諸外国等における対応状況

デジタル空間における情報流通の健全性を巡るリスク・問題について、諸外国等においても、情報伝送 PF サービス等の利用の拡大が今後も見込まれていること、情報伝送 PF 事業者等が同様のアーキテクチャ等によりグローバルにサービスを展開していること、また、生成 AI 等の新たな技術やサービスが進展・普及していること等から、我が国特有の課題ではなく、諸外国等が共通して抱えている課題になっているため、欧米等の国・地域等における法制度等の対応状況を整理する。

Ⅰ　米国

１．法制度に関する動向

（１）情報伝送 PF 事業者の取組への規制に関する議論

米国では、合衆国憲法修正第１条や通信品位法第 230 条により、伝統的に、情報伝送 PF 事業者に対して広範な免責が与えられてきているが、近年、コンテンツモデレーションをはじめとする情報伝送 PF 事業者の取組への規制に関する議論が進みつつある。具体的には、連邦レベルでは、通信品位法の改正案の提出等が実施され、州レベルでは、テキサス州法やフロリダ州法が制定され、その合憲性が争われている。他方、このような情報伝送 PF 事業者の責任に関する議論が進む中、同事業者に対する規制など、政府による介入の妥当性に関しても議論されており、同事業者を巡る規制の在り方について、様々な議論が進んでいる。

連邦最高裁判所の判例[1]において、合衆国憲法修正第１条に基づき、言論・出版の自由（表現の自由）が手厚く保障され、情報伝送 PF サービス上の表現の自由も手厚く保障されている。このような表現の自由については、国家からの自由として理解する見解が支配的である一方で、企業等私人による制約からの表現の自由の保障には消極的であるとされている[2]。そのため、米国における情報伝送 PF 事業者に対する規制については、修正第１条の強力な保護があるために、自由放任（レッセフェール）が基本であり、法規制を行うことはほぼ不可能とされている。特に、米国の合衆国憲法については、欧州連合（以下「EU」という。）と異なり、あくまで国家権力の制御に用いられるものであることから、私人である情報伝送 PF 事業者に対しては、ステイト・アクションの法理[3]を充たさなければ修正第１条は適用されないとされている。

---

[1] New York Times v. Sullivan, 376 U.S. 254 (1964)や United States v. Alvarez, 567 U.S. 709 (2012)など。

[2] 本検討会資料 WG8-3 p.4

[3] 私人の行為が公的機能を果たしている場合や、裁判所が私人の行為を執行している場合、政府ないし州の関与、授権、奨励が存在する場合には、私人の行為を政府ないし州の行為とみなし、憲法によって拘束する、というもの。松井茂記『アメリカ憲法入門〔第7版〕』(有斐閣、2012 年)p.210-215.

また、通信品位法第 230 条[4]により、「双方向コンピュータ・サービスの提供者又は利用者」に対する非常に広範な免責が与えられている。同条(c)(1)では、双方向コンピュータ・サービスのいかなる提供者又は利用者をも、別の情報コンテンツ提供者が提供する情報の発行者又は代弁者として扱ってはならないと定められており、ユーザー等第三者により発信された情報について、情報伝送 PF 事業者に広範な免責が与えられてきている。また、同条(c)(2)では、合衆国憲法上保護されているか否かにかかわらず、提供者が好ましくないと判断する素材へのアクセス又はその利用可能性を制限するために、誠実かつ任意に取られた措置については、双方向コンピュータ・サービスの提供者が免責されている。

以上のように、米国では、情報伝送 PF 事業者に対して広範な免責が与えられてきているが、近年、コンテンツモデレーションをはじめとする同事業者の取組への規制に関する議論が進みつつある。同事業者によるコンテンツモデレーションは、言論の自由として保障されるという見解が依然として有力であるものの、最近では同事業者が持つ社会的権力から表現の自由を確保する必要性を説く学説も有力になっているとされている。また、強すぎる表現の自由の保障が、情報伝送 PF 事業者に対する規制を含め、必要な規制を妨げているのではないかとの問題意識が広がっているとの意見もある[5]。

近年では、情報伝送 PF 事業者の責任を強化する動向も強まっており、通信品位法第 230 条に関しては、2018 年に性的人身取引を促す情報について、同事業者の免責範囲を限定する FOSTA-SESTA (Fight Online Sex Trafficking Act and the Stop Enabling Sex Traffickers Act) が制定されている。また、2020 年には、当時の Twitter 社が米国大統領のツイートに対して注釈を付けたことにより、米国大統領がこの取組に反発をし、米国大統領は「オンラインの検閲の防止に係る大統領令」を制定し[6]、関連する連邦政府機関に対して通信品位法第 230 条の解釈の見直しの検討等を求め、政権交代後においても、民主党及び共和党両党の議員から通信品位法第 230 条の改正論が提案されている。加えて、超党派の議員による改正案[7]も提出されているが、党派対立の影響もあり成立してい

---

[4] 通信品位法第 230 条(c)の条文(和訳)は下記のとおり。
(c)不快感を与える素材の「良きサマリア人」によるブロック及び識別に対する保護
(1)発行者又は代弁者としての取扱いー双方向コンピュータ・サービスのいかなる提供者又は利用者をも、別の情報コンテンツ提供者が提供する情報の発行者又は代弁者として扱ってはならない。
(2)民事責任ーいかなる双方向コンピュータ・サービスの提供者又は利用者も、次の事項を理由として責任があるとみなしてはならない。
(A)当該提供者又は利用者がわいせつな、淫らな、好色な、卑猥な、過度に暴力的な、困惑させるようなその他の好ましくないと判断した素材が憲法上保護されているかどうかにかかわらず、当該素材へのアクセス又はその利用可能性を制限するために誠実に、かつ、任意に取った措置。

[5] 本検討会 WG 第8回 成原准教授の発言は以下のとおり:「他方で、最近では、プラットフォーム事業者の影響力の増大に伴って、こうした新たな社会的権力から表現の自由を確保する必要性を説く学説も有力になっています。さらに、リベラルな法学者を中心に、強すぎる表現の自由の保障が、プラットフォーム規制を含めて、必要な規制を妨げているのではないかと、そういう問題意識も広がっているように感じております。」(議事概要より抜粋)

[6] 大統領令を受け、商務省、司法省及び連邦通信委員会(Federal Communication Commission: FCC)による通信品位法第 230 条の検証がなされたが、トランプ氏が 2020 年の大統領選挙でバイデン氏に敗れたため、当時のパイ FCC 委員長はトランプ政権下での規則の制定を断念するとした。

[7] 例えば、2023 年 12 月時点で、超党派により連邦上院に「プラットフォーム説明責任及び透明化法案」(The Platform Accountability and Transparency Act)が提出された。本法案は、研究者による特定のデータアクセス、自動データ収集に関す

ない状況となっている。

以上のような動きは州レベルでも存在し、一部の州は情報伝送 PF 事業者によるコンテンツモデレーションの在り方を規制する州法を制定し、訴訟において州法の合憲性が争われるなど、コンテンツモデレーションに対する政府の関与の在り方が議論されている。具体的には、フロリダ州やテキサス州といった保守派の影響力が強い州において、同事業者によるコンテンツモデレーションを規制する州法が制定されている。

テキサス州法では、一月当たり米国内で 5,000 万名以上のアクティブユーザーを有する大規模な情報伝送 PF 事業者に対し、ユーザーやその投稿コンテンツを見解に基づき検閲することが禁止されている。さらに、ポリシーや透明性報告書の公表、不服申立ての手続整備も同事業者に義務付けられている。これに対し、同事業者の業界団体である NetChoice は、当該州法は言論の自由を保障した合衆国憲法修正第１条に違反するとし、州法の執行の差止めを求め提訴した[8]。第一審の連邦地裁では、州法の執行が予備的に差止められている。他方、控訴審においては、この規制は同事業者による検閲から利用者の表現の自由を守るための規制であると捉え、必ずしも合衆国憲法修正第１条に違反するものではないと評価され、第一審の予備的差止め命令が停止されている。これに対して、連邦最高裁判所は控訴審の決定を取り消している。

また、フロリダ州法では、年間１億ドル以上の収入、又は月間１億名以上の参加者を有する大規模な情報伝送 PF 事業者に対し、公職選挙の候補者及び報道機関の排除、公職選挙の候補者及び報道機関の投稿したコンテンツを排除することが禁止されている。フロリダ州法においても、合衆国憲法修正第１条違反を理由にした州法の予備的差し止め命令を巡り、係争中[9]の状況となっている。第一審では、州法の一部の執行の予備的差止めが認められ、控訴審においても第一審の予備的差止命令が一部支持されている。しかしながら、控訴審では、コンテンツモデレーション基準の公表など、州法の透明性義務の一部について、違憲となる見込みは認められないとして、原審の予備的差止命令の一部を取り消している。そして、2024 年２月 26 日の連邦最高裁判所では、口頭弁論が行われたところ、連邦政府の訟務長官は、情報伝送 PF 事業者の権力は増大しており、規制する必要があるものの、同事業者のコンテンツモデレーションを規制する必要はなく、州法を支持する必要はないと述べている[10]。

そのほか、情報伝送 PF 事業者の民事責任が問われているケースもある。例えば、ISIS によるテロ被害者の遺族が、ISIS のコンテンツを配信・拡散した同事業者は、テロ攻撃による被害に対して幇助責任等を負うと主張して損害賠償を求めた訴訟[11]では、Google が YouTube で ISIS の広告を承認し、収益を配分していたことから、連邦最高裁判所が通信品位法の解釈を示す可能性が期待されたが、通信品位法の解釈には立ち入らず、同事業者の行為と本件テロ攻撃との間に直接の関係が認められないことなどを理由に、請求

る限界付けられた法的セーフハーバー、開示を通じた強化された透明性という3つの柱で構成されている。

[8] NetChoice, LLC v. Paxton, 603 U.S. __ (2024)

[9] Moody v. NetChoice, LLC, 603 U.S. __ (2024)

[10] 2024 年7月 1 日、フロリダ州とテキサス州のコンテンツモデレーション規制法に関する差止訴訟は、連邦最高裁で破棄差戻しの判断がされている。(https://www.supremecourt.gov/opinions/23pdf/22-277_d18f.pdf)

[11] Twitter, Inc. v. Taamneh, 598 U.S. 471 (2023)、Gonzalez v. Google LLC, 598 U. S. 617 (2023)

は認められないと判示されている。

以上のように、コンテンツモデレーションをはじめとする情報伝送 PF 事業者の取組について、同事業者の責任に関する議論が進む一方で、同事業者に対する規制等、政府による介入の妥当性に関する議論も行われている。また、新型コロナ禍の偽情報対策等の一環で、情報伝送 PF 事業者に対して行われてきた米国政府による「口先介入」に関して、合衆国憲法修正第１条違反を理由にした差止訴訟[12]が係争中となっている。この事件はミズーリ州、ルイジアナ州、そして自らの投稿が削除等された個人が、連邦政府が情報伝送 PF 事業者にコンテンツの削除等について要請を行うことは合衆国憲法修正第１条に違反するなどと主張し、差止めを求めて提訴しているものである。連邦地裁においては、連邦政府機関の職員に、同事業者に対し、コンテンツモデレーションに関する決定の強制や、同事業者のコンテンツモデレーションに関する決定をコントロールすることを禁じる予備的差止命令を発している。これに対し、連邦最高裁判所は予備的差止命令を停止し、2024 年３月 18 日に連邦最高裁判所で本案の口頭弁論が開かれている。2024 年６月 26 日、連邦最高裁判所は米国政府と情報伝送 PF 事業者がやりとりした事実と、投稿内容の制限との間の因果関係が証明されていないと指摘し、また、将来検閲が起きる危険性についても立証不足であり、原告には情報伝送 PF 事業者への要請差止めを求める資格がないとして、破棄差戻しの判断をしている[13]。

（２）なりすまし詐欺行為を抑止するための対策

米国においては、なりすまし詐欺による被害額が 2020 年の３倍以上となる 11 億ドルにのぼると報告されている。こうしたなりすまし詐欺への対策として、2024 年４月１日、米国連邦取引委員会により、政府及び企業のなりすまし行為に関する新たな規則が施行されている。この規則により、政府や企業になりすます詐欺行為を抑止するための対策が強化され、損害を被った場合、被害者が金銭の返還や違反者に対する民事罰の適用を求めて、連邦裁判所に提訴できるようになっている。さらに、同委員会により、この規則について、政府や企業のなりすましに加えて、個人へのなりすましを禁止する規則改正案[14]が公表され、2024 年４月 30 日までパブリックコメントが行われたところである。

（３）刑事罰による対応

米国では、言論の自由が広く認められており、政府が言論の内容を規制することは、わいせつ、児童ポルノ、誹謗中傷、脅迫、喧嘩言葉[15]、虚偽広告等の一部の表現を除き、

---

[12] Missouri v. Biden, 2023 U.S. App. LEXIS 23965.; Murthy v. Missouri, 601 U.S. ___ (2023).
[13] 日本経済新聞「米政府の誤情報対策要請、『検閲』との訴え却下　最高裁」2024 年6月 27 日 (https://www.nikkei.com/article/DGXZQOGN26DXX0W4A620C2000000/)
[14] https://www.ftc.gov/system/files/ftc_gov/pdf/r207000_impersonation_snprm.pdf
[15] 「fighting words」(Chaplinsky v. New Hampshire, 315 U.S. 568 (1942))

基本的に認められていない。言論の自由により保護されない例外的な表現については、法律や判例法により規制されており、民事上・刑事上の責任が発生し得る。

刑事上の責任に関しては、商業目的の誤情報、虚偽の選挙運動資料、爆破予告や虚偽の犯罪報告等につき、法規制が存在する。もっとも、虚偽情報を用いた脅迫行為等が問題とされた裁判例は存在するものの、民事裁判に比べて件数は少ない[16]。

２．その他の取組

（１）大学における取組

米国においては、大学と伝統メディア等における連携・協力が進んでいる。例えば、大学主導で、学生記者が州議会に取材するプログラムが存在する。教育、公衆衛生、住宅、銃の所有等に関する重大な議論が州レベルで発生し、州議会の報道の重要性が向上する一方で、フルタイムの州議会記者の数が 2014 年以来 34%減少している。こうした記者不足に対し、大学主導の報道機関により、学生州議会記者が議会を訪問している。2022 年には約 250 人の学生州議会記者が 1,000 以上の記事を制作し、17 州の 1,200 以上のメディアに提供されるなど、大学主導で記者を育てる仕組みが存在する。

また、米国においては、体系的にジャーナリズムを学べる大学が数多くあり、ジャーナリズム専攻がある大学は 335 校存在するとされている。さらに、地域レベルで人材育成を養成する事例としては、米国バーモンド大学が 2019 年にコミュニティーニュースセンターを立ち上げている。

---

[16] ⑴ United States v. Douglas Mackey, 2023 U.S. Dist. LEXIS 40796 (E.D.N.Y. Mar. 10, 2023)(2023.10.18)
2016 年の大統領選挙において、「特定の候補者の支持者は Facebook や Twitter で特定のハッシュタグを付けて投稿することで投票することができ、また、投票しなければならない」、「特定の候補者の支持者は、当該候補者のファーストネームを特定の電話番号にテキスト送信することで投票することができる」という内容のミームや SNS 投稿を作成して、広めた。懲役 7 か月、執行猶予 2 年、罰金 1 万 5000 ドル（控訴中）。

⑵ United States v. Christopher Charles Perez, 43 F.4th 437 (5th Cir. 2022)(2022.8.3)
食料品店で COVID-19 陽性の人物にお金を支払って商品をなめさせたという虚偽の投稿をした。懲役数か月、執行猶予 3 年、罰金 1000 ドル（量刑について再審理中）。

⑶ United States v. v. Tristan H. Kelly and Cody T. Ritchey, Case No. 5:18-cr-00034-DCR-REW (E.D.Ky.)(2018.9.28)
別人の名前と写真を用いて SNS アカウントを作成し、当該アカウントを使用して当該別人が銃器を使って指定された学校を銃撃すると示唆した。Kelly 氏には禁固 21 か月、Ritchey 氏には禁固 27 か月。

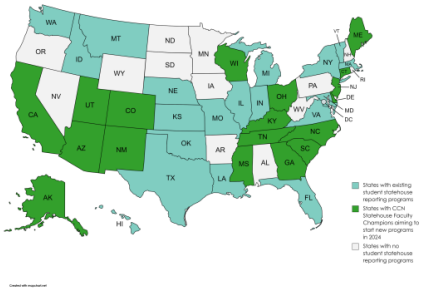
4-1.「制作者」を育て、増やす仕組み～大学主導～

【米国の大学主導の州議会学生記者取材プログラム】

出典：https://www.uvm.edu/ccn/why-student-statehouse-reporting-matters

- 教育、公衆衛生、住宅、銃の所有などに関する重大な議論が州レベルで発生する中、州議会の報道はかつてないほど重要になっている
- ピュー・リサーチ・センターによると、必要性が高まっているにもかかわらず、フルタイムの州議会記者の数は2014年以来34%減少した

- そのギャップを埋めるため大学主導の報道機関が学生州議会記者を議会に投入
- 2022年には、約250人の学生州議会記者が1,000以上の記事を制作し、17州の1,200以上のメディアに提供された。



＜出典：本検討会資料 18－２－２＞

このほか、スタンフォード大学において、ソーシャルメディアに焦点を当てた研究プログラムである「スタンフォード・インターネット観測所[17]」が 2019 年に創設され、2020 年アメリカ大統領選挙等における偽・誤情報等の実態調査等が実施されている[18]。

（２）アメリカ心理学会における取組

アメリカ心理学会（American Psychological Association）は、米国最大の心理学を代表する専門組織であり、会員には、146,000 人を超える研究者、教育者、臨床医、コンサルタント、学生が含まれる。同学会は、心理学の知識の創造、伝達、応用を促進し、社会に貢献することで、人々の生活を向上させることに取り組んでいる。同学会は 2023 年 11 月に誤情報対策に関する共同声明を発表し、当該共同声明とともに、「健康の誤情報を理解し、それに対抗するための心理科学の活用」という報告書が公表されている。同学会は誤情報への効果的な対処法として、８つの具体的なレコメンデーションを提示するとともに、現状の心理科学による誤情報介入の限界を４つの観点で言及している。

【レコメンデーション】

- 訂正なしに誤情報を繰り返すことを避ける。
- ソーシャルメディア会社と協力して、有害な誤情報の拡散を理解し、軽減する。
- 健康的な行動を促進することが既に証明されているツール（例えば、カウンセリン

[17] https://cyber.fsi.stanford.edu/io

[18] 2024 年6月 13 日にウェブメディア「プラットフォーマー」がスタンフォード・インターネット観測所の閉鎖を報道した（https://www.platformer.news/stanford-internet-observatory-shutdown-stamos-diresta-sio/）。スタンフォード大学はこの報道を否定しており、この研究プログラムが新しい体制になることは認めている。

グ、技能訓練、インセンティブ、社会規範）を用いた誤情報是正戦略を用いる。
- 信頼できるソースを活用して、誤情報に対応し、正確な医療情報を提供する。
- 証拠に基づいた方法を何度も繰り返して誤情報を発見する。
- 早い時期からスキルと回復力を身につけることで、敏感な視聴者に接種する、誤情報のプリバンクをする。
- 誤情報の科学的研究のために、ソーシャルメディア会社にデータアクセスと透明性を要求する。
- 健康に関する誤情報の心理学に関する基礎及びトランスレーショナルリサーチ（それに対処する効果的な方法を含む）に資金を提供する。

【誤情報介入の一般的な限界】
- 特に高等教育を受けていない人、農村地域に住んでいる人等のサンプルによる研究が不足していること。
- 効果を維持するために追加接種をいつどのように行うべきかも含め、長期にわたる介入の効果を検討するために、更に長期にわたる研究を実施する必要があること。
- フィールド研究は実施が困難で費用がかかること。
- 既存の研究は、ほとんど排他的に独立した介入の試験に焦点を当てており、介入がお互いを増幅させるのか等、相互作用に関する理解が限られていること。

米国心理学会　報告書（2023年11月）

- 2023年11月、米国心理学会は「健康の誤情報を理解し、それに対抗するための心理科学の活用」という報告書を公表。
- 米国心理学会は誤情報への効果的な対処法として、**８つの具体的なレコメンデーションを提示**。また、**現状の心理科学による誤情報介入の限界を４つの観点で言及**。

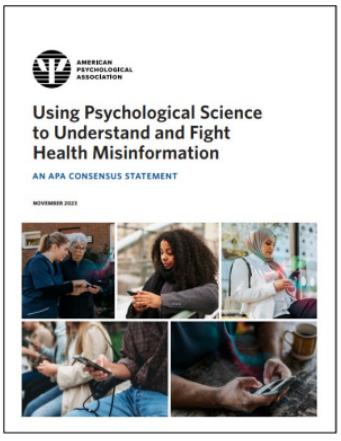
【レコメンデーション】（仮訳）

【出典】米国心理学会のサイト（https://www.apa.org/pubs/reports/health-misinformation）

1. 訂正なしに誤情報を繰り返すことを避ける。
2. ソーシャル・メディア会社と協力して、有害な誤情報の拡散を理解し、軽減する。
3. 健康的な行動を促進することがすでに証明されているツール(例えば、カウンセリング、技能訓練、インセンティブ、社会規範)を用いた誤情報是正戦略を用いる。
4. 信頼できるソースを活用して、誤情報に対応し、正確な医療情報を提供します。
5. 証拠に基づいた方法を何度も繰り返して誤情報を発見する。
6. 早い時期からスキルと回復力を身につけることで、敏感な視聴者に接種する、誤情報のプリバンクをする。
7. 誤情報の科学的研究のために、ソーシャルメディア会社にデータアクセスと透明性を要求する。
8. 健康に関する誤情報の心理学に関する基礎およびトランスレーショナルリサーチ（それに対処する効果的な方法を含む）に資金を提供すること。

【誤情報介入の一般的な限界】（仮訳）

1. 特に高等教育を受けていない人、農村地域に住んでいる人等のサンプルによる研究が不足していること。
2. 効果を維持するために追加接種をいつどのように行うべきかも含め、長期にわたる介入の効果を検討するために、更に長期にわたる研究を実施する必要があること。
3. フィールド研究は実施が困難で費用がかかること。
4. 既存の研究は、ほとんど排他的に独立した介入の試験に焦点を当てており、介入がお互いを増幅させるのか等、相互作用に関する理解が限られていること。



＜出典：本検討会資料５－１－２＞

（３）その他の取組

米国国務省 GEC（Global Engagement Center）と CISA（the Cybersecurity and Infrastructure Security Agency）は、「Harmony Square」というインタラクティブなオンラインゲーム（15 歳以上の利用者が対象）を開発している。当該ゲームは、ソーシャルメディア上で人々を誤解させたり、操作したり、フォロワーを増加させたり、政治的利益のために社会的な緊張を利用したりするために使用される一般的な戦術に対応するようなゲームとなっている。ハーモニースクエアのウェブページには、デジタルリテラシーのスキルを教えるためにゲームを使用するヒントをまとめた教育者向けのツールキットも掲載されている。

