list-style-type: none"> ネット上で、意見の分断が起きている、少数の人が発信しているニュース記事に対してアラートを出す仕組みを開発し実証を行った (2022年)。 TDAI Lab、東京大学鳥海教授、NHKの共同研究、実証。 |

4. 総務省による主な取組状況（上記2. 以外の取組）

偽・誤情報に関する実態調査・分析（リテラシーの重要性）

- 我が国における偽・誤情報に関する実態調査・分析結果によれば、
 - メディアリテラシーが高いほど偽・誤情報と気づく傾向。
 - メディアリテラシー・情報リテラシーが高いほど偽・誤情報を拡散しにくい傾向。

<メディアリテラシー・情報リテラシーと偽・誤情報の真偽判断・拡散行動の回帰分析結果>

コロナワクチン関連の偽・誤情報の真偽判断に対する効果

- メディアリテラシーが1点上昇
⇒ 偽・誤情報と気づく確率が12%増
- 情報リテラシーが1点上昇
⇒ 偽・誤情報と気づく確率が1.8%増

リテラシーが高いほど偽・誤情報と気づく傾向。
特に「メディアリテラシー」はその相関関係が強い。

コロナワクチン関連の偽・誤情報の拡散行動に対する効果

- メディアリテラシーが1点上昇
⇒ 偽・誤情報を拡散する確率が9%減
- 情報リテラシーが1点上昇
⇒ 偽・誤情報を拡散する確率が2%減

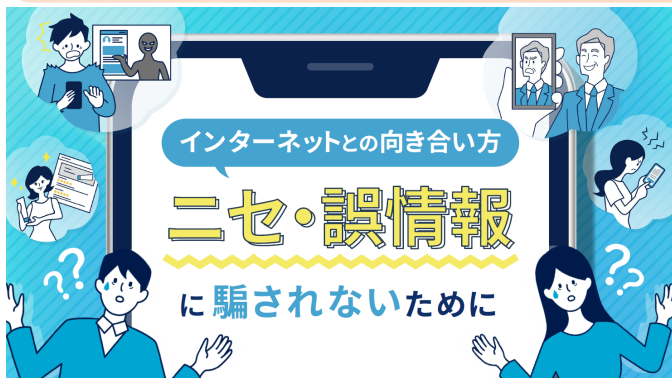
リテラシーが高いほど偽・誤情報を拡散しにくい傾向。
特に「メディアリテラシー」はその相関関係が強い。

- ※ リテラシーを測る設問への回答状況から4点満点で数値化したリテラシーが、1点上がる時に真偽判断・拡散行動の確率がどの程度変化するかを表す。
- ※ 当該分析におけるメディアリテラシー：メディアあるいはメディアメッセージを読み解く能力。メディアリテラシーの種類を「メディアメッセージの構成性」「メディアによる「社会的現実」の構成力」「メディアの商業的性質」「メディアのイデオロギー・価値観伝達」「メディアの様式と言語」「受け手の非画一的解釈性」としている。
- ※ 当該分析における情報リテラシー：「加工されていない生のデータが何かわかる」「筆者の意見が入った文章かわかる」「文章から確実に言えることが何かわかる」から構成される情報を読み解く能力。

偽・誤情報に関する啓発教育教材等の開発

- 総務省において、偽・誤情報に関する啓発教育教材として、「**インターネットとの向き合い方 ～ニセ・誤情報に騙されないために～**」とその**講師用ガイドライン**を開発し、2022年6月に公表。

- (対象者) **若年層～成年層の幅広い年齢層**を対象として作成。
- (所要時間) 60分程度の講義での実施を想定。
- (形式) オンラインでも実施が可能な内容としており、編集が容易なパワーポイント形式にて公表することで、講師の裁量により事例等を追加することが可能。
- (備考) 講座を開催する講師向けに、講師用ガイドラインも用意。各スライドで話す内容を詳細に記載し、読み上げることで講座を実施可能。



PART 01

種類は**2**つ

- 1 **ニセ情報**
ディスインフォメーション
- 2 **誤情報**
ミスインフォメーション

ニセ・誤情報ってなに？

PART 02

「ニセ・誤情報」に気付かない人は「**フィルターバブル**」に陥っている可能性も・・・

私たちはなぜ騙されるのか？

ネットニュース、SNS、検索サービスなどには、**その人が欲しがりそうな情報を分析し同じような情報を表示する「アルゴリズム」**と呼ばれる機能があります。

「アルゴリズム」による偏った情報に囲まれると、あたかも**それが世の中の標準だと誤解**してしまう「**フィルターバブル**」という現象に陥ります。

全員が「嫌い」と言っている

PART 03

3 「ニセ・誤情報」を信じて行動した結果**損害賠償責任を負うケースも・・・**

2019年に話題になったあおり運転事件で「加害者の同乗者」であるというニセ・誤情報を投稿された女性が**膨大な量の中傷被害**を受けました。

女性は**投稿者や拡散させた人物の特定を進め、裁判を行い、すでに損害賠償を命ずる判決**も出されています。

騙されたらどうなる？

PART 04

それでも騙される。だから...

- わからなければ拡散しない
- 誰かを傷つけるなら拡散しない
- 医療・健康情報は安易に拡散しない
- リツイート前にひと呼吸
- 手を止めて「間違いでしょ?」
- 異なる情報は出ていないかチェック

騙されないためには？

PART 04

「ニセ・誤情報」に対抗するため、私たちにできること。まずは基本の4つから。

基本

まずは以下をチェック!

- 情報源はある?
- その分野の**専門家**?
- 他ではどう言われている?
- その画像は**本物**?

情報源はある?
その分野の**専門家**?
他ではどう言われている?
その画像は**本物**?

この4つを詳しく見ていきましょう。(クリック)

まず1つ目、情報源はある?

その情報、どこから発信されたものですか? 信用できますか? 根拠となるモノは今も存在しますか? 消えていませんか? 情報源が「海外」のニュースや論文の場合、あなたはその情報源を確認、理解できますか? これらはすべて確認すべきことです。

特に情報源が海外の場合、わざと間違った翻訳をしたり、一部だけ切り取ったりすることで、内容や結論を捏造している「ニセ・誤情報」がたくさんあります。(クリック)

「インターネットトラブル事例集」の作成・公表

- 子育てや教育の現場での保護者や教職員の活用に資するため、インターネットに係るトラブル事例の予防法等をまとめた「インターネットトラブル事例集」を2009年度より毎年更新・作成し公表。
- 2023年3月31日に同年版を公表。

実際に起きていることでネットの使い方を考えよう！

インターネットトラブル事例集 (2023年版)

総務省 総合通信基盤局 消費者行政第一課
情報流通行政局 情報流通振興課

https://www.soumu.go.jp/use_the_internet_wisely/trouble/

身近な人と一緒に考えよう

とくせいひと かどうかしんり ちゅうもく
ネットの特性と人の行動心理に注目

特に中・高校生の方へ

画面に表示されるいろいろな情報は、あなたの好みや検索ワードに合わせて日々カスタマイズされているんです

見ている情報が全てとは限らない！2つのネットの特性

みなさんは、こんな経験ありませんか？

- 検索した言葉や商品と関連する情報や広告がよく表示される・・・
- 「いいね」「フォロー」をすると似たような情報がオススメされてくる・・・

このようなことが起こるのは、**ネットがあなたの使い方合った情報を選んで見せる仕組み**になっているからです。

用語：フィルターバブル
好みや見たい情報以外が知らないうちにはじかれてしまうこと

用語：エコーチェンバー
似たような意見ばかりに囲まれ考えがより増幅されてしまうこと

情報そのものが誤っていることや意図的に作られたウソの情報ということも。いろいろな情報を積極的にチェックする習慣を持つようしてください

| チェック方法 | |
|--|---|
| <p>情報源はある？</p> <ul style="list-style-type: none"> ☑ その情報はどこから、いつ発信されたものですか？ ☑ 根拠となるモノは今も存在していますか？ ☑ 情報源が「海外の」ニュースや論文の場合、あなたは情報源を確認、理解していますか？ | <p>その分野の専門家？</p> <ul style="list-style-type: none"> ☑ その情報は、専門知識や必要な資格を持った人が責任を持って発信しているものですか？ ☑ その人は過去、ニセ・誤情報を発信して批判されていませんか？ ☑ その人は関連する情報や商品を買っていませんか？ |
| <p>他ではどう言われている？</p> <ul style="list-style-type: none"> ☑ その情報について他の人や他のメディアはどのように言っていますか？ ☑ その人の意見に反論している人はいませんか？ ☑ 別の内容で報じているメディアや、誤りであることを指摘しているメディアはありませんか？ | <p>その画像は本物？</p> <ul style="list-style-type: none"> ☑ 臨場感のある画像が添えられているから？ それだけで「本物」だと判断して大丈夫ですか？ ☑ 画像を検索したら、全く同じ画像がヒットしませんか？ ☑ その画像は過去に撮影された、全く無関係のものではありませんか？ |

① トラブル事例と解決のヒントをこどもにわかりやすくマンガで解説



② スマホやタブレットでも閲覧しやすいバージョンの作成



③ 授業等でご活用いただくための教職員用参考資料

【マンガをよんで考えてみよう！】

このトラブルを発生させないために、どうしたら良いでしょうか？

1 文字だけのコミュニケーションは意外と難しい！

2 画像は本物か？

3 情報源を確認しよう！

4 専門家を確認しよう！

5 他ではどう言われているか？

6 別メディアで報じているか？

7 別の内容で報じているか？

8 別のメディアで報じているか？

9 別のメディアで報じているか？

10 別のメディアで報じているか？

総務省ウェブサイトにおける周知・啓発

- **一般利用者向け**の総務省ウェブページ「上手にネットと付き合いよう！安心・安全なインターネット利用ガイド」にて、特集ページとして、「**ネットの時代におけるデマやフェイクニュース等の不確かな情報**」を掲載。

ネット上には、人を混乱させるためにわざと流された**デマ情報**も。
身近な医療・健康情報、うわさ話やゴシップネタなどにも、**間違った情報**があります！

どのようなメディアでも、間違った情報が流布されることはあります。
特にSNSの場合、誰もが容易に情報発信できることから、正しくない情報も多くあります。
中には、騒ぎを起こすことが目的で発信されたデマも。目にした情報をうのみせず、正確性が判断できない場合には安易に情報を投稿・拡散しないことが大切です。



確認方法

他の情報と比べてみる

ネット検索し、複数の情報を読み比べましょう。本や新聞など、ネット以外で調べるのもお勧めです。

その情報はいつ頃書かれたものか確かめる

元の情報が古いものだった場合、現在とは状況が異なるかもしれないので、注意しましょう。

情報の発信元を確かめる

発信元が明らかであっても、信頼できる人なのか、信頼できるWebサイトなのかを確認しましょう。

一次情報を確認する

その情報が引用や伝聞だった場合は、元になったオリジナルの情報源を探して確かめてみましょう。

フェイクニュースの拡散経験・拡散手段

全体では、「拡散したことはない」人が約7割と最も高い。
一方で、全体の約15%が「拡散した経験がある」と回答。若い年代ほど「拡散した経験がある」と回答した割合が高い傾向が見られた。
拡散手段をみると、「インターネットでは拡散しなかったが、家族や知人に直接伝えて話したことがある」、「Facebookのシェアや、Twitterのリツイートを行ったことがある」が高かった。



フェイクニュース拡散後にとった行動

フェイクニュースを拡散してしまったと気づいた後、ほとんどの人が訂正をしていた。
訂正方法は、「直接伝えた家族や知人に対し、情報が間違っていたことを伝えた」、「Facebookのシェアや、Twitterでのリツイートを解除したり、SNSやブログ上の投稿を削除した」、「SNSやブログ上で、自分が拡散投稿した情報が間違っていたことを知らせる情報を新たに投稿した」の順に多かった。
一方で、「情報が間違っていることに気づきつつ、様々な方法で情報をもっと拡散した」、「情報が間違っていたことに気づいたが、何もなかった」人も存在した。

調査の詳細
https://www.soumu.go.jp/main_content/000693284.pdf

日本におけるフェイクニュースの実態

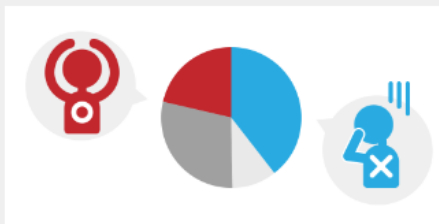
フェイクニュース等の用語の認知状況

「フェイクニュース」という用語について知っている人は全体の約8割。
一方で、「ディープフェイク」や、ファクトチェックという用語を知っている人は全体の約2割。

- ・フェイクニュース：定まった定義はないが、何らかの利益を得ることや意図的に騙すことを目的としたいわゆる「高情報」や、単に誤った情報である「誤情報」や「デマ」などを広く指すもの
- ・ディープフェイク (Deepfake)：AI技術や機械学習の技術を用いて作り出された偽の映像
- ・ファクトチェック：偽情報に対抗するために、情報の真偽を検証する活動のこと



フェイクニュースを見分ける自信



フェイクニュースを見分ける自信について、全体では、「自信がない」人が約4割。全ての年代において「自信がない」が「自信がある」を上回った。

フェイクニュースを見かける頻度



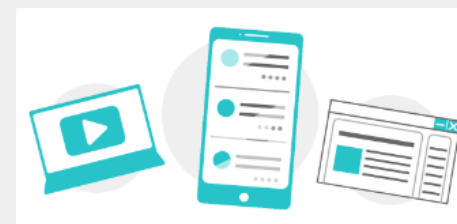
SNSやブログなどでフェイクニュースをみかけた頻度が「週1回以上」となったのは全体の約3割。

見たことのあるフェイクニュースのジャンル



見たことのあるフェイクニュースのジャンルについては、「スポーツ・芸能・文化」(43%)、「社会・事件」(38.3%)、「生活・健康」(26%)、「災害」(25.3%)の順に多かった。

