

# 構成員からのこれまでの主なご意見（第1回～第5回会合）

2024年1月25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

- ※ 本資料は、事務局において、現時点で便宜上整理したものであり、今後の検討等を踏まえ、記載内容や構成等が変更される可能性がある。
- ※ 構成員名に付した括弧付きの数字は、当該ご意見が出された会合の回数を示す。また、黄色ハイライト部分は、前回資料（資料5-2-1）からの追記等による更新を示す。  
例：(1)⇒第1回会合におけるご意見、(2)⇒第2回会合におけるご意見、…

## 目次

① デジタル空間における情報流通の「健全性」の考え方 .....	3
1. 総論：健全性確保に向けた基本理念の在り方 .....	3
2. 表現の自由・知る権利 .....	5
3. 多様性・包摂性 .....	6
4. 法の支配・民主主義 .....	6
5. 公平性・公正性 .....	7
6. 真正性・信頼性 .....	8
7. 安心・安全 .....	11
8. オープン・透明性・アカウントビリティ .....	12
9. プライバシー保護 .....	14
10. グローバル・国際性 .....	15
② デジタル空間における情報流通に関する現在の課題 .....	15
1. 情報の発信・伝送・受信の各段階に共通の課題 .....	15
◆ デジタル空間の情報流通に関するガバナンスの在り方 .....	15
◆ ステークホルダー同士の連携・協力の在り方 .....	20
◆ 生成AI・ディープフェイク技術の進展に伴うリスクへの対応の在り方 .....	24
◆ 技術・研究開発の在り方 .....	29
◆ 研究機関等との連携・協力の在り方 .....	34
◆ 緊急事態（災害、サイバー攻撃など）への対応の在り方 .....	35
◆ 国際連携・協力の在り方 .....	37
2. 主に情報発信段階の課題 .....	38
◆ 発信情報の信頼性を得るためのコスト増への対応の在り方 .....	38
◆ 発信力強化のためのガバナンスの在り方 .....	43
◆ 持続可能なファクトチェック推進のための仕組みの在り方 .....	44
3. 主に情報伝送段階の課題 .....	49
◆ アテンション・エコノミーが引き起こす課題（フィルターバブル、エコーチェンバーを含む）への対応の在り方 .....	49

◆ 偽・誤情報の拡散への対応の在り方.....	55
◆ 広告を巡る課題への対応の在り方.....	60
◆ 事業者の取組（コンテンツモデレーションの方針・体制・実施状況、プロミネンスなど）の透明性・アカウントビリティ確保の在り方.....	62
◆ 事業者の取組（レコメンド、データ取扱い、広告収入の発信者への分配など）の透明性・アカウントビリティ確保の在り方.....	66
4. 主に情報受信段階の課題.....	68
◆ 認知的・社会的バイアスを前提としたリテラシー向上策の在り方.....	68
◆ 発信・拡散主体となり得る受信者側のガバナンス・リテラシー向上策の在り方.....	77
5. その他課題に関するご意見.....	78
③ 各ステークホルダーが果たすべき役割・責務.....	79
1. 総論：各ステークホルダーに期待される役割・責務の在り方.....	79
2. 政府の役割・責務.....	81
3. プラットフォーム事業者の役割・責務.....	85
4. 生成A I・メタバース関連事業者の役割・責務.....	89
5. 通信・放送事業者の役割・責務.....	90
6. 広告主・広告関連事業者の役割・責務.....	92
7. 教育・研究機関の役割・責務.....	92
8. 専門家の役割・責務.....	93
9. 利用者の役割・責務.....	94

## ① デジタル空間における情報流通の「健全性」の考え方

### 1. 総論：健全性確保に向けた基本理念の在り方

- 総務省のこの会の全体的なテーマが、デジタル空間における情報流通の健全性ということになっているが、これを Principles や Practices というところに当てはめたときに、そもそもどういう状態が健全かの議論とある程度の合意というのをしていくのが Principles、原則の議論に相当する。また、健全な状態に持っていくというときに、健全性を誰がどのように担保しているか、また、どのように確認するかということの方法論（Practice）の議論も同時に必要。健全性、偽情報・誤情報に関しても、何をもって偽情報と誰が判断するのかというのは、おそらく非常に難しいグレーゾーンを含んでいるようなものもあるかと思うが、このような点に関し、広く AI ガバナンスという観点も含め、実際にはこの（Principle と Practice の）両方が行き来をしたりすることが重要。【江間構成員(3)】
- 最近 AI ガバナンスでは動きがいろいろ速く、いろいろなところで議論が起きているが、この Principle について、G7 広島 AI プロセスが走っており、OECD AI 原則が最近いろいろと定義も新しく改進したものも出ておりましたが、そこで出しているものや、ユネスコや、国連がハイレベルアドバイザリーボードというのを開始した。【江間構成員(3)】
- 原則について、どのように各国や様々な業界団体で議論していくかとなったときに、ガイドライン、ノンバイディングなものや、又は、規制があるという意味でのレギュレーション、国による法令というところでバイディングなものというのが、2 つ入り交じっているのが現状。日本は総務省と経産省で新事業者ガイドラインが議論をされている。この間出たアメリカの大統領令も基本的にはノンバイディングであるが、事業者とかいろいろなところに、又は、関係省庁にいろいろ考えてくださいという要望を出している。一方、バイディングで一つとても着目をされているところが欧州の AI 規則案で、あまり注目はされていないが、欧州評議会の AI 条約がある。ある程度原則を落とし込んでガイドラインに持っていくところがある。また、Voluntary commitments というところで、様々な組織に対し自主的にガイドラインをつくって、安全性や公平性に対応していこうということを促すような層が新たに結構出てきた。自主ガイドラインを自らつくっていったり、業界で考えていくということが非常に大事になってきている。【江間構成員(3)】
- もっとさらに具体的、テクニカルな点になると、国際的なスタンダード、国際標準や、国内の標準機関が定めているもので、アメリカ NIST の Risk management framework があったり、ISO、CES/CENELEC、IEEE SA や ITU が、技術標準を今いろいろなレイヤーで議論されている。よりプラクティスレベル、事例ベースみたいなところで、GPAI では、実際にいろいろなワーキンググループが動いて、いろいろなツールやアセスメントについて考えていくプロジェクトベースで動いている。また、AI 監査や認証みたいなところは、標準化などと足並みをそろえながら、実際にどのような観点で外部監査、内部監査、

又は第三者認証みたいなことをやっていくのかということを考えていく試みがある。【江間構成員(3)】

- まず、ガイドラインや Voluntary commitments に関して具体的にどのようなお話あるのか、実際にはどういう人たちと一緒に、今、何をテーマとして、どのような領域で議論していくのかみたいなところが、大きくテーマとして考えるべきことかと考えている。【江間構成員(3)】
- AI ガバナンスの Principle を考えていくとき、パーティカルなレギュレーションをつくっていくのか、ホリゾンタルな、いろいろな領域でカバーするようなものをつくっていくのかということ、どちらかというとなパーティカルな、今まで医療、交通、金融など既に規制があったり、あるいは標準があったり、議論があったりする中で、それぞれと二重三重の規制にならないようにきちんと整理をした上で、かつその領域固有の課題を考慮に入れながら、規制だったり、ガイドライン、あるいはレギュレーションをつくっていくということが重要。【江間構成員(3)】
- デジタル空間における情報流通の健全性ということ考えたときに、様々な層において AI ガバナンスの議論と国内外の議論の協調を進めていくことが非常に重要。もともと AI ガバナンスでは Principles to practices と言われていた。まず原則をつくって、これを Practice に落とし込んでいく。しかし、動きの速い技術では、G7 広島 AI プロセスがもう Code of conduct と Guiding principle を同時に出したことからもお分かりのとおり、Principles and practices になってきている。ほぼほぼその両輪を同時に議論していくことが重要。【江間構成員(3)】
- AI ガバナンスに関しては、OECD が色々出ているいろいろな原則とかツールみたいなものを整理して、相互運用可能なものにしていこうという考え方に基づき、いろいろなところが出している原則とか議論とかを、例えば一覧表にしてどういうふうに関連し合っているのか分析。一方で、ハードローとして出てきている AI 法案の影響も非常に大きい。欧州法議会が今まさに起草交渉をしていて、来年の5月には妥結しようと考えている AI 条約に関しても、同じヨーロッパということで、AI 法案はもうできているので、それと反しないというか、Conflict が起きないようにしようということが議論がされているという話はある。このように、Principle を具体的に落とし込んでいくというときには、技術レイヤーの標準化をどういうふうに関連させていくかということが非常に重要。いろいろなところが国単位で出してくるようなものに関しても、例えば日本の新事業者ガイドラインにおいても様々なところでの議論というのを、お互いにウォッチしたりあるいは人同士のインフォーマな情報共有みたいなものもありながら、ネゴシエーションしながらつくっているというところが実態。【江間構成員(3)】
- いろいろな分野に AI が導入されていくとなったときに、基本となるのは既に存在している産業別での議論。安全性に関して非常に重要度が高いようなところに関しては、もちろん既に法令があるところはあるので、それに基づいて議

- 論をしていくという形になる。【江間構成員(3)】
- 様々な基本理念同士がトレードオフ関係にある時に、どのような対応を各ステークホルダーが取ればよいか（あるいは議論すればよいか）の方針が定まっていること。【江間構成員(1)】
  - 健全な情報流通という 이슈は、情報通信政策、メディア政策、消費者保護、知的財産権保護、経済安全保障などを含めて様々な方領域、政策領域が交わる領域。例えばプラットフォームや大規模 AI のような議論は、ほかにも競争政策、情報通信分野の公正な競争という側面も切って切り離せない。本検討会における検討の主題というところからは必ずしも中心ではなく、周辺にあるものであっても、そうした様々な法政策の領域との協調、相互作用というものを念頭に置きながら議論をしていけると良い。【生貝構成員(1)】
  - 検討会で言われている健全性確保という、その健全性というのが何を指すのかというのがはっきりしないといけないと思うので、そこに向けてアメリカ、EU の議論を参考にしながらデジタル空間をどのように嚮導していくのか、そのための概念はまずしっかり立てる必要がある。【水谷構成員(4)】