〈Harmony Square のウェブページ〉

Ⅱ　EU

１．法制度に関する動向

（１）デジタルサービス法（Digital Services Act）

①　総論

EU では、プロバイダの責任を規定してきた「電子商取引指令」（Directive on electronic commerce）[19]について、違法・有害情報に対する情報伝送 PF 事業者等の責任・責務や透明性の在り方が全面的にアップデートされ、「デジタルサービス法」（Digital

[19] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market（2000 年7月 17 日発効）。

Services Act。以下「DSA」という。）が 2022 年に発効されている[20]。

DSA は、デジタルサービスを規制する必要性に対応するため、欧州委員会が 2020 年 12 月に提出した立法構想であるデジタルサービスパッケージの一部である。DSA は、オンライン上の仲介サービス（intermediary service）が EU の域内市場において適切に機能するよう貢献することを目的に、特定の類型の仲介サービス提供者向けのデューディリジェンス義務（due diligence obligations）に関するルールを定めるなどしており（第１条）、EU 域内に事務所が所在する、又は域内に所在するサービス受領者（recipient of the service）に対して提供される仲介サービスに適用される（第２条）。なお、DSA の規律対象となる仲介サービスとは、次の３種類のサービスの総称である（第３条(g)）[21]。

① 「導管サービス」（mere conduit service）：サービス受領者が提供する情報を通信ネットワークで伝送し、又は通信ネットワークへのアクセスを提供するサービス
② 「キャッシングサービス」（caching service）：情報の送信をより効果的に行うことだけを目的として自動的、中間的かつ一時的に情報を保存するサービス
③ 「ホスティングサービス」（hosting service）：サービス受領者から提供され、又は要求された情報の保存を行うサービス

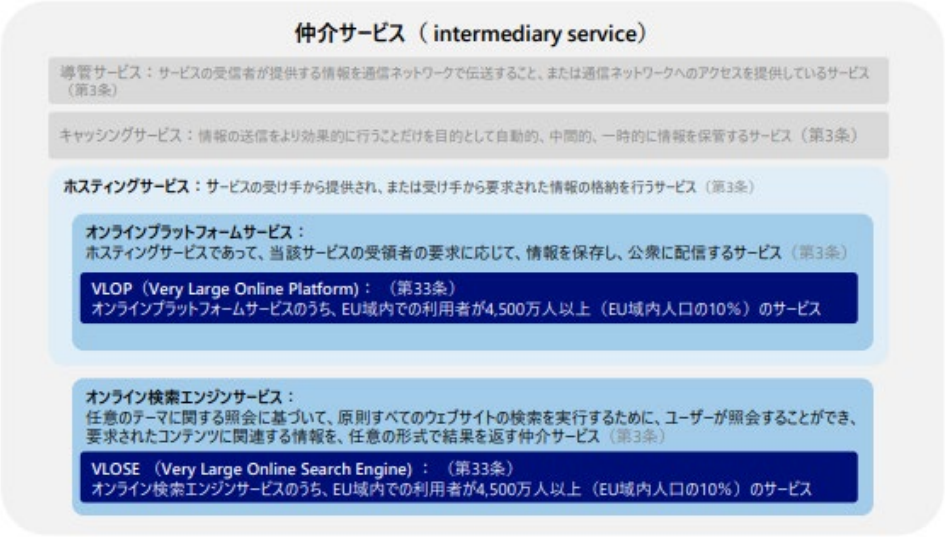
以上のうち、③「ホスティングサービス」の提供者に対しては、仲介サービスの提供者全般に課されている違法コンテンツに関する措置命令を受けた場合の対応結果報告義務（第９条）等に加え、違法コンテンツに関する通報受付体制の整備義務（第 16 条）や削除等の対応をしたコンテンツの発信者への理由通知義務（第 17 条）等の追加的な義務が課されている。さらに、ホスティングサービスのうち、サービス受領者の要求に応じて情報を保存し、公衆に配信するサービスは「オンラインプラットフォーム」（online platform）と定義され（第３条(i)）、その提供者に対しては、広告やレコメンダーシステムに関する透明性確保義務（第 26 条・第 27 条、後述）等を含む、より広範な規律の対象とされている。

加えて、EU 域内で月間アクティブサービス受領者数 4,500 万（EU 域内人口の 10%）以上を有するオンラインプラットフォーム及びオンライン検索エンジンとして、それぞれ欧州委員会に指定された「超大規模オンラインプラットフォーム」（very large online platform。以下「VLOP」という。）及び「超大規模オンライン検索エンジン」（very large online search engine。以下「VLOSE」という。）の提供者に対しては、リスク評価及び軽減措置の実施義務（第 34 条・第 35 条、後述）や危機対応メカニズム（第 36 条、後述）等の特別な義務が課されている。

---

[20] 以下、EU における動向は、本検討会第4回及び WG 第 12 回の生貝構成員の発表、WG 第8回の山本健人構成員の発表、並びに WG 第5回、第8回及び第 12 回の株式会社野村総合研究所の発表を元に記載。

[21] 原則として全てのウェブサイト（又は特定の言語のウェブサイト全て）の検索を実行するためにキーワード、音声、フレーズその他の形式によるクエリをユーザーに入力させ、要求されたコンテンツに関連する情報が含まれる検索結果を任意の形式で返す「オンライン検索エンジン」（online search engine）も仲介サービスの一つとされるが（第3条(j)）、これが3種類のサービスのうちいずれに当たるかは法文上明確でない。



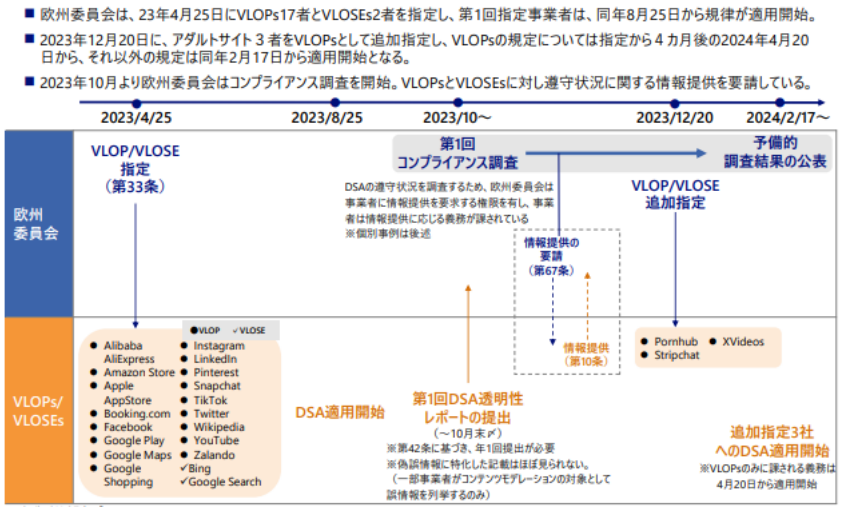
＜出典：本検討会資料 WG５－１－６＞

規律	該当条文		仲介サービス	ホスティングサービス	オンラインプラットフォーム	VLOP・VLOSE
違法コンテンツに関する措置命令・情報提供の命令	第二章	第9条・第10条	●	●	●	●
連絡先（対DSC、対欧州委員会、対関係理事会）、サービス提供者の窓口、法定代理人		第11条・第12条・第13条	●	●	●	●
利用規約の要件		第14条	●	●	●	●
透明性報告義務		第15条	●	●	●	●
利用者への通知・行動の仕組み、情報提供・理由の記載義務		第16条・第17条		●	●	●
刑事犯罪の疑いに関する通知		第18条		●	●	●
内部苦情処理体制・救済の仕組みと法廷外紛争解決		第20条・第21条			●	●
信頼された旗手		第22条			●	●
悪用に対する措置と保護		第23条			●	●
オンライン・プラットフォームのプロバイダーに対する透明性報告義務		第24条			●	●
オンラインインターフェースのデザインと構成		第25条			●	●
オンラインプラットフォームでの広告		第26条			●	●
レコメンダーシステムの透明性		第27条			●	●
未成年者のオンラインでの保護	第三章	第28条			●	●
超大規模オンライン検索エンジン		第33条				●
リスク評価、リスク軽減		第34条・第35条				●
危機対応メカニズム		第36条				●
独立監査（外部リスク監査と公的説明責任）		第37条				●
レコメンダーシステム		第38条				●
オンライン広告の透明性向上		第39条				●
データへのアクセスと精査（当局・研究者）		第40条				●
コンプライアンス機能		第41条				●
透明性に関する報告義務		第42条				●
監督手数料		第43条				●
標準		第44		●	●	●
行動規範、オンライン広告・アクセシビリティの行動規範		第45・第46条・第47条		●	●	●
危機対応への協力		第48条		●	●	●

＜出典：本検討会資料 WG５－１－６＞

この点、オンラインプラットフォーム及びオンライン検索エンジンの提供者において、DSA の規定に基づき、2023 年２月 17 日までに EU 域内の月間アクティブサービス受領者数が公表され、これに基づき、同年４月 25 日に欧州委員会により VLOP の提供者として 17 者及び VLOSE の提供者として２者が指定されている。

同年８月 25 日には、DSA の規定のうち VLOP 及び VLOSE の提供者を規律する条項の適用が開始され、同年 12 月 20 日には新たに VLOP の提供者として３者が追加指定されている。その後、2024 年２月 17 日に、その他の条項を含め DSA の適用が全面的に開始されている。



＜出典：本検討会資料 WG５－１－６）＞

②　違法コンテンツ・利用規約違反情報への対応義務

DSA において、「違法コンテンツ」（illegal content）については、「それ自体、又は製品の販売やサービスの提供を含む活動に関連して、EU 法又は EU 法に準拠している加盟国の国内法に反する情報」と定義されている（第３条(h)）。

DSA には、仲介サービスの提供者に違法コンテンツの削除を直接義務付ける規定は置かれていない一方、DSA 以外の EU 法や、EU 法に準拠している加盟国の国内法に基づき、司法機関や行政機関から、違法コンテンツについて何らかの対応を求める措置命令が仲介サービスの提供者に出される場合があり、そうした措置命令を受けた仲介サービスの提供者に対し、不当に遅滞することなく当該司法・行政機関へ対応結果を報告することが義務付けられている（第９条）。

また、ホスティングサービスの提供者に対して、あらゆる個人や団体から、違法コンテンツと思われる情報が存在する旨の通報を受け付ける体制の整備が義務付けられている（第 16 条）。

以上に加え、オンラインプラットフォームの提供者に対して、違法コンテンツの検知・特定・通報を目的とした専門性と能力を有するものとして各加盟国のデジタルサービス調整機関（後述）に指定された「信頼できる通報者」（trusted flagger）[22]からの通報を優先的に処理するために必要な技術的・組織的措置を講ずる義務（第 22 条）や、明らかな違法コンテンツを頻繁に投稿するサービス受領者に対し、事前警告を発した上で合理的な期間、サービスの提供を停止する義務（第 23 条）が課されている。

違法コンテンツに限らず、仲介サービスの提供者が用意する利用規約に違反した情報を含めた対応に関しては、コンテンツモデレーション [23]の目的で用いられるポリシー、

---

[22] 2024 年6月現在、欧州委員会のウェブサイト（https://digital-strategy.ec.europa.eu/en/policies/trusted-flaggers-under-dsa）上では、フィンランド運輸通信庁（Traficom）によって指定された The Copyright Information and Anti-Piracy Centre (CIAPC)が「信頼できる通報者」として公表されている。

[23] DSA 上、「コンテンツモデレーション」（content moderation）の用語は、「自動化されているか否かを問わず、仲介サービスの

手順、手段及びツールに関する情報を利用規約に含めることが求められている（第 14 条）。また、仲介サービスの提供者に対して、自ら調査して行ったコンテンツモデレーションに関する有意義かつ包括的な情報（違法コンテンツや利用規約違反情報の種類別・検知方法別・具体的な制限措置別の件数のほか、自動化ツールの使用、担当者への訓練や支援のために講じた措置等を含む。）等を記載した透明性レポート（transparency report）を毎年公表することが義務付けられている（第 15 条）。透明性レポートについては、例えば、VLOP 及び VLOSE の提供者に対しては言語ごとの人的リソースに関する情報や使用した自動的手段の正確性に関する指標等をも含めることが求められ、頻度も６か月に１回とされるなど（第 42 条）、段階的な規律が置かれている。さらに、ホスティングサービスの提供者に対しては、削除等の対応をしたコンテンツの発信者への理由通知も義務付けられている（第 17 条）。

③　広告やレコメンダーシステムに関する透明性確保義務

DSA においては、オンラインプラットフォームの提供者に対し、そのオンラインインターフェイス[24]上に掲載される広告[25]について、①広告である旨、②当該広告の名義人、③②以外の資金提供者がいる場合における当該資金提供者、④配信先を決めるために使われる主なパラメータとその変更方法を明確に表示する義務が課されている（第 26 条）。これに加え、VLOP 及び VLOSE の提供者に対しては、更に広告の内容や当該広告のターゲティング対象とされたグループごとの閲覧者数等の情報を含む広告レポジトリを作成し、広告の掲載期間中及び掲載終了後１年間公開することが求められている（第 39 条）。

以上のほか、オンラインプラットフォームの提供者に対して、利用規約中に、①レコメンダーシステム[26]で用いられる主なパラメータ、②利用者において当該主なパラメータを修正し、又は当該主なパラメータに影響を与えるオプションを明記することも義務付けられている（第 27 条）。このうち①については、おすすめされる情報を決定する上で最も重要な基準及びそれらのパラメータが相対的に重要な理由を含め、なぜ特定の情報がおすすめされるのかを説明するものでなければならないとされ、また②については、複数のオプションがある場合、サービス受領者において、いつでも好ましいオプションを選択又は変更できるようにする機能を提供しなければならないとされている。これら

提供者が行う活動であって、特に、サービス受領者が提供する違法コンテンツ又は利用規約違反情報を検知、識別及び対処することを目的とするものをいい、降格、無効化、アクセス不能化、削除など、違法コンテンツ若しくは利用規約違反情報の利用可能性、可視性及びアクセス可能性に影響を与える措置、又はサービス受領者のアカウントの終了若しくは停止など、サービス受領者が当該情報を提供する能力に影響を与える措置を含む。」と定義されている（第3条(t)）。

[24] DSA 上、「オンラインインターフェイス」（online interface）の用語は、「ウェブサイトの全部又は一部を含むあらゆるソフトウェア及びモバイルアプリを含むアプリケーション」と定義されている（第3条(m)）。

[25] DSA 上、「広告」（advertisement）の用語は、「商業目的を達成するためか非商業目的を達成するためかを問わず、法人又は自然人のメッセージを広めるためにデザインされた情報で、特に当該情報を広めることに対して支払われた報酬の対価として、オンラインプラットフォームによってそのオンラインインターフェイス上に表示されるもの」と定義されている（第3条(r)）。

[26] DSA 上、「レコメンダーシステム」（recommender system）の用語は、「サービス受領者によって開始された検索の結果として、又はその他の方法で表示される情報の相対的な順序又は優先順位を決定することを含め、サービス受領者にオンラインインターフェイスでの特定の情報を提案し、又はその情報に優先順位を付けるためにオンラインプラットフォームによって使用される完全又は部分的に自動化されたシステム」と定義されている（第3条(s)）。

に加えて、VLOP 及び VLOSE の提供者に対しては、更にそれぞれのレコメンダーシステムにつき、少なくとも１つはプロファイリングに基づかないオプションを提供することが求められている（第 38 条）。

④　有効なリスク軽減措置の実施義務

DSA においては、VLOP 及び VLOSE の提供者に対し、サービス及びアルゴリズムシステムを含む関連システムの設計・機能、又はそのサービスの利用に起因する EU 域内のシステミックリスクを真摯に特定、分析及び評価することが義務付けられている（第 34 条）。

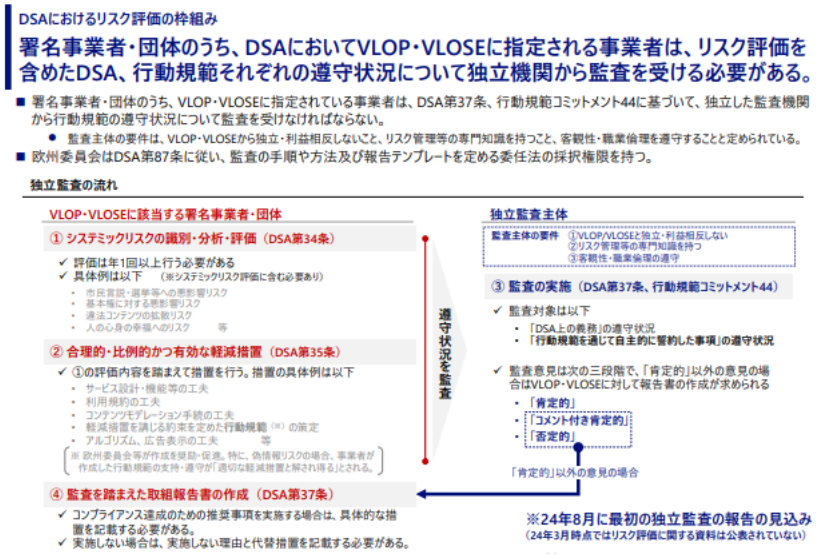
以上のリスク評価については、①VLOP 及び VLOSE を規律する条項の適用開始日（欧州委員会から VLOP 及び VLOSE に指定された旨の通知後４か月が経過する日）までに、②その後少なくとも年１回、③いかなる場合においても過去のリスク評価において特定されたリスクに重大な影響を及ぼす可能性のある機能を展開する前に、それぞれ実施しなければならないとされている。評価すべきシステミックリスクとしては、市民言説・選挙等への悪影響リスク、基本権に対する悪影響リスク、違法コンテンツの拡散リスク、人の心身の幸福へのリスク等が挙げられており、また、リスク評価に当たっては、特に、レコメンダーシステムや関連するアルゴリズムシステムの設計、コンテンツモデレーションシステム、適用される利用規約及びその執行、広告の選択・表示システム、VLOP 及び VLOSE の提供者におけるデータに関する慣行等の要素がシステミックリスクに及ぼす影響の有無・程度を考慮するものとされている。

リスク評価の内容を踏まえ、VLOP 及び VLOSE の提供者に対しては、特定されたシステミックリスクに合わせた合理的、比例的かつ有効な軽減措置を、特に当該措置が基本権に与える影響に配慮しつつ講じなければならないとされている（第 35 条）。軽減措置の具体例としては、サービス設計・機能等の工夫、利用規約の工夫、コンテンツモデレーション手続の工夫、軽減措置を講ずる約束を定めた行動規範の策定、アルゴリズムや広告表示の工夫等が挙げられている。このうち行動規範については、欧州委員会及び欧州デジタルサービス会議により、その策定と参加が奨励・促進されており（第 45 条、前文第 103 項）、特に、偽情報リスクの場合、事業者が作成した行動規範の支持・遵守が「適切な軽減措置と解され得る」とされている（前文第 104 項、後述）。

以上のリスク評価及び軽減措置の実施義務を含む DSA 上の義務の遵守状況（並びに行動規範を通じて自主的に誓約した事項）については、VLOP 及び VLOSE の提供者に対し、独立監査主体[27]による監査を受ける義務が課されている（第 37 条、偽情報に関する行動規範コミットメント 44）。監査意見は①「肯定的」、②「コメント付き肯定的」、③「否定的」の３段階で、「肯定的」以外の意見を受けた VLOP 及び VLOSE の提供者に対しては、取組報告書の作成が義務付けられている。VLOP 及び VLOSE の提供者は、監査意見に付された推奨事項を実施する場合、取組報告書の中に具体的な措置を記載する必要がある。また、推奨事項を実施しない場合は、実施しない理由と代替措置を記載しなければなら

[27] 独立監査主体として認められるのは、①監査対象となる VLOP 及び VLOSE の提供者から独立しており、利益相反しないこと、②リスク管理等の専門知識を有すると認められること、③特に業務規範又は適切な基準の遵守から、客観性及び職業倫理を有すると認められること、という要件を満たす組織とされる。

ないとされている。



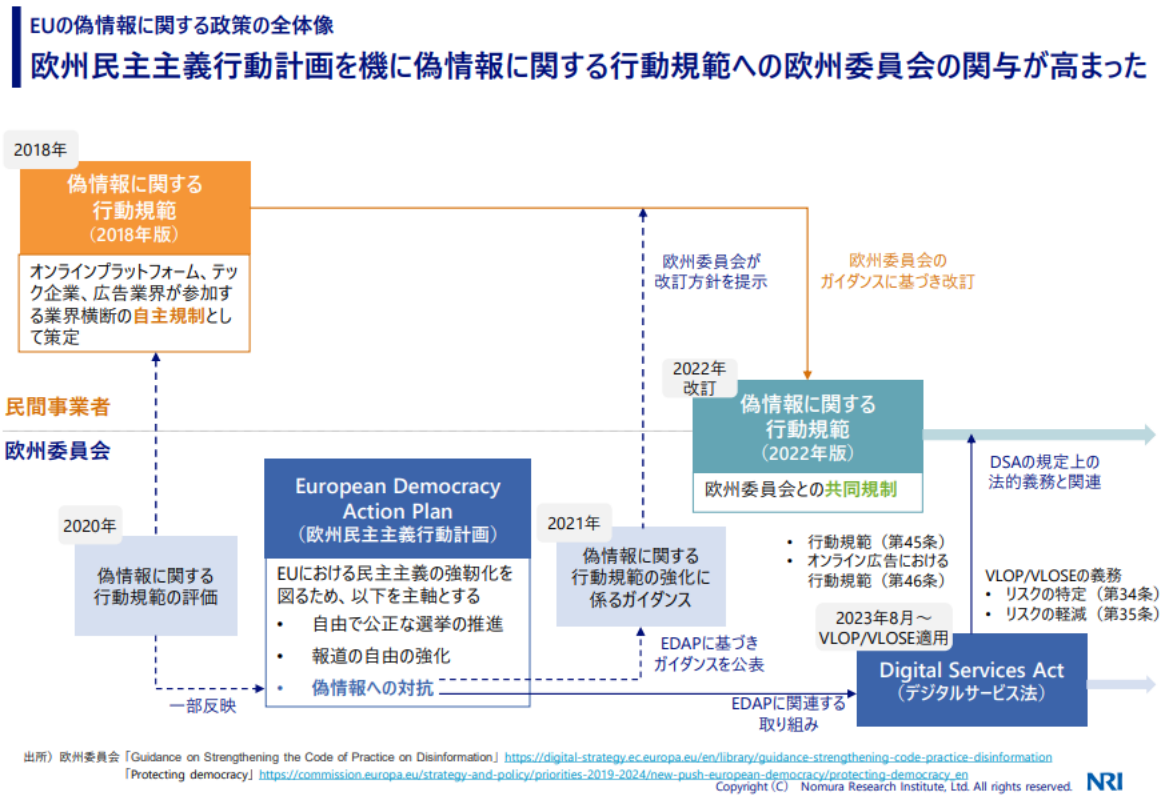
＜出典：本検討会資料 WG12－１－２＞

⑤　偽情報に関する行動規範

上述のとおり、DSA において、欧州委員会及び欧州デジタルサービス会議により行動規範の策定と参加が奨励されており、行動規範の遵守は VLOP 及び VLOSE の提供者に課せられるリスク軽減義務の一環に位置付けられている。特に偽情報リスクの場合、行動規範の支持・遵守は適切な軽減措置と解され得る一方、行動規範への不参加は、DSA の義務違反の考慮要素となり得るとされている（前文第 104 項）。