フェイクニュースに接することが多いと思う情報源



フェイクニュースに接することが多いと思う情報源について、「SNS」が全体で約6割であり、全ての年代において最も高かった。

YouTubeクリエイターによる啓発動画の制作・配信、啓発イベントの開催

- YouTubeクリエイター 9 者（チャンネル登録数：延べ1,210万人）の協力により、**フェイクニュース対策啓発動画**を作成・公開。
- **G7関連イベント「フェイクニュースと日本」**の開催と合わせて動画を2023年 4 月25日に公開。

1. YouTubeによる啓発動画キャンペーン

YouTubeクリエイター 9 者（チャンネル登録数：延べ1,210万人）の協力により、フェイクニュース対策啓発動画を作成・公開。

2023年 4 月25日に動画を公開し、28日間（5月22日時点）で再生回数759万回、高評価数41万、コメント数4,500件を記録。

啓発動画キャンペーンURL：https://youtube.com/playlist?list=PLQntWbrycbJcpM6aVvc5gnP_HMxPF5weB

2. G7関連イベント「フェイクニュースと日本：私たちにできること・社会としてできること」

開催日時：2023年 4 月26日（水） 15:00-18:00

形式：対面とYouTubeライブ配信のハイブリッド

主催：国際大学グローバル・コミュニケーション・センター（GLOCOM）

後援：グーグル合同会社、経済産業省、総務省、デジタル庁

協力：スマートニュース株式会社、スローニュース株式会社、一般社団法人セーファーインターネット協会、一般社団法人ソーシャルメディア利用環境整備機構（SMAJ）、特定非営利活動法人ファクトチェック・イニシアティブ（FIJ）、ヤフー株式会社、LINE株式会社

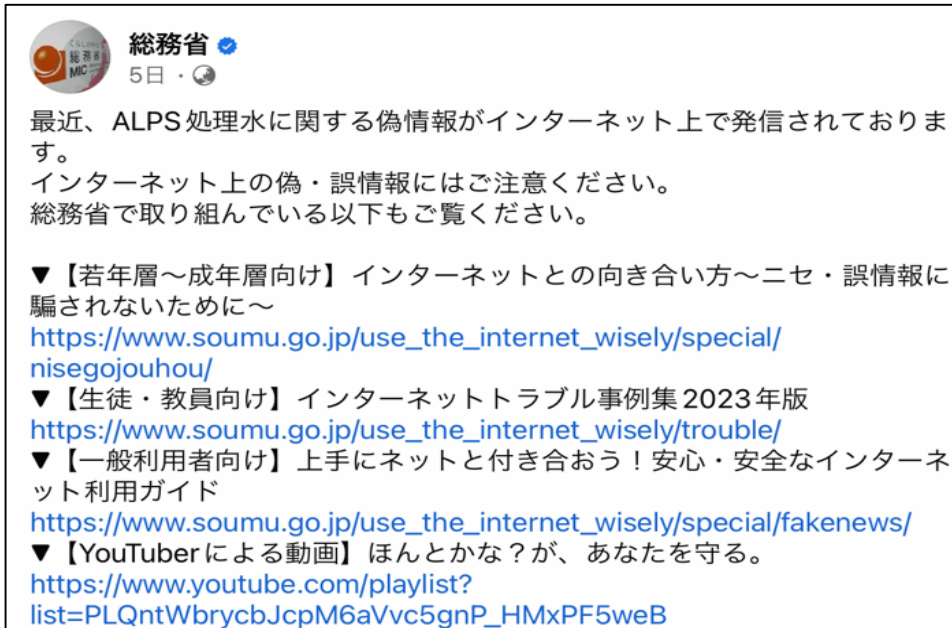
※背景と議論するトピック

新型コロナウイルスパンデミック、ロシアのウクライナ侵攻、選挙などで多くのフェイクニュースが拡散。フェイクニュースを見聞きした人の約4人に3人が誤りだと気づいていない。本イベントでは、人々がどのように気を付ければよいのか、産官学民がどのように対抗すればよいのか議論。

イベント動画URL：<https://www.glocom.ac.jp/events/information/8716>

総務省SNSアカウントによる情報発信

- 2023年夏頃、外務省や経済産業省等において、ALPS処理水の海洋放出等に関する偽情報の拡散が確認。
- 総務省SNSアカウント（X（旧twitter）、FB、Instagram）において、インターネット上の偽・誤情報に関する一般的・基本的な注意と、各種啓発コンテンツの紹介を情報発信。



最近、ALPS処理水に関する偽情報がインターネット上で発信されております。
インターネット上の偽・誤情報にはご注意ください。
総務省で取り組んでいる以下もご覧ください。

▼【若年層～成年層向け】インターネットとの向き合い方～ニセ・誤情報に騙されないために～
https://www.soumu.go.jp/use_the_internet_wisely/special/nisegojohou/

▼【生徒・教員向け】インターネットトラブル事例集2023年版
https://www.soumu.go.jp/use_the_internet_wisely/trouble/

▼【一般利用者向け】上手にネットと付き合いおう！安心・安全なインターネット利用ガイド
https://www.soumu.go.jp/use_the_internet_wisely/special/fakenews/

▼【YouTuberによる動画】ほんとかな？が、あなたを守る。
https://www.youtube.com/playlist?list=PLQntWbrycbJcpM6aVvc5gnP_HMxPF5weB



外務省
Ministry of Foreign Affairs of Japan

報道発表

外務省のものとする偽文書に関する報道について

令和5年8月14日
英語版 (English)

- 8月14日、外国の自称「メディア」により、7月8日に東京電力が実施した処理水のタンク群の調査の結果、放射能濃度が基準を大幅に超過したため、プラスチックの交換によってALPS処理水の希釈を加速し、安全基準を満たすことが検討されているかのごとき内容が記載された、外務省の公電とされる文書に関する報道がなされました。
- 上述の内容は事実無根です。また、外務省の公電とされる文書は全くの偽物です。
- 外務省幹部とされる人物のALPS処理水に関する発言を記したとされるメモ、IAEA事務局の内部文書に見せかけた文書に加え、外務省の公電を偽造するなど、手段を選ばない悪意のある偽情報の拡散は、我々の社会が基礎を置いている自由や民主主義といった普遍的価値に対する脅威であるほか、被災地の復興を妨げ、復興に向け努力する被災地の人々の感情をも大きく傷つけるものであり、政府として、偽情報やその流布に断固として反対します。

(参考) ALPS処理水
ALPS（多核種除去設備（Advanced Liquid Processing System））等により、トリチウム以外の放射性物質について安全に関する規制基準値を確実に下回るまで浄化した水。さらにALPS処理水は、その後十分に希釈され、トリチウムを含む放射性物質の濃度について安全に関する規制基準値を大幅に下回るレベルにした上で、海洋放出されることが想定されている。

基本 まずは以下をチェック!

- 情報源はある?
- その分野の専門家?
- 他ではどう言われている?
- その画像は本物?

応用 さらに以下もチェック!

- 「知り合いだから」という理由だけで信じているのでは?
- 表やグラフも疑ってみた?
- その情報に動機はある?
- ファクトチェック結果は?



経済産業省
Ministry of Economy, Trade and Industry

ALPS処理水の海洋放出に関する偽情報について

2023年9月2日

- ALPS処理水の海洋放出に関し、「核汚染水」との用語を用いて西村経済産業大臣の8月25日の記者会見動画を加工・改ざんする形で、日本の水産物輸出の取扱いを誤解させるような、事実無根の情報がネット上で発信されています。
- ALPS処理水海洋放出に関する悪意のある偽情報の拡散は、被災地の復興を妨げ、復興に向け努力する被災地の人々の感情をも大きく傷つけるものです。
- 日本政府は、今後もALPS処理水について、高い透明性をもって、科学的根拠に基づく丁寧な情報提供を続けるとともに、偽情報やその流布に断固として反対します。

【出典】総務省SNSアカウント、外務省ウェブページ (https://www.mofa.go.jp/mofaj/press/release/press1_001532.html)
経済産業省ウェブページ (<https://www.meti.go.jp/press/2023/09/20230902002/20230902002.html>)

ICTリテラシー向上に向けたこれまでの総務省の施策

- 総務省は、電気通信サービスの消費者利益の保護の観点や、青少年インターネット環境整備法に基づき、これまでも、青少年が安全に安心してインターネットを利用できるための環境整備に向けて、インターネット・リテラシーの向上のための青少年・保護者に対する普及啓発活動等を中心に実施。
- 学校教育外においても、青少年がSNSをはじめとしたオンラインサービスをスマートフォンなどの端末を通じて日常的に利用する機会が増えていることや、自分の好みの情報、自分と似た意見ばかりに触れやすくなるオンラインサービスの特性も踏まえ、適切なICT活用のために必要なリテラシー啓発施策の在り方を見直す必要。
- 上述の状況・課題は成年層や高齢層にも当てはまることや、世代に応じた課題が存在。

| | 【若年層向け施策】 | 【成年層向け施策】 | 【高齢層向け施策】 |
|-------------------|--|--------------------|--------------|
| ICTの適切な活用のための施策 | | ※ | デジタル活用支援推進事業 |
| ICT利用時の危険回避のための施策 | e-ネットキャラバン(小中高生等向け出前講座) ILAS (青少年がインターネットを安全に安心して活用するためのリテラシー指標) インターネットトラブル事例集 偽・誤情報に関する啓発教育教材 「インターネットとの向き合い方～ニセ・誤情報に騙されないために～」 啓発サイト「上手にネットと付き合いおう！～安心・安全なインターネット利用ガイド～」 青少年が安全に安心してインターネットを利用できる環境の整備等に関する法律 | 十分にアプローチできていなかった領域 | |

※ 関係省庁において、DXを推進するデジタルスキルの養成に係る取組を実施。

ICTリテラシーの向上に向けた2023年度の取組①

- リテラシーの全体像のフレームとして、デジタル社会において身につけるべき5つの能力を定義し、4段階の習熟度を設定。
- 作成したリテラシー全体像のフレームを活用することにより政府・民間事業者の取り組むべき施策・サービスの明確化や目標設定、有効性の評価が可能となる。

【短期的取組事項】

習熟度に応じた身に付けるべき能力の具体化の検討、社会全体のリテラシー習熟度を計測する方法の検討。

これからのデジタル社会において身に付けるべき能力（素案）

- a. データや情報、デジタルコンテンツを検索、評価、管理する能力
- b. デジタル空間において安全を確保する能力
- c. デジタル技術を通じて他者や社会と関わる能力
- d. デジタルコンテンツの作成・編集に関する能力
- e. デジタル技術の利用に当たっての課題解決やデジタルツールを用いた課題解決に関する能力

リテラシーの全体像のフレーム

| 習熟度 身に付けるべき能力 | Lv.1 人に助けってもらえばできる (自分自身で生活に必要なデジタルサービスを十分に活用できない) | Lv.2 自分自身でデジタル技術を利用できる (生活に必要なデジタルサービスを活用できる) | Lv.3 基礎的なリテラシーを一通り理解 (主体的に学ぶ、善し悪しを判断、インターネットの特性を理解) | Lv.4 デジタル空間の公共性を踏まえてデジタルを生活の中で使いこなす (発信者としての責任、公共への貢献) |
|--|--|---|---|--|
| a. データや情報、デジタルコンテンツを検索、評価、管理する能力 | | | | |
| b. デジタル空間において安全を確保する能力 | | | | |
| c. デジタル技術を通じて他者や社会と関わる能力 | | | | |
| d. デジタルコンテンツの作成・編集に関する能力 | | | | |
| e. デジタル技術の利用に当たっての課題解決やデジタルツールを用いた課題解決に関する能力 | | | | |

⇒各能力の具体化を検討

5つの能力に対して、
4段階の習熟度を設定

ICTリテラシーの向上に向けた2023年度の取組②

- 世代に共通するICT活用に当たっての課題について、以下のとおり整理。

(1) インターネット上の情報流通に関する仕組みの理解

- ・ デジタル空間における様々なシステム・ネットワーク等の技術的な仕組みの理解が不十分である。
- ・ SNSや検索エンジンで表示される情報に関する「アルゴリズム」について、世代を問わず、その仕組みや効果、影響に関する知識が十分ではない。
- ・ 「認知バイアス」により、無意識のうちに合理的ではない行動、偏った判断をすることがある。

(2) 情報を理解するリテラシー（事実と意見、推測、判断、行動の切り分け等）の習得

- ・ 事実の提示、推測、判断、行動の論理の切り分けが理解できないという問題がある。
- ・ かなりの人が間違った情報に触れており、そのうち2割程度しか偽・誤情報と見分けられなかったという民間の調査結果もあり、偽・誤情報に対する弱さは各世代であまり変わらない。

(3) インターネット上の情報を熟慮する機会の確保

- ・ インターネット上では「アテンション」が重要視されることを背景に、反射的な思考や反応が重視される環境にある。
- ・ 情報過多の時代であるからこそ、情報を熟慮し、十分な分析や検証を行う機会を確保する必要がある。

(4) デジタル空間における情報発信者としての意識や社会参加への意識の醸成

- ・ デジタル空間では、誰でも容易に発信ができ、他人を傷つけてしまう機会も増え、責任が発生するようになっており、情報発信者であることの認識が求められる。
- ・ メディアを使いこなした上で、自分のメッセージを届けて、社会に働きかけるトレーニングが必要。

 **以上の共通課題に対するリテラシー向上のためのコンテンツの開発、検証を実施。**

開発、検証にあたっては、公共施設やプラットフォーム事業者との連携を実施。

(例) ワークショップ形式の双方向型講座の検証、無関心層へもリーチするコンテンツの作成

ICTリテラシーの向上に向けた2023年度の取組③

【中長期的取組事項】

- 優先すべきセグメントとして青少年層、保護者層、高齢者層と設定。
- セグメントによって、よく接触するメディア、強化すべき領域等が異なると考えられることから、各セグメントに固有の特徴に関する調査・分析を行った上で、各セグメントの特徴を踏まえたコンテンツ、届け方について検討を深めることが必要。

構成員からの主な御意見

- ・ 年長者の経験が機能しない世界において、各世代へのアプローチが必要であり、普及啓発の場所・方法について具体的な検討が必要。
- ・ 青少年に対しては、「小さな世界で失敗できる」環境の確保が必要。
- ・ 保護者と子どもの関係や、高齢者にどう学習機会を設定して誘因するかを考える必要であり、公共図書館や公民館、生涯学習センター等を活用していくことが重要。