## 2. 表現の自由・知る権利

- 誰もが自由に情報発信できるという民主主義の実践機会を増やして、その質の向上（信頼できる取材手法、飽きさせない編集、引きつける語り、最新テクノロジー活用など）を図ることが、結果的に情報流通の健全性確保につながるのではないかと。【脇浜構成員(1)】
- 偽情報とされるものの中には、意図的に作られた偽情報だけでなく、悪意はないが間違っている情報や、事実関係は間違っていないが異なる文脈で使われることで誤った印象を植え付けるものなどが含まれており、その境界はしばしば不明確であり、何が偽情報で何がそうでないかを判別することは容易ではない。そのような状況で、①何を削除するか、②どのくらいの数を削除するか、といったことについて法制度を作ったり、統一的な基準を設けたりすることは必ずしも適当ではない。プラットフォーム事業者に過度の削除圧力をかけることは、当該プラットフォームに情報を投稿する利用者の表現の自由とプラットフォーム事業者自身の表現の自由を共に脅かすことにつながる。【森構成員(1)】
- 今後の対応にあたっての基本的な考え方として、情報受領者（国民）の「自由に、さまざまな意見、知識、情報に接し、これを摂取する機会」（最大判昭和 58 年 6 月 22 日民集 37 卷 5 号 793 頁）、特に信頼性の高い情報へのアクセス保障が重要。【水谷構成員(1)】
- 消費者において、表現の自由、知る権利についての理解ということも、正しい情報の見極め、情報発信の責任ということも同時に身に付ける必要。【増田構成員(1)】
- コンテンツ・モデレーションの自由という言葉を使ってしまったが、少し詳細に言うとアメリカでは DPF 事業者の行うコンテンツ・モデレーションを、私自身はイコールで結ぶのはあまり賛成ではないが、報道機関の編集の自由、

編集判断の自由と重ねて議論している論稿や下級審の判例などがあったりする。また逆に、それに対しては、プラットフォームをコモンキャリアと位置づけて、だから編集判断はしてはならないという議論をしている下級審判断もあって、今、ここが対立している。研究者の中でもコンテンツ・モデレーションを表現の自由の観点からどう位置付けるか、様々に議論されているので、今後の推移を見守らなければいけないと思う。ただ少なくとも修正1条がDPF事業者がコンテンツをモデレーションする行為に関係してきそうだというのはアメリカの現状ではある。【水谷構成員(4)】

### 3. 多様性・包摂性

- 現在の「構造」下では、アテンションを奪うための強力なレコメンダーシステム（UX）によって、他律的に情報を「偏食」させられているような状況が拡大（さまざまな情報を主体的に摂取できるという「知る自由」の実質的な制約）。この「偏食」によって、偽情報等に対する「免疫」も低下しているとすれば、「情報の健康」というコンセプトを打ち出し、情報摂取行動の大切さを意識できるようリテラシー教育を行うことが重要（意識変化の重要性）。【山本構成員(1)】
- DE&I (Diversity, Equity and Inclusion) の立場から、特定の状況において不利益を被る人がいないかどうかを検討すること（状況によっては情報にアクセスができない人たちが出てくる可能性があるなど、前提条件として置かれている事項を想定すること。例えばスマートフォンを持っていることが想定されていないか、日本語が読めない場合があることを想定しているか等）。【江間構成員(1)】

### 4. 法の支配・民主主義

- 効果を急ぐと、とかく私を含め技術に感心がある方はできるからやろうというところに行ってしまうところがある。しかしながら、今起きている問題は非常に複雑で、立場によって見解が大きく異なる。その中で、世の中で起きている問題をまずきちんと定義し、分解し、対応していくのかという冷静さが必要。直近でも、ディープフェイクを用いた偽情報の流通の問題提起が行われているところだが、こういった社会の安全・民主主義を脅かすというものに対処していくためにも理念や手続が可能な限り民主的なもの、ないしは立法がどうしても必要ということであれば、立法事実を明確化する、特定していくことが必要。【クロサカ構成員(1)】
- デジタル空間の情報流通の健全性の確保は、現代社会における民主主義の維持や自由な言論の場の確保に不可欠。【森構成員(1)】
- この検討の基本理念として、ヨーロッパの憲法学者を中心にデジタル立憲主義というような概念が論文等で議論をされているところであり、既存の憲法学が培ってきたある種の価値などをデジタル空間の技術等々に合わせて翻訳し、それを目標に制度設計を行っていくということと理解。例えば、EUのデジタルサ

ービス法や、民主主義行動計画といった環境形成におけるある種の基本理念になっている。この点、我が国では、国民の、又は、情報受領者の知る自由とか、知る権利という概念にポイントが置かれる。この検討会の主軸でもあるが、やはり信頼性の高い情報へどうやってこのジャングルの中でアクセスを保障していくかということが重要。【水谷構成員(1)】

- 朝日新聞で北九州市立大の山本健人先生がデジタル立憲主義というヨーロッパの議論を紹介されていたが、そこでの議論が参考になるのではないか。日本では最高裁が自由に様々な意見、知識、情報に接し、これを接種する機会というようなことを指摘しているの、これを確保することが一つポイントになると思う。【水谷構成員(4)】
- (欧州では) 12月12日にディフェンス・オブ・デモクラシーパッケージというのが出ており、欧州民主主義行動計画というのに基づいてDSAなども議論されている部分もあったかと思うが、直近でそういう比較的大きな補完の動きなどがあったりしたとも受け止めている。関連して2022年には実践規範の強化がされていて、その中でコンテンツ・モデレーションなどが強化されている部分がある。【落合構成員(4)】
- 欧州民主主義行動計画等との関わりという意味でいうと、まず一つは、特に来年、欧州議会の選挙があり、その防衛パッケージといったようなことで、特に第三国からのリプレゼンティティブや介入に非常に力点を改めて置いた行動を加盟国にも促す部分が大きいのかなと思う。【生貝構成員(4)】
- DSAの背景にある基本的価値について、最近の論文で、『ブリュッセル効果』を書いたアニュ・ブラッドフォードが、EUのDSAとか、AIAを、EUのデジタル憲法だと評しているのを見た。【水谷構成員(4)】
- 基本的価値という意味だと、デジタル立憲主義は、憲法的価値というものをデジタル空間でも私的なプラットフォームも含めて実現していくかというところは、デジタルサービス法の中でも非常に出てきているのを感じている。わけでも、非常に興味深いのが、利用規約に基づくコンテンツ・モデレーションポリシーの実施に当たっては基本権に配慮しなければならないといったような、大変真正面のことなども含まれていたりして、まさしく憲法的価値というものとデジタル空間を改めて考えるきっかけにもなるのかなと思う。【生貝構成員(4)】

## 5. 公平性・公正性

- AIガバナンスに関し、特に、Fairness、Accountability、Transparencyみたいなことは、2~3年前から重要なテーマになっている。【江間構成員(3)】
- アルゴリズムのバイアスといったような意味だと、まさにこれが多分システムリスクとか、そういうものの軽減措置の中にも今後、恐らく様々な形で具体的に組み込まれてくるのだと思う。システムリスクの対象は非常に広く、基本権だとかジェンダーだとか、そういったような問題も含めてということなので、まさにちゃんとそれに合わせた形でデータのバイアス、アルゴリズム

のバイアスみたいなのもちゃんと調整していかないといけないということが、今後その条項の具体化の中でいろいろ出てくるんだろうと思う。【生貝構成員(4)】

- コンテンツ・カルテルと競争法の議論について、競争法の話とどこまではっきりリンクしているかは私も追い切れていないが、ただ、ソーシャルメディアの企業が今、寡占状態にある、それによりカルテル的な存在になりつつあるというところは確実に議論の背景にはある。【水谷構成員(4)】
- 競争というファクターをどう考えるかということも一つの大きな関連するイシューとして考えていくことの価値というのは大きいのだろう。それは例えば伝送レイヤーといったようなものも非常にデジタル経済独特の競争のメカニズムといったようなものが働く、それが例えばプラットフォームエコノミーといったような状況を現に生み出している。さらにこれからそれがデータの集積ということによって、またこの競争の状況つまり情報流通を媒介するまさに基盤の競争状況といったようなものが非常に大きな変革のさなかにあるといたしましたときに、まさしく競争法でいうところの市場と言論の自由市場でいうところの市場といったようなもの、この2つは違ったものでありながらも、その関係性をどう考えていくかといったようなことがやはりこの中でも時々意識される必要がある。【生貝構成員(5)】

## 6. 真正性・信頼性

- 『ジャーナリズムの原則』は欧米のメディアの中では非常に重要な側面。『ジャーナリズムの原則』がどうして優れているかというと、民主主義とニュースの関係を非常に鮮やかに論理立てて説明をしているからである。「ジャーナリズムの一番の目的は、市民が自由を守り、自治を行うために必要な情報を提供することである」が、これは、この本ができたアメリカでもみんな分かっているだろうとって長らく言語化されてこなかった。2001年に初版が出たが、1997年頃から、実はプロジェクトが始まっており、元ニューヨーク・タイムズのビル・コヴァッチら2人が、ジャーナリストとメディア研究者などを招集し、若いジャーナリストたちが自分たちが何のために仕事をしているのかという目的意識を失っているのを、それを何とかしたいと言ってディスカッションを始めたところから始まっている。300人以上のジャーナリストに対し、あなたがジャーナリストとして一番大切にしているものは何か、ジャーナリストとしての責任は何だと思うかというような姿勢や価値を問うような質問をたくさん投げかけ、それらの回答を集めて編み上げたのがこれらの言葉になっている。【奥村構成員(3)】
- 例えば、日本の議論と結構違うのは、中立という言葉がこの原則の中で出てこない。中立は理念としてはあるが、現実としては実践できないという考え方だからである。それで終わりではなく、ニュースの消費者である一般市民とのフォーラムを十数回開き、これがあなた方が必要としているものか、ニュースメディアに求めるものかということは何度も問いかけたところ、この本には、「ニ



ユースに携わる者が知っているべきで一般市民が当然期待していいもの」として、10の原則は精緻に編み出され、2007年、2014年、2021年に版を改めて出版されているが、著者たちはその度にこのデジタル、ミスインフォメーションの世界で、この原則や表現がまだ通用するのかということについて、厳しく問いかけを行って、残ってきたのがこの10の表現。①Journalism's first obligation is to the truth. (ジャーナリズムの第一の責務は真実を求めることである。)、②Its first loyalty is to citizens. (まず、市民に対し忠実であるべきである。)、③Its Essence is a discipline of verification. (その本質とは、検証を実行できる能力(規律)である。)、④Its practitioners must maintain an independence from those they cover. (ジャーナリズムに携わる者は、取材対象からの独立を維持しなくてはならない。)、⑤It must serve as a monitor of power. (権力を監視する機能を果たさなくてはならない。)、⑥It must provide a forum for public criticism and compromise. (公共の問題に関する批判や、歩み寄りを行う討論の場を提供しなくてはならない。)、⑦It must strive to make the significant interesting and relevant. (重大な出来事を興味深く、社会的に意味のあるものにするよう務めなければならない。)、⑧It must present the news in a way that is comprehensive and proportional. (ニュースをわかりやすく、偏らないものとして示さなければならない。)、⑨Its practitioners have an obligation to exercise their personal conscience. (それに携わる者は、自らの良心を実践する義務がある。)、⑩Citizens have rights and responsibilities when it comes to the news as well - even more so as they become producers and editors themselves. (市民の側も、ニュースをよりよいものにしていくことについて、権利と責任がある - 彼らも記者や編集者になれるようになった現在はなおさらである。)