偽情報リスクに関しては、DSA 以前の 2018 年 10 月に、オンライン上の偽情報の拡散に対処するための自主規制の枠組みとして、オンラインプラットフォーム、大手テック企業、広告業界の代表を署名事業者・団体とする「偽情報に関する行動規範」（Code of Practice on Disinformation。以下「2018 年版行動規範」という。）が公表されていたが、「欧州民主主義行動計画」（European Democracy Action Plan。以下「EDAP」という。）[28]（2020 年 12 月）等を機に欧州委員会の関与が高まり、共同規制の枠組みへと発展している。

[28] 欧州委員会委員長の政策ガイドラインで発表された「欧州委員会作業計画」の主要なイニシアティブの1つであり、EU 全体で市民に力を与え、より強靭な民主主義を構築することを目的としている。



＜出典：本検討会資料 WG８－１＞

2018 年版行動規範は、政府機関（欧州委員会）、民間事業者（オンラインプラットフォームや広告事業者）、関連団体（伝統メディア、市民団体、ファクトチェック機関、研究者）がそれぞれに異なる役割[29]を与えられたマルチステークホルダーフォーラムにおける議論等の連携・協力を経て策定され、①「広告掲載の精査」、②「政治広告と争点ベースの広告の透明性」、③「サービスの完全性」、④「消費者のエンパワーメント」、⑤「研究者コミュニティのエンパワーメント」の５つの分野にわたるコミットメントを含むものとなっている。

欧州委員会において、2020 年９月に 2018 年版行動規範に対する評価が実施され、（ⅰ）EU 内における偽情報に対するプラットフォームのポリシーの透明性の向上に貢献したこと、（ⅱ）EU 内で偽情報に対抗するために署名事業者・団体が講じた措置について以前よりも公に精査が行われるようになったことを成果として評価する一方、(a) 報告の質について、提供される情報が署名事業者・団体間で異なり、一貫していないこと、(b) 特に「広告掲載の精査」に関し、報告されたポリシーには偽情報対策に特化していないものがあったこと、(c)「サービスの完全性」に関し、加盟国レベルでの透明性が不十分であったこと等が欠点として指摘されている。

[29] 具体的には、欧州委員会がフォーラム全体の枠組み・運営をサポートする（なお、フォーラムの議長は欧州委員会を構成する通信ネットワーク・コンテンツ・技術総局（DG-CONNECT）が務めた。）。一方、オンラインプラットフォーム、広告事業者及び広告主で構成されるワーキンググループが行動規範や KPI の策定作業を担い、これに対して伝統メディア、市民団体、ファクトチェック機関及び研究者で構成される相談委員会が助言・コメントを行った。

	コミットメント	成果
行動規範の欠点	A）広告掲載の精査	• オンラインでの偽情報の拡散に寄与する経済的要因の阻止が目的であるが、署名事業者・団体が報告したポリシーは偽情報対策に特化していないものがあった。 例）誤解を招く広告、根拠のない商業的な主張、詐欺的な商慣行の制限 • プラットフォーム自身のサービスへの争点ベースの広告の掲載に関する措置について、検証可能な証拠が提供されておらず、独立したファクトチェック団体との連携が不十分であった可能性がある。
	B）政治広告と争点ベースの広告の透明性 ※欧州民主主義行動計画でEUレベルでの有料政治広告の透明性を高める解決策を検討	• 争点ベースの広告の特定と開示について、署名事業者・団体間でアプローチが異なり、措置の効果を評価するための具体的な情報と検証ツールを提供していない。 • 広告が組織的に共有される際に、ラベルが表示される続ける確実な手段を検討していない。 • ユーザーや研究者が広告ライブラリにアクセスする際、APIの機能制限やリポジトリの完全性、検索可能な情報の質に懸念が残る。 • 政治広告の統一的な登録・認証手続きがない。
	C）サービスの完全性	• 署名事業者・団体が報告した操作技術等への対策はグローバルで提供されるため、EUへの影響と関連性を正確に評価できず、EU加盟国レベルでの透明性は不十分。 • 検知された偽情報キャンペーンに対するユーザーの関与のレベルについて更なる透明性が必要。
	D）消費者のエンパワーメント	• 提供されたツールがEU全体ですべての言語で完全かつ平等に利用できると判断できない。 • 信頼できる情報源とのユーザーの関与向上等への効果を示すデータが提供されていない。 • ツールやファクトチェック等の取り組みの有効性に関する透明で体系的な分析がないため、取り組みが時宜にかなわず、最適でない可能性がある。 • ユーザーが偽情報の可能性にフラグを立て、行動の結果について適切な情報を得るための使いやすく統一された手順がない。 • ファクトチェック団体との協力の効果について、比較可能な評価指標が提供されていない。 • 一部のプラットフォームが開発した信頼性に関する指標において検索サービスやレコメンダーシステムにおける詳細な情報は入手できない。
	E）研究者コミュニティのエンパワーメント	• すべてのEU加盟国と言語を完全にカバーした、プラットフォーム間で一貫したファクトチェックサービスの利用が実現するためのファクトチェック団体との協力モデルが検討されていない。

出所）欧州委員会(2020)「Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement」

＜出典本検討会資料 WG８－１＞

以上の評価を踏まえ、また、EDAP に基づき、欧州委員会において、2021 年５月に「偽情報に関する行動規範の強化に係るガイダンス」（European Commission Guidance on Strengthening the Code of Practice on Disinformation）が公表されている。このガイダンスでは、署名事業者・団体が 2018 年版行動規範の欠点にどのように対処すべきかが示され、2018 年版行動規範の署名事業者・団体及び新たに署名事業者・団体となることを希望する者において、同ガイダンスを踏まえ、行動規範の改訂に向けた検討を行う場としての総会（Assembly）や、分野ごとの課題グループ（Issue Group）を中心した改訂作業が進められた。

総会においては、2018 年版行動規範の策定プロセスと同様、民産学官のマルチステークホルダーによる関与のもとで議論が行われた。具体的には、欧州委員会が総会での議論をファシリテートし、議題提供を含む必要な組織的又はロジスティックな支援が提供された [30]一方、課題グループにおける具体的な起草作業においては、主にオンラインプラットフォームや広告事業者が担い、これに対して欧州委員会や総会の構成員によって招集された専門家やオブザーバが意見を述べるものとされた。

こうした議論を経て、2022 年６月に「2022 年強化された偽情報に関する行動規範」（The Strengthened Code of Practice on Disinformation 2022。以下「2022 年版行動規範」という。）が公表された。2022 年版行動規範については、2018 年版行動規範の大まかな構成を引き継ぎつつ、コミットメントの対象分野として、新たに①「ユーザーのエンパワーメント」[31]、②「ファクトチェック団体のエンパワーメント」、③「透明性センター」[32]、④「常設タスクフォース」が定められ、全体で 44 のコミットメント、129 の

[30] なお、総会の議長は、2018 年版行動規範におけるマルチステークホルダーフォーラムと同様、DG-CONNECT が務めた。
[31] 2018 年版行動規範では「消費者のエンパワーメント」。
[32] 前述（3－2－1－2）の DSA に基づく「透明性レポート」の公表義務とは別途、署名事業者・団体が共通ウェブサイトとして設置する「透明性センター」において、半年ごとにコミットメントの遵守状況を公開するものとされている。直近では 2024 年3月 26

措置（measures）から構成されている。どのコミットメントに署名するかは各事業者・団体が決定し、コミットメントへの実効性を確保するのも各事業者・団体の責任とされている。

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

2022年版行動規範の構成と項目一覧

赤太字：2018年版からの追加箇所

目次	小項目
1.前文（PREAMBLE）	• –
2.広告掲載の精査 （SCRUTINY OF AD PLACEMENTS）	• 偽情報の収益化廃止 • 偽情報を含む広告への取り組み • 関係者との協力
3.政治広告 （POLITICAL ADVERTISING）	• 政治広告と争点広告の共通理解 • 政治広告や争点広告の効率的な表示 • 政治広告または争点広告の検証コミットメント • 政治広告や争点広告のためのユーザー向け透明性コミットメント • 政治または争点広告のリポジトリと、政治または争点広告データにアクセスするためのアプリケーション・プログラミング・インターフェース（API）の最小限の機能 • 市民社会のコミットメント • 継続的な協力
4.サービスの完全性 （INTEGRITY OF SERVICES）	• 許されない操作行動の共通理解 • AIシステムの透明性義務 • 協力と透明性
5.ユーザーのエンパワーメント （EMPOWERING USERS） ※2018年版では「消費者のエンパワーメント」	• メディア・リテラシーの向上 • サービスのアーキテクチャの「安全な設計」、透明性の高いポリシー、レコメンダーシステムの説明責任 • ユーザーが偽情報を識別できるようにする • 有害な虚偽及び誤解を招く情報にフラグを立てる機能 • メッセージングアプリの偽情報を抑制するための対策
6.研究者コミュニティのエンパワーメント （EMPOWERING THE RESEARCH COMMUNITY）	• 偽情報研究のため署名事業者・団体データの開示とアクセス • 研究目的のデータへのアクセスに関するガバナンス体制のさらなる精査 • 研究者との協力 • 研究機関の透明性とデータ共有
7.ファクトチェック団体のエンパワーメント （EMPOWERING THE FACT-CHECKING COMMUNITY）	• ファクトチェック・コミュニティとの協力 • 署名事業者・団体のサービスにおける事実確認の利用と統合 • ファクトチェッカーによる関連情報へのアクセス • ファクトチェッカーの基準
8.透明性センター（TRANSPARENCY CENTRE）	• –
9.常設タスクフォース（PERMANENT TASK-FORCE）	• –
10.行動規範のモニタリング（MONITORING OF THE CODE）	• –

出所）欧州委員会(2023)「The Strengthened Code of Practice on Disinformation 2022」

＜出典：本検討会資料 WG８－１＞

署名事業者・団体については、2018 年版行動規範では 16 団体である一方、2022 年版行動規範にはオンラインプラットフォーム事業者・団体、広告関連事業者・団体、ファクトチェック団体・関連サービス事業者、その他 NGO を含めた市民・業界団体・関連事業者など 43 団体[33]が署名・参加している（2024 年１月 17 日現在）。

日にレポートが公開されており、次回レポートの公開は 2024 年9月に予定されている。

[33] 当初はX（旧 Twitter）も署名していたが、2023 年5月に脱退。

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

2022年版行動規範の署名事業者・団体

■ 2024年1月17日時点の情報として、欧州委員会公式HP上で公開されている全43団体。（※アルファベット順）

凡例：
オンラインプラットフォーム事業者・団体
広告関連事業者・団体
ファクトチェック団体・関連サービス事業者
その他（市民・業界団体・関連事業者）

#	団体名
1	ActiveFence
2	Adobe
3	AI Forensics
4	Alliance4Europe
5	Avaaz
6	Clubhouse
7	Crisp
8	Debunk EU
9	Demagog
10	Democracy Reporting International
11	DoubleVerify
12	DOT Europe
13	Ebiquity
14	European Factchecking Standards Network
15	European Association of Communication Agencies (EACA)
16	Faktograf
17	FIDU (Italian Federation for Human Rights)
18	Globsec
19	Google
20	IAB Europe (Interactive Advertising Bureau Europe)
21	Kreativitet & Kommunikation
22	Les Surligneurs
23	Logically
24	Maldita.es
25	Meta
26	Microsoft
27	Newsback
28	Newtral
29	NewsGuard
30	PagellaPoltica
31	Reporters without Borders (RSF)
32	ScienceFeedback
33	Seznam
34	The Bright App
35	The Daily Ledger
36	The GARM Initiative
37	The Global Disinformation Index
38	TikTok
39	Twitch
40	Vimeo
41	VOST Europe
42	WhoTargetsMe
43	World Federation of Advertisers (WFA)

出所）欧州委員会(2023)「Signatories of the 2022 Strengthened Code of Practice on Disinformation」
https://digital-strategy.ec.europa.eu/en/library/signatories-2022-strengthened-code-practice-disinformation

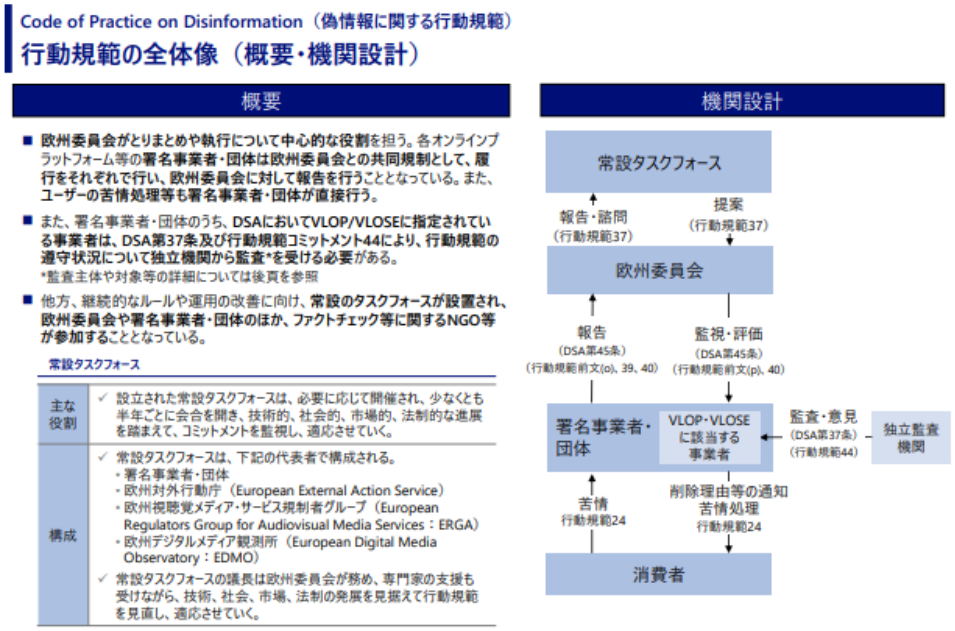
＜出典：本検討会資料 WG８－１＞

なお、2022 年版行動規範における「偽情報」及び「誤情報」の定義については、EDAP の定義に依拠している。すなわち、「偽情報」（disinformation）は「欺き、又は経済的若しくは政治的利得を確保する意図を持って拡散され、公共への損害を生じさせ得る虚偽の、又は誤解を招くコンテンツ」、「誤情報」（misinformation）は「有害な意図を持たずに共有されながらも、その効果は未だ有害である虚偽の、又は誤解を招くコンテンツ」とそれぞれ定義されている。特に「誤情報」については、虚偽の情報を善意で友人や家族に共有する場合が例示されている。また、EDAP の定義に加え、2022 年版行動規範においては、誤解を招く広告、誤報、風刺及びパロディ、明白に確認されている党派性の強いニュース及び解説が、いずれも「偽情報」に含まれないことが明記されている。

2022 年版行動規範のガバナンスに関しては、欧州委員会がとりまとめや執行について中心的な役割を担っており、署名事業者・団体自身と欧州委員会による共同規制というべき体制がとられている。すなわち、署名事業者・団体は、履行をそれぞれで行い、ユーザーの苦情処理等についても直接行うこととされている（コミットメント 24）一方、署名事業者・団体から欧州委員会に対して報告が行われるとともに、欧州委員会は署名事業者・団体を監視・評価するものとされている（前文(o)・(p)、コミットメント 39・40[34]）。また、継続的なルールや運用の改善に向け、欧州委員会が議長を務め、署名事業

[34] 署名事業者・団体は、コミットメント 39 において「当規範の公約を確実に履行するために、十分な財務的・人的資源を投入し、適切な社内手続きを実施すること」、コミットメント 40 において「署名事業者・団体は、実施期間終了後（当規範署名後6ヶ月）1ヶ月以内に、前文にあるベースライン報告書を欧州委員会に提出する」こととされている。なお、DSA 上も、欧州委員会は署名事業者・団体による行動規範の不遵守に対して必要な措置を講ずるよう求める権限を与えられている（第 45 条）。

者・団体のほか、欧州対外行動庁、ファクトチェック等に関する NGO 等が参加する「常設タスクフォース」が設置され、少なくとも半年ごとに開催する会合の中で、専門家の支援も受けながら、技術的、社会的、市場的、法制的な進展を踏まえて、行動規範を見直し、適応させていくものとされている（コミットメント 37）。さらに、DSA において VLOP 及び VLOSE の提供者である署名事業者・団体においては、DSA 上、行動規範の遵守状況についても独立監査主体による監査を受ける義務を負い、DSA を通じた欧州委員会による執行・監督（後述）も受けることとされている。



＜出典：本検討会資料 WG８－１）＞

⑥　危機対応メカニズム

DSA において、欧州委員会は、欧州デジタルサービス会議からの勧告に基づき、VLOP 及び VLOSE の提供者に対し、危機的状況に起因するリスクを軽減するための一定の措置を講ずるよう求めることができるものとされている（第 36 条）。具体的には、危機（crisis）が発生した場合に、欧州委員会は、①そのサービスが危機に著しく寄与しているかどうかを評価すること、②その寄与を防止、排除又は制限する措置を講ずること、③評価結果を欧州委員会に報告することのうち一つ又は複数の行動をとることを VLOP 及び VLOSE の提供者に求めることができるとされている。

なお、ここでいう「危機」については、「EU 域内又はその重要な部分において、公共の安全又は公衆衛生に対する重大な脅威をもたらす異常事態が発生した場合」に発生したものとみなされ（第 36 条第２項）、具体的には、武力紛争、テロ行為、地震やハリケーンなどの自然災害、パンデミックや公衆衛生に対する国境を越えたその他の深刻な脅威から生ずる可能性があるものとされている（前文第 91 項）。

以上のうち②の具体的な措置としては、コンテンツモデレーションプロセスの適合、コンテンツモデレーションに専念するリソースの増加、利用規約や関連するアルゴリズムシステム及び広告システムの適合、「信頼できる通報者」との協力の強化、啓発措置の実施、信頼できる情報の促進、オンラインインターフェイスのデザインの適合等が例示されている（前文第 91 項）。VLOP 及び VLOSE の提供者は自ら措置の内容を決定することができるが、措置を講ずる義務には拘束力があり、欧州委員会は措置を監視し、措置が効果的かつ適切であるかどうかを評価するために VLOP 及び VLOSE の提供者と「対話する」(engage in a dialogue) ことができるものとされている。また、欧州委員会は、VLOP 及び VLOSE の提供者が講じた措置に効果がない、又は不適切であると判断した場合は、VLOP 及び VLOSE の提供者に対し、措置の見直し又は措置の適用を中止するよう要求することができるとされている。

以上の危機対応メカニズムについては、2022 年２月のロシアによるウクライナへの侵攻に伴い追加されているが、欧州委員会の権限への限定が小さいことに対する市民団体からの批判を受け、危機対応の最長期間を３か月に設定するに至った経緯がある。

この他、欧州デジタルサービス会議は、欧州委員会に対し、オンライン環境における危機的状況に対処するための自主的な危機プロトコルの作成の開始を勧告することができ、欧州委員会は、「オンラインプラットフォームが違法コンテンツや偽情報の急速な拡散に悪用された場合、又は信頼できる情報を迅速に発信する必要が生じた場合」（前文第 108 項）に、当該勧告に基づき、自主的な危機プロトコルを作成することができるものとされている（第 48 条）。

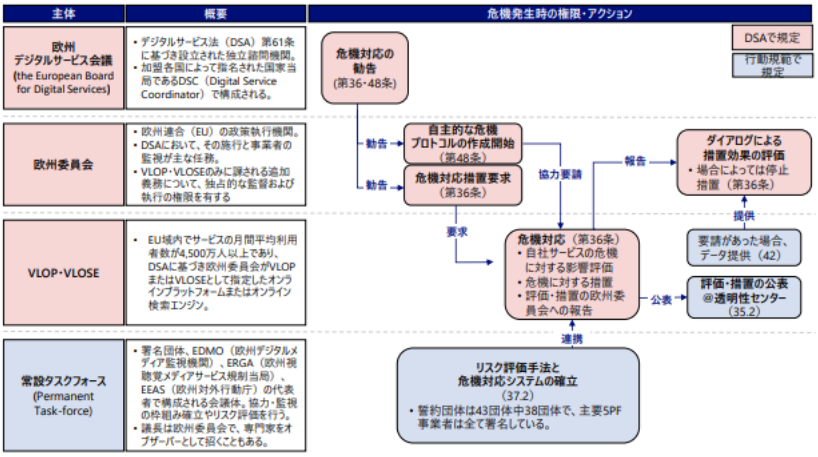
欧州委員会は、VLOP 及び VLOSE の提供者や、適切な場合には他のオンラインプラットフォームやオンライン検索エンジンの提供者が、こうした自主的な危機プロトコルの作成、試験、適用に参加することを奨励し、促進するものとされている。自主的な危機プロトコルについては、①公的機関やその他の信頼できる機関からの危機に関する情報を目立つよう表示、②危機管理専用窓口を設置、③他の規定の義務遵守のためのリソースを危機的状況から生ずるニーズに適合させることのうち少なくとも１つを含む必要があるとされている。

なお、以上の危機対応メカニズムが発動され、若しくは自主的な危機プロトコルが作成された、又はそれらに向けた検討が開始された事例は確認されていない（2024 年３月現在）。

また、以上に加え、2022 年版行動規範においては、危機に対応したユーザーインターフェイスの設計や、危機対応システムの構築、措置の公表についてのコミットメント及び措置が定められている。