各セグメントに関する調査・分析（R5年度に調査を実施）

- ・ よく接触するメディアやその利用時間、利用目的、デジタルツールの活用状況といった特徴
- ・ ICTの活用に伴い、遭遇する可能性の高い課題（ネット依存、誹謗中傷、偽・誤情報、フィッシング詐欺 など）

対象層の特徴を踏まえたコンテンツ

- ・ 各セグメントに届きやすいチャネルを想定したコンテンツの在り方（動画、テキスト、e-ラーニングなど）
- ・ 各セグメントの課題に応じた、取扱う領域の強弱

対象層の特徴を踏まえた届け方

- ・ 各セグメントに届きやすいチャネルの活用、連携
 - PF事業者との連携
 - 教育の場の活用（e-ネットキャラバンなど）
 - 社員研修の場の活用
 - デジタル活用支援推進事業の活用
- ・ 双方向的な学び

上記を踏まえ、令和6年度予算要求

ICTリテラシーの向上に向けた2023年度の取組④

- 生成AIの特徴・課題を整理し、リテラシーの全体像において、どの能力を重点的に取り組むべきか検討する。
- 生成AIの活用に当たり、重点的に身に付けるべき能力の要素を盛り込んだコンテンツの開発と届け方を整理する。

生成AIの特徴・課題の整理（案）

・AIが提示する回答は、確率的に最もらしいものである

・AIが取り込むデータや情報によって、回答に偏りが生じる

・最終的にAIの回答の真偽を判断する能力が必要である

→課題を深掘りする必要



| | Lv.1 人に助けてもらえ ばできる | Lv.2 自分自身でデジ タル技術を利用でき る | Lv.3 基礎的なリテラ シーを一通り理解 | Lv.4 デジタル空間の公 共性を踏まえてデ ジタルを生活の中 で使いこなす |
|--|--------------------------|-----------------------------------|-----------------------------|--|
| a. データや情報、デ ジタルコンテンツ を検索、評価、 管理する能力 | | | | |
| b. デジタル空間に おいて安全を確 保する能力 | | | | |
| c. デジタル技術を通 じて他者や社会と 関わる能力 | | | | |
| d. デジタルコンテン ツの作成・編集 に関する能力 | | | | |
| e. デジタル技術の 利用に当たって の課題解決やデ ジタルツールを用 いた課題解決に 関する能力 | | | | |

リテラシーの全体像のうち、重点的に取り組むべき能力を検討

※ AIによる指導、評価、認定等を批判的に捉える視点についても配慮

短期的に取り組む事項

- 生成AIの特徴及び課題を整理し、リテラシー全体像のうち重点的に取り組むべき能力を検討する。
- 生成AIを賢く使うために身に付けるべき能力を踏まえ、コンテンツに盛り込むべき要素を整理する。
- 上記を踏まえ、コンテンツを開発する。併せて、コンテンツの対象者や届け方、効果測定についても検討する。

ICTリテラシーの向上に向けた2024年度の取組

- リテラシー向上に関する検討会においてとりまとめられたロードマップに基づき、各年齢層の特徴を踏まえた幅広い世代を対象とした新たなリテラシー向上のためのコンテンツ開発を実施。

幅広い世代を対象とした新たなリテラシー向上の推進

短期的に取り組む事項（R5年度実施）

リテラシーの全体像と指標の作成

- 今まで利活用を含めて整理はなされていない、リテラシーの全体像の枠組みを整理（5つの能力と習熟度レベルのフレーム設定）
⇒R5年度中に各能力の具体化とそれを測る指標を検討

世代共通課題

- 共通課題の整理（・インターネット上の情報流通に関する仕組み等の理解・情報を理解するリテラシー（事実と意見、推測、判断等）の習得等）
⇒R5年度中に青少年層を対象とするe-ネットキャラバン、PFerと連携してコンテンツを開発

優先して対応すべき対象層に関する特徴分析

- 優先して対応すべき対象層の特徴に関する調査、分析を実施し、各対象層の課題を抽出 ● 特徴を踏まえた届け方についても検討

- ⇒R5年度内
- リテラシー全体像、指標作成
 - 共通課題向けコンテンツの開発
 - 対象層の特徴分析

中長期的に取り組む事項（R6年度実施予定）

優先すべき対象層の特徴を踏まえた対策

青少年層

保護者層

高齢者層

AIリテラシー及び偽誤情報の特徴を踏まえた対策

対象層の特徴を踏まえた対策

対象層の特徴を踏まえた対策

対象層の特徴を踏まえた対策

- コンテンツ開発 ● 有効な届け方の検証 ● コンテンツ開発 ● 有効な届け方の検証 ● コンテンツ開発 ● 有効な届け方の検証

教える人材の育成/関係者の取組の連携・協働推進

- 各対象層の教える人向けコンテンツ開発 ● 各関係者の取組のマッピングを踏まえた、効果的な届け方の検証

- ⇒R6年度内
- 生成AI、偽誤情報への対応を含めた、各世代の特徴を踏まえたコンテンツの開発
 - 教える人材向けコンテンツの開発
 - 効果的な届け方のための連携プロセスの検証

ICT活用のための新たなリテラシー向上のコンテンツ開発及びインターネットの安心・安全利用の促進のための周知啓発を進め、幅広い世代のリテラシー向上を推進する。

Ⅱ 地域のDXの推進を支える情報通信環境の整備

8 誰1人取り残さないための取組

[2] 幅広い世代を対象としたICT活用のためのリテラシー向上推進

（1）幅広い世代を対象としたICT活用のためのリテラシー向上推進

- ・SNS利用者の低年齢化、生成AIの普及等を背景にSNSを利用した迷惑行為、**AIにより生成されたデマ画像の拡散等のインターネット上におけるリテラシーに関する問題が顕在化**してきている。
- ・総務省ではリテラシー向上に関する検討会においてとりまとめられたロードマップに基づき、各年齢層の特徴を踏まえた幅広い世代を対象とした新たなリテラシー向上のためのコンテンツ開発を実施。併せて、青少年インターネット環境整備法に基づく、インターネットの安心・安全な利用の促進のための周知啓発を継続して実施。



【予算】幅広い世代を対象としたICT活用のためのリテラシー向上推進事業

2.2億円（4年度2次補正 0.9億円 5年度 0.7億円）

V 国際競争力の強化、国際連携の深化に向けた先導的取組の推進

16 国際競争力の強化に向けたAIなどの科学技術・イノベーションの推進

[2] 生成AI等の開発力・リスク対応力強化等に向けた取組の推進

（1）我が国における大規模言語モデル（LLM）の開発力強化に向けたデータの整備・拡充及びリスク対応力強化

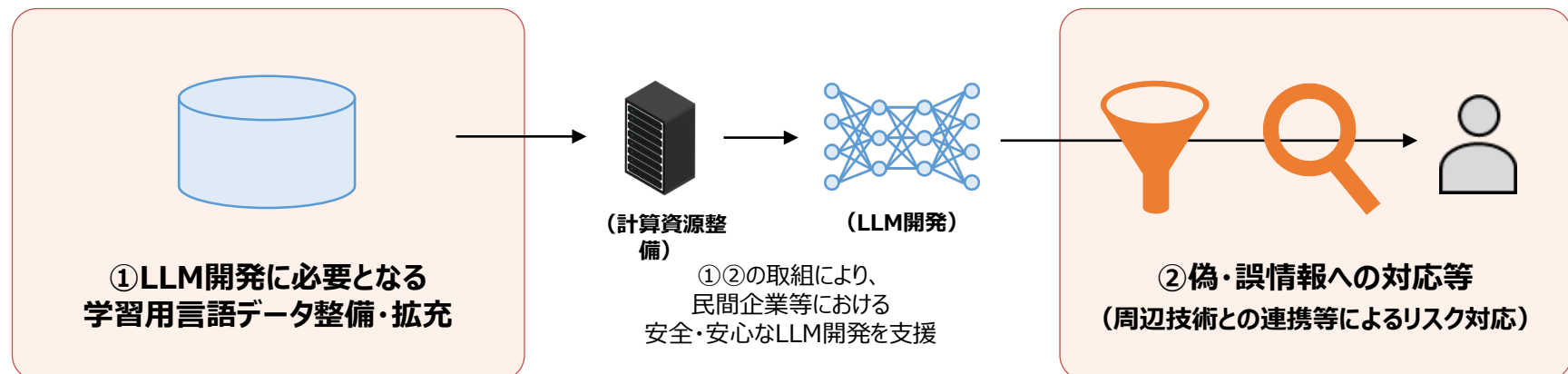
- ・国立研究開発法人情報通信研究機構（NICT）において、民間企業等におけるLLM（※）開発に必要となる大量・高品質で安全性の高い日本語を中心とする学習用言語データを整備・拡充し、法的課題を検討・整理した上で我が国のLLM開発者等にアクセスを提供。

※ Large Language Model：大規模言語モデル

- ・**ディープフェイクに代表される偽・誤情報等、生成AIに起因する様々なリスクに対応するための技術の開発・実証を実施し、成果を関係事業者等に展開。**

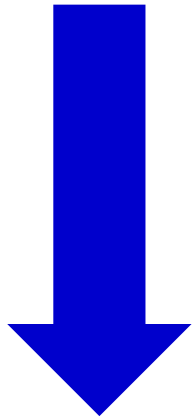
【予算】我が国における大規模言語モデル（LLM）の開発力強化に向けたデータの整備・拡充及びリスク対応力強化 【新規】