【奥村構成員(3)】

- 2014年にはソーシャルメディアの発達に伴って10項目めが追加された。ソーシャルメディアでメディアにフラットに働きかけられるようになったニュースの消費者も一定の責任があるだろうという考え方。少なくともアメリカやイギリスなど欧米の先進国では、明文化はされていなくてもこのような大原則が社会で承認され、ニュースメディアが存立しているという構造と理解している。そうすると、メディアの倫理規範というようなものは、どうやってその価値を守るかということに焦点が移る。要するにHowの議論。そして、それを公開して、読者、消費者と共有して品質保証をしていくという考え方で、程度の差こそあれ、クオリティメディアはほとんどやっていると言っても過言ではない。

【奥村構成員(3)】

- 米国のポインター研究所というジャーナリズムの研究所がセントピーターズバーグというフロリダ州にあるが、そこに附属している機関として、国際ファクトチェックネットワークがあり、ヨーロッパよりは、アメリカ大陸とイギリス、それからアジアを中心にしてファクトチェックの振興などを行っている。こちらにシグナトリーという認証ファクトチェック団体があるが、その資

格を与えるプロセスについて、ファクトチェックはとにかく信用されなければ成立しないため、認証は定評のある大手メディアでも時間をかけて行われる。しかし、世界各国の国内事情も違うため、大きな5つの原則を提示している。これらに合致した活動をしているかどうかということ、外部の人が判断する。具体的には、①「非党派制と公平性」として、どの党派にも偏らないというのが1番目、②「情報源の基準と透明性」として、どのように情報収集を行っているか、ファクトチェックにどのようなソースを使ったのかということが必ず明示されているかを、記事1本1本審査し、③「資金源と組織の透明性」として、Conflict of interestを生じないために、資金減をきっちり明示をして、どこから幾らもらってどう使っているかということ公開しているかどうか、④「検証方法の基準と透明性」として、どのような手順をとったかについて、その後どのように処理をしたか、ファクトチェックにはレーティングというのがあるが、白か黒かとすっぱり判断できるというより、グレーのものが非常に多いが、どのような形でそのグレーという判定を出したのかという根拠を問われ、⑤「オープンで誠実な訂正方針」として、真実やファクトは日々変わる中、新しい技術がどんどん明らかになり、情勢が変わってくると、ファクトチェッカーはそれにスピーディーに対応し、ここはこのように間違っていたのでこのように訂正しますというようなことを言わなければならないという原則となっている。【奥村構成員(3)】

- インターネットの社会になってくる中で、いろいろな情報源であったり、どういう過程でこの情報がつくり出されているかが相対化されてきたり、個人であっても信頼性のある情報を発信できるような方も出てきたりするであろうという中では、信頼性担保に関する取組を行っていて、それをこういう検証をしてより向上させているという、ある種のディスクロージャーとそれに対する説明が重要になってくる。そういうものをしっかりやられているものは比較的信頼性が高いというか、受け手側の評価もされやすくなる。信頼性の高いものは何なのかといったときに、適切に情報公開を行って一定の手続を定めている人たちが信頼性の高い方々なんだというふうに見ていくのがよりよい。【落合構成員(3)】
- 健全な情報空間の鍵になるのは情報的健康、インフォメーションヘルスだというようなことを申し上げており、要するに健康とのアナロジーで捉えてきたところがあるが、プレバンキングというのが例えば心理的な予防接種であるとか、心理的な耐性というものをつけ、しかも、そういった研究がWHOと共同でやられているというところで、非常に健康問題のような形で議論されているんだなということを知り、方向性が私自身も見えてきた。【山本構成員(4)】
- メディア自由法については、一つはコンテンツ・モデレーションとまさに重要な情報を発信するメディアの関係というものをどう考えていくか。これは非常に重要であることに加えて、コンテンツ・モデレーションだけに限らず、いかに信頼できる、まさにジャーナリズム、情報というものを多元的な形で欧州全体でしっかり確保していくのか。そのことというのは、まさに健全な情報空間

といったようなものの実現の中で非常に重要だろうなと思っている。【生貝構成員(4)】

- リテラシー向上検討会は、デジタル・シチズンシップというものを提唱している。それは結局のところユーザーとしてのリテラシーとかだけじゃなくて、やっぱりインターネット全体とかビジネスモデル全体について理解を持つことが重要。これがデジタル・シチズンシップ。そこにおけるキーワードは、もちろんアテンション・エコノミー、フィルターバブル、エコーチェンバーということになる。【森構成員(5)】

## 7. 安心・安全

- 児童のデータの保護の重要性も国外でよく聞くところであり、日本の法制度上、やや手当が薄い部分であるというのは否めない。ヨーロッパは GDPR の中に児童の保護に配慮した規定などがあるし、デジタルサービス法の中でもオンラインプラットフォーム事業者に対する義務の中でプロファイリングに基づく広告を児童に対して行ってはならないことや、イギリスのチルドレンズ・コード、アメリカでもカリフォルニア州の法令などがある。こうした動向を踏まえて、日本も法的な対応を検討するという余地はあるのではないか。【石井構成員(1)】
- 青少年保護、利用者保護は非常に重要。【増田構成員(1)】
- 昨今では、セイフティ&セキュリティというのが、生成 AI に絡めて重要な議論される項目として立ち上がってきているところもある。【江間構成員(3)】
- EU の DSA だとか規制法の話を知っていると、EU の法制構成全体は、偽・誤情報、ターゲティング広告、プロファイリング、ダークパターン、それから生成 AI、全て関係するところだと思うが、健全な情報環境を実現するために社会の誘導を、社会が変なほうに誘導されないようにする、そういう視点を持った規制に動いているのかなというように感じた。マニピレーションという言葉が使われていたかと思うが、個別の個人情報保護とか消費者保護とか、そういうプレイヤーの話ではなく、もっと大きな、もう一つ大きな観点の守るべき価値というもの、それは社会全体が何かしらのものに御誘導されないようにするという、そういう方向性の議論がなされていると理解していいのかなと思う。【石井構成員(4)】
- 比較的これまで、個別の救済だとかに焦点を当ててきたところから、もうプラットフォームが社会全体に影響を与えるものになってきたところ、システムミクマクなマクロな問題にこのデジタルサービス法というのが西側の中だとある種、初めて真正面から取り組んでいるルールになってくるのかなと思う。そうしたときに、イノベーションが起こり続けるからこそ、リスクはすごくいろいろあり得る。そうした時に、マニピレーションや操作というのは間違いなくあるリスクだし、あるいは、誰もその意思がなくてもカスケード的な現象って起こったりすることもあるし、あるいは我々一人一人の本性というのがいろいろな差別を引き起こすといったようなことも当然あると聞いた時に、それをある種、アンプリファイさせないだとか、そういった、ある種、フューチャーブ

- ルーフィングに、リスクに対処しているのかなという状況。【生貝構成員(4)】
- 2001年プロ責法の成立。掲示板管理者、プラットフォームの損害賠償責任を限定する。それから発信者情報開示制度をつくるという大きな新しい制度ができた。【森構成員(5)】
  - 2008年、この年、青少年インターネット利用環境整備法。これは携帯電話事業者にフィルタリングサービスを提供する義務を課するというものですがけれども、フィルタリングの対象となる情報というのはもちろん権利侵害情報には限られない。名誉毀損とかプライバシー侵害には限られないわけなんですけれども、ですので広い違法有害情報ということ。【森構成員(5)】
  - 2018年になりますと、権利侵害情報のところ、総務省ではないんですけども知財本部にインターネット上の海賊版対策に関する検討会議が設置されました。ブロッキングについての検討が行われるということ。総務省におかれましても、もちろん通信の秘密との関係がありますので注目をされていた。【森構成員(5)】
  - 2020年8月にインターネット上の誹謗中傷への対応の在り方に関する緊急提言、これはプラ研の緊急提言。もともと偽情報と利用者データの2つを所管していたわけですがこれもこれがこの権利侵害情報について緊急提言をする。この背景は5月のプロレスラー木村花さんの自殺を受けたものですがけれども、これによってプラ研の守備範囲が右2つ（偽情報と利用者データ）から権利侵害情報のほうに拡大してくる。3つにクロスオーバーするということ。【森構成員(5)】
  - 2022年、権利侵害情報に関してはプラ研に誹謗中傷ワーキングが設置。偽情報に関しては、ICT活用のためのリテラシー向上に関する検討会が設置。【森構成員(5)】
  - 権利侵害情報について同じ場所で議論すべきと言っているわけではない。別々のところでもいいと思うが、共通軸、対応の観点において同じものがあるということは意識した上で、しっかり連携し情報共有して進めるべき。【森構成員(5)】

## 8. オープン・透明性・アカウントビリティ

- 透明性を高めることで、アテンションの獲得“だけ”を狙って、偽情報を放置したり、エコーチェンバーを放任したりするような企業（情報的健康を無視するような企業）が市場において適切に批判されるようになれば、ビジネスモデル自体（＝構造）が変化していく可能性がある。【山本構成員(1)】
- 重要なのは透明性の確保。目指すべき社会をしっかりと考えて、具体的な透明性・アカウントビリティの確保を促していく、プラットフォーム事業者に促していくということがとても重要。例えば、どういうデータを公開する必要があるとかということをしつこりと定義づけし、その定義の果てに得られた結果が、しっかりエビデンスベースで有効な対策を検討していくことができるという状態にしておく。これが社会としては重要。【山口構成員(2)】
- AIガバナンスに関し、特に、Fairness、Accountability、Transparencyみたいな