分野	コミットメント・措置	内容
ユーザーのエンパワーメント	コミットメント22 措置7	**（ユーザーインターフェースの設計）** • 関連署名団体は、公共や社会が特に関心を持つ話題や**危機的状況**において、利用者を権威ある情報源に導くような製品や機能（情報パネル、バナー、ポップアップ、地図やプロンプト、信頼性指標など）を設計し、適用する。
透明性センター	コミットメント35 措置4	**（危機対応措置の公表）** • **危機的状況**において、署名団体は透明性センターを利用し、**危機**に関連して講じられた具体的な緩和措置に関する情報を公表する。
常設タスクフォース	コミットメント37 措置2	**（危機対応システムの構築）** • 署名団体は、タスクフォースにおいて、特に（これに限定されるものではないが）以下の業務に取り組むことに同意する： ▪ 選挙や**危機**のような特殊な状況下で使用するリスク評価手法と迅速な対応システムを確立する。 ▪ 選挙や**危機**のような特別な状況下で各団体と協力・調整する。 ▪ 行動規範の「約束と措置」を実施するために整合化された報告テンプレート、精緻化された報告方法、モニタリングのための関連データ開示について合意する。 ▪ 今後のモニタリング・サイクルを通じて、整合化された報告テンプレートの質と有効性、およびモニタリング目的のデータ開示の形式と方法を見直し、必要に応じて適合させる。 ▪ サービスレベルと構造に関する指標や、これらの指標を測定するために提供されるデータについての品質と有効性に対する評価に協力する。 ▪ 構造指標を洗練、テスト、調整し、国家レベルでそれらを測定するメカニズムを設計する。 ▪ 悪意ある行為者が用いるTTPのリストに合意のもと、公表・更新し、本規範の第IV章に沿って、それらに対抗するための対策の基本要素、目標、ベンチマークを定める。
監視体制の強化	コミットメント42	**（情報・データの提供）** • 関連する署名団体は、選挙や**危機**のような特別な状況において、欧州委員会の要請があれば、タスクフォースによって確立された迅速な対応システムに従い、特別な報告書や定期的な監視の中の特定の章を含む、相応かつ適切な情報やデータを提供することを約束する。

＜出典：本検討会資料 WG12－１－１＞



＜出典：本検討会資料 WG12－１－１）＞

⑦　執行・監督

各加盟国内における DSA の執行・監督については、各加盟国が所管当局として指定したデジタルサービス調整機関（Digital Services Coordinators。以下「DSC」という。）が担うものとされている。ただし、VLOP 及び VLOSE の指定とその提供者に対する執行・監督については、欧州委員会が独占的な権限を有しており、各 DSC と欧州委員会は、DSA を一貫して効率的に適用するために緊密に協力し、相互に援助し合うものとされている。

このように、DSA の執行・監督については、各加盟国の DSC と欧州委員会が担うものとされているが、さらに、DSC の代表と欧州委員会で構成される独立した諮問会議として欧州デジタルサービス会議（European Board for Digital Services）が設置され、DSC と欧州委員会を支援するものとされている。

なお、2024 年 5 月 29 日現在、EU 加盟国 27 か国中 21 か国が DSC に指定した機関を公表しており、そのうち 17 か国は通信系の当局が DSC に指定されている。

各国におけるDSC（Digital Services Coordinators）の任命状況（2024年5月29日現在）

■ 2024年5月29日現在、EU加盟国27か国中21か国がDSCに任命した機関を公表している。
■ 公表している21か国のうち、17か国が通信系の当局をDSCに指名している。

：通信系当局
－ ：公表なし

加盟国	DSC任命機関
オーストリア	Austria Communications Authority（オーストリア通信局）
イタリア	Authority for Communications Guarantees（通信規制庁）
ベルギー	－
ラトビア	－
ブルガリア	Communications Regulation Commission（通信規制委員会）
リトアニア	－
クロアチア	Croatian Regulatory Authority for Network Industries (HAKOM)（クロアチアネットワーク産業規制庁）
ルクセンブルク	Competition Authority（競争庁）
キプロス	Cyprus Radiotelevision Authority（キプロスラジオテレビ局）
マルタ	Malta Communications Authority（マルタ通信局）
チェコ共和国	Czech Telecommunication Office（チェコ通信局）
オランダ	Authority for Consumers and Markets（ACM）（オランダ消費者市場庁）
デンマーク	Danish Competition and Consumer Authority（デンマーク競争・消費者庁）
ポーランド	－

加盟国	DSC任命機関
エストニア	－
ポルトガル	National Communications Authority (ANACOM)（国家通信局）
フィンランド	Finnish Transport and Communications Agency (TRAFICOM)（フィンランド運輸通信庁）
ルーマニア	National Authority for Management and Regulation in Communications (ANCOM)（通信管理規制機関）
フランス	Regulatory Authority for Audiovisual and Digital Communication（視聴覚およびデジタル通信規制局）
スロバキア	－
ドイツ	Federal Network Agency for Electricity, Gas, Telecommunications, Post and Railways（ネットワーク庁）
スロベニア	Agency for Communication Networks and Services of the Republic of Slovenia（スロベニア通信ネットワーク・サービス庁）
ギリシャ	Hellenic Telecommunications and Post Commission（通信・郵便委員会）
スペイン	National Commission for Markets and Competition（国家市場競争委員会）
ハンガリー	National Media and Infocommunications Authority（国家メディア情報通信局）
スウェーデン	Post and Telecom Authority（郵便・電気通信局）
アイルランド	Media Commission（メディア委員会）

出所）欧州委員会「Digital Services Coordinators」https://digital-strategy.ec.europa.eu/en/policies/dsa-dscs -

Copyright (C) Nomura Research Institute, Ltd. All rights reserved. NRI

特に VLOP 及び VLOSE の提供者の義務について、遵守状況に疑念を抱いた欧州委員会は、情報提供の要請（第 67 条）、インタビューの実施（第 68 条）、立入検査（第 69 条）をはじめとするコンプライアンス調査を行うことができるとされている。コンプライアンス調査後も DSA 違反の疑念が残る場合、欧州委員会は法的手続を開始でき（第 66 条）[35]、この法的手続で違反認定がなされる可能性があり、かつサービス受領者に重大な損害が及ぶ緊急の可能性がある場合、欧州委員会は明白な違反の証拠に基づき、VLOP 及び VLOSE の提供者に対し、相応かつ一時的な暫定措置を命じることを決定できることとされている（第 70 条）。

[35] 例えば、2023 年 12 月には X（旧 Twitter）、2024 年2月には TikTok に対する法的手続が開始された。

参考：VLOPに対する欧州委員会からの情報提供の要請と法的手続きの状況（1/2）

	VLOP（事業者名）	経緯	提供を要請した情報・論点
2023年10月12日	X	• 違法コンテンツや偽情報（特にテロリスト、暴力的なコンテンツ、ヘイトスピーチ）の拡散が疑われるとの指摘を欧州委員会が受け、開始 • ハマスによるイスラエルへのテロ攻撃が発生	• 危機対応プロトコル（第48条）の発動と機能に関する質問（10月18日期限） • その他の義務への遵守状況に関する質問（10月31日期限） ※違法コンテンツに関する通知、苦情処理、リスク評価、特定されたリスクへの軽減措置に関する方針や対応を含む
10月19日	TikTok	• ハマスによるイスラエルへのテロ攻撃が発生し、偽情報（特にテロリスト、暴力的なコンテンツ、ヘイトスピーチ）の拡散が疑われる	• 危機対応における違法コンテンツや偽情報の拡散に関するリスク評価（第34条）と軽減措置（第35条）の詳細（10月25日期限） • オンライン上の未成年者の保護に関する規定（第28条）への遵守状況（11月8日期限）
	Facebook Instagram（Meta）	• ハマスによるイスラエルへのテロ攻撃が発生し、偽情報（特にテロリスト、暴力的なコンテンツ、ヘイトスピーチ）の拡散が疑われる	• 選挙の完全性を保護するためのリスク評価（第34条）と軽減措置（第35条）（11月8日期限） • 危機対応における違法コンテンツや偽情報の拡散に関するリスク評価（第34条）と軽減措置（第35条）の詳細（10月25日期限）
11月9日	YouTube（Google） TikTok	• －	• 未成年者保護の義務（第28条）のうち、精神的・身体的健康へのリスクの評価と軽減措置の詳細（11月30日期限） • 未成年者のサービス利用に関する義務（第28条）の遵守状況（11月30日期限）
11月10日	Facebook Instagram（Meta） Snapchat（Snap）	• －	• 未成年者保護の義務（第28条）のうち、精神的・身体的健康へのリスクの評価と軽減措置の詳細（12月1日期限） • 未成年者のサービス利用に関する義務（第28条）の遵守状況（12月1日期限）
12月1日	Facebook Instagram（Meta）	• Instagramにおける自作の児童への性的虐待画像の流通	• 未成年者保護の義務（第28条）のうち、精神的・身体的健康へのリスクの評価と軽減措置の詳細（12月22日期限） • Instagramにおける「おすすめ」システム（第27条）と潜在的に有害なコンテンツの増幅（第34条）に関する情報（12月22日期限）

出所）欧州委員会 「Supervision of the designated very large online platforms and search engines under DSA」 https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses
Copyright (C) Nomura Research Institute, Ltd. All rights reserved.

参考：VLOPに対する欧州委員会からの情報提供の要請と法的手続きの状況（2/2）

	VLOP	経緯	提供を要請した情報・論点
12月18日	X	• リスク管理、コンテンツモデレーション、ダークパターン、広告の透明性、研究者のデータアクセスに関してDSA違反の可能性を評価するため、**法的手続き**を開始 **※DSAにおける初の事例** • 透明性レポート（11月3日公表）や10月12日の情報提供の要請に対する回答に基づき決定	• プラットフォーム上の情報操作への対抗措置（特にコミュニティ・ノート機能、利用規約上のプラットフォーム操作ポリシーの適用）の有効性 • 資格のある研究者*に対する公開データへのアクセス付与義務に関する不備（データのスクレイピング禁止は第40条12項に違反） *第40条8項の条件を満たすとDSCが承認した研究者のみが対象 • 広告に関する情報を取得するためのツールの提供義務に関する不備
2024年1月18日	2023年4月25日指定VLOP17者（Twitter、Wikipedia以外）	• 各国やEUの選挙等のイベントを控え、プラットフォーム上の違法コンテンツを継続的な監視が特に重要	• 資格のある研究者に対する公開データへのアクセス付与義務（第40条12項）の遵守状況（2月8日期限）
2月19日	TikTok	• 未成年者の保護、広告の透明性、研究者のデータアクセス、中毒性の高いデザインや有害コンテンツのリスク管理に関してDSA違反の可能性を評価するため、**法的手続き**を開始	• システムデザインに起因するリスク評価と軽減措置の不備（第34条1項・2項、第35条1項） • 不適切コンテンツへのアクセス防止のための年齢確認ツールの不備（第28条1項） • 「おすすめ」システムにおける未成年者のプライバシーと安全を確保する措置（第27条）の不備 • 資格のある研究者に対する公開データへのアクセス付与義務（第40条12項）に関する不備 • 広告に関する情報を取得するためのツールの提供義務に関する不備（第39条1項）
3月1日	Meta	• －	• FacebookとInstagramが広告運用、レコメンダーシステム、および当該サブスクリプションサービスオプションの導入に関するリスク評価に関する義務を遵守するために講じた措置に関する情報提供（3月22日期限） • その他以前に要請された選挙や未成年者の保護に関する回答に対する追加情報の提供要請も含む（3月18日期限）

出所）欧州委員会 「Supervision of the designated very large online platforms and search engines under DSA」 https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses
「Commission decision initiating proceedings pursuant to Article 66(1) of Regulation (EU) 2022/20651」 https://ec.europa.eu/newsroom/dae/redirection/document/98715
Copyright (C) Nomura Research Institute, Ltd. All rights reserved.

＜出典：本検討会資料 WG５－１－６＞

最終的に、VLOP 及び VLOSE の提供者が DSA の関連規定に違反し、又は暫定措置等に適合していないと判断した欧州委員会は違反の決定を行い(第 73 条)、当該 VLOP 及び VLOSE の提供者に対し、前会計年度における全世界年間売上高の６％を超えない制裁金を課すことができることとされている（第 79 条)。

（２）AI 法（Artificial Intelligence Act）

EU では、世界初の AI に関する包括的な法的枠組みとして、AI 法（Artificial Intelligence Act）が 2024 年５月に成立されている。AI 法は、AI システムに対し、リスクに応じて、①禁止される AI 行為、②ハイリスク AI システム、③特定の AI システムの透明性義務、④自主的な行動規範の４類型に分けた規律のほか、⑤生成 AI を含む汎用目的 AI（General Purpose AI）に対する特別な規律が設けられている。

このうち②ハイリスク AI システムについては、既存の EU 法で適合性評価義務の対象となる機械、玩具やレジャー用船舶等に用いられる AI のほか、バイオメトリクス（遠隔生体識別・感情識別)、インフラ管理・運用や司法又は民主主義プロセス等の新たに指定された領域における AI が含まれている。これらについては、リスクマネジメントシステムの構築、データとデータガバナンス等の要求事項が定められているほか、公的機関等の配備者には基本権影響評価の実施と当局への提出が求められている。

なお、AI 法の審議プロセスの中で、DSA の VLOP 及び VLOSE の提供者が用いるレコメンダーシステムに用いられる AI をハイリスク AI システムに含む提案も行われたが、最終的には含まれないこととされている。その結果、レコメンダーシステムがもたらすリスクについては、DSA に基づく透明性確保義務やリスク評価及び軽減措置の実施義務等を通じた対応が図られることになっている。

1．AI法案：ハイリスクAI

- ハイリスクAIのカテゴリー
  - 既存EU法での適合性評価義務対象：機械、玩具、レジャー用船舶、リフト、爆発性雰囲気装置、無線機器、圧力機器、索道設備、個人用保護具、ガス機器、医療機器
  - AI法での新たな指定：バイオメトリクス（遠隔生体識別・感情識別）、インフラ管理・運用、教育や職業訓練での学生や希望者の評価や受入れの合否、雇用、労働管理、自営業へのアクセス、重要な民間・公共サービス（公的支援金給付、融資、緊急対応措置）、法執行、移民・亡命・国境管理、**司法又は民主主義プロセス**
    - ※審議プロセスの中でデジタルサービス法のVLOP/VLOSEが用いるレコメンダーシステムを直接ハイリスクAIに含む提案も行われたが、最終的には含まれず
- ハイリスクAIシステムの要求事項
  - リスクマネジメントシステムの構築、データとデータガバナンス、技術文書、記録保持、透明性と利用者への情報提供、人間による監視、正確性・堅牢性・セキュリティ
  - 要求事項具体化手段としての整合規格
- ハイリスクAIシステム提供者の義務：上記要求事項の遵守確保等
- ハイリスクAI配備者の義務：**基本権影響評価**の実施と当局への提出

＜出典：本検討会資料 WG12－２＞

また、③特定の AI システムの透明性義務に関しては、ディープフェイク[36]を生成する AI システムの配備者に対し、当該コンテンツが人為的に生成・操作されたものであることを開示する義務が課されているほか、汎用目的 AI を含むコンテンツ生成 AI 提供者に対し、AI システムの出力が機械可読形式でマークされ、人為的に生成・操作されたことを検出できることを保証する義務が課されている。この義務については、AI 法の前文で、DSA において VLOP 及び VLOSE の提供者に評価と軽減措置の実施が義務付けられるシステミックリスク、特に偽情報を通じたものを含む民主的プロセス、市民的言説及び選挙プロセスに対する実際又は予見可能な悪影響のリスクを特定・軽減することとの関連性が明記されており、AI 法は DSA を補完する関係にあるとされている。

[36] AI 法において「ディープフェイク」(deep fake)の用語は、「実在の人物、物体、場所、団体又は出来事に類似し、真正又は真実であると人に誤認させるであろう、AI で生成され、又は操作された画像、音声又は映像コンテンツ」と定義されている。

1．AI法案：特定のAIシステムの透明性義務

- 自然人と直接対話するAIシステム提供者の開示義務
- 感情識別・生体識別システム配備者の本人通知義務
- ディープフェイク生成AIシステム配備者：当該コンテンツが人為的に生成・操作されたものであることを開示する義務
- 汎用目的AIを含むコンテンツ生成AI提供者：AIシステムの出力が機械可読形式でマークされ、人為的に生成・操作されたことを検出できることを保証する義務
  - 前文120「本規則において、特定のAIシステムの提供者および配備者に課される、当該システムの出力が人為的に生成または操作されたものであることを検知し、開示することを可能にする義務は、**規則（EU）2022/2065（※デジタルサービス法）の効果的な実施を促進**するために特に関連する。これは、特に、**人工的に生成または操作されたコンテンツの拡散から生じる可能性のあるシステミックリスク、特に偽情報を通じたものを含む民主的プロセス、市民的言説および選挙プロセスに対する実際または予見可能な悪影響のリスク**を特定し、軽減するための、非常に大規模なオンラインプラットフォームまたは非常に大規模なオンライン検索エンジンのプロバイダーの義務に関して適用される。」

＜出典：本検討会資料 WG12－２＞

同様に AI 法と DSA の補完関係を示すものとして、⑤生成 AI を含む汎用目的 AI に対する特別な規律に関し、汎用目的 AI モデル（10^25FLOPs 以上等）の提供者に対し、システミックリスクの評価・軽減や重大インシデントへの対応の文書化と当局への報告等の義務が課されている。それぞれの義務は整合規格により具体化されるとされており、それまでは欧州委員会主導で策定する行動規範（codes of practice）の遵守が求められている。

（３）メディア自由法（Media Freedom Act）

EU では、EDAP の一環として、域内におけるメディアの自由・多元性・独立性の保護を目的としたメディア自由法（Media Freedom Act）が欧州委員会により 2022 年９月に採択され、欧州理事会と欧州議会の間で 2023 年 12 月に暫定合意へと至った後、2024 年３月に欧州議会により可決されている。

メディア自由法においては、EU 域内市場におけるメディアサービスのための共通の枠組みを確立し、ジャーナリスト及びメディアプロバイダーを政治的介入から保護するための措置が導入されている。

具体的には、①編集の独立性の保護（加盟国によるメディアサービス提供者の編集決定への介入の禁止及び取材源に関する情報提供の強要の禁止など）、②公共メディア（public service media）の独立性の確保、③メディア多元性のテスト（media pluralism test：メディアの市場集中度がメディアの多元性と編集の独立性に与える影響についての加盟国による評価など）、④透明性のある国家の広報活動（transparent state advertising：公共機関による広報活動の割当について、透明性と被差別性を確保するための基準に従うことなど）のほか、⑤オンラインにおけるメディアコンテンツの保護（偽

情報のようなシステミックリスクを伴わないコンテンツに関し、VLOP が自社のポリシーに沿わない特定の合法メディアコンテンツを取り下げる場合には、事前にその理由をメディア提供者に対して通知することの義務付けなど）等の規定が置かれており、DSA とともに、EU における情報伝送 PF サービスを中心としたデジタル空間における情報流通の多元性を確保するための方策としても重要な位置付けを有している。

（４）データ法（Data Act）

EU では、データの利用を促進するために策定された 2020 年の欧州データ戦略の一環として、2023 年９月に施行されたデータガバナンス法（Data governace Act）を補完するものとして、データ法案が提案され、2023 年 11 月 9 日に成立したところである。同法は 2024 年 1 月 11 日に発行し、原則として、2025 年 9 月 12 日に施行される。欧州データ戦略は、データの単一市場を設立し、未利用データの公開をすることで研究者、行政機関等の利益のためにデータ流通の自由化を目指すものである。この目的を実現するため、データ法においては、例えば、公共利益や法的義務の遂行のために公的機関がデータの要求をした際は、データ保有者はデータを利用できるようにしなければならないこととされている。

データ法の適用範囲として、事業者間、事業者・消費者間、事業者・政府間のいずれにも適用される。特に事業者・政府間においては、緊急でない状況では、公的機関は非個人データのみを要求することができるが、公的緊急事態（洪水や山火事など）に対応するために必要な場合は、データをその保有者から要求することができるものとされている。

なお、公的機関は公的緊急事態にデータ保有者から得たデータについては、その必要性がなくなり次第、受領データを消去（データ共有された研究機関等も含む）し、データ保有者に対してその消去したことを通知しなければならないものとされている。

（５）ドイツにおける刑事罰による対応

ドイツでは、偽・誤情報の発信又は媒介を直接的に禁止する法制度は存在しないが、該当する行為に対して刑事罰が適用され得る。

偽・誤情報の発信又は媒介と関連性のある犯罪の例

犯罪名	条文	備考
犯行を行う旨の脅迫により公共の平和を乱す罪	126	公共の平和を乱すのに適した方法で、殺人等の重大な犯罪や、重大な犯罪ではないものの公共の危険に関わる犯罪等を行う旨の脅迫を行う行為（1項）及び当該行為が差し迫っているように装うこと（2項）を規制
民衆扇動罪	130	公共の平和を乱すのに適した方法で、特定の集団に対して憎悪を扇動する行為、暴力的・恣意的措置を呼びかける行為及び特定の集団に属する者をそのことを理由に侮辱又は中傷する行為を規制
信条冒涜罪	166	公共の平和を乱すのに適した方法で、他者の宗教やイデオロギー等に対する嫌悪を示すコンテンツを広める行為等を規制
侮辱罪	185	侮辱行為を規制 事実に関する真実性の証明があったとしても、適用を受けることがある（192条）
悪評の流布罪	186	特定の人物の価値を貶め又は当該人物に関する世評に悪影響を及ぼすような事実で、真実と証明できないものを主張し、又は広める行為を規制
名誉毀損罪	187	真実を知りながら、特定の人物について虚偽の事実を主張し又は広めることによって、当該人物の価値を貶め、当該人物に関する世評に悪影響を及ぼし又は当該人物の信用を危険に晒す行為を規制 故意犯である点において上記悪評の流布罪とは区別される
写真撮影等によるプライバシー・人格権侵害罪	201a	プライベートな場所にいる特定の人物の写真や画像、特定の人物の無力さを示す写真や画像又は極めて侮辱的な態様で死者を描写した写真や画像を許可なく撮影・送信すること、上記の写真や画像を使用すること等を規制

＜偽・誤情報の発信又は媒介と関連性のある犯罪の例（資料 WG16－３）＞

ドイツにおける特徴として、重大な犯罪や、公共の危険にかかわる犯罪を行う旨の脅迫行為や、その犯罪行為が差し迫っているように装うことを広く処罰する規定（ドイツ刑法第 126 条）や、特定の民衆に対する憎悪を呼び起こすような行為やナチスの行為を礼賛したり、ナチスが行った行為は嘘であるといった情報を流布したりする行為を処罰する規定（同法第 130 条）が存在する。また、プライバシー・人格権侵害について、刑事罰の対象となっている点も特徴的である（同法第 201 条 a）。

また、政治家及び議員である原告の写真とともに「（社会的）統合は、あなたがドイツ人としてトルコ語を学ぶことから始まる」という文言を含む Facebook 上の特定の投稿及びこれと同一・本質的に類似する投稿について、原告が被告（現 Meta 社）に対し、当該投稿の差止めと損害賠償を請求した事件において、裁判所は、当該投稿が原告が上記発言をしたかのような印象を与える虚偽の引用であり、原告の人格権（誤った、改ざんされた、歪曲された発言の再現から保護される権利）を侵害していると判断した上で、被告のようなプラットフォーム事業者を間接的な妨害者と捉え、過去の裁判例[37]を引用し、