（LLM開発から利用までのプロセス 及び 実施施策）



5. AI（特に生成AI）を巡る主な議論

総務省におけるAIのルール作りに対する検討・取組の経緯



2016年4月 G7情報通信大臣会合（高松）

- ・ 開発原則のたたき台を提示

2016年10月 AIネットワーク社会推進会議 設置

2017年3月 国際シンポジウム「AIネットワーク社会推進フォーラム」開催

2017年7月 「国際的な議論のためのAI開発ガイドライン案」作成・公表

2018年7月 「AI利活用原則案」作成・公表



2019年3月 「人間中心のAI社会原則」策定

2019年5月 OECD（経済協力開発機構）「AIに関する理事会勧告」採択

2019年6月 G20「G20 AI原則」採択

2019年8月 「AI利活用ガイドライン」策定・公表



2023年4月 AI戦略チームスタート

2023年5月 AI戦略会議スタート

2023年4月 G7デジタル・技術大臣会合（高崎）

2023年5月 G7広島サミット

- ・ 広島AIプロセスの開始

2023年6月 AIネットワーク社会推進会議 再始動

- ・ AIガイドラインの見直し・統合を議論

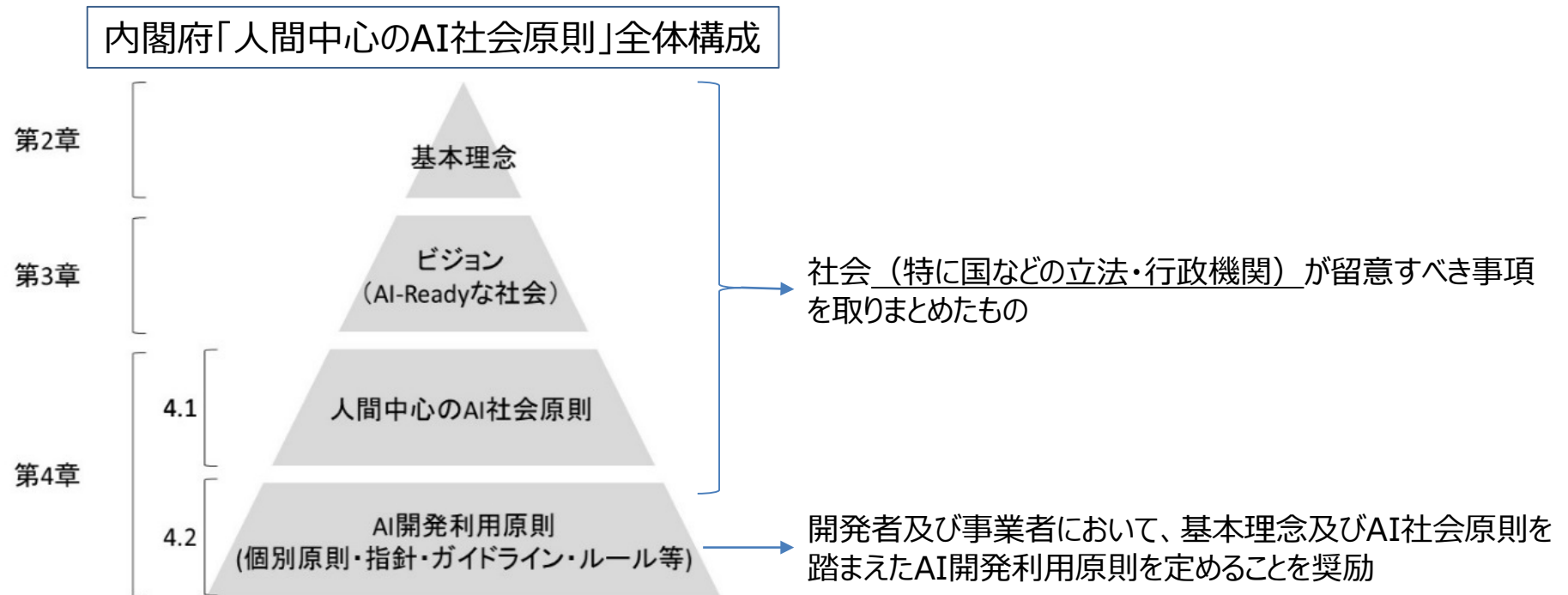
「人間中心のAI社会原則」

2018年5月 内閣府に「人間中心のAI社会原則検討会議」を設置

2019年2月「人間中心のAI社会原則会議」に移行

2019年3月 統合イノベーション戦略推進会議において「人間中心のAI社会原則」を決定・公表

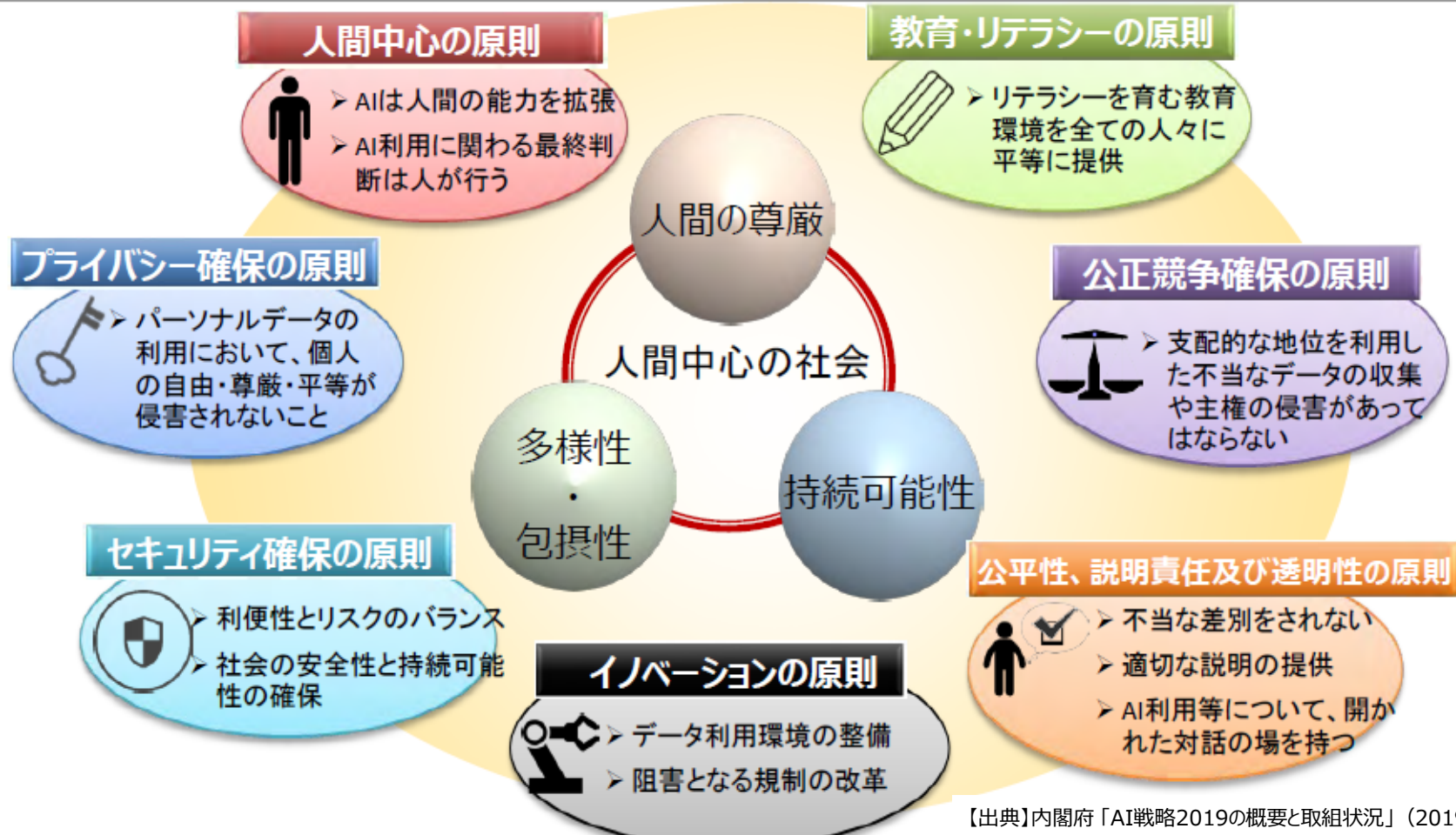
- 当該会議（議長：須藤修先生）では、産学民官のマルチステークホルダーにより、AIをより良い形で社会実装し共有するための基本原則となる「人間中心のAI社会原則」について検討。
- 「人間中心のAI社会原則」は、社会（特に国などの立法・行政機関）が留意すべき原則をまとめたもの。



「人間中心のAI社会原則」

【人間中心のAI社会原則】

- 世界でAIの倫理的側面に関する議論が進展
- AIに関する人々の不安を払拭し、積極的な社会実装を推進するため、我が国としての原則を3月に策定
- 今後、AI社会原則に関する多国間の枠組みを構築



【出典】内閣府「AI戦略2019の概要と取組状況」（2019年11月）

OECDにおけるAI原則（2019年5月）

- AIの開発・実装が進む中、予測可能で、安定的かつ柔軟な開発・利用環境が求められていることから、「人間中心」の考え方を基本に、AIに関わる全ての人に適用される実用的な指針が必要との考えの下、2019年5月に、OECDがAI原則を策定・公表。

| 原則 | 説明 |
|--------------------|---|
| 包摂的な成長、持続可能な開発及び幸福 | AIに関わる全ての人、人間にとって有益な成果を追求するために、AIの責任ある管理・運用に積極的に取り組むべき。 |
| 人間中心の価値及び公平性 | AI開発・運用者は、法の支配、人権及び民主主義的価値観を尊重すべき。その文脈に適合したメカニズムと予防措置を備えるべき。 |
| 透明性及び説明可能性 | AI開発・運用者は、AIシステムへの一般的な理解やAIの影響を受ける人の理解を促進するため、意味のある情報を提供すべき。 |
| 頑健性、セキュリティ及び安全性 | AI開発・運用者は、AIシステムの入力データ、処理過程及び決定に関し、検証可能なものとすべき。また、体系的なリスク管理を行うべき。 |
| アカウントビリティ | AI開発・運用者は、AIシステムの適切な作動や上記の原則を尊重していることについて、アカウントビリティを果たすべき。 |

推奨される政府の取組

研究開発
への投資

AIに関する
エコシステム
の発展

イノベーションと
競争を促すための
政策環境

人材育成
労働市場変革
への備え

国際協力

生成AIやディープフェイク技術を利用した巧妙な偽・誤情報の拡散の事例

| 年 | エリア | 概要 |
|------|-------|---|
| 2019 | マレーシア | 閣僚が同性と性行為（マレーシアでは禁止）をしている動画が偽造され、拡散。 |
| 2021 | 日本 | 東北地方で発生した地震直後に、官房長官が記者会見で笑みを浮かべる偽画像が拡散。 |
| 2021 | 欧州 | ロシアの議員のディープフェイク動画と気付かずに欧州の議員がビデオ電話会議を実施。 |
| 2022 | 欧州 | ロシアによるウクライナへの軍事侵攻に際し、ゼレンスキー大統領が市民らに投降を呼びかける内容の偽動画が拡散。 |
| 2022 | 日本 | 台風の発生時に、静岡県で住宅が水没したとする偽画像が作成され、SNSに投稿・拡散。 |
| 2023 | 米国 | 政治活動家が、バイデン大統領が第三次世界大戦の開始を告げる動画を作成。作成者はAIで作成した旨を説明したが、多くの人が説明をつけないまま動画を拡散。 |
| 2023 | 米国 | 国防総省の近くで爆発が起きたかのように見せかけた偽画像が拡散。株価が一時的に下落する事態に発展。 |
| 2023 | 日本 | 生成AIを使って民放の实在のアナウンサーの声や動きを再現し、投資情報サイトへの登録を促す内容の偽動画がSNS上で拡散。 |
| 2023 | 日本 | 総理大臣の声と画像を使って卑猥な言葉で語りかけているように見せかけ、实在する民放の報道番組で放送されているかのように編集された偽動画がSNS上で拡散。 |



生成AIを巡るG7での議論動向

【G7群馬高崎デジタル・技術大臣会合の閣僚宣言（仮訳）（抄）】（2023年4月30日）

47. 生成AI技術が国や分野を超えてますます顕著になっていることを踏まえ、これらの技術の持つ機会と課題を早急に把握し、これらの技術が発展する中で、安全性や信頼性を促進し続ける必要があると認識している。我々は、**AI ガバナンスや著作権を含む知的財産権の保護、透明性の促進、外国からの情報操作を含む偽情報への対処方法や、責任ある形での生成 AI を活用する可能性といったテーマを含む生成 AI に関する G7 における議論を引き続き行うための場を設けることを計画**している。これらの議論は、専門知識を活用し、政策展開の影響に関する分析を検討する OECD や、関連する実践的なプロジェクトを実施する GPAI などの国際機関を活用する必要がある。

【G7広島首脳コミュニケ（仮訳）（抄）】（2023年5月20日）

38. (略) **人工知能 (AI)**、メタバースなどの没入型技術、量子情報科学技術、その他の新興技術などの分野において、デジタル経済のガバナンスは、**我々が共有する民主的価値に沿って更新し続けられるべきである**。これらは、公正性、説明責任、透明性、安全性、オンラインでのハラスメント、ヘイト、虐待からの保護、プライバシー及び人権の尊重、基本的自由、そして個人データの保護を含む。(略)

我々は、国や分野を超えてますます顕著になっているAIの機会及び課題について直ちに評価する必要性を認識し、OECDなどの国際機関が政策展開の影響に関する分析を検討し、人工知能グローバルパートナーシップ (GPAI) が実践的なプロジェクトを実施することを奨励する。この観点から、我々は、関係閣僚に対し、生成AIに関する議論のために、包摂的な方法で、OECD及びGPAIと協力しつつ、G7の作業部会を通じた、**広島AIプロセスを年内に創設**するよう指示する。これらの議論は、ガバナンス、著作権を含む知的財産権の保護、透明性の促進、偽情報を含む外国からの情報操作への対応、これらの技術の責任ある活用といったテーマを含み得る。

「AIに関する暫定的な論点整理」

- ✓2022年11月、大規模言語モデル「GPT 3」をベースにした対話型言語モデル「Chat GPT」や、テキストを入力すると画像を生成する「プロンプト型画像生成AI」、テキストから作曲を行うAIなど、多様な用途でのAIが登場。
- ✓国内では、**AI戦略チーム**（村井英樹総理補佐官の下、関係省庁の実務者級（審議官・課長級）のチームを編成）、**AI戦略会議**（座長：松尾 豊（東京大学大学院工学系研究科教授））において、我が国における生成AIをめぐる課題について検討を行い、2023年5月26日、「AIに関する暫定的な論点整理」を取りまとめ

基本的な考え方

- 国際的なルール構築に向けた主導的役割の発揮
 - **AIに国境はなく、国際的な共通理解・ルールづくり・相互運用性が重要。我が国は、「広島AIプロセス」などを通じ、議論をリードすべき**
- リスクへの対応と利用
 - 生成AIの開発・提供・利用を促進するためにも、生成AIの懸念やリスクへの適切な対応を行うべき。いわば、「ガードレール」の設置が必要。
- 多様な関係者を巻き込んだ迅速かつ柔軟な対応
 - 政府は、広島AIプロセスなどの検討スケジュールも踏まえつつ、マルチステークホルダーを巻き込んだ、迅速かつアジャイルの対応が求められる。

AIの利用

- 生成AIは、デジタル化を加速させ、我が国全体の生産性向上のみならず、様々な社会課題解決に資する可能性がある。
- AI利用を加速するため、医療や介護・行政・教育・金融・製造等のデータ連携基盤の構築・DFFT構想の具体化・人材育成・スタートアップの事業環境整備を進めるべき。
- 政府機関が一体となって、機密情報漏洩のリスクなどに配慮しつつ、率先して生成AIの利用可能性を追求することが重要。
- 幅広い世代が生成AIの恩恵を享受できるよう、**スキル・リテラシーを身に付けることが大切。**

リスクへの対応

- リスク対応の基本方針
- 透明性と信頼性
 - AI開発者・提供者には、現行法令やガイドラインに則り、積極的な情報開示を求めたい。
 - □政府は、主要なAI開発者・提供者に対して、透明性・信頼性の確保を直接働きかけることも検討すべき。
 - **生成AIの普及を踏まえ、既存のガイドラインに関して、必要な改定などを検討する必要がある。**その際、諸外国における検討とも協調し、第三者認証制度や監査制度等も参考とすべき。
 - **顕在化したリスクを低減するような技術の研究開発・普及を奨励することも望ましい。**
- 懸念されるリスクの具体例と対応

AI開発力

- 生成AIに関する基盤的な研究力・開発力を国内に醸成することが重要。**政府は、AI開発におけるインフラとも言うべき、計算資源とデータの整備を行うことが最も重要。**
- 計算資源を活用するための電力調達が課題。地方のデータセンターの活用を含め、電力を有効活用する方策の検討が必要。
- 開発に用いることのできる日本語を中心とするデータの整備・拡充を進めるべき。**
- また、生成AI自体の開発は、市場原理を最大限尊重した、民間活力を十分活用した従来型ではない開発促進策が期待される。世界からトップ人材が集まる研究環境の構築も期待される。

広島AIプロセス 閣僚級会合

開催日程

令和5年9月7日（木）20:00～21:00（オンライン開催）

参加国等

- ・議長として松本総務大臣が参加。
- ・G7各国（加、仏、独、伊、米、英）、EU、国際機関（OECD、GPAI）が参加。

主な成果

- ◆ 本会合の結果、G7広島AIプロセスの中間的な成果として、「**G7広島AIプロセス G7デジタル・技術閣僚声明**」を採択。
- ◆ 閣僚声明において**以下の項目について合意**。

○OECDレポートに基づく優先的な課題、リスク及び機会に関する理解

- G7共通の優先的な課題・リスクとして、透明性、偽情報、知的財産権、プライバシーと個人情報保護、公正性、セキュリティと安全性等が例示。また、機会として、生産性向上、イノベーション促進、ヘルスケア改善、気候危機の解決への貢献等が例示。

○高度なAIシステム（基盤モデルや生成AIを含む。以下同じ。）に関する国際的な指針(guiding principles)及び行動規範(code of conduct)

- 高度なAIシステム技術の進歩に鑑み、**AI開発者を対象とする国際的な行動規範の策定が国際社会の喫緊の課題の1つ**であることを認識。AI開発者を対象とする行動規範を策定する基礎として、以下の項目で構成される指針の骨子を策定。**AI開発者を対象とする指針と行動規範のG7首脳への提示を目指す。**
- 年内に、**開発者を含む全てのAI関係者向けの国際的な指針を策定。**