- なことは、2~3年前から重要なテーマになっている。【江間構成員(3) ※再掲】
- G7の広島AIに関しても、IGFの京都でいろいろ事業者の方をお招きして議論があったが、そこでTransparency、透明性というのが非常に重要であるというせりふがあった。一方、アカウントビリティ、説明責任や、あるいは日本語でもう少しきちんと言うならば、説明をする責任ではなくて誰が何か問題があったときに責任をとるのかということを一応考えなければいけないということで、その点に関してはそれぞれの関係者が果たすべき役割や責任というもの、サプライチェーンの中で考えいくことが重要。【江間構成員(3)】
  - DPF 事業者内部のものと、DPF 事業者と政府間のものという、二つの不透明さを背景に、アメリカの連邦議会で、プラットフォーム説明責任及び透明化法案というような、内容的に見ると部分的にアメリカ版 DSA と言ってもいいかなというような法案が出てきたりしているが、これは法案レベルなので実際に成立するかどうか、かなり怪しいところだと私自身は思うが、この法案は、以前から議論されてきたが、今回もまた出ている。【水谷構成員(4)】
  - DPF 規制について、EUはむしろ、コンテンツの管理をもっとやれ、デジタルサービス法でもシステミックリスクをちゃんと分析評価して、軽減措置をちゃんととりなさいというような規制が入っていたりするが、それとアメリカの規制の議論は全く異なる側面がある。一方で DPF 事業者の透明性、不透明さ、彼らのコンテンツの管理のプロセスの不透明さをめぐる議論というのは、実は世界的に共通した話題としてアメリカやそれ以外でもあがったりしている。【水谷構成員(4)】
  - 透明性というのはあくまで目的ではなくて手段だと考えている。ではそれで何を達成するのかという点で、一つは DPF 事業者がユーザーや市民社会に対する legitimacy を醸成する役割や、あるいは、有名な裁判官の言葉だが、日光は最高の消毒液であるというようなことが情報公開とかの文脈でよく言われるため、透明性を高めることによって内々で行われている不誠実な対応に対する抑止効果が出るだろうというようなことが考えられる。また、規制を何かしら入れた場合の効果測定のため、モニタリングするためにも透明性は要るし、あるいは、AI等の技術利用がどういうリスクをもたらすかということも実は分からないところがあるので、これもモニタリングするためには透明性が重要になってくる。【水谷構成員(4)】
  - 透明性確保についてもプラットフォーム・ガバナンスとして、市場とか、いろんな圧かにプラットフォームはさらされているが、それらがうまく機能しないようであれば、何かしら法的な義務というものもきちん議論しなくてははいけないし、もう一つ、透明性レポートをあげてもらったら終わりではなくて、透明性は手段であるので、その目的を達成するために外部からきちんとモニタリングを入れることも重要。【水谷構成員(4)】
  - 我が国がどうするかというのは難しいが、一つは我が国の最も重要な価値をどう実現するかといったときに、まず一つは議論ができる前提として透明性というものをしっかり確保する。何が起きているか分からないのでは、まさに議

論のしようがない。そして、もう一つの論点としてのリスクの評価と軽減という意味では、これはできる限り公表された情報に基づいて公論をして、議論をできるだけメディアを含めてしていく中で、そういったようなことをデジタルプラットフォーム様がちゃんと取り込んでいけるような、そういったプロセスというものが極めて重要だろうと思う。【生貝構成員(4)】

## 9. プライバシー保護

- 個人の権利保護の観点で、例えば、関係者への開示・訂正・削除等の権限など個人に対しどういう権限が確保できるようにするかも重要。また、意図的な発信者に対してはどうか対応していくのか。リテラシー向上は、個人の権利確保の前提になってくる。【落合構成員(1)】
- 生成 AI のデータの正確性は偽・誤情報対策にも重なる部分はあるが、プロファイリングにせよ、偽・誤情報にせよ、個人が自由に行えるはずの選択を歪めるという行為に対し、いかに対策を講じるかという点が情報流通の健全性を支える上で重要。【石井構成員(1)】
- 2010年に諸問題研、ここで第2次提言が公表。この中で、ライフログを利用したターゲティング広告について配慮原則を公表する。ここが利用者データ保護の大きなポイントだと思うんですけども、こんなことが書いてある。「ライフログ活用サービスは揺籃期にあり、事業者に過度の負担をかけることは避けるべきであることから、規制色の強い行政等によるガイドライン化を避けて、穏やかな配慮原則を策定することとし」ということですが、新しいサービス、揺り籠の時期にあるサービスで、あまり強い規制をかけないというのは、これは全く見識だと思えますけれども、ただ、これは2010年ですので13年ぐらいたっておりますから、もはや揺り籠の時期でないということも明らか。【森構成員(5)】
- 2012年になりまして、スマートフォン・プライバシー・イニシアティブの公表。スマートフォンのアプリの中の情報収集モジュールによってスマホの情報が広告事業者等に送信されてしまうというようなことがありまして、「こんなに怖いスマホのアプリ」とか、そういう報道がなされまして、これによって、スマートフォンそれ自体あるいはアプリの普及に支障があることを心配してこういうガイドラインをつくったということ。これはまさに今で言うところの外部送信についてのガイドラインであった。【森構成員(5)】
- 2015年、総務省ではありませんけれども個人情報保護法の改正がパーソナルデータ検討会において行われた。その議論の中でパーソナルデータ検討会の技術検討WGで、端末等識別子をこれを個人情報保護法の規制の下に置こうという提案がなされたわけですがけれども、最終的にはそうはならなかった。検討会では特に反対の意見はなかったんですけどもうまくいかなかったというところ。【森構成員(5)】
- 2018年、偽情報と利用者データについてはっきりクロスオーバーが起こる。2つの問題が統合的に検討されることになり、プラ研がこの年に始動。開催の趣

旨のところを見ますと、「本研究会では、近年、プラットフォーム事業者が大量の利用者情報を活用してサービスを提供していることを踏まえ、利用者情報の適切な取扱いの確保の在り方等について検討を行います。」とされているが、最初から偽情報の問題も 이슈 になっていまして、最初の成果物である中間報告書、それから最終報告書どちらにも偽情報のことが記載されている。【森構成員(5)】

- 2021 年になるとプラ研に利用者情報 WG が設置。またその中間取りまとめが公表されまして、その中で「cookie や位置情報等を含む利用者情報の取扱いについて具体的な制度化に向けた検討を進める」とされている。要するに、電気通信事業法を改正して外部送信の規制をしっかりと法律に入れようということになった。このプラ研での取りまとめを受けて、その下のところ電気通信事業ガバナンス検討会が法改正の中身を具体的に検討するということになりませんが、やはりここでも事業者団体の反対と規制レベルの後退みたいなことが報道されたという経緯がある。【森構成員(5)】
- 2022 年、利用者データのところはとうとう電気通信事業法が改正され、ここに利用者情報に関する規制というものが入ってくる。具体的には外部送信の規制、それから特定利用者情報の規制というものもできた。特定利用者情報の規制は、LINE の利用者データが中国の委託先で閲覧可能になっていた事件を受けたもの。一定の重要なユーザー情報については、これをしっかりと管理してもらいたい。そのために内部的なルールをつくって、また外部に公表する方針もつけて公表しなさいと、そういう規律。【森構成員(5)】

## 10. グローバル・国際性

- 欧州 AI 規制法案、欧州評議会 AI 条約、G7 広島 AI プロセス等を含め、国際的な指針や行動規範、法令等の履行状況の確認、情報共有の仕組み等をいかに構築していくか。特に国際的に相互に確認できるような評価法をいかに確立できるか。【江間構成員(1)】
- 欧米が AI ガバナンスの議論を先導しているとはいえ、ほかの国々との関連性というのを全く無視していいというわけではないし、むしろそことのハーモニーをいかに考えていくかということが重要。おそらく UN のハイレベルのところ(国連 AI 諮問機関)は、その辺も考慮に入れながら、国際的な多様性というところも考慮して、いろいろな国の人たちが入ってこられるような枠組みとは一体何なのかということ議論する場所になると考えている。【江間構成員(3)】

## ② デジタル空間における情報流通に関する現在の課題

### 1. 情報の発信・伝送・受信の各段階に共通の課題

#### ◆ デジタル空間の情報流通に関するガバナンスの在り方

- 政府だけで対策するわけにもいかないし、企業、例えば、メディアやプラットフォームが何かをすれば、それで全てが片付くことでもない。そういう意味では、法制度、メディアやプラットフォームを中心とする企業側、個人のリ

テラシー向上、権利の確保など、各者における相互の対策・対話の促進が重要。そのフレームワークとしてアジャイル・ガバナンスがあり、環境等のリスク分析を行い、その結果、対応すべき目標が何か次第に変わっていく可能性があることを認識しつつ、ガバナンスモデルをアップデートすることが重要。【落合構成員(1)】

- 技術を誰でも使いやすくなるような社会環境を整備していくという意味で、法律の側であったりだとか、プラットフォーム、メディアなどに求めていく対応を整理していくことが比較的効率的。【落合構成員(2)】
- OP の利用においては、技術を中立的かつ汎用的に使えるということ、あと言論に対するガバナンスを構築するということを峻別する必要。OP では、OP そのものは、このメディア又はこの記事、このコンテンツはいいよ、悪いよということ言うのではなく、何らかの規律をコミュニティに適用したい。例えば、業界団体のような主体が OP を使って、自らの規律に基づいて情報の確からしさを検証可能な状態にしていく。言わば自主規制、ないしは、緩やかな共同規制のようなガバナンスを模索している。【クロサカ構成員(1)】
- ガバナンス設計の観点から、今まで憲法学は個人に表現の自由があり、そこにある種の規制がかかることを個人の権利ベースで議論をしてきたが、放送制度の例があるように、デジタル環境のいろいろなステークホルダーやアクターの機能等を見据えて構造・システム全体を把握しながら制度設計をしていく視点が求められる。特に、プラットフォームのモデレーションやキュレーションにおいては、どうしても事後的にこの情報が削除されていないとか、何でこの情報が削除されたんだという事後のエラーをチェックするということに関心がいきがちだが、それだけではなく、そうしたモデレーションやキュレーションのシステムがどのように設計されてきたのかというシステム全体に目配りをするガバナンス設計が必要。表現者中心から環境中心アプローチ、権利ベースからシステムミックスな視点に基づくガバナンス設計。【水谷構成員(1)】
- 広く市民やメディアを含む様々なステークホルダーが健全な情報流通というものの在り方を議論し、主体的に関与し続けることが可能な仕組みや環境というものをどのようにつくっていくかが大変重要。個別の 이슈や技術に関わるガイドラインなどを作るプロセスに加え、いかにして様々ステークホルダーが関与し続けられる情報環境をつくっていくのか。もともとアメリカでプラットフォーム企業の自主的な取組として始まった透明性レポートは、EU のデジタルサービス法の中でも法的な位置づけを得て、また、同法の中では、コンテンツ・モデレーションに対する個別の理由というものをちゃんと説明して匿名化した結果を公表したり、研究者などへのプラットフォームのデータ提供などを通じて、広く様々な主体が情報流通の現状というものを正しく深く理解・議論し、必要があれば、個人個人がそうした情報流通の在り方に異議を申し立てることができる環境をつくっている。そうしたことを念頭に、マルチステークホルダーに基づく情報空間のガバナンスの在り方というものの、そして、その実現手段のオプションというものを様々な選択肢の中から広く考えていくことが