- プラットフォーム事業者やサービスプロバイダーは、権利を侵害するコンテンツを事前に発見して対応する積極的義務を負うものではない
- もっとも、権利を侵害する特定のコンテンツにつき通知を受けた場合にはこれを非表示にする等の対応行う義務を負う
- さらに、かかる通知の内容が具体的であり、これによって問題となっている権利侵害を容易に理解することができる場合には、プラットフォーム事業者において、通知から理解した権利侵害の内容に基づき対象となるコンテンツを特定した上で（調査義務）、対応を行う義務を負う。通知は上記の程度に具体的であれば足り、通知に際して必ずしも URL を指定しなければならないものではない

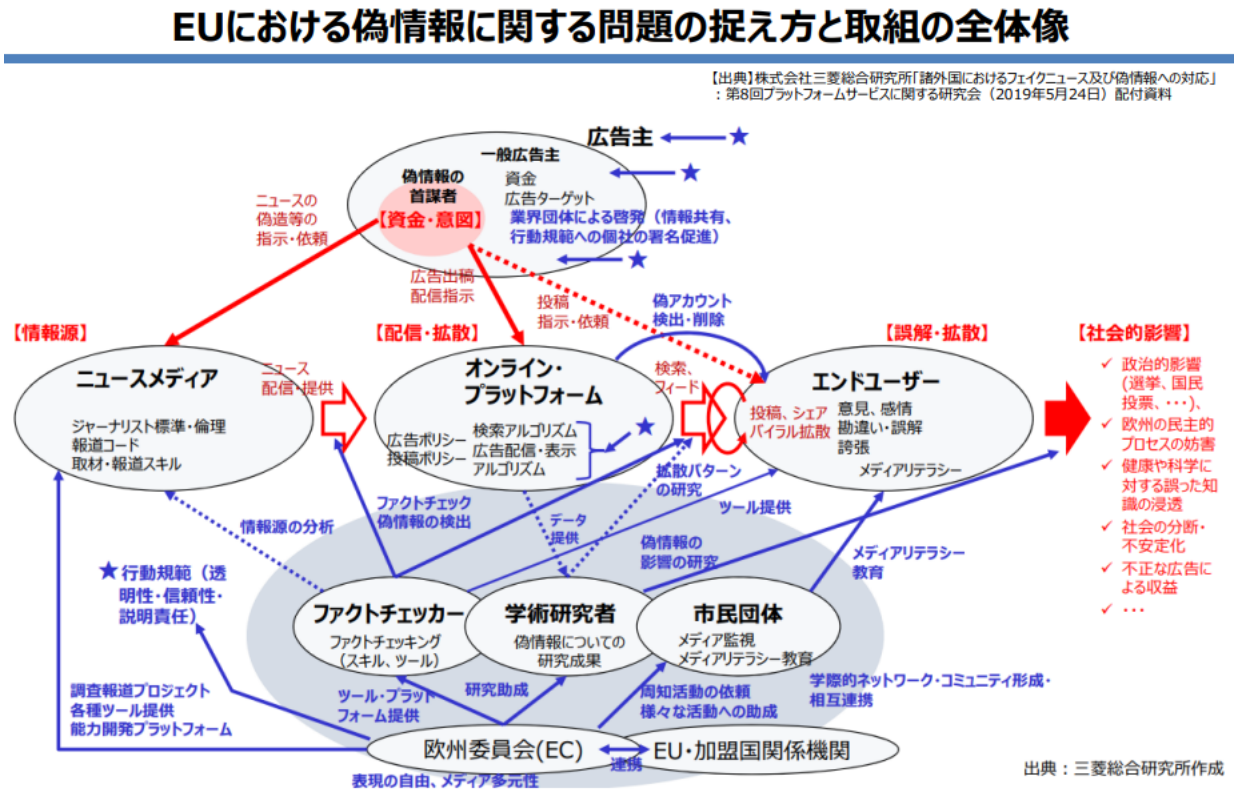
[37] BGH, judgment of August 17, 2011 - I ZR 57/ 09 - pencil perfume; BGH, judgment of July 24, 2018 - VI ZR 330/17 Rn. 36 - trial reporting

と判示した。

２．その他の取組

（１）EU における偽・誤情報に対するマルチステークホルダーによる連携・協力関係

EU では、デジタル空間における偽・誤情報の流通・拡散に関する問題に対する取組として、それぞれ直接的な情報発信・伝送・受信主体となるニュースメディア、オンラインプラットフォーム、エンドユーザーのみならず、これらの情報流通を資金面で支える広告主、さらにはファクトチェッカー、学術研究者、市民団体といったステークホルダーも、それぞれファクトチェック（スキルやツールの提供・共有）、偽情報についての研究成果の共有、メディアの監視やメディアリテラシー教育等を行っており、法制度以外の面でも多様な連携・協力関係が生まれている。



＜出典：本検討会資料５－２－２＞

特に、研究者へのデータ提供関係では、オンラインプラットフォームや検索エンジンによって展開されるアルゴリズムシステムの影響に関する研究をサポートする科学的・技術的専門知識を提供するため、2023 年４月に欧州委員会の補助機関として「欧州アルゴリズム透明化センター」（European Centre for Algorithmic Transparency。以下

「ECAT」という。）が設立されており[38]、30名を超える研究スタッフにより構成されている[39]。

ECATでは、学術界、市民社会、国家行政機関、産業界の研究者や実務家からなる国際的なコミュニティと連携しつつ、技術的、倫理的、経済的、法律的、環境的な視点を統合した学際的なアプローチがとられており、プラットフォーム事業者に対する評価・調査やアルゴリズムシステムの短期・中期・長期の社会的影響の研究、公正・透明で説明性のあるアルゴリズムアプローチに向けた実践的な方法論の開発、アルゴリズムによる透明性に関する知識の共有と国際的なステークホルダーとの議論の促進など、多様な方法で欧州委員会や各国におけるDSAの施行をサポートしている。

（２）欧州デジタルメディア監視機関（European Digital Media Observatory）における取組

欧州デジタルメディア監視機関（European Digital Media Observatory。以下「EDMO」という。）[40]は、欧州大学院（European University Institute）を中心とする事業体が、EU からの資金提供を受けながら、ファクトチェックを専門とする機関、学術研究機関、伝統メディア、ジャーナリスト、テック企業やプラットフォーム事業者、広告関係団体その他の関係ステークホルダーによる連携・協力のためのハブとなることを想定して運営され、偽情報に対抗するコミュニティを支援するプロジェクトである。

EDMOは、上記の関係ステークホルダーが相互に積極的に交流することを促進し、かつ、政策立案者に支援を提供することを通じ、偽情報に対抗するための各ステークホルダーの行動を協調させることに寄与することが期待されている。

[38] 本検討会資料14－1。
[39] https://algorithmic-transparency.ec.europa.eu/team_en（2024年6月21日時点）。
[40] https://edmo.eu/。

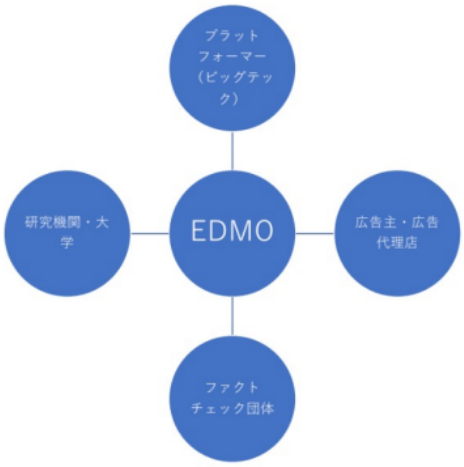
欧州デジタルメディア観測所(EDMO)の事例：牧野（2022）の現地調査などより
(The European Digital Media Observatory)

・欧州における偽情報対策のステークホルダーを束ねる連携組織
ステークホルダー＝ファクトチェック組織、学術・研究機関、メディア・ジャーナリスト、テック企業・PF、広告業界など

・2020年6月、EUI（欧州大学院大学）内に設立

・主たる任務：偽情報対策の実践基準(Code of Practice)の改訂、署名ステークホルダーの遵守モニタリング

・署名は34企業・団体　Meta、Google、Twitter（現X）、ByteDance(TikTok)、Microsoft、IAB（インターネット広告団体）など



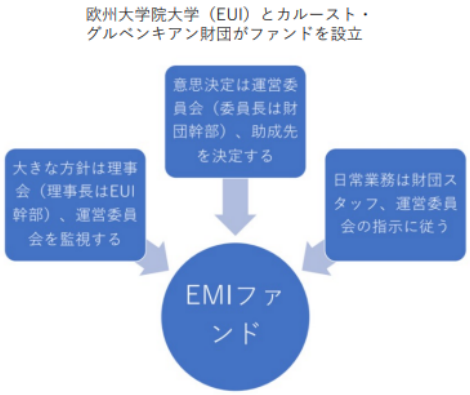
＜出典：本検討会資料 20－２－１＞

EDMO の主たる任務の一つは、2022 年版行動規範の改訂及び署名ステークホルダーの遵守状況のモニタリングであり、これを含め、EDMO の具体的な活動としては、①欧州内のファックトチェック機関をマッピングし、クロスボーダーな共同活動や専門的なトレーニングモジュールを発展させることでそれらのファクトチェック機関を支援すること、②偽情報に関する査読済みの学術論文のグローバルレポジトリを作成し、定期的に更新することを含め、欧州レベルでの偽情報に関する研究活動をマッピングし、支援し、かつ協調させること、③意識の向上、オンライン上の偽情報に対するレジリエンスの構築及びメディアリテラシーキャンペーンの支援を目的とした情報や教材をメディア関係者、教員及び市民に提供する公開ポータルを構築すること、④偽情報をよりよく理解するために活動する学術研究者によるプラットフォームのデータへのセキュアかつプライバシーが保護されたアクセスを確保するための枠組みを設計すること、⑤偽情報の拡散及び影響を抑止するためにオンラインプラットフォームが導入しているポリシーの監視について政府機関を支援すること、の５本柱で成り立っている。

上述のとおり、EDMO は EU から資金提供を受けているものであるが、ファクトチェック活動やメディアリテラシー活動への支援は EU から独立した運営ができるように工夫された欧州メディア情報基金（EMI ファンド）を通じて行われるものとされている。

欧州デジタルメディア観測所(EDMO)の事例：牧野（2022）の現地調査などより
(The European Digital Media Observatory)

・EUから財政援助を受けている
・EDMOとEUの関係
「政治に左右されず、独立性・公平性を維持するのは極めて大切」（ゴーリ事務局長、2022）
「EDMOは、EC（欧州委員会）を含む公的機関から完全に独立した統治機構を有している」（ECホームページ）
https://digital-strategy.ec.europa.eu/en/policies/european-digital-media-observatory
・ファクトチェック活動やメディアリテラシー活動への支援はEMIファンド（欧州メディア情報基金）を通じて行われる。EMIファンドは独立した運営ができるように工夫（右図）



▼ファクトチェックに限らない包括的な偽情報対策
▼公的機関からの独立：財政援助を受けているが実践的に可能かどうか

＜出典：本検討会資料 20－２－１＞

（３）Factisk.no（ノルウェー）における取組[41]

ノルウェーでは、2017 年から、主要６メディアが協力して運営するファクトチェック団体である Factisk.no が活動している。

Factisk.no における最近の取組として、2022 年２月に発生したロシアによるウクライナ侵攻を契機に、12 メディア 32 記者の共同作業により、ロシア関係の偽・誤情報を分析する OSINT[42]プロジェクト（Factisk.Verifiservar）が開始された。このプロジェクトにおいては、参加メディアの規模等に応じ、記事の執筆・発表、ファクトチェックの実施、ノウハウの共有等の役割分担が細かく決められており、ファクトチェック結果は全報道メディアに開放される。

この取組は、特定の分野について競争を保留して、社会に流通する情報の量や多元性を高めることについて、ノルウェーの伝統メディアの間で一種の共通認識が存在することの表れであるとの指摘がある。

Ⅲ　英国

１．法制度に関する動向

英国では、個人にとってより安全なオンラインサービスの提供を確保することを目的

[41] 本検討会第3回会合(2023 年 11 月 27 日)における奥村構成員ご発表資料参照。
[42] Open Source Intelligence の略で、一般に公開されている情報源からアクセス可能なデータを収集・分析等する諜報活動を指す。

として、違法又は子供に有害なコンテンツや活動によるリスクを特定・軽減・管理する義務等をオンラインサービスの提供者に対して課すオンライン安全法（Online Safety Act, 2023）が 2023 年 10 月 26 日に制定されている。

オンライン安全法の監督・執行については、英国情報通信庁（Office of Communication。以下「OFCOM」という。）が担っており、OFCOM に対し、行動規範の策定・公表が義務付けられているほか、ガイダンスを発行し情報伝送 PF 事業者が同法の定める義務を遵守できるよう支援することが義務付けられている。

他方、情報伝送 PF 事業者等に課されている主な義務としては、①違法コンテンツに関するリスク評価の義務、②違法コンテンツに関する安全義務（違法コンテンツを速やかに除去するためのシステム及びプロセスの導入等）、③コンテンツ報告及び不服申立てに関する義務（違法コンテンツや子供に有害なコンテンツを容易に報告できる仕組みの設定等）、④子供に関するリスク評価の義務、⑤子供の保護に関する安全義務等がある。

さらに、ユーザー間サービスを提供する事業者（例：SNS、写真・ビデオ共有デバイス、インスタントメッセージサービス、オンラインゲームサービス 等）で、一定以上の規模を持つ事業者に対しては、追加義務として、⑥ユーザーエンパワーメントに関する義務、⑦詐欺的広告の防止義務、⑧本人確認の義務が課せられている。

上記の①～⑧の義務違反に対する罰則としては、1,800 万ポンド、当該企業の最終事業年度における全世界売上高の 10 %のいずれか高い額を上限とする制裁金が課せられる。

事業者に課される義務・違反時の罰則の概要

項目	内容	条項
事業者に課される主な義務	● 違法コンテンツ（Illegal content）に関するリスク評価の義務：違法コンテンツが個人に及ぼすリスク（9条5項(b)the level of **risk of individuals** who are users of the service encountering the following by means of the service）に関する評価を実施し、サービスに重要な変更を加える場合を含めて、アップデートを実施 ● 違法コンテンツに関する安全義務：サービスのデザイン及び運用に関して、違法コンテンツに関するリスクを効果的に管理・軽減する方策の採用、違法コンテンツを速やかに除去するためのシステム及びプロセスの導入 、利用規約又は公表文書において違法コンテンツからの保護に関する措置の開示 ● コンテンツ報告及び不服申立てに関する義務：ユーザーや影響を受ける個人が違法コンテンツや（子供がアクセス可能な場合）子供に有害なコンテンツを容易に報告できる仕組みの設定、ユーザー等からの不服申立て手続の整備及び不服申立てに対する対応の実施	● 9条、26条 ● 10条、27条 ● 20条、21条、31条、32条
子供がアクセスする可能性の高いサービスに対しての追加義務	● 子供に関するリスク評価の義務：コンテンツが子供に及ぼすリスクに関する評価を実施し、サービスに重要な変更を加える場合を含めて、アップデートを実施 ● 子供の保護に関する安全義務： サービスのデザイン及び運用に関して、子供に及ぼすリスクを効果的に管理・軽減する方策を採用、コンテンツのリスクから子供を保護するためのシステム及びプロセス（「ユーザー間サービス」では年齢認証・推計）の導入 、利用規約又は公表文書において子供の保護に関する措置の開示	● 11条、28条 ● 12条、29条
特定のカテゴリーのサービスに対しての追加義務	特定のカテゴリーのサービスは今後、別途定められる予定の規則によって規定される ● ユーザーエンパワーメントに関する義務：ユーザーエンパワーメントに関する評価を実施し、サービスに重要な変更を加える場合を含めて、アップデートを実施、大人のユーザーがコンテンツコントロールを行える仕組みを採用、可能な最も早い機会において、デフォルト設定を維持するか、変更するかを選択できるシステム及びプロセスの導入、利用規約において利用可能なコントロール機能及び直近のユーザーエンパワーメントに関する評価の要旨を開示 ● 詐欺的広告の防止義務：詐欺的広告を速やかに除去するためのシステム及びプロセスを導入し、そのために用いている技術を利用規約において開示 ● 本人確認の義務：大人のユーザーに対して、サービスの利用に本人確認が不要の場合でも、本人確認のオプションを付与し、利用規約において開示	● 14条、15条 ● 38条 ● 64条
違反時の罰則	● 1,800万ポンド、または、当該企業の最終事業年度における全世界売上高の10%のいずれか高い額を上限とする制裁金が課される可能性がある	● スケジュール13の4条1項、5条3項

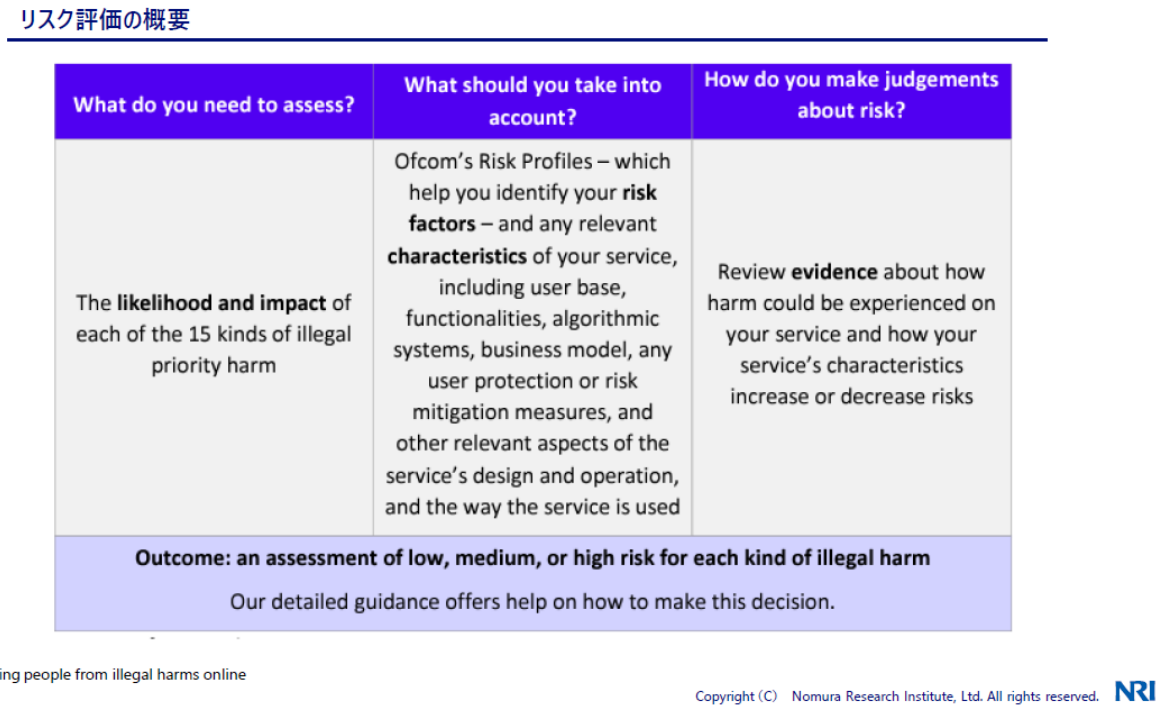
https://www.legislation.gov.uk/ukpga/2023/50/contents/enacted
https://www.theverge.com/23708180/united-kingdom-online-safety-bill-explainer-legal-pornography-age-checks
西村・あさひ法律事務所「英国オンライン安全法（Online Safety Act）の解説-その適用範囲と要対応事項の概要-」

Copyright (C) Nomura Research Institute, Ltd. All rights reserved. NRI

＜出典：本検討会資料 WG12-１-２＞

また、上記の義務のうち①違法コンテンツに関するリスク評価の義務については、情

報伝送 PF 事業者自身が実施するとされており、OFCOM からの提示案[43]では、４つのステップに分けて実施することが提唱されている。現状ではリスク評価の義務はまだ施行されていないものの、OFCOM によるガイダンスの最終版が公表されてから３か月以内にリスク評価を行うことが課されており、内容に関する周知が先行して OFCOM からされている。



リスク評価の概要

What do you need to assess?	What should you take into account?	How do you make judgements about risk?
The **likelihood and impact** of each of the 15 kinds of illegal priority harm	Ofcom's Risk Profiles – which help you identify your **risk factors** – and any relevant **characteristics** of your service, including user base, functionalities, algorithmic systems, business model, any user protection or risk mitigation measures, and other relevant aspects of the service's design and operation, and the way the service is used	Review **evidence** about how harm could be experienced on your service and how your service's characteristics increase or decrease risks
Outcome: an assessment of low, medium, or high risk for each kind of illegal harm Our detailed guidance offers help on how to make this decision.		