- | | |
|--|---|
| ・高度AIシステムの適切な安全対策及び導入前の社会的リスクの考慮 | ・サイバーセキュリティ及びインサイダー脅威対策を含む強固なセキュリティ管理措置への投資 |
| ・高度AIシステム導入後の脆弱性の特定と低減に向けた努力 | ・電子透かし技術等のAIが生成したコンテンツを利用者が識別できる仕組みの開発及び導入 |
| ・モデルの能力、限界、適切・不適切な利用領域の公表 | ・社会、環境、安全のリスクを軽減するための研究・投資の優先的な実施 |
| ・AI開発者と政府、市民社会、学界との間での責任ある情報共有 | ・気候危機等の世界最大の課題に対処するための高度なAIシステムの優先的な開発 |
| ・プライバシー・リソース及びAIガバナンス・リソース等のリスク管理計画及び低減手法の開発及び開示 | ・国際的に認知された技術標準の開発及び整合性確保の推進 |

○偽情報対策に資する研究の促進等のプロジェクトベースの協力

- OECD、GPAI、UNESCO等の国際機関と協力し、AIによって生成された偽情報を識別するための最先端の技術的能力に関する研究の促進等、プロジェクトベースの取組を推進することを計画。

○上記の取組を進めるに当たっては、幅広いマルチステークホルダーの意見を採り入れることに合意。

広島AIプロセス国際指針・行動規範①

- AI開発組織向けの国際指針について、2023年10月9日のIGFでのG7非公式会合で基本合意（EU・米・日がそれぞれパブコメ）。
- 2023年10月30日、広島AIプロセスに関するG7首脳声明を発出し、AI開発組織向けの国際指針と国際行動規範について歓迎し、公表。

【指針と対応する行動規範（例）の概要】

1. 高度なAIシステムの市場投入前及び、高度なAIシステムの開発を通じて、AIライフサイクルにわたるリスクを特定、評価、低減するための適切な対策を実施する。

（行動規範例）市場投入前の「レッドチーミング」などの内部及び独立外部テストによるリスクの特定と低減

リスク例：化学・生物兵器の開発等に係るハードルを下げるリスク、有害な偏見や差別を社会等にもたらすリスク、偽情報助長やプライバシー侵害など民主主義的価値や人権に対するリスク

2. 市場投入後に脆弱性、インシデント、悪用パターンを特定し、低減する。

（行動規範例）コンテストや賞金などを活用した、第三者及び利用者による問題や脆弱性の発見と報告の促進

3. 十分な透明性の確保や説明責任の向上のため、高度なAIシステムの能力、限界、適切・不適切な利用領域を公表する。

（行動規範例）安全性・セキュリティ・社会や人権に対するリスクに関する評価、AIモデルの能力や限界等を含んだ透明性報告書や使用説明書の公表

4. 産業界、政府、市民社会、学术界を含む関係組織間で、責任ある情報共有とインシデント報告に努める。

（行動規範例）安全性・セキュリティ・信頼性を確保するため、情報共有のための基準・メカニズム・ベストプラクティスを開発し採用

5. リスクベースのアプローチに基づいたAIのガバナンスとリスク管理ポリシーを開発、実践、開示する。特に高度AIシステムの開発者向けの、プライバシーポリシーやリスクの低減手法を含む。

（行動規範例）個人データ、ユーザーのプロンプトや出力を含めたプライバシーポリシーの開示

職員が自らの責務や組織のリスク管理慣行を熟知するための方針・手順・訓練の確立

6. AIのライフサイクル全体にわたり、物理的セキュリティ、サイバーセキュリティ及び内部脅威対策を含む**強固なセキュリティ管理措置**に投資し、実施する。

(行動規範例) 情報セキュリティに関する安全運用措置等による「モデルウェイト」やアルゴリズムの保護
最も貴重な知的財産や企業秘密を保護するための強固な内部脅威検知プログラムの確立

7. AIが生成したコンテンツを利用者が識別できるように、**電子透かしやその他の技術等、信頼性の高いコンテンツ認証および証明メカニズム**を開発する。またその導入が奨励される。

(行動規範例) 電子透かしや証明システムなど、AI生成コンテンツであることを利用者が判断できるためのツールやAPIの開発
AIと接していることを利用者が認知できるようなラベリング表示メカニズムの導入

8. 社会、安全、セキュリティ上の**リスクの低減のための研究**を優先し、効果的な低減手法に優先的に投資する。

(行動規範例) 民主的価値の確保や人権の尊重等に関する研究の実施、協力や投資
環境及び気候への影響を含むリスク低減ツールや積極的リスク管理作業への投資

9. 気候危機、健康・教育などの、**世界最大の課題に対処**するため、高度なAIシステムの開発を優先する。

(行動規範例) 国連SDGsの進捗を支援するためのAI開発を支援

10. **国際的な技術標準**の開発と採用を推進する。

(行動規範例) 電子透かしを含む国際的な技術標準とベストプラクティスの開発や利用に貢献

11. 適切なデータ入力措置と**個人情報及び知的財産の保護**を実施する。

(行動規範例) プライバシーや知的財産を尊重するための安全措置の実施、適用される法的枠組みの尊重

6. 諸外国の動向

米国・豪州における最近の主な動向

【米国における通信品位法230条等】

- 通信品位法(1996年)第230条では、プロバイダは、①第三者が発信する情報について原則として責任を負わず、②有害なコンテンツに対する削除等の対応（アクセスを制限するため誠実かつ任意にとった措置）に関し、責任を問われないとされ、**言論の自由を重視する立場から、プロバイダには広範な免責が容認。**
- **偽誤情報流通に対してオンラインプラットフォームの責任を高めるため通信品位法230条の改正を伴う検討**はバイデン政権や議会において実施。2022年9月、ホワイトハウスは「競争及び技術プラットフォームの説明責任の強化」を公表。議会において230条改正を伴う法案が提出（2023年2～3月には4件）。
- 2021年には中間選挙が実施。州レベルで選挙の完全性を担保するために専門組織整備や情報発信などの取組が実施。また、COVID-19、ディープフェイクを禁止する例など偽誤情報に対抗するための州法等が検討。

【豪州における通信法改正案等】

- 2022年12月、**DIGI（デジタル産業団体）**は更新版「**偽・誤情報に関するオーストラリアの行動規範（ACPDM）**」を公表。小規模プラットフォームの参加、偽情報を拡散しないようアルゴリズムやデジタル広告への対応等が追加。2023年5月、3回目の「**透明性レポート**」を公表。
- 2023年7月、プラットフォーム事業者の取組に関し、**オーストラリア通信メディア庁（ACMA）**により、**行動規範の効果等に関する助言として、2回目の報告書（1回目は2021年6月）**が公表。
- 2023年1月、政府は「New disinformation laws」制定と、**オンラインの偽・誤情報と闘うためにデジタルプラットフォームに対する新たな権限をACMAに付与**することを公表。同7月、**Communications Legislation Amendment Bill（通信法改正案）**に関するパブリックコメントを同8月まで実施。

英国・ドイツ・フランスにおける最近の主な動向

【英国におけるオンライン安全法案】

- 2023年10月に「**Online Safety Act（オンライン安全法）**」が成立。
- ①プラットフォーム事業者に対する**オンライン上の違法コンテンツ・行為への対処を求める法定の注意義務（自社サービスに関するリスク評価や表現の自由とプライバシーへの配慮、透明性確保等）**、②注意義務の遵守のための**執行権限をOfcomに付与**、③**Ofcomは注意義務の履行方法の概説を行動規範として作成・提示等**を規定。

【ドイツにおけるネットワーク執行法】

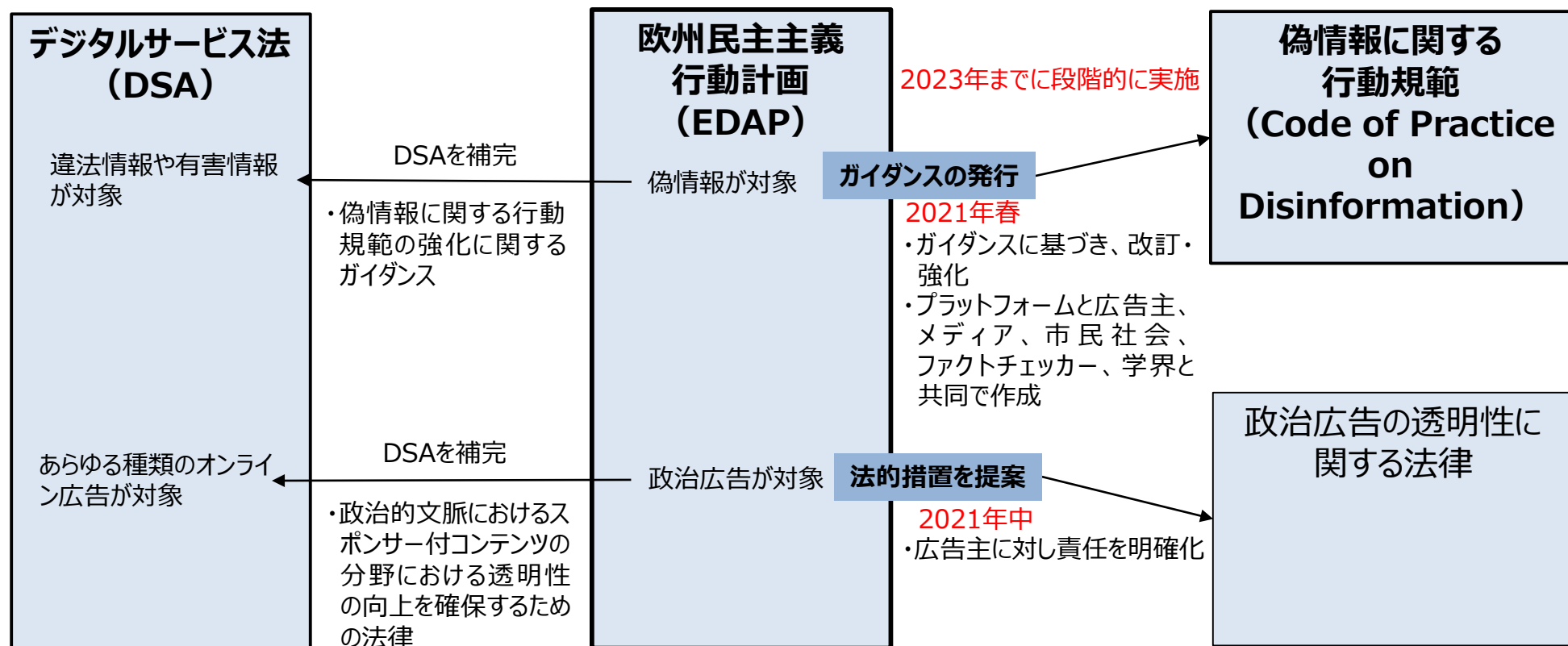
- 2017年10月に「**NetzDG（ネットワーク執行法）**」が発効。利用者数200万人以上のプラットフォーム事業者は、**刑法違反コンテンツに関し、違反報告数、削除件数、違法投稿防止のための取組、社内体制等に関する透明性レポートを半年に一度公開**する義務。
- **2021年の改正**では、**異議申立の機会の確保**に関する規定と**異議申立件数に関する透明性確保義務**や、**公共目的の研究者に対するデータ提供**に関する規定が追加。

【フランスにおけるインターネット上のヘイトコンテンツ対策法等】

- 2020年5月、**オンライン・ヘイト等を含む違法コンテンツ**（例えば、生命・人格に対する侵害等の称揚、出自又は特定の民族・人種・宗教を理由とした差別・憎悪・暴力の扇動、児童ポルノの拡散）への対処を目的に**インターネット上のヘイトコンテンツ対策法**が議会下院で可決。その後、憲法院に付託され合憲性審査が行われた結果、**コンテンツの過剰削除を促進する可能性があるとして、違法コンテンツの削除等義務などの条項が違憲**と判断。当該条項が削除され、2020年6月に同法は公布・施行。
- 2018年11月、**選挙時におけるフェイクニュース対策のための「情報操作との戦いに関する法律」**が成立。選挙期間内に、検察官や候補者等から求めを受けた裁判官はプラットフォーム事業者に対し送信防止措置を命令。同事業者は、アルゴリズムの透明性確保、虚偽情報を大量に拡散する偽アカウント対策、拡散メカニズムの情報提供等などの協力とこれらの実施方法の公表が義務付け。

EUにおけるデジタルサービス法と欧州民主主義行動計画

- 偽情報への取組として、自主的な取組である「偽情報に関する行動規範」に加え、「デジタルサービス法 (DSA)」が加わり、2段構造。
- 「欧州民主主義行動計画 (EDAP)」はDSAを補完。DSAに示された措置をEDAPにおいて具体化。
- EDAPの3つの柱のうち、「自由で公正な選挙の促進」、「偽情報への対抗措置」がDSAと関連。



欧州民主主義行動計画（EDAP）の概要①

- 「欧州民主主義行動計画（EDAP）」が2020年12月2日に公表。
- 3つの柱「自由で公正な選挙の促進」、「メディアの自由の強化」、「偽情報への対抗措置」から構成。上記ごとに、委員会としての行動が提示。

| 項目 | 内容 |
|------|--|
| 背景 | ・過激主義と両極化の高まりと、人々とその選出された代表者との距離感によって、EUとその加盟国の民主主義は課題に直面している状況にある。 |
| 目的 | ・EU市民に権限を与え、EU全域でより強靱な民主主義を構築する ・3つの柱「自由で公正な選挙の促進」、「メディアの自由の強化」、「偽情報への対抗措置」から構成される |
| 特徴 | ・「自由で公正な選挙の促進」のために、政治広告に関する法的措置を提案（2021年予定） ・「メディアの自由の強化」のために、ジャーナリストや市民社会を戦略的訴訟（SLAPP；Strategic Lawsuit Against Public Participationの略、スラップ訴訟）から守るためのイニシアチブを提示 ・「偽情報への対抗措置」を行う（以降、別説明有）。 ・欧州選挙の1年前の2023年までに行動計画を段階的に実施。以降、達成状況や見直しの必要性についての評価をおこなう |
| 公表日 | 2020年12月2日 |
| 実施主体 | ・EU |