## 重要。【生員構成員(1)】

- ミスインフォメーションや何かを判定する過程でいうと、なんだかよく分からなくて、ミスインフォメーションなのかどうかも分からないようなものがポワッと浮いている時間がすごく長い方が多い。そして、世の中にはミスインフォメーションは多分なくならなくて、ファクトチェッカーの数が追いつかなくて、恐らくこれからどんどん増えていってしまうことになる、もう少しプロセスで考えたい。要するに、何かポワッと不透明で不安定で不安で、すごく分からないようなものが続くところで、社会をどうやって制御するかということを考えていかなければならない。【奥村構成員(2)】
- メディアの倫理規範の価値をどうやって守るかのHowについて、例えば、ジャーナリストは株の取引をしない。不祥事や新製品の開発等会社の市場的な評価を上下させる情報をいち早くゲットできるため、当たり前といえば当たり前のこと。ただし、例えば、APでは、「ビジネス関連の記者とエディターは株を持つてはいけない」、それから、自分の取材領域の関連企業の株は持つてはいけないと書いてある。BBCでは、少なくとも自分の金融商品がどれぐらいの利益を上げたかを上司に説明しなければならぬと書いてある。例外は、親の資産を相続したとか、親戚の会社の株を引き受けなければならぬ場合としか書かれておらず、あとは上司に全て報告している。ニューヨーク・タイムズでは、持つていい金融商品は自分でコントロールできないものに限ると書いてあり、ニュースに関わるスタッフは自分で管理するタイプの金融商品は持たないと書いてある。そして、やむを得ず保有している金融資産が現在担当のニュースで利益相反を起こす可能性があるとその記者が判断すれば、必ず上司に報告して判断をおおぎ、場合によっては担当替えをさせられるということを受入れると書いてある。記者が上司への報告義務を怠ったり、上司が判断をしなかったりということで、責任の所在がかなり明確になる。そして、さらに責任の所在を明確にするため、スタンダードエディター、オピニオンエディターやマネージングエディターというようなポジションを説明し、この人たちが一体どんなところで何の判断をするかというようなことを全て書いている。【奥村構成員(3)】
- シグナトリーという認証ファクトチェック団体の資格を与えるプロセスにおいて、IFCNが定める5原則に合致しているかを外部のアセサーが判断する。シグナトリーに申請があるとアセサーに一斉に連絡が来る。この申請の担当になりたいアセサーは手を挙げると4万円ぐらいの謝礼がもらえるが、その申請メディアと数か月にわたり、多分10回ぐらいのやり取りをしなければいけない。アセサーはファクトチェック団体がやっている日々の記事を調べ、それで基準に合致しているかを、説明やレポートを受けながらやり取りをしていく。この承認のプロセスはアセサーの名前とともに記録が公開されている。日本でも今年、3つのシグナトリーが誕生している。なるべく情報を出す人たちがちゃんとした行動をとっているということを保障し、ニュースやファクトチェックの安心というようなものを社会的に担保している。【奥村構成員(3)】

- AI はただ単に AI を開発している企業がそのままサービスを消費者に見えやすい形で提供しているというわけではなく、国や組織をまたいでいる。AI 開発者、AI 提供者、AI 利用者がいるというすごくシンプルな図は、例えば、Google やマイクロソフトは、マイクロソフトは開発者で OpenAI とも連携しているが、比較的、提供されているサービスがどのような人たちに使われているかという関係性が分かりやすいというような海外の企業もある一方、日本は主にすごくサプライチェーンが長いという特徴がある。実際受けているアプリケーション等は、特に LLM を提供しているのは全然違う企業だが、実際に生成 AI を使って、そのサービスやアプリまでやっていくとなると、本当に長いサプライチェーンになってくるので、実際、どこがアルゴリズムを開発しているのか、誰がデータを提供しているのか、又はどういうふうにサービスが提供されているのかは非常に複雑になってきているので、何か問題あったときに誰がどう責任をとるのかということが非常に難しくなっている。その整理をする必要が出てきているというのが、基本的にガバナンスをめぐる一つの大きな論点。そのような中においては、ステークホルダーを特定した上で、事業者間での円滑なアジャイルなガバナンスをしていく上という観点からすると、きちんと契約を取り交わした上で、契約自体に何か寡占や、社会的なパワーバランスによる不公平とかが起きていないかということは、もちろん市場が監視することは重要ではあるが、特に透明性をもって、可能な範囲での透明性、情報提供ということを互いにやっていくということが大事。【江間構成員(3)】
- 法律やガイドラインだけが規律というか、人々や技術や AI を提供している組織に要求をする方法ではない。法令等は国家による強制、エンフォースがあるが、業界団体がつくるガイドラインやポリシーはエンフォースしないがそれなりに守ることのインセンティブをうまくつくることにより、影響力があるかもしれない。日本は特に事業者ガイドライン等をつくっても、かなり真面目に取り組む企業も多くあるので、エンフォースしない、ある種のソフトローでも、ある程度の規律を守るような仕組みが構築できるのではないかと。一方、国や組織をまたいで AI の開発や提供利用が行われている中、それぞれの国又は国際的な関係性の中で、法令やガイドラインの相互認証等、又はどのような透明性をもって対応しているのかを確認し合うことは非常に難しいので、もう少しスピーディーな、学会の基準や慣習又は市場による監視、ESG 投資で投資家がこういう製品や商品をつくることを望むという要望を出すことにより、ある一定程度のコントロールや規律をかけていくことや、日本は炎上リスクという評判を非常に重視する国民、国柄があるので、ある種の人々がちゃんとした行いをすると、規律といったようなまさにフーコー的なディシプリンが見られているかもしれないという観点から対応しなければならない。様々な規律を使いながら AI ガバナンスをある程度、アジャイルに行っていくことが重要。【江間構成員(3)】
- 実際に AI システムがきちんと動いているのか、又は生成 AI とか様々、問題が出てくるときにどうやって対応していくのか、健全性を確保していくのかとい

うことにおいて、一つの方法論として、様々なところで AI Audit、AI による AI サービスシステムの監査というのが取り上げられている。もっとも、これはそんなに簡単なことではない。Principle とか Practice とか様々なものがある中で、では一体そのサービスの何を監査していくのかといったときに、その対象となる立証命題をブレークダウンして考えていく必要がある。例えば、公平性ですと、AI システムの出力に不適切なバイアスがかかっていないかというような形まで落とし込んで、それを、システムを見ていたり、あるいはサービスを提供する組織の運用の仕方とかを見ていくという形になる。それぞれの立証命題がどのようにしてブレークダウンできるのかといったことの基準は様々で、各企業や組織においても、何を重視するのか、どういうことを重視するのかというのはかなり多様になっているので、全てのものを一遍に見ていくことはなかなか難しいところであり、それぞれがどういう優先順位で見えていくのかということもかなり議論した上で監査をしていくということが重要。【江間構成員(3)】

- AI サービスやシステムを監査するという点においても、どのレイヤーを監査していくのかということが、もう少し整理する必要がある。アルゴリズム監査とかデータの中身の監査というようなことが一般的に思いつくが、従来の AI システムがソフトウェアだけではなく、例えば自動運転とかハードの中に組み込まれていくようになったときに、そちらのハードウェア自体のデバイスの性能とこちらのアルゴリズムの関係性ということも考えなければならない。また、どのようにしてサービスが提供されているのか、情報が適切に開示されているのかに関しても、監査というのは入ってくる。【江間構成員(3)】
- システムやサービスだけではなく、どちらかというとマネジメントのほう、内部統制はどのように効いているのかということも監査するということもあり、どちらかというと技術的にアルゴリズムの監査をしていくというよりは統制をどのようにされているかという観点から監査をしていくという、できることからやっていくような、この辺からいくのではないかと。ただ、どのタイミングで監査するのは非常に難しいと感じており、AI は PoC レベルや新規開発のところがあり、それぞれをリリースした後にさらにまた追加開発をしたりで、常にこの時点でもう一旦製品がフィックスされることはなかなかない可能性もある。そのときに、監査をした時点と実際に動いているという点で AI の出力結果が、精度が異なるみたいなケースも出てきて、タイミングがいろいろとある中で、どのように誰がどの段階で、内部監査の人と外部監査の人が場合によっては協力することも必要になってくるかもしれない。【江間構成員(3)】
- 誰が監査をするのかというところでも、AI の技術の知識もあれば監査の知識もあるみたいなスーパーマンを 1 人想定するということは少し難しいので、チームで監査していくことが重要。また、何もかも責任を全て外部監査で任せすぎてしまうと、そういうシステムを監査したくないというような監査人不足になってしまう可能性もあるので、誰が監査するのかどうやって監査するのかということも含めて、ある種の監査人のリテラシーだとか教育というののも一緒に必

要になってくる。【江間構成員(3)】

- AI 監査というのは、制度的な要因もあれば社会的な要因、そして技術の複雑性という様々な要因がからまっており、非常に難しい。生成 AI の監査をめぐる議論というのも、著作権とか偽情報・誤情報、感情操作、いろいろな方が議論の俎上に上っているが、それを適切に評価する基準開発というのがまだまだ未熟な段階において、それを監査するというところまで行くのは少しラグがある。一方、アルゴリズムそのものというよりは、インターフェースをどのように設計しているのかも監査対象になり得て、生成物に対して適切な表示がされているのかどうかということも監査の対象となってくる。【江間構成員(3)】
- 日本の場合、事業者が割と真面目に事業者ガイドラインを守るとか、あとレビューリジョンリスクを恐れるという話があるが、AI ベンダーは外国企業が多い中、果たして日本人だからという国民性にどの程度依拠しているのかと思う。【森構成員(3)】

- デジタルサービス法の行動規範がどういうものなのかといったようなものについて、リスク低減義務というものを行動規範で具体化していく、法と行動規範の関係のさせ方が非常に工夫された形になっていて面白く、招請を適切な説明なしに拒否した場合や、あるいは、ディスインフォメーション対応としてリスク評価をやらないといけないような説明もあったりする。【生貝構成員(4)】
- 行動規範のアップデートに関しては、本当に頻繁にアップデートできるのが行動規範の良いところで、例えばディープフェイクの問題は、あの頃に改めて初めて出てきたところもあったので、そういうものも含めたアップデートかなという認識。【生貝構成員(4)】
- 異議申立ての仕組みが DSA には入っていて結構目玉だと思うが、この個人の異議申立てという仕組みというのは、事後的な救済という意味で、法律家からするとすごくなじみ深い。だが、データベース自体がもう 8 億あるという話で、量的な観点から考えても、それを利用するというのは非常に限られてくるユーザーになってくると思う。それと関連して、個別事案の事後的な救済は、例えばコンテンツ・モデレーションに用いられているアルゴリズムのバイアスの問題だとか、根っこの部分に関係してくる問題には対処し切れないのではないかなと思う。【水谷構成員(4)】