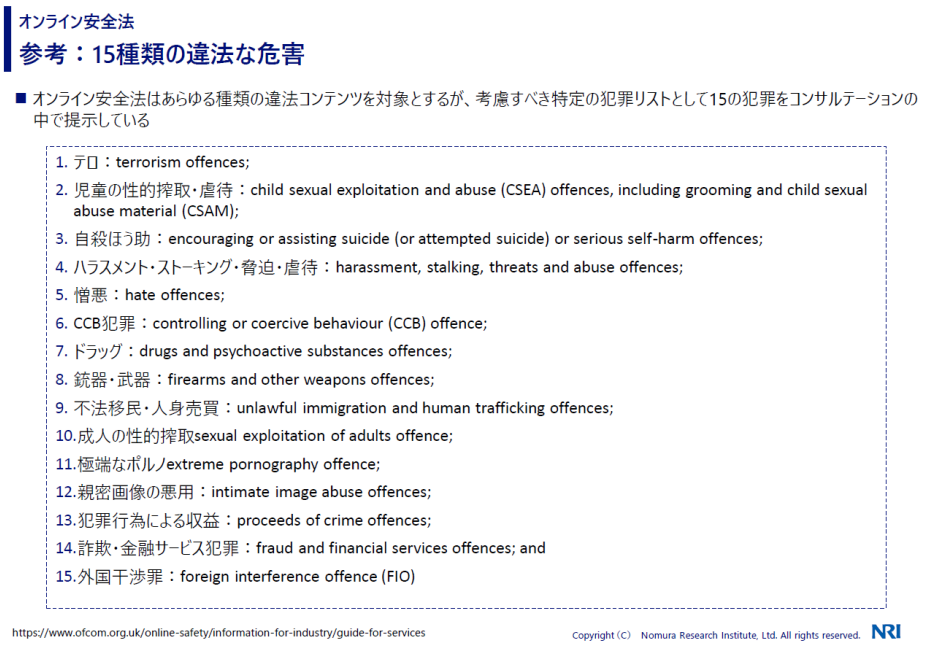
出所：Protecting people from illegal harms online

Copyright (C) Nomura Research Institute, Ltd. All rights reserved. NRI

＜出典：本検討会資料 WG12-１-２＞

オンライン安全法上、違法コンテンツとは、テロ、児童の性的搾取、自殺勧奨、自傷行為、ハラスメント、ヘイトクライム、支配行為、薬物犯罪、武器関連犯罪、入国管理法違反、人身売買、成人の性的搾取、過激なポルノ、親密な画像の乱用、犯罪収益、詐欺、外国干渉などを対象とするコンテンツとされている。違法コンテンツの定義の中に、偽情報は含まれていない。

[43] OFCOM は3つの段階に分けて具体的な施行に向けたガイダンスと行動規範を公表するとしており、2023 年 4Q から、順次コンサルテーションを公表している。



＜出典：本検討会資料 WG12-１-２）＞

前述の義務のうち⑦詐欺的広告の防止義務については、「第１類サービス」として例えば、大規模な SNS 等、そして、「第２A 類サービス」として、例えば、大規模な検索エンジン等に対して、①詐欺的広告から成るコンテンツに個人が遭遇する事態を防ぐ義務、②詐欺的広告から成るコンテンツが存在する時間を短くする義務、③指摘を受けた際に速やかに削除する義務、が課せられている。

なお、オンライン安全法上の偽・誤情報対策に関するものとしては、①偽・誤情報のアドバイザリー委員会の設置、②新たな虚偽通信罪の規定、③OFCOM に対するメディアリテラシー義務が定められている。

①偽・誤情報のアドバイザリー委員会は、OFCOM が指名した委員によって構成されている。同委員会は、OFCOM に対して、規制対象サービスの当該サービス上の偽・誤情報への対処に関する助言や、偽・誤情報に関して規制対象サービスに課す透明性レポートやメディアリテラシーに対する OFCOM の権利行使に関する助言をすることが求められており、設置から 18 か月以内の報告書の公表、また、その後定期的に報告書を公表することが義務付けられている。

②新たな虚偽通信罪については、虚偽であると知っている情報を、情報が心理的又は身体的危害を与えることを意図していた場合及びその情報を送信することについて合理的な理由がない場合に違反となるとされている。

③OFCOM に対するメディアリテラシー義務については、規制対象サービスを利用する際に自分自身や他人を守ることができる方法に関して、一般市民の認識と理解を高めるための措置を講じる義務とされている。

２．その他の取組

（１）ケンブリッジ大学等の取組

プレバンキングについて、ケンブリッジ大学の研究チームを中心にWHOやGoogle等との共同研究による大規模な実証研究が行なわれている。その中で、ゲーム形式の「BAD NEWS」や「GO VIRAL!」といった、誤情報を広めるためによく使われるテクニックをあらかじめ学ぶことができるようなコンテンツの開発も行なわれている。

＜出典：本検討会資料４－１＞

また、Google、BBCと共同研究を行ない、行動心理学に関する予備知識がなくてもプレバンキングを実践できるよう、学術的研究を実践的なハウツーガイドとしてまとめた「A Practical Guide to Prebunking Misinformation」が作成されている。

＜出典：本検討会資料４－１＞

Ⅳ　大洋州地域

１．法制度に関する動向

（１）オーストラリア

オーストラリアでは、2023 年に、偽・誤情報対策強化に向けた改正法案が提案され、オーストラリア通信メディア庁（Australian Communications and Media Authority。以下「ACMA」という。）の自主規範に関する執行権限の強化や、偽・誤情報に関する定義の詳細化が盛り込まれている。

ACMA の自主規範に関する執行権限の強化については、ACMA に情報収集、行動規範の策定の命令・執行、業界標準の策定・執行の権限を付与することが規定されている。

また、偽・誤情報に関する定義について、誤情報は、①虚偽、誤解を招く、又は欺瞞的なデジタルサービスを利用して広められたコンテンツ、②コンテンツがデジタルサービスを通じてオーストラリアのエンドユーザーに提供されている、③デジタルサービス上でのコンテンツ提供が重大な害（serious harm）を引き起こすか、又はその一因となる可能性が合理的に高い、④コンテンツが誤情報の目的で除外されない [44]と定義されている。その上で、偽情報については、（ｉ）上記の誤情報の基準を満たすコンテンツであることに加え、（ⅱ）コンテンツが、自動化されたプロセスや外国からの干渉を含め、欺罔的意図をもって流布されていること（意図的又は悪意を持って流布された偽情報のコンテンツを捕捉する）と詳細にされている。この改正法案は 2023 年６月 25 日に公開され、2,265 件の一般からの意見と約 20,000 件のコメントが寄せられており、2024 年３月時点では、議会への法案提出は未定となっている。

２．その他の取組

（１）オーストラリア

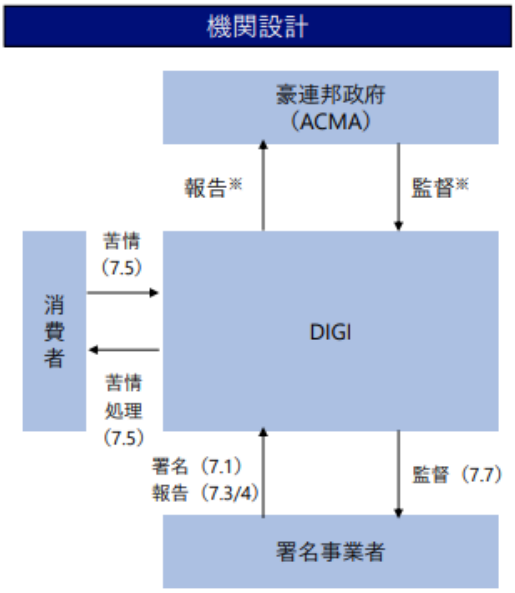
オーストリア政府が公表した行政文書 [45]において、主要な情報伝送 PF 事業者に対して、偽情報とニュースの質に関する自主的な行動規範を策定するよう要請するとともに、ACMA がその規範を監督し、情報伝送 PF 事業者による対策の妥当性と偽情報の広範な影響について政府に報告するという執行の枠組みが規定されている。

そして、2019 年にオーストラリア政府から産業界に対して行動規範の策定が依頼され、これに対し、デジタル産業団体（The Digital Industry Group Inc。以下「DIGI」という。）が草案作成に手を挙げ、2021 年２月、DIGI から「偽・誤情報に関するオーストラリアの行動規範」（The Australian Code of Practice on Disinformation and Misinformation:ACPDM)が公表されている。2021 年５月には署名企業８社(Adobe、Apple、Facebook、Google、Microsoft、Redbubble、TikTok、Twitter）により、最初の透明性レポートが公表されている。

---

[44] 除外される目的:娯楽、パロディ、風刺を目的として善意で制作されたコンテンツ、専門的なニュースコンテンツ、連邦・州・準州・地方政府によって認可されたコンテンツ、認定教育プロバイダーによって、又はそのために制作されたコンテンツ

[45] Regulating in the digital age Government response and implementation roadmap for the Digital Platforms Inquiry



＜出典：本検討会資料 WG８－１＞

以上の行動規範には、DIGI に１年後に行動規範の見直しを行う義務が含まれており、DIGI はこれに基づきコンサルテーションを行い、2022 年 12 月 22 日、パブリックコンサルテーションへの対応レポートとそれを踏まえた更新版の行動規範（以下「2022 年行動規範」という。）が公表されている。

2022 年版行動規範では、より小規模な情報伝送 PF の参加の促進、「harm（害）」の定義の更新、偽情報を広めるレコメンドシステムやデジタル広告への掲載を広告主に思い止まらせるためのコミットメントの追加等の変更が行われており、「指導原則」をもとに、偽情報や誤情報の定義、目標と達成すべき結果、具体的な措置が定められている。

文書	目次	概要
本編	1.前文 （Preamble）	• 制定の経緯や目的を記載
	2.指導原則 （Guiding Principles）	• 全体を通じた原則を記載
	3.用語法（Glossary）	• 用語の定義を記載
	4. 本規範のスコープ、適用、開始 （Scope, application and commencement of this Code）	• 規範の適用範囲等を記載
	5. 目標と措置 （Objectives and Measures）	• 達成目標とそのための結果（アウトカム）、取るべく具体的な措置を記載
	6. プラットフォーム特有の措置 （Guidance on platform-specific measures）	• 措置の導入にあたり、署名事業者・団体が考慮すべき事項を記載
	7. 規範の運用 （Code administration）	• 加入や脱退、透明性レポートの提出、苦情の処理、事務局等を記載
APPENDIX 1	Opt-in Nomination Form	• オプトインの表明に関する書式を記載
APPENDIX 2	Annual Transparency Report	• 透明性レポートの書式を記載

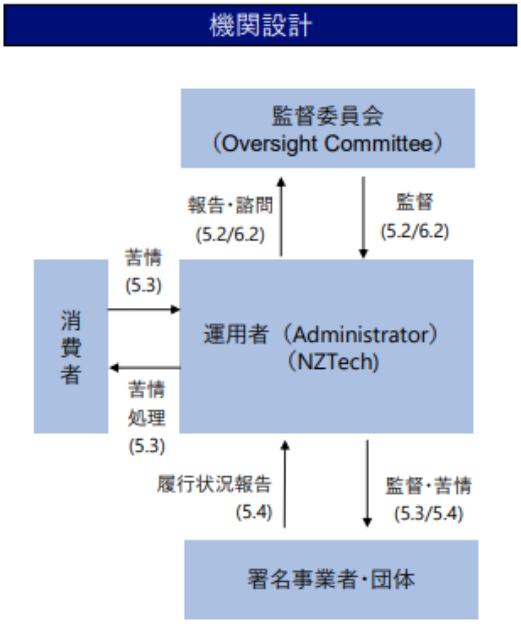
＜出典：本検討会資料 WG８－１＞

（２）ニュージーランド

ニュージーランドでは、急速に進化するインターネットのグローバルな性質に対し、安全性に対しても同様に柔軟で俊敏なアプローチが必要であり、効果的なオンライン安全対策には、専門知識とパートナーシップを活用し、透明性を確保することが必要であるとされ、行動規範が策定されている。2024 年３月時点で、Meta (Facebook and Instagram)、Google (YouTube)、TikTok、Twitch 及び X (twitter)の５社が署名している。

具体的には、オンライン上の安全性確保に関する取組みを行い、20 年以上の歴史を持つ NGO である Netsafe において、1,600 以上の加盟者を持つニュージーランドのテック産業団体である NZTech の協力を得て、マオリ族のアドバイザー、政府、市民社会、一般市民と協議して行動規範が策定されている。

また、NZTech が中心となり、履行確保や消費者からの苦情処理等を担っており、その執行については、NZTech、Netsafe、署名事業者・団体、マオリパートナー、市民社会、その他関連する合意された利害関係者（政府、学識経験者など）の代表を含む様々な利害関係者で構成される監督委員会により監督されている。政府機関の関与は監督委員会への委員派遣にとどまっている。



＜出典：本検討会資料 WG８－１＞

（３）EU・オーストラリア・ニュージーランドの行動規範の比較

まず、偽・誤情報の定義について、偽情報については、オーストラリア及びニュージーランドでは、ほぼ同じ定義が採用されている。EU では、不正な行為による伝播という要素がなく、風刺やパロディ等の除外がある。また、誤情報については、虚偽・誤解を招く表現、ユーザーによる伝播、害を起こす意図がない点は３か国で共通している。オーストラリア及びニュージーランドでは、害をもたらす合理的可能性の要素が加わり、オーストラリアでは風刺等の除外があると定義されている。

次に、共通している点として、①広告収益のはく奪を狙い、特に偽情報による収益のはく奪に向けた措置（コンテンツの適格性の判断やその判断メカニズムの導入）や、関係者との協力を求めている点、②偽・誤情報の拡散防止に向け、ポリシー・手順等の公表・導入・維持、完全性やセキュリティを損なう行動の禁止又は管理について、これらの導入及び定期的な報告を求めている点、③関連ステークホルダーのエンパワーメント（ユーザーに対しては、ユーザーが適切なコンテンツを識別できるよう、レコメンドシステムに関する透明性や選択肢の提示、ファクトチェック機関等からの確認ラベルの表示の導入、研究者に対しては、研究の支援や研究者とのイベントの開催等、ファクトチェック機関に対しては、ファクトチェック機関との連携関係の構築）を求めている点、④コミットメントに基づく措置や実際に行われた内容について、対外的に公表されることを求めている点がある。

他方、異なっている点として、①政治広告（政治広告の共通定義策定、有料コンテンツとの区別、政治・争点広告のレポジトリの確保、API の設定、市民社会へのコミットメント)、②サービスの完全性（AI システムの透明性、偽・誤情報のフラグ付け、強制措置の不服申立、メッセージアプリの抑制措置)、③研究者へのエンパワーメント（データアクセスとそのガバナンス体制)、④ファクトチェック機関のエンパワーメント（ファクトチェック機関によるデータアクセス、ファクトチェック機関の基準）については、EU の行動規範が最も包括的であり、EU にのみ規定がある項目となっている。

	EU	オーストラリア	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	署名事業者・団体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）
行動規範の運用主体	欧州委員会	DIGI （同上）	NZTech （非営利業界団体）
公表時期	2022年６月16日 ※2018年10月に自主規制を公表、欧州委員会からの2020年9月の評価、2021年5月のガイダンスを踏まえ、22年6月に公表	2022年12月22日（最終更新日） ※2021年2月22日に適用開始、同年10月11日更新	2022年７月25日
対象とする事業者・サービス	署名事業者・団体（オンラインプラットフォーム事業者、広告関連事業者・団体、ファクトチェック団体・関連サービス事業者やその他の市民・業界団体・関連事業者など） （加盟を希望する団体は、コミットメントと措置に同意する登録用文書に記入し、常設タスクフォースに提出することで、本規範に参加できる） * Google, Meta, Microsoft, TikTok等の43者が署名済（24年2月時点）	署名事業者・団体（"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。"） * Google, Meta, Microsoft, X(Twitter), TikTok, Redbubble, Apple, Adobe, Twitchの9者が署名済（24年2月時点）	署名事業者・団体（"デジタル・プラットフォームは、当規範の署名事業者・団体となることができ、いつでも加盟することができる"） * Google, Meta, X(Twitter), TikTok、Twitchの5者が署名済（24年2月時点）
運用状況	署名事業者・団体が半年ごとに透明レポートを公表 最新版として、2023年7月分（対象期間：2023年1月1日～6月30日）までが公表されている 2024年1月分（2023年7月1日～2024年1月31日）は未公表（24年2月時点）	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）	2023年に各社が透明性レポートを提出しているほか、2024年1月には独立レビューの結果が公表。

＜出典：本検討会資料 WG８－１ ＞

Ⅴ OECD

OECD では、GPAI(Global Partnership on Artificial Intelligence)や IEEE(Institute of Electrical and Electronics Engineers) 等と協力し、生成 AI で作られた偽情報・ディープフェイク・その他の操作されたコンテンツによる、公共の言説の歪曲、陰謀論の創出・拡散、選挙への干渉、市場の歪曲、暴力の扇動といったリスクを低減し、生成 AI の悪用・誤用に対するレジリエンスを構築すべく、生成 AI 時代の信頼に関するグローバル・チャレンジを設立することを 2023 年７月に発表されている[46]。

グローバル・チャレンジにおいては、技術者、政策立案者、研究者、専門家、開発者など多様なステークホルダーと連携・協力し、生成 AI による偽情報の更なる拡散に対抗するための効果的なアプローチに関する具体的な知見を得て、国際的に展開し得る確立された手法を生み出すべく、信頼を促進する革新的なアイディアを検証することとされている。具体的には、生成 AI による偽情報の拡散に対抗する①アイディアを公募し、採択された提案にリソースを配分した上で、②プロトタイプの開発、評価を得た取組には更なる追加リソースを配分し、③パイロット・プロジェクトの実施・展開を行い、優れた取組は表彰することとされている[47]。なお、本取組は、広島 AI プロセス包括的政策枠組みにおけるプロジェクト・ベースの協力に位置付けられている[48]。

[46] https://oecd.ai/en/wonk/global-challenge-partners
[47] https://globalchallenge.ai/
[48] https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document02.pdf

さらに、OECD では、偽情報・誤情報に関する各国の取組、OECD の関連レポートを各国間で共有すること等を通じ、政府の取組を比較し適切かつ効果的な対応についての知見を得て、外国からの干渉に対抗し情報のインテグリティや民主主義の強化を図るべく、公共ガバナンス委員会（Public Governance. Committee：PGC）により、「OECD DIS/MIS Resource Hub」が 2022 年 11 月に設置されている[49]。そして、専門家会合及び 15 加盟国からなるステアリンググループにおける検討、公開カンファレンスの開催を経て、その成果として、「Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity」が 2024 年３月に公表されている。報告書の提言の主な内容としては、①情報源の透明性、アカウンタビリティ、多元性を強化する政策の実行として、ローカル・ジャーナリズムを含む多様かつ独立したメディアの促進や、オンラインプラットフォームの説明責任及び透明性の強化、②偽情報に対する社会のレジリエンスの涵養として、偽情報を認識し対抗する個々人の批判的思考力の強化や、社会全体を動員した包括的かつエビデンスに基づく政策の実行、③情報空間の清廉性を保持するための制度設計や政策手法の更新として、公的機関による一貫したビジョンとアプローチを担保する、権限の明確化と基本的な自由を尊重した上での、監督権限や調整メカニズム、戦略枠組み、能力構築プログラムの適切な発展・実施や、民主主義国家間の国際協力が挙げられている。

こうした動向も踏まえ、OECD 閣僚理事会において、2024 年５月、「人工知能に関する理事会勧告」（Recommendation of the Council on Artificial Intelligence）が改定されている。具体的には、法の支配、人権並びに民主主義的及び人間中心の価値観を尊重するための具体的な取組として、AI によって増幅された誤情報や偽情報への対処が盛り込まれている。また、安全上又はセキュリティ上のリスクをもたらすことがないよう、情報インテグリティを強化するメカニズムを整理すべき旨が盛り込まれている[50]。その上で、閣僚声明においては、OECD が、社会、民主主義的価値及び制度に対する信頼への主要なリスクである偽情報、誤情報に焦点を当てつつ、AI に関する国際的な議論に積極的に貢献することを期待し、OECD のデジタル政策に関連するものを含め、偽情報及び誤情報に対抗するための政策措置に関する広範な活動を歓迎する旨が盛り込まれている[51]。

Ⅵ　国連

１．情報インテグリティに関する行動規範

国連では、国連事務総長報告書（2021 年９月）[52]「私たちの共通の課題」（Our Common

---

[49] その後、情報の清廉性や民主主義の強化をより前面に打ち出すべく、名称を'the OECD Hub on Information Integrity'と改称。
https://www.oecd.org/stories/dis-misinformation-hub/#:~:text=The OECD Hub on Information,to tackling mis- and disinformation

[50] https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449

[51] https://www.mofa.go.jp/mofaj/files/100664931.pdf(仮訳は https://www.mofa.go.jp/mofaj/files/100664930.pdf)

[52] https://www.unic.or.jp/news_press/info/42716/

Agenda）で挙げられた 12 のコミットメント[53]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月に公表されている。当該報告書では、「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及び基準を策定」が目標として規定されている。

また、同年 6 月には「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示されている。具体的には、情報インテグリティに関する行動規範の基礎となる９つの原則が提示され、ポリシーブリーフにおいては、ほとんどのデジタルプラットフォームでは、何らかの自主規制、モデレーション、監視の仕組みが導入されているが、コンテンツ削除の方針と実践に関する透明性は、依然として課題となっていることが指摘されている。

【９つの原則】

① Commitment to information integrity（情報インテグリティへの取り組み）
② Respect for human rights（人権の尊重）
③ Support for independent media（独立系メディアへの支援）
④ Increased transparency（透明性の向上）
⑤ User empowerment （ユーザー・エンパワーメント）
⑥ Strengthened research and data access（研究とデータアクセスの強化）
⑦ Scaled up responses（対応の拡大）
⑧ Stronger disincentives（より強力な阻害要因）
⑨ Enhanced trust and safety（信頼性と安全性の向上）

【出典】国連「Common Agenda」ウェブページ（https://www.un.org/en/common-agenda/policy-briefs）等

＜出典：本検討会資料５－１－２＞

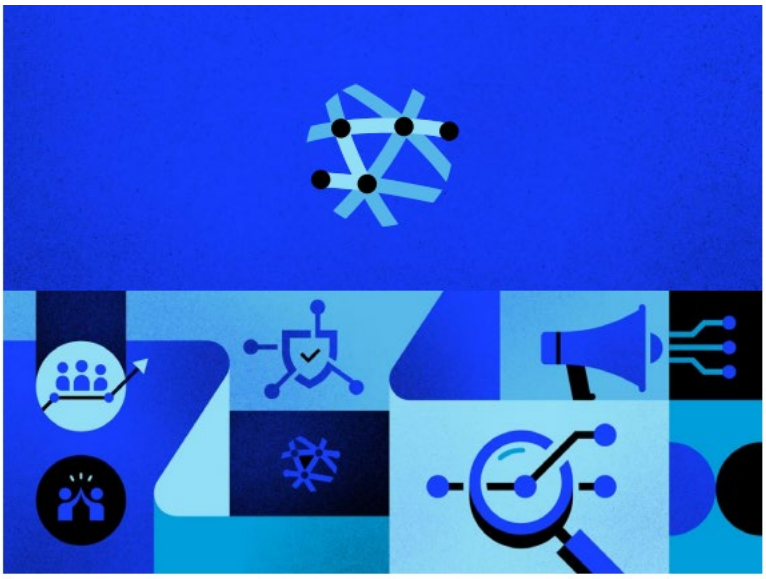
当該情報インテグリティを受け、2024 年９月の「未来サミット（Summit of the Future）」[54]までに「デジタルプラットフォーム上の情報インテグリティに関する行動規範」を作成するべく、世界中で国別・地域別のコンサルテーションを実施した[55]。行動規範は、情報インテグリティへの取組、人権の尊重、独立系メディアの支援、透明性の向上のレコメンデーションに基づくことができるとされている。

２．情報インテグリティのための国連グローバル原則

国連は 2024 年６月 24 日、誤情報、偽情報、ヘイトスピーチの拡散から生じる危害を

[53] https://www.unic.or.jp/files/Our_Common_Agenda_KEY-PROPOSALS-ACROSS-THE-12-COMMITMENTS_J_FINAL.pdf
[54] https://www.unic.or.jp/files/our-common-agenda-summit-of-the-future-what-would-it-deliver_J.pdf
[55] https://www-un-org.translate.goog/en/information-integrity/code-of-conduct?_x_tr_sl=en&_x_tr_tl=ja&_x_tr_hl=ja&_x_tr_pto=sc

抑止するための緊急行動に向けた提言である「情報インテグリティのための国連グローバル原則」を発表している[56]。



＜出典：United Nations「United Nations Global Principles For Information Integrity」＞

この原則は、人々に自らの権利を要求する力を与えることを目指しており、何十億もの人々が偽りの物語や歪曲、嘘にさらされているこの時代において、これらの原則は表現及び意見の自由に対する権利など人権に強く根差した、今後進むべき確かな道を示すものとされている。

またこの原則は、加盟国、民間セクター、若者のリーダー、メディア、学術界、市民社会との幅広い協議の結果であり、そこに記された提言は、人権、平和な社会、そして持続可能な未来を擁護する、より健全かつ安全な情報空間を育むことを意図されており、提案には以下が含まれている。