（出典）各種資料よりみずほ情報総研作成
European Democracy Action Plan: making EU democracies stronger 他
https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2250
https://ec.europa.eu/info/strategy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan_en#countering-disinformation
https://ec.europa.eu/info/sites/info/files/qanda_20_2251_en.pdf

欧州民主主義行動計画（EDAP）の概要②

- 3つの柱のうち、「偽情報への対抗」関連では、以下がポイント。
 - ✓ 偽情報の発信者に対するコストを科すための取組（EU内外の発信者とも）
 - ✓ プラットフォーマーが署名した「Code of Practice on Disinformation（偽情報に関する行動規範）」の見直し、co-regulatory framework（共同規制）化。
 - ✓ 上記に伴い、オンラインプラットフォームに向けて、「行動規範を強化するためのガイダンス（guidance to enhance the Code of Practice）」を発行（2021年春）。その後、EUは新たな行動規範の実施状況をモニタリング。
 - ✓ EU内外の偽情報へのメディア・リテラシー向上プロジェクトへ支援と資金提供。

| 項目 | 内容 |
|------------|--|
| 対抗措置が必要な背景 | ・偽情報の急増によって、 <u>民主主義制度を不安定にし、市民の信頼を損なう可能性</u> がある。誤情報、偽情報、海外からの情報操作に対抗するために異なった対応が求められる。 |
| 偽情報への対抗措置 | ・行動計画の4.偽情報への対策に下記について記載されている。 ・加害者にコストを課すことを含む、情報空間において外国からの干渉に対抗するためのEUの既存の「ツールボックス」を改善する ・ <u>デジタルサービス法（Digital Services Act）</u> に従い、現在の「偽情報に関する行動規範」を、 <u>オンラインプラットフォームの義務と説明責任の共同規制の枠組みへと変革</u> ・前述の偽情報に関する <u>行動規範を強化し、その実施をモニタリングするより堅牢なフレームワークを構築するためのガイダンスを2021年春に発行</u> する。ガイダンスは、プラットフォームがどのように対策を強化すべきかを示す |

（出典）各種資料よりみずほ情報総研作成

European Democracy Action Plan: making EU democracies stronger 他

https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2250

https://ec.europa.eu/info/strategy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan_en#countering-disinformation

https://ec.europa.eu/info/sites/info/files/qanda_20_2251_en.pdf

【出典】プラットフォームサービスに関する研究会第24回会合（2021年3月17日）参考資料6（みずほ情報総研作成資料）

偽情報に関する行動規範（Code of Practice on Disinformation）の概要

- 「偽情報に関する行動規範」が2022年6月16日に公表。「欧州民主主義行動計画（EDAP）」における柱「偽情報への対抗措置」をさらに強化した規範。
- 規範内容について、欧州委員会が公表したガイダンスに応えていると欧州委員会は評価。

| 項目 | 内容 |
|------|--|
| 背景 | ・EUにおける政治分野の偽情報対策の一環として行われた、2018年の民間の自主規範である「偽情報に関するEU行動規範」の策定がされるも、成果が部分的であったことから、これを強化する計画として「欧州民主主義行動計画」がその行動規範の強化するためのガイダンス（偽情報に関する行動規範の強化についての欧州委員会のガイダンス）を21年に公表し、その成果物として22年に公表。 |
| 目的 | ・主要なソーシャルネットワークや広告者および広告業界による自主規範の強化 |
| 特徴 | ・規範署名者は21年に公表した「偽情報に関する行動規範の強化についての欧州委員会のガイダンス」の実行を移すことが求められる。 ・8つの分野における44のコミットメントを揚げ、コミットメントごとに計128の具体的な措置を定めている。 ・8つの分野とは①非収益化②政治広告の透明性③サービスの完全性の確保④ユーザーを強化する取組⑤研究者を強化する取組⑥ファクトチェッカーのコミュニティの強化⑦透明性センターおよびタスクフォース⑧モニタリングの枠組みの強化 ・上記の措置にはその達成度を測るための「定性的報告要素」と「サービスレベル指標」がそれぞれ設けられている。 |
| 公表日 | 2022年6月16日 |
| 実施主体 | ・EU |

偽情報の行動規範の強化に関する欧州委員会のガイダンス（目次） ※仮訳

| | | |
|---|---|----------------------------------|
| 1. 導入 | 6. サービスの保全 | 8.1.5 他の利害関係者のためのデータへのアクセス |
| 2. COVID-19のモニタリング – 結果と学び | 6.1 許容されない操作的行動に関する共通の理解 | 8.2 署名者と研究者の間の協力枠組み |
| 3. 対処されるべき水平的な課題 | 6.2 許容されない操作的行動を制限する強化されたコミットメント | 8.3 ファクトチェッカーとの協同 |
| 3.1 規範の目的を達成するための強化されたコミットメント | 6.3 コミットメントの調整、協力と透明性 | 8.3.1 協力の形 |
| 3.2 スコープの拡大 | 7. ユーザ・エンパワーメント | 8.3.2 署名者のサービスにおけるファクトチェックの利用と統合 |
| 3.3 参加者の拡大 | 7.1 メディアリテラシー強化措置へのコミットメント | 9. 規範のモニタリング |
| 3.4 特定のコミットメント | 7.2 「安全なデザイン」へのコミットメント | 9.1 KPI（重要業績評価指標） |
| 3.5 欧州デジタルメディア観測所（EDMO） | 7.3 レコメンドシステムのアカウントビリティ | 9.1.1 サービスレベル指標 |
| 3.6 迅速なアラートシステム | 7.4 信頼できる公益情報の可視性 | 9.1.2 構造指標 |
| 4. 広告配置のセキュリティ | 7.5 誤った又はミスリードなコンテンツと交流した（交流してきた）ユーザーにして発行される警告 | 9.2 モニタリング枠組み |
| 4.1 偽情報の非収益化 | 7.6 有害な誤情報にフラグを立てる機能 | 9.2.1 定期報告 |
| 4.2 関係者間の協力の改善 | 7.7 情報に基づいたオンラインナビゲーションのための指標の利用可能性 | 9.2.2 透明性センター |
| 4.3 偽情報を含む広告に対処するというコミットメント | 7.8 メッセージングアプリの偽情報を抑制するための措置 | 9.2.3 永続的なタスクフォース |
| 5 政治広告と課題ベースの広告 | 8. 研究及びファクトチェックコミュニティのエンパワーメント | 10. 結論及び次のステップ |
| 5.1 政治及び課題ベースの広告への効率的なラベル付け | 8.1 偽情報に関する研究のための署名者のデータへのアクセス | |
| 5.2 政治及び課題ベースの広告の検証及び透明性 | 8.1.1 データへのアクセスの一般的な枠組み | |
| 5.3 メッセージングプラットフォームにおける透明性 | 8.1.2 匿名化された及び非パーソナルデータへのアクセス | |
| 5.4 政治広告のターゲティング | 8.1.3 追加的な精査を要するデータ（含パーソナルデータ）へのアクセス | |
| 5.5 アプリケーション・プログラミング・インターフェース（API）のために改善された広告レポジトリと最低限の機能 | 8.1.4 EDMOの役割 | |

欧州委員会・豪州・ニュージーランドの行動規範ガイドンスと行動規範※2023年8月時点の公開情報をもとに総務省作成

| | 欧州委員会 | 豪州 | ニュージーランド |
|---------------|--|---|--|
| 自主的な行動規範 | “2022 Code of Practice on Disinformation” (偽情報に関する行動規範) | Australian Code of Practice on Disinformation and Misinformation” (偽・誤情報に関する豪州の行動規範) | Aotearoa New Zealand Code of Practice for Online Safety and Harms |
| 行動規範の作成主体 | 署名者らによるタスクフォース | Digital Industry Group Inc. (DIGI) (非営利業界団体) | Netsafe (独立オンライン安全団体) |
| 公表時期 | 2022年6月16日 | 2022年12月22日(最終更新日) | 2022年7月25日 |
| 対象とする事業者・サービス | 署名者(広告エコシステムのプレーヤー、広告主、アドテック企業、ファクトチェッカー、振興または専門プラットフォーム、市民社会、偽情報に特化した専門知識を持つ第三者機関など) * Google, Meta, Microsoft, TikTok等44者が署名済 | 署名者(“本規定は、オーストラリアのユーザーに製品やサービスを提供し、エンドユーザーがオンラインで偽・誤情報を伝播するリスクを特定し、かつ/またはオンラインで偽・誤情報の伝播を減らすことに貢献できるデジタルプラットフォーム*が署名することができる。”) *「オンライン検索エンジン、ソーシャルメディア、デジタルコンテンツ事業者」(ACCC(豪州競争消費者委員会)の定義) * Google, Facebook, Microsoft, Twitter, TikTok, Redbubble, Apple, Adobeが署名済 | 署名者(“デジタル・プラットフォームは、当規範の加盟者となることができ、いつでも加盟することができる。”) * Netsafe, NZTech, Meta, Google, Amazon, Twitter, TikTokの6者が署名済 |
| 運用状況 | 2023年2月、30の署名者が実施状況に関する最初の基調報告を公表。次回報告は7月予定(8月現在未公開) | 2021年から各社は透明性レポートを提出(直近は2023年5月に公開済)。ACMA(オーストラリア通信メディア庁)が実施状況についてモニタリングを実施し、措置の十分性等について数度報告(直近は2023年7月)。 | 公開情報は特になし |
| 行動規範の政府ガイドンス | “European Commission Guidance on Strengthening the Code of Practice on Disinformation”(偽情報に関する行動規範の強化に関する欧州委員会ガイドンス) | “Regulating in the digital age - Government Response and Implementation Roadmap for the Digital Platforms Inquiry”(“デジタル時代の規制 - デジタルプラットフォーム調査に対する政府の対応と実施ロードマップ”) | |
| 概要 | 欧州委員会のモニタリングや評価に基づき、関連するステークホルダー等との議論を踏まえてガイドンスを作成。行動規範のギャップや不足等に対処し、より透明性があり、安全で信頼できるオンライン環境を構築するために対策強化する方法について委員会の見解を示したもの。 | 政府は主要なデジタル・プラットフォームに対し、偽情報やニュース・コンテンツの信頼性シグナルに関する懸念に対処するためにプラットフォームが何を行うかをまとめた自主的な行動規範を策定するよう要請。 | 特になし |
| 行動規範指針の作成主体 | 欧州委員会 | オーストラリア政府 | |
| 公表時期 | 2021年5月 | 2019年12月12日 | |

デジタルサービス法（DSA）の概要

- 2022年11月一部施行。超大規模オンラインプラットフォーム／検索エンジンに対し、**違法・有害コンテンツ（偽情報を含み得る）の拡散による重大な社会的リスクを軽減する措置の実施等**を義務付け。
- リスク軽減措置の実施に関する評価に際し「**偽情報に関する行動規範**」への署名と遵守が参照され得る。