#### ◆ ステークホルダー同士の連携・協力の在り方

- 誰がステークホルダーとなるか、ステークホルダーの多様性に配慮しているか（「利用者」が「開発者」と同義なこともあれば、利用者自身の立場や状況も多様である場合も想定にいれるべき）。【江間構成員(1)】
- デジタル空間の拡大・深化の流れが加速し、デジタル空間と物理空間がシームレスにつながりインフラ化していくがゆえに（サイバー）攻撃に対する防御の事前評価・対策だけではなく、インシデント発生時における対応も各ステークホルダーが連携して行うこと。インフラ化しているデジタル空間は特定のステークホルダーだけでは対応できない状況になる可能性もある。【江間構成員(1)】

- AI 監査に関し、内部監査のような第一線、第二線と言われているような企業内での監査と、外部の標準化だったりとか、あるいは外部監査実施者による監査といったいろいろな関係者がいる。様々なステークホルダーを特定していきましようというときに、具体的に企業内部の人たちというのも、ある種重要なファクター。【江間構成員(3)】
- AI 技術の発展もあり、言語の壁、特に、物理的な境界がほぼない中で、偽・誤情報そのものに関しても、これが偽かどうかという明快な線引きができない中でどのように捉えるかも課題。どうしても研究者としては明示的な偽情報が分かりやすいので、そちらの方を先に分析してしまうこともある。明示的な偽・誤情報とは言えないものの、ある種今後の懸念材料になり得るものがいろいろとネットワーク上の中にあって、そういったものをどのように位置づけ、単一の偽・誤情報だけではなくて、ネットワーク全体として総合的に捉えていくかということも大事。単一的な取組には限界があり、多様な観点から社会全体としての偽・誤情報に柔軟に対応できるような基盤の底上げの検討の一助になれば。【澁谷構成員(1)】
- 偽・誤情報をプラットフォーム事業者側の問題、あるいはリテラシーを中心とした一人一人の個人の問題として捉えるということは限界があって、社会経済的な背景、偽情報の需要のメカニズムやエコシステムの解明など、いろいろな多面的なアプローチというのが大事になってくる。また、単一のプラットフォームごとの流通状況の把握ということも非常に限界があり、外部 URL も非常にいろいろ共有されており、様々なプラットフォーム間の情報流通の状況の把握ということも課題。【澁谷構成員(2)】
- 「構造」や意識の変化には時間が必要。その間、「構造」や認知心理学等を踏まえたベスト・プラクティスを共有し（認知心理学等を踏まえたファクトチェック記事の流通など）、その導入・実装についてインセンティブを付与することが重要。【山本構成員(1)】
- 信頼できる情報源が喪失してきている。メディアへの信頼度が世界各国で低下している中、アテンション・エコノミーが一方で支配的になってきている。その中で、情報生態系全体で人々が適切に正しい情報にアクセスできるような仕組みを考えていく必要。そのためには、ステークホルダー間の連携の強化は外せないことであり、ベストプラクティスの共有、連携の強化、さらに、各ステークホルダーが一堂に会する場というものをぜひつくってほしい。【山口構成員(1)】
- マイクロソフトは偽動画判定ソフトを開発して報道機関に提供すると報じられているが、こういったメディア企業とプラットフォーム事業者の連携は欠かせない。【山口構成員(2)】
- 重要なのは、ステークホルダー間の連携で、偽・誤情報対策に結局特効薬はない。根絶はもちろん不可能。なぜかという、インターネットが普及前から我々の世界にあったもの。しかしながら、問題を改善することはできるはず。自由・責任・信頼があるインターネットを築くためにも、ステークホルダー間の連携

が必須。例えば、メディアとプラットフォーム事業者と業界団体と教育関係者、アカデミア、そういった人たちが対等の立場で参画して議論を重ねられる会議体が考えられ、ベストプラクティスや技術の共有、偽・誤情報の傾向情報の共有、又は具体的な対策の議論・連携や幅広い役割をそういう会議体に期待される。例えば、ディスインフォメーション対策フォーラムは、そういう意図に近かったが、主な構成員は有識者であり、対等な立場での議論は結構難しかったというふうに感じている。もっといろんな業種の人たちが一堂に会して、第三者からいろんな意見交換できるのかというのは期待されるところで、もちろんディスインフォメーション対策フォーラム第2弾をやるという手もあるし、また別の形でやるという手もあると思うが、そういう場がまたできるといい。【山口構成員(2)】

- 媒介者の役割と、あとは個人も含めたいろいろなタイプの発信者がいる中で、どうするとより問題が少ない可能性が高いような情報が広く拡散されていく可能性が高まるか、そこを追求していくことが論点。この点、プラットフォーム等の拡散について重要な役割を持っている方々についてはどうしてもらおうといいのかという点と、個人に対して、もし誤った情報を拡散する可能性がある場面で、なるべく思いとどまってもらう対策として何が考えられるのか、を全部総合して考えていくことになるが、最終的にはいろいろな方策を組み合わせ、今後の情報格差の状況を、変化を見ながら対策を少しずつチューニングしていくことを今後議論していく方向と思う。【落合構成員(3)】
- 生成 AI による偽・誤情報は特定のプラットフォームで閉じることなく、多様なメディアを通じて流通する傾向にある。生成 AI による偽・誤情報拡散によるリスクアセスメントも確立していない状況であるので、生成 AI 事業者、プラットフォーム事業者、通信・放送事業者などが連携してインシデントやリスクの共有を行い、連携しながら対応を行っていくことが必要。【越前構成員(1)】
- メディア同士はコラボレーション、協力しなければならないという風潮が世界的に高まっている。協力できる分野、一定のスキルなど共有できる部分を探し出して合意して、実行に移す。一つの出来事でも複数のメディアの評価があったほうがいい場合もあり、分野をちゃんと特定しなければならない。【奥村構成員(3)】
- ファクトチェックのコラボレーションの特徴としては、①強力なリーダーシップがあってスタートし、必要性や価値についての共通認識があること、②経営規模・実力に応じた負担を分担することの合意として、マウントも劣等感も感じない環境で、自分たちが損だ得だというようなことを一定程度外視しなければならない分野をちゃんと決めて、その部分は腹をくくるということ、③エンジニア、研究者やビジネスとの連携により、活動を社会的なものに広げ、社会的な認知を拡大していくことである。日本のメディアはコラボレーションがあまり得意ではない。そもそも、引用する文化が非常に希薄であり、他社のスクープを引用して速報するというような例はあまり見たことがない。数日後とかに非常に小さな扱いで、小さく伝える事例の方が非常に多い。【奥村構成員(3)】

- ファクトチェックの世界では国単位、あるいは南米とかアフリカなどでは共通の言語を話す国、全ての中でのメディア間の連携が進んでいる。東アジアは言語がすごく違うので、非常に苦しい環境でもある。これはミスインフォメーションが大量に発生し、ファクトチェッカーの数が足りないからである。世界のファクトチェッカーの中ではとにかく Duplication は避けましょうというのが合い言葉になっている。要するに、同じ問題を重複して、複数の者がファクトチェックしなくていいじゃないかという考え方で、ファクトチェッカーの集団をセクターとして捉えて、それ全体で仕事をしようという認識である。例えば、フィリピンの #FactFirstPH という試みは、2022 年の大統領選挙等のミス／デイスインフォメーションの対策のための連携で、140 以上のニュースメディア、市民団体、ビジネス団体、企業、弁護士等法律家の団体や研究者でグループになった。中心となったのはノーベル平和賞を取ったラップラーのマリア・レッサさん。この仕事は4つのレイヤーに分かれ、ファクトチェックをし、それを社会のどのようなネットワークで広めていくかというメッシュというプロセスがあり、それを研究者が記録・分析し、アカウントビリティとして、それを記録・分析し、論文としても発表するということで、ファクトチェック記事を900本近く出して、そして学術論文を20本近く出しているという大きな成果を上げている。ノルウェーでは、2017年から、主要の6メディアが協力してファクトチェックをシェアする団体を運営していたが、2022年2月からウクライナ戦争が始まったので、主にロシアのミスインフォメーションを分析するというOSINTに関するプロジェクトをスタートさせた。Verifiservar は要するに空いているバーという意味で、要するに、コロナで閉まってしまったバーを拠点にして、12のメディア、32人の記者が分担して仕事をしており、ファクトチェックをする労力の分担が細かく決められており、メディアの実力とか人員とか収益とかでポジションが決められており、0.4というポジションもある。この0.4のポジションの記者はどうしているかというと、別に何か2週間に1回来るとかではなく、記事を最終的に出すというような責任までは負えないが、ただファクトチェックには加わるとか、このようなプラットフォームがあると自分のメディアにそういうスキルを学んで帰れるというメリットもあるので、そのようにして合意をし、特定の分野について競争を保留して、情報のボリュームを社会的に厚くしていこうという共通認識ができています。また、ブラジルの大統領選挙でも、コンフィルマ 2022 というプロジェクトがあり、5つの主要メディアが参加し、そして、Meta がやっている WhatsApp も参加した。特筆すべきは、Meedan（ミダン）というアメリカとイギリス、南米、インド等にネットワークをつくる非営利のエンジニア集団があるが、そちらの Tipline というアプリを提供してもらい、そちらを重点的に使ったという点で、ユーザーが質問をチャットボットに投稿すると、既にファクトチェックの結果があれば、5社のファクトチェックの中からそれをピックアップしてその人に答えを返してくれるというチャットボットで、優れているのは、その答えがなかった場合、その5社に対し、Tipline Inbox というが、こんな質問が来ているが、どこかの社が

ファクトチェックしないかと、各社にそれを送るという点。それを何日かかかって各社がファクトチェックをすると、ちゃんとチャットボットでフィードバックをしてくれるということで、33万件を超える質問を処理したというのが成果。【奥村構成員(3)】