・各国政府、テクノロジー企業、広告主、メディアその他のステークホルダーは、いかなる目的であっても、偽情報とヘイトスピーチの利用、支持、または増幅を自粛すべきである。
・各国政府は、情報への迅速なアクセスを提供し、自由で存続可能、かつ独立した、複数のメディアがいる環境を保証し、ジャーナリストや研究者、市民社会に対する強力な保護を確実なものにすべきである。
・テクノロジー企業は、すべての製品の設計において安全性とプライバシーを確保する

[56] https://www.un.org/en/information-integrity

と共に、オンライン上で標的とされる集団のニーズに特別な注意を払いつつ、国や言語を越えて一貫したポリシーと資源を適用すべきである。これらの企業は、危機対応力を高め、選挙をめぐる情報インテグリティを支援する施策を講じるべきである。

- AI テクノロジーの開発に携わるすべてのステークホルダーは、すべての AI アプリケーションの設計、展開、利用において確かな安全、安心、責任、倫理と、人権の擁護を保障するために、緊急かつ即時の、包摂的で透明性のある施策を講じるべきである。
- テクノロジー企業は、ビジネスモデルの領域を策定する上で、プログラマティック広告に依存することなく、また、ユーザーの反応を引き出すことを人権よりも優先しないように、また、同様にプライバシーや安全よりも優先しないようにし、ユーザーがオンライン上の体験や個人情報に関してより幅広く選択し、管理できるようにすべきである。
- 広告主は、テクノロジー業界に対して、故意でなかったとしても広告予算が偽情報や憎悪の資金源となったり、人権を損なったりすることのないよう、デジタル広告のプロセスの透明性を求めるべきである。
- テクノロジー企業と AI 開発者は、意味のある透明性を確保し、ユーザーのプライバシーを尊重しつつ研究者や学術界がデータにアクセスできるようにし、公開性のある独立監査を委託し、産業の説明責任に関する枠組みを共同策定すべきである。
- 各国政府、テクノロジー企業、AI 開発者、広告主は、子どもたちの保護とエンパワーメントに向けた特別な施策を講じると共に、政府は親や保護者、教育者たちにリソースを提供すべきである。

３．インターネット・ガバナンス・フォーラム（IGF）京都 2023

インターネット・ガバナンス・フォーラム（以下「IGF」という。）では、インターネットに関する諸課題について、国連主催のもと、政府、民間、技術・学術コミュニティ、市民社会等のマルチステークホルダーが対等な立場で対話を行うインターネット政策の分野では最も重要な会議の１つであり、2023 年 10 月に、IGF 京都 2023 が開催され、偽・誤情報等をテーマとするセッション等が様々開催されている。例えば、Day0（2023 年 10 月 8 日）において、総務省が協力や主催する偽・誤情報関連セッションとして、ハイレベルリーダーズセッションとワークショップ２つのセッションが開催されている。

インターネット・ガバナンス・フォーラム（IGF）京都2023

※「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」第1回会合資料（資料1－3）より抜粋

- インターネットに関する諸課題について、**国連主催**のもと、**政府、民間、技術・学術コミュニティ、市民社会等のマルチステークホルダーが対等な立場で対話**を行う**インターネット政策の分野では最も重要な会議**の１つ。
- 今般、**日本がホスト国として2023年会合を開催**（2005年の設立以来、毎年1回開催）。

＜開催概要＞

鈴木大臣メッセージ

岸田総理メッセージ

- 日時：2023年10月8日（日）～10月12日（木）（５日間）
  ※ホスト国主催の「Day0」（10/8）と国連主催の「Day1-4」（9-12）で構成
- 場所：国立京都国際会館（京都府京都市）
- 参加者：**9,279人以上**
  （うち現地参加＝6,279人※、オンライン＝3,000人以上（国連発表）（※)現地参加者数は史上最多
  岸田総理、鈴木大臣や**各国閣僚、議員**が参加
- メインテーマ：**"The Internet We Want – Empowering All People"**
  （「私たちの望むインターネットーあらゆる人を後押しするためのインターネット」）

Day1 Music Night

➢ **閣僚級ハイレベルセッションや各国議員のパーラメンタリートラック、ユーストラック含め、約300セッションが開催**

<サブテーマ>
①AI・先端技術　②インターネットの分断回避　③サイバーセキュリティ、サイバー犯罪、オンラインの安全性　④データガバナンス・トラスト
⑤デジタルデバイド・包摂性　⑥グローバルデジタルガバナンス・協力　⑦人権及び自由　⑧持続可能性・環境

＜出典：本検討会資料５－１－２＞

まず、ハイレベルセッションでは、偽・誤情報対策について、生成 AI による影響、これまでの取組を踏まえた教訓、被害をうけるコミュニティを守るための対応、規制上の課題、情報の生産者・消費者としての在り方について議論され、偽・誤情報による影響としては、例えば、次が挙げられた。

・ 認知過程への侵食、自由な意思の喪失、感情の兵器化
・ 女性、難民、人種的・民族的少数者、LGBTQ+等に対するヘイト
・ 社会の分極化、選挙プロセスの操作、民主主義への脅威

また、偽・誤情報対策としては、例えば、次の必要性が示された。

・ 「デジタル立憲主義」や表現の自由等の人権保護等による共通の原則や価値の共有
・ 国際的な規範
・ エコシステム全体や構造的問題としての対応
・ マルチステークホルダーによる連携・協力、国際的な連携・協力
・ ファクトチェックの推進、国際機関とファクトチェック機関の連携
・ メディアやジャーナリストの役割
・ 「情報的健康」等を通じたリテラシーの向上、市民社会のエンパワーメント
・ 「Originator Profile」等、デジタル署名・電子透かし等による来歴管理等の技術的な対応
・ 「Digital Service Act」等の行動規範や法制等による AI やデジタルプラットフォームにおける透明性の確保等のための規制枠組み

さらに、2024 年には、複数の民主主義国家において選挙が実施されるなど、今後、偽・誤情報対策は一層重要になっていくとの認識が共有された。

●モデレータ
・デボラ・スティール（アジア太平洋放送連合代表）
●パネリスト
・山本 龍彦（慶應義塾大学教授）
・マリア・レッサ（2021年ノーベル平和賞受賞者、フィリピン・Rappler共同創設者）
・ベラ・ヨウロヴァー（欧州委員会委員）
・ランディ・ミッシェル（米国・国家安全保障会議）
・ニック・スゾール（Meta監督委員会委員）
●レスポンデント
・ネザル・パトリア（インドネシア・通信情報技術副大臣）
・ポール・アッシュ（ニュージーランド・サイバーデジタル首相特別代表）

＜出典：本検討会資料５－１－２＞

次に、ワークショップにおいては、モデレーターにより、Existing Practices against Disinformation（EPaD）が紹介され、G7 各国等における偽・誤情報対策に関する民間企業や市民団体を含むマルチステークホルダーによる既存プラクティスが共有された。

また、各パネリストにより、偽・誤情報対策に関する取組として、「MAFINDO's Role in Today's Information Ecosystem」（アリボウォ氏）、Rappler による「Fighting Disinformation：Spreading the value of truth-telling」（チャイ氏）、Microsoft による取組（マデリン氏）、「Misinformation and Disinformation in Our Society」（山口氏）について発表された。

さらに、プラットフォーム事業者やテック企業等の民間企業、メディア・ ジャーナリスト・ファクトチェック機関、法律家、アカデミア、個人・市民社会や政府等のマルチステークホルダーによる地域や国境を越えた連携・協力の重要性が共有された。例えば、次の必要性が示された。

- 情報エコシステムにおけるプラットフォーム事業者・メディア・個人や政府等のステークホルダー間の連携・協力、市民社会やジャーナリズムとテック企業との連携
- デバンキングのみならず、偽・誤情報が出現する前の未然の取組であるプレバンキング
- 調査能力や人材共有等によるジャーナリストやニュースルームの強化
- メディアによるファクトチェック、市民社会によるメッシュ、アカデミアによるリサーチ、法律家によるアカウンタビリティの 4 層によるファクトチェック・ファクトベースの報告
- ファクトチェックを広げるための地方メディアとの連携、ファクトチェックの拡散の提供
- メディア情報リテラシー教育、こども等情報の受け手側への対応
- 漫画やインフルエンサー等による若年層へのリーチ等オーディエンスや時代に合わせた取組
- 民間のテック企業による貢献・責任として、テクノロジーの提供やラベル付け等による信頼できるニュースの発信、EU の行動規範等に基づくコンテンツモデレーショ

ン、そして、ゲーム等を通じたリテラシー教育等による開かれた民主主義プロセスの保護や情報エコシステムの提供
- インターネットを超えた社会全体の情報生態系の問題として捉えた対応
- 生成 AI の普及により誰もが偽・誤情報を作成・拡散できる時代が到来し、人の目によるチェックでは限界がある状況における技術による 対応
- グローバルだけでなくローカルな透明性の確保
- 偽情報の作成者に収入が入らないようにすること
- このセッションのようなグローバルなコラボレーション
- ファクトチェック機関の国際的な協調
- 各国におけるベストな情報やトレーニングの共有
- 偽・誤情報を生成等する AI に関する国際的なルール作り

●モデレータ
・古田 大輔（日本ファクトチェックセンター編集長）
●パネリスト
・アリボウォ・サスミト（インドネシア・MAFINDO創設者）
・チャイ・F・ホフィレナ（フィリピン・Rappler共同創設者）
・マデリン・シェパード（Microsoftデジタルセーフティー代表）
・山口 真一（国際大学GLOCOM主幹研究員/准教授）

＜出典：本検討会資料５－１－２＞

Ⅶ　東南アジア

１．ASEAN

東南アジア諸国連合（以下「ASEAN」という。）では、情報（放送）担当大臣会合（ASEAM Ministers Responsible for Information。以下「AMRI」という。）において偽情報対策関係の議論が行われており、正確で信頼性を有する情報の流通を確保することなどを目的とした「メディアにおけるフェイクニュースや偽情報対策における政府情報の管理に関する ASEAN ガイドライン」が採択されているほか、AMRI に設置されている情報（放送）担当高級実務者会合（Senior Officials' Meeting Responsible for Information。以下「SOMRI」という。）に「ASEAN フェイクニュースタスクフォース（TTFN : Task Force on Fake News)」が設置されている。

デジタル分野の政策については、デジタル大臣会合(ASEAN Digital Ministers Meeting。以下「ADGMIN」という。)において協議されており、2024 年会合の共同声明では、責任ある安全な方法で新興技術の開発に拍車をかけるための政策と指針を確立する必要性を踏まえ、「人工知能（AI）のガバナンスと倫理に関する ASEAN ガイド」（ASEAN Guide on Artificial Intelligence（AI）Governance and Ethics[57]）が承認されている。同ガイドは、ASEAN における AI ソリューションの信頼できる展開を支援する実用的かつ実行可能なツールとして作成され、その中で、生成 AI のリスクの一つとして、虚偽又は誤解を招

[57] https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf

く情報の生成や大規模な伝搬の可能性等が挙げられており、生成 AI の有するリスクへの対処の１つとして、生成 AI の生成したコンテンツとそうでないコンテンツを区別するために電子透かしを付与することが有用であると言及されている。また、ADGMIN では、デジタル分野で ASEAN として推進していくべき方針をまとめた「ASEAN デジタルマスタープラン[58]」などの策定が行われている。

また、マレーシアの Sebenarnya、シンガポールの Factually、タイの Anti- Fake News Center の政府系ファクトチェック団体が存在し、国内伝統メディアがファクトチェックの対象になっている。

ファクトチェック組織の独立性について

政府系「ファクトチェック」の研究①　マレーシア、シンガポール、タイ（※）

マレーシアのSebenarnya（2017～）、シンガポールのFactually（2012～）、タイのAnti- Fake News Center（2019～）はいずれも政府運営のファクトチェックサイト

Fake Newsの発信源（%）	Sebenarnya	Factually	AFNC
国内ソーシャルメディア	78	61	57
国内伝統メディア	6	37	34
海外ソーシャルメディア	0	0	1
海外伝統メディア	1	3	1
政府組織	0	0	2
政府の行動や政策関連ポストの割合(%)	**Sebenarnya**	**Factually**	**AFNC**
政府の行動や政策に言及	64	87	46
その他の情報提供	36	13	54
政府関連ポストにおける政府擁護の割合（%）	**Sebenarnya**	**Factually**	**AFNC**
政府・国家の擁護	55	53	80
擁護的な特徴なし	45	47	20

（※）L.　Schuldt （2021） 'Official Truth in a War on Fake News:Govermental Fact-Checking in Malaysia, Singapore, and Thailand'

▼国内伝統メディアがファクトチェックの対象になっている
▼政府関連が46～64%、政府関連での「政府・国家の擁護」は53～80%
▼ただし筆者は過剰なプロパガンダではないと結論。政府による簡潔な否定発表にとどまるからという理由

FIJ

＜出典：本検討会資料 20－２－１＞

２．インドネシア

インドネシアでは、通信情報省（KOMINFO）がインターネット上の偽・誤情報対策を担当している。同省では、2020 年～2024 年を期間とする戦略計画が策定され、同計画において、①メディアリテラシー向上を図るワークショップの開催、②SNS を通じた政府公式アカウントによる情報発信、③偽・誤情報をデマと認め、通信情報省 HP 等でデマである旨を発信するとともに偽・誤情報に対抗する事実関係・シナリオをまとめ発信などの偽・誤情報への対策が盛り込まれている。

また、通信情報省では、偽・誤情報のデマ認定の際には関係省庁と連携するとともに、大統領府の担当間で月２回の会合を開催している。

[58] ASEAN-Digital-Masterplan-2025.pdf

法制度については、プラットフォーム規制として、2020 年プラットフォーム事業者に対して公序良俗に反する不法コンテンツの削除や特定ユーザーに関するデータ開示を強制できる制度が導入されている。

ファクトチェック関係では、ファクトチェック NPO 団体の MAFINDO が、Google や Meta、TikTok などのプラットフォーム事業者やコロナ禍では WHO や UNESCO、UNICEF などの国際機関との連携、メディア機関等とのファクトチェック関係の情報共有のためのウェブサイト（CekFakta.com）を通じた情報共有等の実施などの各ステークホルダーとの連携協力を進めている[59]。

３．マレーシア

マレーシアでは、王立警察及び通信マルチメディア委員会（Malaysian Communications and Multimedia Commission。以下「MCMC」という。）が偽・誤情報対策を所管している。

リテラシー向上の取組として、MCMC が公開している偽・誤情報を含む市民のリテラシー向上のためのウェブサイトである Klik Dengen Bijak[60]（Click Wisely）等が提供されている。

ファクトチェック関係では、MCMC と様々な機関との連携により 2017 年に設置された「Sebenarnya」が、国内 SNS や伝統メディアの情報についてファクトチェックを行っているほか、政府傘下の国営通信ベルナマが 2020 年に設置した「MyCheck Malaysia」が国際ファクト・チェッキング・ネットワーク（以下「IFCN」という。）のガイドラインの範囲内で、信頼できる事実に基づくニュースを作成することで偽・誤情報の拡散防止を図っている。

また、インドネシアの MAFINDO などの国外のファクトチェックネットワークとキャパシティ・ビルディングやベストプラクティスの共有等で協力している。

法制度については、偽情報関連の法令として、1998 年通信マルチメディア法（Communications and Multimedia Act 1998[61]）が存在し、同法第 211 条は、他人を困らせ、虐待し、嫌がらせをする意図で、不適切なネットワーク設備又はネットワークサービスを利用して虚偽のコンテンツを作成又は流布することが禁じられている。

４．フィリピン

フィリピンでは、偽・誤情報対策は大統領府広報部が主導し、関係省庁が協力する形で推進されている。

リテラシー向上の取組としては、大統領府が実施しているオンラインプラットフォームにおける誤報や偽情報の拡散に対抗できるようにすることを目的とした「メディアと情報のリテラシー（MIL：Media, Information Literacy）」キャンペーンを実施している。

---

[59] IGF 京都 2023「Sharing “Existing Practices against Disinformation（EPaD）”」におけるプレゼンを元に記載。

[60] https://www.mcmc.gov.my/en/media/announcements/klik-dengan-bijak

[61] https://www.mcmc.gov.my/en/legal/acts

ファクトチェック関係では、ラップラーというニュースメディアが中心としてなり実施した＃FactsFirst PH という 2022 年の大統領選挙などの偽・誤情報対策のための連携の試みがあり、140 以上のニュースメディア、市民団体、ビジネス団体、企業、弁護士など法律家の団体や研究者が参加した。具体的には、ファクトチェック団体のファクトチェック記事を研究者が分析・記録などを行い、論文として公表していく取組であり、同取組を通じてファクトチェック記事を約 900 本、学術論文を 20 本近く出す等の成果を挙げた。

また、ラップラーは若者へのリーチを目的として漫画などのコンテンツを作成し、インフルエンサーとの連携等を行っている[62]。

法制度については、偽・誤情報に対する包括的な法律は存在しないが、2012 年サイバー犯罪法がオンライン上での詐欺やなりすましなどの犯罪行為に適用されている。

５．シンガポール

シンガポールでは、シンガポール情報通信省（MCI：Ministry of Communications and Information）において、国立図書館庁（NLB：National Library Board）と連携・協力し、リテラシー向上の取組として、偽情報をよりよく見分けられるようにするための支援施策である SURE Campaign[63]が 2013 年より実施されている。

法制度については、偽情報関連の法令として、オンライン虚偽情報及び情報操作防止法（Protection from Online Falsehoods and Manipulation Act 2019[64] ：POFMA）とオンライン安全法（Online Safety（Miscellaneous Amendments）Act 2022[65]）が存在し、前者においては、虚偽情報の拡散の禁止や虚偽情報の訂正指示等が規定されており、後者においては、オンライン通信サービス（Online Communication Services ：OCS）上にある悪質なコンテンツを規制し、オンライン安全性を高めることによりシンガポールのユーザーを保護することを目的として制定されている。

ファクトチェック関係では、2012 年に開設された政府の公式ウェブサイトである「Factually」において国内 SNS や伝統メディアの情報についてファクトチェックを行っている。

６．タイ

タイでは、デジタル経済社会省（Ministry of Digital Economy and Society 。以下「MDES」という。）と国家放送通信委員会（The National Broadcasting and

---

[62] IGF 京都 2023「Sharing “Existing Practices against Disinformation（EPaD）”」におけるプレゼンを元に記載。

[63] 「SURE」という名称はニュースの信頼性を評価するため個人が留意すべき4つの概念 Source、Understanding、Research、Evaluate の頭文字をとったもの。学校（SURE for School）、職場（SURE for Work）、社会（SURE for Life）の3つのアプローチにより、各対象者に合ったプログラムや教材の提供、講演会やワークショップの開催を実施。
https://sure.nlb.gov.sg/about-us/sure-campaign/

[64] https://sso.agc.gov.sg/Acts-Supp/18-2019

[65] https://sso.agc.gov.sg/Acts-Supp/38-2022/Published/20221221?DocDate=20221221&WholeDoc=1

Telecommunication Commission。以下「NBTC」という。）が偽・誤情報関係を担当している。

リテラシーの取組として、NBTC が実施しているデジタルリテラシーを高める活動である Media Information and Digital Literacy（MIDL）Campaign が実施されている。

ファクトチェック関係では、MDES に設置されているフェイクニュース対策センター（Anti Fake News Centre。以下「ANTC」という。）が SNS や国内伝統メディア等のチェックを通じた偽情報の流通防止の取組や偽情報の分析、特定を実施している。ANTC は法律執行機関とも連携し、偽情報を拡散する者を起訴するための情報提供を行っている。

行政機関やファクトチェック機関間での連携も進んでおり、NTBC において、シンガポールの情報通信メディア開発庁（Infocomm Media Development Authority ： IMDA）やマレーシアの MCMC と偽情報対策の対応について議論が進められている。

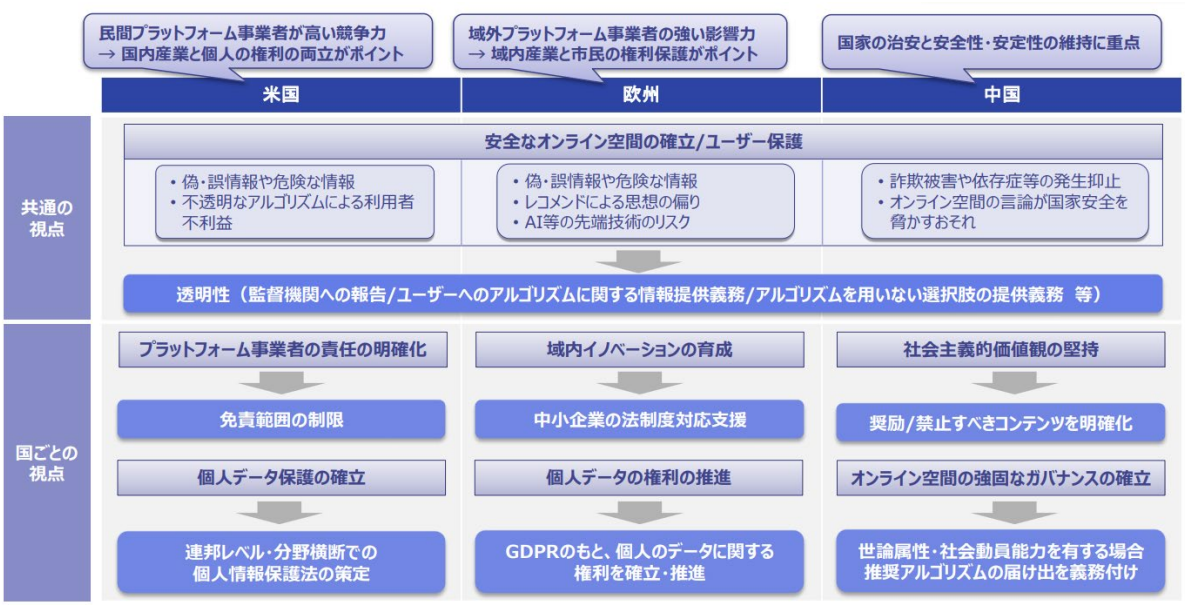
法制度については、偽情報関連の法令として、コンピュータ関連犯罪法（Computer-related Crime Act 2017 ）が存在し、同法第 14 条において、国民に損失を与える可能性の高い偽情報をコンピュータ上で拡散させた場合、5 年以内の懲役か 10 万バーツ以内の罰金、又はその両方が課され得ると規定しているほか、第 20 条において、デジタル経済社会大臣に対し、インターネット上で平和及び秩序又は国民の良俗に反する情報の流布を認めた場合には、同委員会の同意の下、当該情報の流通を停止し、又は当該データをコンピュータ・システムから削除するための法的措置を取る権限の付与について規定している。