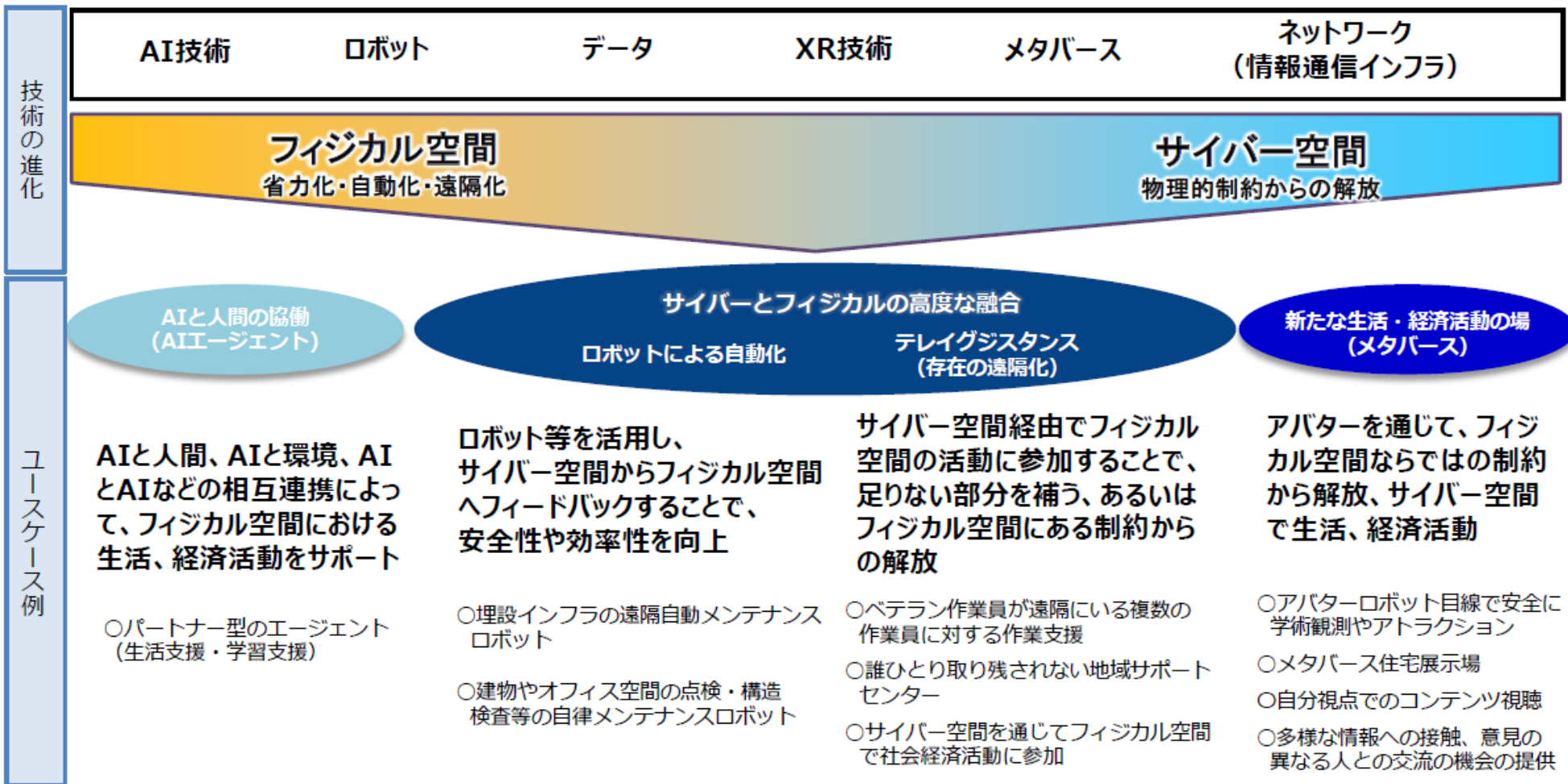
| | |
|--|--|
| 事業者の特性に応じた規制 | 仲介サービス提供者、ホスティング事業者、オンライン・プラットフォーム、超大規模オンライン・プラットフォーム/超大規模オンライン検索エンジン、に段階的な義務を規定 |
| 超大規模サービス(※)に対する重点的な規定 ※ 4500万人以上(EU人口の10%に相当する)の利用者にリーチするサービス | 超大規模オンライン・プラットフォーム(VLOP)/超大規模オンライン検索エンジン(VLOSE)に対して、透明性の確保や透明性報告の追加的義務、データへのアクセス・アルゴリズムの説明、システミック・リスクの評価・軽減措置実施、危機のプロトコル作成、欧州委員会による独占的監督権限(調査、暫定措置、モニタリング、前年度の総売上高の6%を上限とする罰金等)、監督手数料負担(サービスの規模に比例し、全世界の年間純利益の0.05%を超えない)を規定 |
| 零細・小規模事業者への配慮 | 一定の義務免除(透明性報告義務、オンライン・プラットフォームの義務)、新規制適用の猶予期間延長(いずれもVLOPに該当する場合を除く) |
| 事業者の透明性・説明責任を中心とした対応 | 説明や情報公開、関係者による情報へのアクセス権(利用規約での説明、透明性報告、データへのアクセス・アルゴリズムの説明(VLOP)、リスク評価・軽減措置実施(VLOP)等) |
| 違法・利用規約違反コンテンツへの対応システム | 通報制度の整備、信頼できる専門機関(信頼された旗手)との連携、対応結果(削除やアクセス制限等)の理由の通知、異議申立制度の整備、違法・無根拠な通報・異議申立の受付停止、それらの実施・対応状況についての透明性報告等 |
| 未成年者、障害者、少数者等を含むユーザ保護 | ダークパターン・ユーザ操作等の禁止、未成年者にわかりやすい説明、未成年者に対するプロファイリングを用いた広告の禁止、特別カテゴリー情報(例:性的指向、宗教、民族性など)によるプロファイリングを用いた広告禁止等 |
| 深刻なリスク・犯罪行為、緊急事態への対応 | 取引事業者のKYBC*1、違法な製品・サービスを知った場合の利用者への通知*1、刑事犯罪の疑いの当局への通知(ホスティング事業者)、危機対応メカニズム(VLOP)等 *1 消費者が取引者と遠隔契約を締結することを可能にするオンライン・プラットフォームの場合 |

7. その他参考

2030年頃の来たる未来の姿

情報通信審議会「2030年頃を見据えた情報通信政策の在り方」最終答申（令和5年6月23日）

- 我が国では、少子高齢化による労働人口の減少等による国内市場の縮小、災害の激甚・頻発化、公共インフラの老朽化などの様々な課題が山積。国際情勢もめまぐるしく変化しており、国の政策、経済活動や国民生活にも影響。
- 一方、AI,ロボット等の情報通信技術の進化と普及により、**省力化・自動化・遠隔化、サイバー空間での新たな生活・経済活動が可能となることにより、様々な制約から解放され、全国どこにいてもそれぞれのライフスタイルやニーズ等にあった豊かな生活を実現するなど、Society5.0の実現が期待される。**



2030年頃に向けて、我が国が求められる変化、情報通信政策の検討の方向性

情報通信審議会「2030年頃を見据えた情報通信政策の在り方」最終答申（令和5年6月23日）

- コロナ禍でのデジタル化の進展等により、国民生活や経済活動における情報通信の果たす役割やその利用に伴うセキュリティの確保が一層重要。特に、コンテンツ・サービス・端末・機器のレイヤーにおける海外のプラットフォーム事業者やベンダーの存在感が高まり、また、近年の米中の緊張関係等の国際情勢の変化を背景とした情報通信分野のサプライチェーンリスクといった課題が顕在化。
- 2030年頃を見据えて、Society5.0の実現及び経済安全保障の確保を図る観点から、今後の情報通信政策の在り方について諮問。**本年6月に最終答申**。社会経済環境の変化やICTの進展を通じ、デジタル空間と実空間とが高度に融合・一体化し、デジタル空間が新たな社会の一形態にもなり、これまでの生活空間が拡張される未来が想定。我が国における豊かかつ健全な情報空間の確保等が重要。

我が国が向き合う課題【利用者視点】

2. 健全なサイバー空間の確保

⑪サイバー空間の分断・偽誤情報の増加・深刻化

- 国家による介入、ビッグ・テック企業へのデータ集中、フィルターバブルやエコーチェンバーなど、ネットの分断が深刻化。
- アテンション・エコノミーや偽・誤情報を使った情報戦など、偽・誤情報の規模が拡大。リテラシーの低い人が偽・誤情報を拡散しやすい。

我が国がなすべきこと

民間取組、国際社会との連携

- 民間の自主的な取組（プラットフォーム事業者による違法・有害情報や広告の削除基準・条件の公表等の適切な対応、透明性・アカウントビリティ確保、ファクトチェック等）を基本とし、事業者からのエビデンスを含んだ説明を踏まえた国の対策を検討。
- 全世代に対するリテラシー向上の取組の実施。
- インターネットは、引き続き自由で、分断のない、国境を越えてグローバルに流通可能な環境を目指すべき。
- 自由・責任・信頼があるインターネットを築くため、各ステークホルダー間の連携による社会全体での取組、国際連携の強化。

未来に備えた取組

各政府戦略等における偽・誤情報対策に関する記載

● 経済財政運営と改革の基本方針2023（令和5年6月16日閣議決定）

第2章 新しい資本主義の加速 2. 投資の拡大と経済社会改革の実行 （デジタルトランスフォーメーション（DX）、AIへの対応）

- ・ **デジタル空間の誤情報等**への対応を行う。

● 新しい資本主義のグランドデザイン及び実行計画2023（令和5年6月16日閣議決定）

4. AI

（中略）生成AIの技術進歩と普及によって、AIがどのようなデータをどのように処理しているのかが分からない、**巧妙な偽情報が拡散するなど、AIが社会にもたらす不安やリスクへの懸念が世界的に高まっており**、安全保障上のリスクも指摘されている。そこで、AIに関する国際的な議論、多様なリスクへの対応を進めるとともに、AIの利用の促進、AIの開発力の強化を図る。

（1）国際的な議論とリスクへの対応

個人情報の不適切な利用、セキュリティに関する不安、**偽情報・誤情報による混乱**、著作権侵害など、AIに関する多様なリスクに関して、国際的な議論や技術の急激な変化も踏まえ、「広島AIプロセス」のスケジュールも念頭に検討し、対応していく。

● デジタル社会の実現に向けた重点計画（令和5年6月9日閣議決定）

第1 安全・安心で便利な国民の生活や事業者の活動に向けた重点的な取組

10. 国家安全保障戦略等に基づく取組等の推進

インターネット上の偽情報への対策について、憲法によって表現の自由及び通信の秘密が保障されていることを前提に、海外動向等も踏まえつつ、**産学官民の多様なステークホルダーと連携し、プラットフォーム事業者の自主的な対応の促進と当該取組の透明性・アカウントビリティ確保のための枠組み・方策の検討**や、ユーザのICTリテラシー向上の推進、**研究開発の推進**など、総合的な対応を進める。

● 統合イノベーション戦略2023（令和5年6月9日閣議決定）

（1）先端科学技術の戦略的な推進 ① 重要技術の国家戦略の推進と国家的重要課題への対応 （重要技術の国家戦略の強化）

AI 国際的な議論とリスクへの対応

また、機密情報の漏洩や個人情報の不適正な利用、プライバシーに関するリスク、**偽情報などが社会を不安定化・混乱させるリスク**、サイバー攻撃が巧妙化するリスク、著作権侵害のリスクなど、**AIに関する多様なリスク**に関して、「**広島AIプロセス**」などのスケジュールも念頭に、国際的な議論の動向も踏まえ、必要な対応を検討していく。

「デジタル空間における情報流通の健全性」に関連する規定等

【デジタル社会形成基本法（令和3年法律第35号）】

第十条（個人及び法人の権利利益の保護等）

デジタル社会の形成に当たっては、高度情報通信ネットワークの利用及び情報通信技術を用いた情報の活用により個人及び法人の権利利益、国の安全等が害されることのないようにされるとともに、高度情報通信ネットワークの利用及び情報通信技術を用いた情報の活用による**信頼性のある情報の自由かつ安全な流通の確保**が図られなければならない。

【岸田総理大臣の政策スピーチ（「自由で開かれたインド太平洋（FOIP）のための新たなプラン」（抄）】 （2023年3月20日）

偽情報の拡散は、人々の政治的自己決定を妨げ、国家の自律性を脅かす各国共通の課題です。**自由で公正な情報空間**を確保すべく、偽情報対策の知見を地域に広げるためのワークショップなどを年内に開催します。

【サイバーセキュリティ戦略（抄）】（2021年9月28日）

サイバー空間を「自由、公正かつ安全な空間」とすることにより、基本法に掲げた目的に資すべく、国は、これまで2度にわたり、我が国のサイバーセキュリティに関する施策についての基本的な計画として、サイバーセキュリティ戦略を策定してきた。

先に述べた時代認識を踏まえれば、その目的、そしてサイバー空間に対する考え方はいささかも変わるものではない。むしろ、その確保が危機に直面する中で、「**自由、公正かつ安全なサイバー空間**」を確保する必要性はこれまで以上に増しているとの認識が深められるべきである。

コンテンツモデレーションにおける透明性と説明責任に関するサンタクララ原則

- コンテンツモデレーションにおける透明性及び説明責任に関する原則として、人権団体や学識経験者から構成されるグループによって2018年に策定、2021年に改訂。
- FacebookやGoogle、Twitterなど*12のプラットフォーム企業が支持。
*Apple、Facebook、GitHub、Google、Instagram、LinkedIn、Medium、Reddit、Snap、Tumblr、Twitter、YouTube。
- 規制の雛形を意図したものではなく、規制当局が政策形成においてモデレーションで考慮すべき事項を知るための「ガイド」であるとされる。

基本原則 (Foundational Principles)

- ① 人権及びデュープロセス (Human Rights and Due Process)
- ② 理解しやすいルール及びポリシー (Understandable Rules and policies)
- ③ 文化的能力 (Cultural Competence)
- ④ コンテンツモデレーションへの国家関与 (State Involvement in Content Moderation)
- ⑤ 確実性及び説明可能性 (Integrity and Explainability)

運用原則 (Operational Principles)

- ① 数値 (Numbers)
- ② 告知 (Notice)
- ③ 異議申立 (Appeal)

政府及びその他の国家機関のための原則 (Principles for Governments and other State Actors)

- ① 企業の透明性に対する障害の排除 (Removing Barriers to Company Transparency)
- ② 政府自身の透明性向上 (Promoting Government Transparency)

未来のインターネットに関する宣言（2022年4月）

- 2022年4月28日、主催国である米国や初期パートナー国である日本、オーストラリア、カナダ、欧州連合（EU）、英国を含む賛同国・地域が開かれた、自由で、グローバルで、相互運用可能で、信頼性のある、安全な、未来のインターネットへの支持を呼びかける「未来のインターネットに関する宣言」に賛同。

A DECLARATION for the FUTURE of the INTERNET

We are united by a belief in the potential of digital technologies to promote connectivity, democracy, peace, the rule of law, sustainable development, and the enjoyment of human rights and fundamental freedoms. As we increasingly work, communicate, connect, engage, learn, and enjoy leisure time using digital technologies, our reliance on an open, free, global, interoperable, reliable, and secure Internet will continue to grow. Yet we are also aware of the risks inherent in that reliance and the challenges we face.

We call for a new Declaration for the Future of the Internet that includes all partners who actively support a future for the Internet that is open, free, global, interoperable, reliable, and secure. We further affirm our commitment to protecting and respecting human rights online and across the digital ecosystem. Partners in this Declaration intend to work toward an environment that reinforces our democratic systems and promotes active participation of every citizen in democratic processes, secures and protects individuals' privacy, maintains secure and reliable connectivity, resists efforts to splinter the global Internet, and promotes a free and competitive global economy. Partners in this Declaration invite other partners who share this vision to join us in working together, with civil society and other stakeholders, to affirm guiding principles for our role in the future of the global Internet.

RECLAIMING THE PROMISE OF THE INTERNET

The immense promise that accompanied the development of the Internet stemmed from its design: it is an open "network of networks", a single interconnected communications system for all of humanity. The stable and secure operation of the Internet's unique identifier systems have, from the beginning, been governed by a multistakeholder approach to avoid Internet fragmentation, which continues to be an essential part of our vision. For business, entrepreneurs, and the global economy as a whole, interconnection promises better access to customers and fairer competition; for artists and creators, new audiences; for everyone, unfettered access to knowledge. With the creation of the Internet came a swell in innovation, vibrant communication, increased cross-border data flows, and market growth—as well as the invention of new digital products and services that now permeate every aspect of our daily lives.