- 2017年に行ったヒアリングのやり取りに関し、今でもそんなに大きな変化はないと思うが、東日本大震災、福島第一原発の報道について、主要メディア10数社にインタビューした。振り返って、どのような教訓を得て、ニュースメディアの組織内でどのように共有して、これからあるかもしれない南海トラフ等にどうやって対処していくつもりなのかを聞いた。当時、テレビも新聞もどこを見ても同じというような、欲しい情報がないというような批判があった。自分に必要な情報がいつ伝えられるか分からないし、どこにあるか分からないから待ってられない。当時のツイッター等では、例えば、電力やガスや水道や鉄道等の分野に分け、このニュースメディアを見ればそれについては全部分かるというようなことはしてもらえないのか、テレビは1日に数時間でもいいからそういう放送をしないのかとか、というようなこともあった。そういう分担やコラボするアイデアはどうかということも質問に入れたが、ほとんど各社が、あり得ない、今この段階ですることは考えていないという反応だった。消極的な理由は、精神論ではなく、もう少し深いところとして、多少自分たちの数値化できないようなスキルに契機をしているのかもしれない、例えば、電力や水道等インフラに分けたときに、あのとき一番の焦点が東京電力だったが、ある一定の期間の取材を放棄して、一つの社に電力の取材を任せてしまった後、遅れをとってしまうのではないかと、ある一定の分野を自分たちが専ら担当することになると、取材のノウハウが流出してしまうのではないかとというようなこと。ただ、共有できる取材のノウハウ等、絶対秘密を守り通したいという人脈や情報源というのは区別できるはずで、実際、報道実務家フォーラムがあり、早稲田大学のジャーナリズム大学院等が中心になって、報道の優れた実例をその当事者が自分で発表し、そして日本全国からメディアの人が集まって聞くような取組があるので、多分手のうちは明かすということができないわけではなさそう。また、コラボレーションを調整する労力の負担も考えておかなければならない。誰がどのような形でやるのかということをちゃんと考えておいた方がいい。プラットフォーム上でやり取りされる情報の多くはニュースであるということを考えると、コラボレーションのモデルや枠組み、基準を示す責任も期待もニュースメディアにあると思っているので、このような動きが少し、このような場で加速することを期待。【奥村構成員(3)】

- デジタル空間における人材育成は全体にかかってくる論点。【落合構成員(5)】

#### ◆ 生成AI・ディープフェイク技術の進展に伴うリスクへの対応の在り方

- 生成AIについては、学習済みのAPIが広く公開されており、誰もが最新の生成AIを使ってコンテンツを創作できる状況になりつつある。国内でも面白半分で偽・誤情報を作成、共有し、関係各所に大きな混乱を招いたケースだけで



はなく、明確な意図を持って視聴者を誤解させる目的で作成したケースも出てきている。このような脅威に対し、偽・誤情報の自動検知について、国内でも実導入に向けた動きがある。【越前構成員(1)】

- 特に生成 AI を用いて創作される偽・誤情報の拡散を技術的な手段でどのように防ぐか。生成 AI が生成する画像、映像、音声、文書のクオリティーは日々進化している状況であり、一部の偽・誤情報は人間が見たり聞いたりしてもリアルなのかフェイクなのか全く判別がつかなくなっている状況。【越前構成員(1)】
- 国内では生成 AI の脅威が 2021 年頃から深刻化。例えば、詐欺・詐称として、海外では、フェイク音声で企業の幹部になりすまして現金を搾取（2019 年）やフェイク顔でイーロン・マスクになりすまして Zoom 参加（2020 年）が発生していたが、国内では、機械学習モデルで生成・配布したサンプル顔画像を利用企業が自社の宣伝に不正利用（2021 年）したり、思考誘導・世論操作として、海外では、架空の人物になりすまして株価操作を目論む（2019 年）事案があったが、国内では、官房長官の地震直後の記者会見の表情を改ざん（2021 年）、そして、特定個人に対する名誉毀損・いじめとして、国内では、Deepfake によるアダルトビデオ公開・逮捕（2020 年）や、海外では、娘のライバルを蹴落とすため、母親がライバルの Deepfake 生成（2021 年）する事案があった。【越前構成員(3)】
- AI の音声詐欺は、海外でもものすごい問題になっていて、日本ではまだ少ないが、米国では 14%以上、インドだと 20%程度、要は個人を狙った、特にシニア層を狙った AI の音声の詐欺は、脅威になっている。【越前構成員(2)】
- 顔の生成に特化した生成 AI が作り出した顔画像をよく見ていただくと、髪の毛の一本一本、生え際も虹彩も、本物と見まがうようなレベルでつくられている。これは、私のような研究者が見ても非常によくできていると感じる。要は、こういったものがソーシャルメディアに出てきたときに、一般の方に真贋の見分けがつくかと、AI 製なのか否かというのも極めて厳しいんじゃないかと感じている。【越前構成員(3)】
- 違う顔に置き換えた、Face swap という技術によるものがあるが、通常、動画だと時間軸で粗が見えやすいが、非常になめらかに、これは 4 年前に Google から公開されたデータセットだが、他の顔に置き換えるということができてしまっている。【越前構成員(3)】
- 顔・音声・身体・自然言語など大量の高品質な人間由来の情報を AI が学習し、本物と見まがうようなフェイクメディアの生成が 2018 年から可能になっている。もともと Deepfake は顔を swap するということから出てきた技術だが、そういったものが 2018 年ぐらいから出てきて、GROVER と呼ばれるフェイクニュースの生成モデルというのも出てきたという状況。実はもう 2019 年頃から、海外ではこういったフェイク音声で現金を搾取した事例や、架空の人物になりすまして株価操作を目論む事例が起きていた。例えば、2020 年には、イーロン・マスクの一枚の顔写真を使った表情操作と呼ばれる技術で、こういったものがビデオコミュニケーション（zoom）に使われて参加者はびっくりしているとい

う状況のビデオができています。また、ウクライナ大統領の Deepfake によるロシアへの降伏呼びかけ(2022年)や、最近問題となっている拡散モデル(Stable Diffusion等)を用いた偽・誤情報の拡散(2022年)がある。これはプロンプト一つで非常に高品質な画像ができてしまうということで、大きな社会問題になっている。あと、ChatGPTを用いたマルウェア作成やフィッシングメール作成もプロンプトをうまく工夫するとできてしまうという問題が顕在化してきている。【越前構成員(3)】

- 顔を対象としたフェイクメディアの生成手法について概観すると、大きく分けて5つのタイプがある。①顔全体の合成で、ノイズ(潜在変数)を用いて、実世界に存在しない顔画像全体を生成するというもの。特に StyleGAN が有名だが、顔の生成に特化したモデルが出てきている。プロンプトから顔画像を生成(Stable Diffusion) & LoRA によるファインチューン(2023年)。②顔の属性操作で、例えば、インプットはリアルな顔だが、属性、髪の色や性別を自然な形で AI によってスタイル変換するという手法も出てきている。また、実は顔全体の生成において、今年になって、Stable Diffusion などの拡散モデルを対象とした効率的なファインチューニングの方法が出てきており、今までの拡散モデルは、例えば、洪水や災害など一般的な事象を表すような自然画像などをつくるのは簡単だが、特定の人物を対象としたチューニングを行うということも、今年になって非常に簡単にできるようになってきている。チューニングした結果を使って、プロンプトを入れることで簡単に特定の人物に特化した、そういったフェイク映像というのが簡単にできてしまう。ダウンロードしたパラメータ群と、例えば、Stable Diffusion を組み合わせて、特定の人物に特化した、拡散モデルというのができてしまうことが、現在、大きな問題になっており、これにより、プロンプトのみで、つくれる画像の可能性が極めて高まってしまっているというのが大きな問題。③顔映像、画像の表情操作は、攻撃者の表情とターゲットの顔画像や映像を合成して、攻撃者の表情と同期したターゲットの顔映像を生成するもので、最初に出てきたのはコンピュータービジョンのトップカンファで出てきた Face2Face と呼ばれるもの。ターゲットであるリアルな顔映像にリアルタイムでソースである攻撃者の表情を転写してビデオを生成しようというものが2016年に提案された。Face2Face は、ソースでビデオを使っていたが、一枚の写真を使って表情を転写するという手法もその後出てきている。これは現状、極めて大きな問題となっていると我々は考えており、例えば今、国内で非常に盛んである、スマホを使ったインターネットバンキングのアカウント開設に使う顔認証、eKYC と呼ばれるものに対する脅威となっている。④顔映像の話し方操作は、リップシンクと呼ばれるもので、例えば、リファレンスの映像であるオバマ元大統領がしゃべっている映像に違う音声を、異なるタイミングで録音したオバマさんの音声を入力すると、その音声に合わせてオバマさんの口元が自然な形で同期する方法を提案している。リップシンクは現在、国内で大きな問題になっており、口元のみを生成してフェイクをつくるということだが、自然な形で口元が合成されている。⑤顔の入替え、

Face swap についてはよく知られた手法だが、近年は非常に効率化し、一枚の写真を使って元の映像の顔を置き換えるという方法も提案されている。【越前構成員(3)】

- JST から予算を頂き、CREST プロジェクトを 2020 年 12 月から開始しており、大阪大学の馬場口先生と東京工業大学の笹原先生が主たる共同研究者になる。この申請を提案した時はまさに COVID-19 の真ただ中で、インフォデミックと呼ばれる社会に恐怖や混乱を引き起こす不確かな情報が氾濫していた状況だった。例えば、科学的根拠のないフェイクニュースや、ソーシャルディスタンスを守って並んでいる画像ですが、望遠カメラによって特定の方向から撮影することで、意図的に密集状態を演出するようなものも広義にはフェイクではないかと考えた。素材自体はリアルだが、意図的なディレクションによって大衆の世論を誘導するというものも、広義にはフェイクメディアとして我々は考えた。2020 年初頭にプロジェクト申請したが、当時から今後 AI を使った愉快犯や攻撃者が、多様なフェイクメディア (FM) を駆使してインフォデミックを意図的に発生させる可能性があるということ考えた。具体的には、3 つのタイプとして、①メディアクローン型 FM は、ディープフェイクのような、本物に限りなく近いが本物でない、生成 AI がつくったようなフェイク、②プロパガンダ型 FM は、世論操作のために、メディアを意図的に加工するようにつくられたフェイクメディア、③人間ではなく、AI を誤動作や誤判定させるものに特化した敵対的サンプル型 FM も広義の意味ではフェイクメディアと我々は考えた。【越前構成員(3)】
- リスク認識については、定性的だけではなく、定量的なものや、研究として実際どういう実態があるのかを把握していくことが重要。その中で、データの利用や、生成 AI 等の AI がどういうふう利用されているのかの実態を見ていく。自動判定技術による対応可能性、これもどうしても後追いになるところがあるが、今後も継続的に基礎資料となるような材料を集めていくことが必要。【落合構成員(1)】
- 生成系 AI の普及によるコンテンツ爆発の影響は、偽情報にも及んでおり、AI によって生成される偽情報の問題や、AI による偽情報判定の問題も、本検討会の中心的課題。【森構成員(1)】
- AI とプラットフォームとデータの 3 つは重要で密接不可分なものである中、生成 AI や基盤モデルの開発者・提供者の役割が非常に重要視されてくる。そこで作られた偽動画等が流通するのは、SNS 等のプラットフォーム上であることが多い。デジタルサービス法の中では、AI 規則とは別に、リスク軽減措置の中で行動規範の形で、生成 AI コンテンツに対しプラットフォーム上でのラベリングを求める議論がある。生成 AI とそれを流通させるプラットフォーム、それぞれの責務・役割というのは何なのかという両面から考えていく必要がある。この検討会がプラットフォームの問題と、そして生成 AI などの新しい技術を併せて取り扱う枠組みができていくことは大変適切。【生貝構成員(1)】
- 生成 AI とプライバシー保護に関する国内外の動向を踏まえると、イタリアの