Ⅷ　その他

１．レコメンデーションやアルゴリズムに関するルール整備

「レコメンデーションやアルゴリズムに関するルール整備の諸外国動向調査」[66]によると、米国、EU 及び中国の各国ともに、有害・危険な情報や詐欺等から利用者を保護するため、デジタル空間の安全性を向上させる課題意識が見られている。そして、それに対応するため、情報伝送 PF 事業者に対して、透明性要件（監督機関への報告、ユーザーへのアルゴリズムに関する情報提供義務、アルゴリズムを用いない選択肢の提供義務等）を定めるなどの対策が施行又は検討されており、各国の社会情勢や法整備の状況などを踏まえ、それぞれの視点から対策が取られている。

[66] 2024 年3月現在、株式会社みずほリサーチ＆テクノロジーズへ委託事業



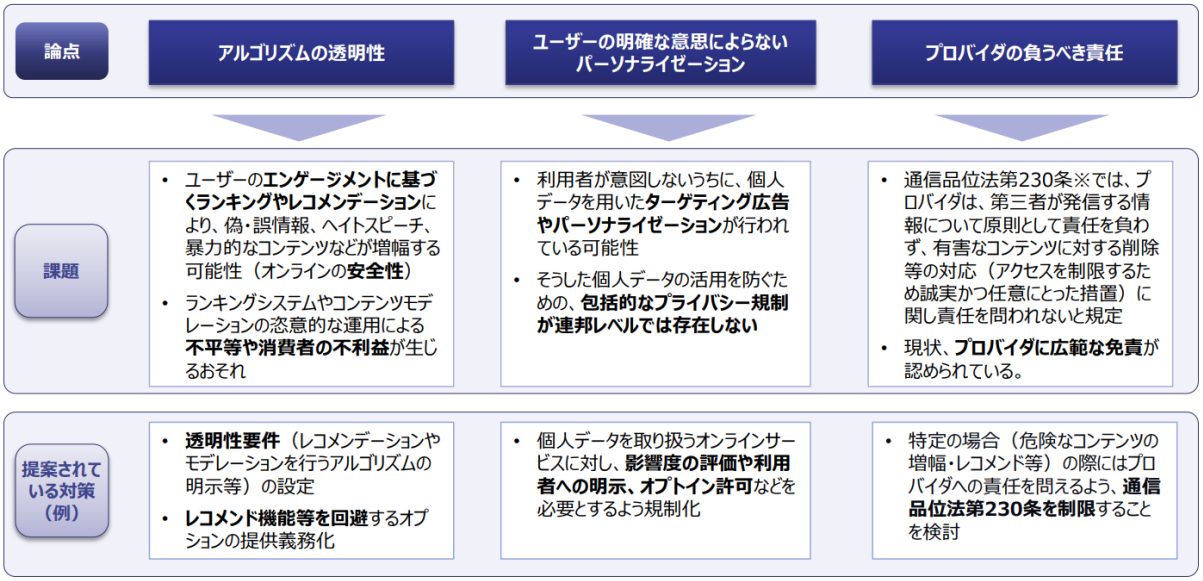
＜出典：本検討会資料 WG14－１＞

米国では、情報伝送 PF サービスで使用されているアルゴリズムを規制する全国的な法令等は、現時点では確認されていないが、各州で制定される州法や、合衆国議会に提案されている法案等では、①アルゴリズムの透明性確保、②個人データの利用に係る消費者の権利保護、③プロバイダの責任範囲拡大等を視野に入れた議論が行われている。

以上のうち①については、適格性判定アルゴリズムの公正性の調査を行うよう指示するとともに、事業者に対し、アルゴリズムの透明性要件を定めるアルゴリズム公正法（案）（Algorithmic Fairness Act of 2020）や、アルゴリズムをどのように使用しているのか、またこれらのアルゴリズムを動かすためにどのようなデータを収集しているのかをユーザーに説明することを義務付けるアルゴリズムの正しさ及びオンラインプラットフォーム透明性法（案）（Algorithmic Justice and Online Platform Transparency Act of 2021）が提案され、ユーザーのエンゲージメントに基づくランキングやレコメンデーションにより、偽・誤情報、ヘイトスピーチ、暴力的なコンテンツなどが増幅する可能性や、ランキングシステムやコンテンツモデレーションの恣意的な運用による不平等や消費者の不利益が生じるおそれに対し、透明性要件の設定や、レコメンド機能等を回避するオプションの提供義務化が検討されている。

②については、個人情報を使用する事業者に対し、利用者への通知や削除要求への対応等を求めるカリフォルニア州消費者プライバシー法（California Consumer Privacy Act （CCPA））が成立するなど、利用者が意図しないうちに、個人データを用いたターゲティング広告やパーソナライゼーションが行われている可能性や、個人データの活用を防ぐための包括的なプライバシー規制が連邦レベルでは存在しないことに対し、個人データを取り扱うオンラインサービスに対する影響度の評価や利用者への明示、オプトイン許可などを必要とするよう規制化が提案されている。

③プロバイダの責任範囲拡大については、現状、プロバイダに広範な免責が認められていることに対し、危険なコンテンツの増幅・レコメンド等の特定の場合にはプロバイダへの責任を問える通信品位法第 230 条を制限することが検討されている。

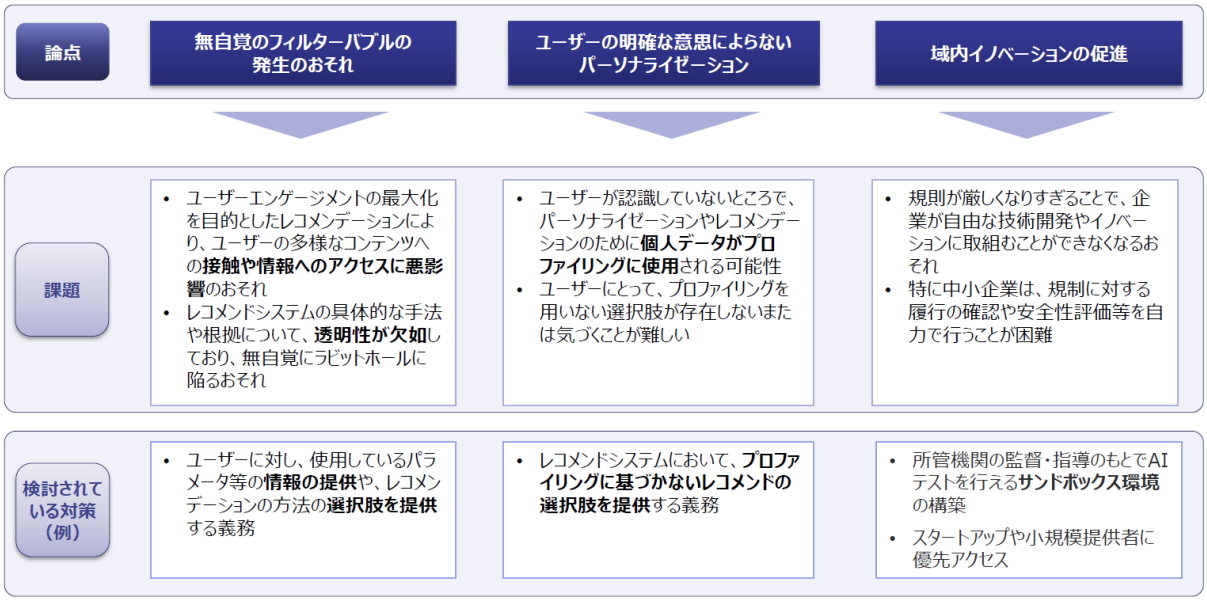


＜出典：本検討会資料 WG14－１＞

EU では、DSA において、欧州市民の保護のため、安全性や透明性の仕組みを情報伝送 PF 事業者に求めており、AI 法では、安全性確保と同時に、AI 活用やイノベーションの推進など、域内の産業振興も含めた政策が取られている。

特に、ユーザーエンゲージメントの最大化を目的としたレコメンデーションにより、ユーザーの多様なコンテンツへの接触や情報へのアクセスに悪影響のおそれがあることや、レコメンドシステムの具体的な手法や根拠について透明性が欠如していることに対し、使用しているパラメータ等をユーザーに情報提供することや、レコメンデーションの方法の選択肢を提供する義務が課されている。また、ユーザーが認識していないところで、パーソナライゼーションやレコメンデーションのために個人データがプロファイリングに使用され、ユーザーにプロファイリングを用いない選択肢が存在しない又は気づくことが難しいことに対し、プロファイリングに基づかないレコメンドの選択肢を提供する義務が課されている。

他方、規制が厳しくなりすぎることで、企業が自由な技術開発やイノベーションに取組むことができず、特に中小企業で規制に対する履行の確認や安全性評価等を自力で行うことが困難になるおそれに対し、所管機関の監督・指導のもとで AI テストを行えるサンドボックス環境の構築し、スタートアップや小規模提供者に優先アクセスさせることが課されている。



＜出典：本検討会資料 WG14－１＞

中国では、「ネットワーク情報コンテンツ環境ガバナンス規定」（Regulation on the Governance of the Online Information Content Ecosystem）[67]により、デジタル空間に携わる各主体に対する規定を定め、安全かつ国家の安定・社会主義的価値観の醸成に有益なガバナンスの確立が目指されている。

「インターネット情報サービスのアルゴリズム・レコメンデーション管理規定」（Internet Information Service Algorithmic Recommendation Management：IISARM）では、ユーザーに対しアルゴリズムに関する情報やアルゴリズムを用いたレコメンデーションを使用しないオプションを提供することなど、ユーザー保護に焦点をあてた規定が策定されている。アルゴリズムの適用が普及する中で、アルゴリズムによる差別や依存症の誘発など未成年者のユーザーに心身の健康を害するコンテンツが表示されたり、ITサービスに不慣れな高齢者が詐欺等のトラブルに巻き込まれたりするケースが増加していることに対し、ユーザーがアルゴリズムについて知る権利やアルゴリズムを選択する権利の保護要件を規定されている。

[67] 2019/12/20 公表、2020/3/1 施行

＜出典：本検討会資料 WG14－１＞

２．その他の取組

（１）偽・誤情報検知等を目的とする研究開発

諸外国（一部日本含む）では、偽・誤情報検知等を目的に ICT ツールの研究開発が行われている。その目的を分類すると５つあり、(a)コンテンツの検証、(b)リソースの信頼性検証、(c)フェイクニュース検出、(d)コンテンツ検証作業支援、(e)その他に分類される[68]。

No.	ツール名	目的	内容
1	SYNTHETIQ VISION（日本）	・コンテンツ検証（動画、画像）	・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	・コンテンツ検証（動画、画像）	・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	・コンテンツ検証（動画、画像）	・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）	・コンテンツ検証（動画、画像）	・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	・コンテンツ検証（動画）	・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。 ・インテルとビ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	・コンテンツ検証（動画）	・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	・コンテンツ検証（画像）	・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	・コンテンツ検証（画像）	・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。

＜出典：プラットフォームサービスに関する研究会（第 45 回）資料３＞

[68] https://www.soumu.go.jp/main_content/000882501.pdf

No.	ツール名	目的	内容
9	Sphere	• コンテンツ検証（テキスト）	• 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	• 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）	• 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	• リソースの信頼性検証	• 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）	• リソースの信頼性検証	• 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	• リソースの信頼性検証（作成者）	• 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	• 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	• 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	• フェイクニュース検出	• 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」	• フェイクニュース検出	• 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。

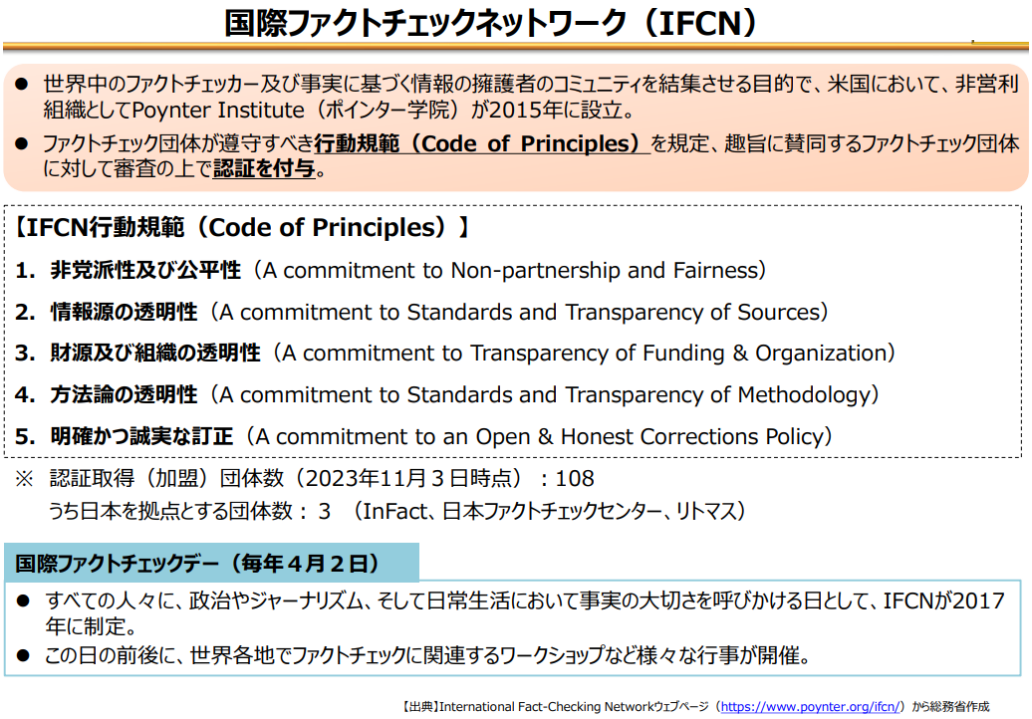
＜出典：プラットフォームサービスに関する研究会（第 45 回）資料３＞

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）	• コンテンツの検証作業支援	• 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（FII）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	• コンテンツの検証作業支援 • コンテンツ検証（画像）	• 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	• コンテンツの検証作業支援 • フェイクニュース検出	• 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」	• コンテンツの検証作業支援 • コンテンツ検証（テキスト）	• 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	• ライブファクトチェック結果共有	• 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）	• フェイクニュースへのアラート	• 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。

＜出典：プラットフォームサービスに関する研究会（第 45 回）資料３＞

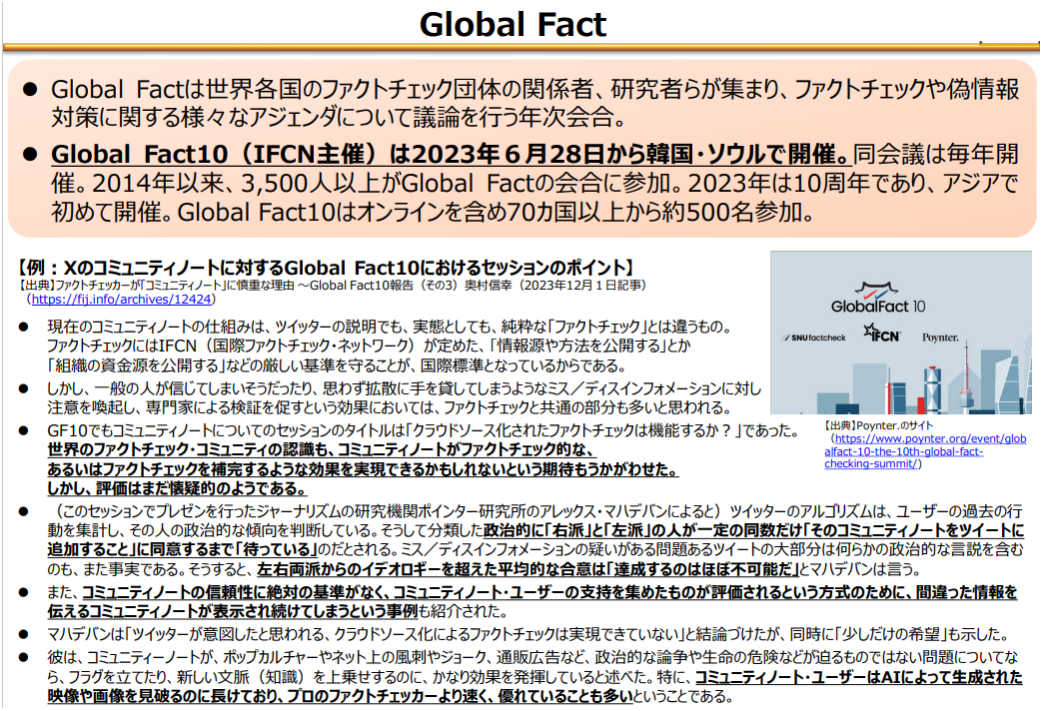
（２）国際的なファクトチェックの推進

IFCN はファクトチェック活動の原則についての綱領「Code of Principles」を定めており、「非党派制と公平性」「情報源の基準と透明性」「資金源と組織の透明性」「検証方法の基準と透明性」「オープンで誠実な訂正方針」という５つの原則がある。当該原則は 2016 年９月に制定され、国際標準のファクトチェックの原則として認知されている[69]。IFCN は 2017 年から当該綱領に基づいて加盟団体の審査を開始し、2022 年４月時点で、IFCN には 100 以上の団体が加盟し、20 以上の団体が更新審査中となっている[70]。



＜出典：本検討会資料５－１－２＞

国際的なファクトチェックに関する会合については、Global Fact や APAC Trusted Media Summit がある。Global Fact は世界各国のファクトチェック団体の関係者、研究者らが集まり、ファクトチェックや偽情報対策に関する様々なアジェンダについて議論を行う年次会合である。Global Fact10 は 2023 年で 10 周年であり、2023 年６月に初めてアジア（韓国・ソウル）で開催された。Global Fact10 はオンラインを含め 70 か国以上から約 500 名が参加した。



＜出典：本検討会資料５－１－２＞

APAC Trusted Media Summit は、2018 年より開催されている、アジア太平洋地域全体で誤情報と戦うジャーナリスト、ファクトチェッカー、教育関係者、活動家、政策立案者が集まる年次会議である。会議目的は、関係者間でネットワークを構築し、ベストプラクティスを共有することにある。Google News Initiative が主催であり、共同主催者は ANNIE、Aliansi Jurnalis Independen、DataLEADS、Taiwan FactCheck Center の４団体がある。

＜出典：本検討会資料５－１－２＞

（３）個人レベルの介入研究

近年の重要な取り組みの一つとして、個人レベルの介入に関するツールボックスが公開された。これは、世界中の最新の研究を含む 81 の科学論文で取り上げられた介入に関する最新の説明から構成され、主要な介入９種類の概念的な概要と、介入を裏付ける実証的証拠の要約が提供されている。研究者だけでなく、政策立案者、教育者、一般の人々がアクセスでき、誤情報問題のさまざまな側面に対処するためのオンラインツールボックスが提供されている[71]。

（４）The Debunking Handbook

誤情報対策として、学術的知見をもとにした実践的提言がなされた「The Debunking Handbook 2020」では、誤情報に関する心理的特徴、それらにもとづく訂正の効果を挙げるための留意点が整理されている。市民、政策立案者、ジャーナリスト、その他実業家などを対象に 2015 年以降に誤情報に関する心理学分野における学術的実績のある研究者 22 名を選定し、心理的特徴と対策案の根拠となるエビデンスを集め、それぞれに対し、

[71] Kozyreva, A., Lorenz-Spreen, P., Herzog, S.M. et al. Toolbox of individual-level interventions against online misinformation. Nat Hum Behav (2024). https://doi.org/10.1038/s41562-024-01881-0

「エビデンスの強度」と「デバンクにおける重要性」を評定し、評定結果を分析するなど、様々な研究者等による連携・協力が実施されている。

（５）ブラジルにおける取組

ブラジル大統領選挙では、５つの主要メディア[72]により、非営利エンジニア集団Meedan[73]のネットワークを活用しながら、ファクトチェックが推進されている。

Meedan は、ユーザーがある質問を Meedan 上のチャットボットに投稿した際、既にファクトチェック結果がネットワーク上にある場合、質問者にその結果を返してくれるシステムであるところ、この Meedan が提供した共同フィードは、仮に質問に対するファクトチェック結果がない場合、５つの主要メディアに対して、ファクトチェックを促す仕組みになっており、33 万件を超える質問が処理されている。

（６）カナダにおける取組

カナダでは、メディアリテラシーウィーク [74]と呼ばれるデジタルメディアリテラシーを向上させる毎年恒例のキャンペーンが実施されている。この期間はカナダ国内の学校、図書館、博物館、コミュニティグループなどによるイベントやアクティビティが開催され、2022 年のメディアリテラシーウィークではオンタリオ州オタワの公共図書館で図書館主催の司書が講師を務めるワークショップが実施されている。

メディアリテラシーウィーク 2024 は、2024 年 10 月 21 日から 10 月 25 日までの開催が予定されている。

Ⅸ　小括

デジタル空間における情報流通の健全性を巡るリスク・問題については、情報伝送 PF 事業者等が同様のアーキテクチャ等によりグローバルにサービスを展開していること等から、我が国特有の課題ではなく、諸外国等が共通して抱えている課題になっている。

課題への対応状況については、各国で区々であり、例えば、米国においては、伝統的に情報伝送 PF 事業者に対して広範な免責が与えられている一方、情報伝送 PF 事業者の取組への規制に関する議論が近年進みつつある州もある。また、民間による自主的な取組が進んでいる国もあり、オーストラリアやニュージーランドでは、情報伝送 PF 事業者が民間主導の行動規範に参画し、ステークホルダーにおける連携・協力が進められている。

---

[72] Agencia Lupa,Aos Fatos,Projeto Comprova,Estadao Verfaca,Unlverso Online

[73] https://meedan.com/

[74] https://mediasmarts.ca/media-literacy-week

そのような中で、EU は共同規制として、DSA において行動規範の策定と参加を奨励しており、2022 年版行動規範には、オンラインプラットフォーム事業者、広告関連事業者・団体、ファクトチェック団体・関連サービス事業者、その他 NGO を含めた市民・業界団体・関連事業者など 43 団体が署名・参加するなど、様々なステークホルダーが参画している。

加えて、各国では法制度以外の面でも多様な連携・協力関係が生まれている。例えば、欧州では、様々なステークホルダーによる偽・誤情報への取組として、ニュースメディア、オンラインプラットフォーム、エンドユーザーのみならず、広告主、ファクトチェッカー、学術研究者、市民団体といったステークホルダーも、それぞれファクトチェックスキルやツールの提供・共有、偽情報についての研究成果の共有、メディアの監視やメディアリテラシー教育等が行われている。また、米国では大学機関と伝統メディアが連携・協力する取組もある。

また、ASEAN 諸国においては、ファクトチェックに関して、マルチステークホルダーによる連携・協力（例えば、インドネシアやフィリピン）のみならず、各国間での連携（例えば、インドネシアとマレーシア間の協力）が進むほか、リテラシー向上に関するキャンペーン等が実施されている。

さらに、国連においても行動規範を作成する取組が進められており、インターネット・ガバナンス・フォーラム（IGF）ではマルチステークホルダーによる取組が進められる等、国際機関等における連携・協力も進んでいる。