Over the last two decades, however, we have witnessed serious challenges to this vision emerge. Access to the open Internet is limited by some authoritarian governments and online platforms and digital tools are increasingly used to repress freedom of expression and deny other human rights and fundamental freedoms. State-sponsored or condoned malicious behavior is on the rise, including the spread of disinformation and cybercrimes such as ransomware, affecting the security and the resilience of critical infrastructure while holding at risk vital public and private assets. At the same time, countries have erected firewalls and taken other technical measures, such as Internet shutdowns, to restrict access to journalism, information, and services, in ways that are contrary to international human rights commitments and obligations. Concerted or independent actions of some governments and private actors have sought to abuse the openness of Internet governance and related processes to advance a closed vision. Moreover, the once decentralized Internet economy has become highly concentrated and many people have legitimate concerns about their privacy and the quantity and security of personal data collected and stored online. Online platforms have enabled an increase in the spread of illegal or harmful content that can threaten the safety of individuals and contribute to radicalization and violence. Disinformation and foreign malign activity is used to sow division and conflict between individuals or groups in society, undermining respect for and protection of human rights and democratic institutions.

OUR VISION

We believe we should meet these challenges by working towards a shared vision for the future of the Internet that recommit governments and relevant authorities to defending human rights and fostering equitable economic prosperity. We intend to ensure that the use of digital technologies reinforces, not weakens, democracy and respect for human rights; offers opportunities for innovation in the digital ecosystem, including businesses large and small, and, maintains connections between our societies. We intend to work together to protect and fortify the multistakeholder system of Internet governance and to maintain a high level of security, privacy protection, stability and resilience of the technical infrastructure of the Internet.

We affirm our commitment to promote and sustain an Internet that is open, free, global, interoperable, reliable, and secure and to ensure that the Internet reinforces democratic principles and human rights and fundamental freedoms; offers opportunities for collaborative research and commerce; is developed, governed, and deployed in an inclusive way so that underserved and underserved communities, particularly those coming online for the first time, can navigate it safely and with personal data privacy and protections in place; and is governed by multistakeholder processes. In short, an Internet that can deliver on the promise of connecting humankind and helping societies and democracies to thrive.

The Internet should operate as a single, decentralized network of networks—with global reach and governed through the multistakeholder approach, whereby governments and relevant authorities partner with academics, civil society, the private sector, technical community and others. Digital technologies reliant on the Internet, will yield the greatest dividends when they operate as an open, free, global, interoperable, reliable, and secure system. Digital technologies should be produced, used, and governed in ways that enable trustworthy, free, and fair commerce; avoid unfair discrimination between, and ensure effective choice for, individual users; foster fair competition and encourage innovation, promote and protect human rights; and, foster societies where:

- Human rights and fundamental freedoms, and the well-being of all individuals are protected and promoted;
- All can connect to the Internet, no matter where they are located, including through increased access, affordability, and digital skills;
- Individuals and businesses can trust the safety and the confidentiality of the digital technologies they use and that their privacy is protected;

- Businesses of all sizes can innovate, compete, and thrive on their merits in a fair and competitive ecosystem;
- Infrastructure is designed to be secure, interoperable, reliable, and sustainable;
- Technology is used to promote pluralism and freedom of expression, sustainability, inclusive economic growth, and the fight against global climate change.

PRINCIPLES TO PROMOTE THIS VISION

The partners in this Declaration intend to uphold a range of key principles, set out below, regarding the Internet and digital technologies; to promote these principles within existing multilateral and multistakeholder fora; to translate these principles into concrete policies and actions; and, work together to promote this vision globally, while respecting each other's regulatory autonomy within our own jurisdictions and in accordance with our respective domestic laws and international legal obligations. These principles are not legally binding but should rather be used as a reference for public policy makers, as well as citizens, businesses, and civil society organizations.

PROTECTION OF HUMAN RIGHTS AND FUNDAMENTAL FREEDOMS

- Dedicate ourselves, in conducting and executing our respective domestic authorities, to respect human rights, including as reflected in the Universal Declaration of Human Rights, as well as the principles of the rule of law, legitimate purpose, non-arbitrariness, effective oversight, and transparency, both online and offline, and call upon others to do the same.
- Promote online safety and continue to strengthen our work to combat violence online, including sexual and gender-based violence as well as child sexual exploitation, to make the Internet a safe and secure place for everyone, particularly women, children, and young people.
- Promote safe and equitable use of the Internet for everyone, without discrimination based on sex, race, color, ethnic, national or social origin, genetic features, language, religion or belief, political or any other opinion, membership of an indigenous population, property, birth, disability, age, gender identity or sexual orientation.
- Reaffirm our commitment that actions taken by governments, authorities, and digital services including online platforms to reduce illegal and harmful content and activities online be consistent with international human rights law, including the right to freedom of expression while encouraging diversity of opinion, and pluralism without fear of censorship, harassment, or intimidation.
- Protect and respect human rights and fundamental freedoms across the digital ecosystem, while providing access to meaningful remedies for human rights violations and abuses, consistent with international human rights law.
- Refrain from misusing or abusing the Internet or algorithmic tools or techniques for unlawful surveillance, oppression, and repression that do not align with international human rights principles, including developing social score cards or other mechanisms of domestic social control or pre-crime detention and arrest.

A GLOBAL INTERNET

- Refrain from government-imposed Internet shutdowns or degrading domestic Internet access, either entirely or partially.
- Refrain from blocking or degrading access to lawful content, services, and applications on the Internet, consistent with principles of Net Neutrality subject to applicable law, including international human rights law.
- Promote our work to realize the benefits of data free flows with trust based on our shared values as like-minded, democratic, open and outward-looking partners.
- Promote cooperation in research and innovation and standard setting, encourage information sharing regarding security threats through relevant international fora, and reaffirm our commitment to the framework of responsible state behavior in cyberspace.

INCLUSIVE AND AFFORDABLE ACCESS TO THE INTERNET

- Promote affordable, inclusive, and reliable access to the Internet for individuals and businesses where they need it and support efforts to close digital divides around the world to ensure all people of the world are able to benefit from the digital transformation.
- Support digital literacy, skills acquisition, and development so that individuals can overcome the digital divide, participate in the Internet safely, and realize the economic and social potential of the digital economy.
- Foster greater exposure to diverse cultural and multilingual content, information, and news online. Exposure to diverse content online should contribute to pluralistic public discourse, foster greater social and digital inclusion within society, bolster resilience to disinformation and misinformation, and increase participation in democratic processes.

TRUST IN THE DIGITAL ECOSYSTEM

- Work together to combat cybercrime, including cyber-enabled crime, and deter malicious cyber activity.
- Ensure that government and relevant authorities' access to personal data is based in law and conducted in accordance with international human rights law.
- Protect individuals' privacy, their personal data, the confidentiality of electronic communications and information on end-users' electronic devices, consistent with the protection of public safety and applicable domestic and international law.
- Promote the protection of consumers, in particular vulnerable consumers, from online scams and other unfair practices online and from dangerous and unsafe products sold online.
- Promote and use trustworthy network infrastructure and services suppliers, relying on risk-based assessments that include technical and non-technical factors for network security.
- Refrain from using the Internet to undermine the electoral infrastructure, elections and political processes, including through covert information manipulation campaigns.

- Support a rules-based global digital economy which fosters trade and contestable and fair online markets so that firms and entrepreneurs can compete on their merits.
- Cooperate to maximize the enabling effects of technology for combating climate change and protecting the environment whilst reducing as much as possible the environmental footprint of the Internet and digital technologies.

MULTISTAKEHOLDER INTERNET GOVERNANCE

- Protect and strengthen the multistakeholder system of Internet governance, including the development, deployment, and management of its main technical protocols and other related standards and protocols.
- Refrain from undermining the technical infrastructure essential to the general availability and integrity of the Internet.

We believe that the principles for the future of the Internet are universal in nature and as such we invite those who share this vision to affirm these principles and join us in the implementation of this vision. This Declaration takes into account, and expects to contribute to, existing processes in the UN system, G7, G20, the Organisation for Economic Cooperation and Development, the World Trade Organization, and other relevant multilateral and multistakeholder fora, the Internet Corporation for Assigned Names and Numbers, Internet Governance Forum, and Freedom Online Coalition. We also welcome partnership with the many civil society organizations essential to promoting an open, free, global, interoperable, reliable, and secure Internet, and defending fundamental freedoms and human rights online. Partners in this Declaration intend to consult and work closely with stakeholders in carrying forward this vision.

★ ★ ★ ★ ★

G7広島首脳コミュニケ（2023年5月）

- 2023年5月に開催されたG7広島サミットにおいて、5月20日、G7首脳より、広島首脳コミュニケが発出。
- G7首脳において、「民主的制度に対する信頼を損ない、国際社会における意見の対立を招く偽情報を含む外国からの情報操作及び干渉に対処することにより、情報環境を保護する」というコミットメントを再確認。
- その上で、「事実に基づく、質の高い、信頼できる情報の普及が確保されるよう取り組み、デジタル・プラットフォームがこのアプローチを支援するよう呼びかける」ことを宣言。

Human Rights, Refugees, Migration, Democracy（抄）

47. We reaffirm our shared belief that democracy is the most enduring means to advance peace, prosperity, equality and sustainable development. We reaffirm our commitment to protecting the information environment by supporting media freedom and online freedom, including protection from online harassment and abuse, internet shutdowns and disruptions, as well as addressing foreign information manipulation and interference, including disinformation, which is designed to undermine trust in democratic institutions, and sow discord in the international community. (...) We will work towards ensuring that fact-based, quality and trustworthy information is promoted, and call on digital platforms to support this approach. (...)

（仮訳）人権、難民、移住及び民主主義

47. 我々は、民主主義が平和、繁栄、平等及び持続可能な開発を促進するための最も揺るぎない手段であるとの我々の共通の信念を再確認する。我々は、オンラインでのハラスメントや虐待、インターネットの遮断や分断からの保護を含む、メディアの自由及びオンラインの自由を支援し、民主的制度に対する信頼を損ない、国際社会における意見の対立を招く偽情報を含む外国からの情報操作及び干渉に対処することにより、情報環境を保護するという我々のコミットメントを再確認する。（略）

我々は、事実に基づく、質の高い、信頼できる情報の普及が確保されるよう取り組み、デジタル・プラットフォームがこのアプローチを支援するよう呼びかける。（略）

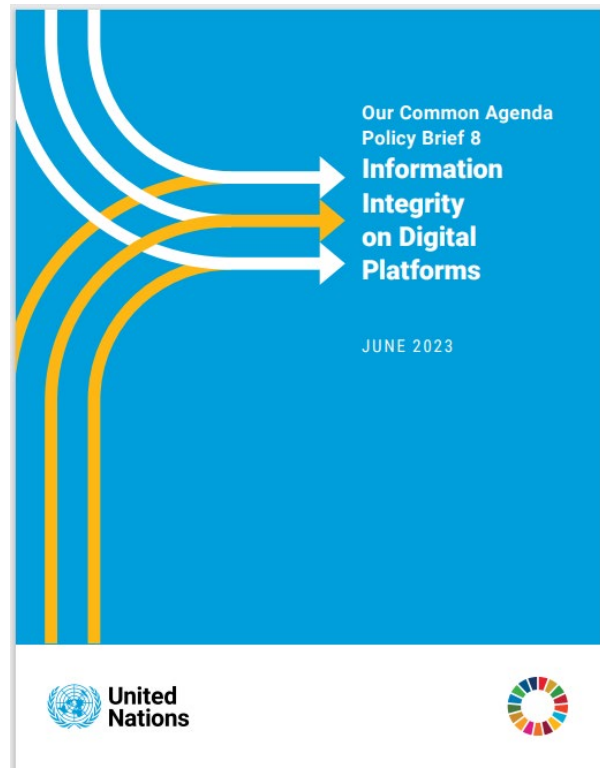
国連「デジタルプラットフォームにおける情報インテグリティ」（2023年6月）

- 2021年9月の国連事務総長報告書「私たちの共通の課題（Our Common Agenda）」で挙げられた12のコミットメントの一つである「7. デジタル分野での協力を改善する（Improve digital cooperation）」について、グローバル・デジタル・コンパクト（GDC）の詳細が記載された国連事務総長報告書が、2023年5月に公表。
- 2023年5月の当該報告書における「E. デジタルにおける信頼と安全」において、「偽情報、ヘイトスピーチ、その他の有害なオンラインコンテンツに対処するため、デジタルプラットフォームとユーザーに対する強固な説明責任基準および基準を策定」が目標として規定。
- また、同6月、「デジタルプラットフォームにおける情報インテグリティ（Information Integrity on Digital Platforms）」が公表され、デジタル・プラットフォーム上の情報インテグリティに対する行動規範を通じた世界的な行動の枠組みが提案。



7. デジタル分野での協力を改善する

- **グローバル・デジタル・コンパクト**により：
 - すべての学校を含め、すべての人々をインターネットにつなぐ
 - インターネットの断片化を回避する
 - データを保護する
 - オンラインで人権を適用する
 - 差別や誤解を招くコンテンツに対する説明責任基準を導入する
 - 人工知能(AI)の規制を強化する
 - グローバルな公共財としてのデジタルコモンズ



オンライン上の情報インテグリティに関するグローバル宣言（2023年9月）

- 2023年9月20日、カナダとオランダにより、「オンライン上の情報インテグリティに関するグローバル宣言（the Global Declaration on Information Integrity Online）」が立上げ。日本のほか、米国、英国、仏国、独国、豪州、ニュージーランド等を含む32か国が参加。
- 参加各国に対して、次の通り、具体的な措置をとることを求めるとともに、産業界に対しても一定の取組を期待。

【参加各国による具体的な措置】

The Declaration calls upon all participating states to take concrete steps to:

1. Abstain from and condemn state-led disinformation campaigns
2. Respect, promote and fulfill the right to freedom of expression
3. Implement relevant legislation in compliance with international law
4. Avoid stifling freedom of expression under the guise of countering disinformation
5. Promote stronger civic education online and digital literacy
6. Support independent media, news and journalism
7. Take active steps to address disinformation targeting groups in vulnerable situations

【産業界による一定の取組】

Invites industry to:

1. Respect the rule of law and commit to not abusing human rights
2. Enhance transparency in advertising, algorithm and content moderation processes
3. Provide clear and timely redress mechanisms to users
4. Support for independent researchers and fact-checkers
5. Abide by the UN Guiding Principles on Business and Human Rights