監督機関が ChatGPT を一時停止したという事案が有名だが、各国の規制機関が調査を行ったり、国際的な議論の場でも情報発信を数多く行っている。入力するデータ、出力する結果について、それぞれ正確性の問題が出てくるというのは、プライバシーの問題でもあり同時に、偽・誤情報対策ともつながる部分であろう。【石井構成員(1)】

- AI 技術が発展して生成 AI が普及する中で、誰もがディープフェイク、つまり偽画像や偽動画を生み出すことができるディープフェイクの大衆化が起きた。高度な技術を持っていなくても、誰もが偽画像や偽映像を簡単に作れる時代がやってきた。これにより偽・誤情報が爆発的に増加することが予想される。2016年以降、我々の社会は with フェイクと言ってもいいような時代だった。常にフェイク情報が、偽・誤情報が我々の身近にあった。それが、AI 技術が発展したことによってさらに爆発的に増加する。まさに with フェイク 2.0 といえるような時代に来たと言える。既に災害、戦争などの有事に社会を混乱させる投稿、イスラエルとハマスの衝突での偽動画や、明確に政治的意図を持って流された情報とか詐欺行為とか、様々なものが既に利用されていて、今後裁判での証拠画像、映像の捏造、こういったことも懸念されている。【山口構成員(1)(2)】
- 世論工作も大衆化。つまり誰もが非常に簡単に偽動画、偽画像、偽テキストを作って世論工作をしたり社会を混乱させる力を手に入れた。この世論工作はビジネスにもなり、ある組織が大量のアカウントを作成し、AI で SNS への投稿を自動作成して、それを投稿して世論工作をすることをビジネスとしてやっている。既に幾つかの国の選挙では使われ、日本語によるものであることが指摘。AI で生み出された顔写真を使ったアカウントで偽情報を拡散しているという事例がある。またイスラエル、ハマスの衝突でも大量の偽画像、偽動画が拡散。さらに SNS や動画共有サービスが普及し、それと生成 AI を組み合わせることで、国際世論の工作や誘導が簡単になっている。つまり、AI で偽画像を作っても、それが一部の人にしか届かない、国内で閉じている場合は国際世論の誘導はできない。しかし、グローバルなプラットフォームが当たり前の時代だからこそ、それと組み合わせることで、より一層問題が大きくなって、国際世論の工作の誘導まで積極的にやられている。【山口構成員(1)(2)】
- メディアも生成 AI の渦に飲み込まれていることが懸念。例えば、アメリカ国防総省の近くで爆発が起きたとする偽の画像がネットに拡散、これをブルームバグを装ったアカウントが投稿したことによって株価が下落するということが米国で起きた。さらに、インドの主要テレビ局が誤って放送してしまった。このように、メディアのふりをした偽・誤情報の拡散、メディアが偽画像、偽映像を取り上げてしまうような誤報道が今後増えるという懸念がある。また生成 AI を報道に活用することで、誤報が今後かなり増えてくることも懸念。【山口構成員(2)】
- マイクロソフトのブラッド・スミス社長も、AI で最も懸念しているのはディープフェイクだと述べている。ある米国の調査では 63.5% の人が既に AI が作ったテキストか人間が作ったテキストか識別できていないことが分かっている。

絵画コンテストで AI で作った絵画が優勝したこともあるが、この先技術が発展し、数年以内にテキスト、画像、映像、あらゆる領域で、人の目では検証できないレベルになると予想。【山口構成員(2)】

- AI による偽画像、偽映像度が蔓延することによって、事実の写真、映像についても AI が作ったんじゃないかと疑う必要が出てくる。この技術に対抗するにはやはり技術的な対抗で、AI が生成したかどうか判断する技術の発展も欠かせない。それも時に誤判定する。実際その誤判定が問題なった事例が、既にイスラエル、ハマス衝突である。こういった中で信頼のよりどころがなくなることが懸念。【山口構成員(2)】
- 偽情報の生まれる背景にある経済と政治という 2 つの動機は、生成 AI 時代になってもおそらく変わらない。その規模がより拡大していく。誰でも簡単にビジネス、あるいは世論操作に参入できるようになる。【山口構成員(2)】
- 個人レベルとシステムレベルの相互作用、Human-Computer Interaction について、今後、人間の認識を形成する考え、信念といったものがコンピューターとのインタラクションに非常に影響されたり、深く関わってくる。ここでは search engine manipulation effect とあるが、恐らくはもっと生成 AI を頼るようになっていくだろうと思われ、もっとこの問題は深刻に展開するのではないか。【森構成員(4)】
- 最近 AI 法案がトリログで、合意をして非常に関心を集めているところだが、AI 法案を含めて今までの AI 規制論は大体 2 つの問題意識に分けられると思う。ビフォア生成 AI って、要するに焦点リスクは製品安全とプロファイリングの 2 つだった。それが情報を生成する AI になって、本格的に情報法の領域に関わるようになってきた。それは今、デジタル情報法で問題にしてきたような偽・誤情報への対策、情報環境全般への影響ということ。それがまた、直接的に国家安全保障の問題になってきたりするわけだが、言ってみれば議論の内容というのは、デジタルプラットフォームに関して議論されてきたものと近づいてきている部分が、少なくともアフター生成 AI についてはあるのかなと思う。【生貝構成員(4)】
- 全体として生成 AI の利用ということで、それに伴うリスクへの対応は重要。【落合構成員(5)】

#### ◆ 技術・研究開発の在り方

- 情報技術の進化に伴って重要な社会構造の変化が生じている部分について見ていくことが重要。以前は有体物を信頼してという状態であったが、どうしても無体物、データ、アルゴリズムなどが中心になってくるような部分が出てくる。判断の主体が、以前は人がほぼ判断をしていたという状況であったのが、AI などを含めたシステムの影響も非常に大きくなってきている。どうしても予測可能性だったり、統制がどこまでできるのかという領域が出てくる。今回の健全性については、サイバーフィジカル空間の融合による社会構造の変化が Society5.0 となる中、かなりサイバー側に寄っている部分もあり、世の中のフ

ィジカルな実態と組合せをしてデジタル化をしていく分野よりは、より一層早い取組が求められる領域であると思われるため、まず、リスク認識がどこにあるのかを考え、これを理解していくことが重要。【落合構成員(1)】

- 信頼できる情報源の確保という観点から、技術の利用をどうできるのか。オリジネーター・プロファイリングのほか、海外のメーカーやプラットフォームなども情報発信源の特定などに関して取組をしている。検証可能性をどう確保していくかは重要だが、国際的な相互運用性、国際議論の中でどういうふうに役割を果たしていくのかも非常に重要。また、サイバーセキュリティの関係で、意図的な攻撃者が存在するということを念頭に置いて対策をしていくことが重要。【落合構成員(1)】
- 技術の進化のスピードはとても速い。それに対して、しっかり法制度をつくっていただくというのは大事だが、そこにどうしても対応の速度の差があるが、その差を覚悟することが重要。【後藤構成員(1)】
- 私どもの大学でも、情報の健全性に取り組んでいる研究や学生がおり、この5年ぐらい非常に盛ん。例えば、グローバル・ディスインフォメーション・キャンペーンという国際法的な問題というグローバルな安全保障上の問題の観点、次に、情報セキュリティ心理学の観点でどう分析できるのか、それから、情報の拡散実態を実際に調べる技術とその報告など、多角的に研究開発が進められている。つまり、1つの技術ではなくて、いろいろな研究開発が必要だろう。【後藤構成員(1)】
- 防御手法について、我々のグループは世界で最初に顔の入替えや表情操作において、AIを用いてリアルかフェイクか判断可能な手法を提案。この論文は2018年12月にフォレンジックス系のワークショップで発表したものだが、引用が千回を超えており、この分野において最初に出された論文ということはこのコミュニティの中では認知されている。原理はそれほど難しくなく、シンプルな4層のCNNというニューラルネットワークを使い、当時は学習するデータがほとんどなかったため手作業でデータを集めて来て、リアルとフェイクのラベルをつけてモデルを学習させていたが、このモデルに関しては、中間層を見てみると、大体口元とか目の辺りのアーティファクトと呼ばれるようなノイズを見て、AIが与えられた顔映像に対してリアルかフェイクか判断しているというのが推察される。翌年、これを改善した方法を提案し、少しインテリジェントな方法になっているが、先ほどの手法は4層のCNNだけでリアルかフェイクを判定していたが、複数の目利きというかネットワーク、カプセルと呼ばれるニューラルネットワークを用いて、異なる3つのカプセルにより、リアルかフェイクを判断している。映像の右列の3つの映像について、違うところがアクティベートしているのが分かるかと思うが、要するに複数の目利きが参加することで、いろいろな観点から真贋判定してもらい、それを上位のカプセルで真贋判定する。これは何を意識しているかということ、新たな攻撃、新たなフェイク手法が入ってきたときも、いろいろな目利きがあるので、未知の攻撃に対してもそれなりに強いということ意識してつくった方法で、Face swapも非常に

高い精度でリアルかフェイクが判断できている。また、表情操作、Face2Faceについて、同じ人物だが口元が違うというものについても精度高く真贋判定ができています。さらに、我々のグループは、真贋判定と同時に、どの部分が改ざんされたかを推定する手法を世界で最初に提案した。当時顔のフェイクの生成手法は1通りではなく、複数出てきており、改ざんされた領域を見ることで、どの手法でつくられたか推測したいというニーズに基き、この手法を提案した。例えば、映像をフェイクと判定し、さらにこういった領域が改ざんされているということを同時に推定し、Face swapのスモーキングによって生成されたものということがこの改ざん領域を見て分かる。【越前構成員(3)】

- 最終的には人間中心の健全なサイバー社会を実現するために、多様な FM への対処と意思決定支援が重要。例えば、高度な FM 検出技術として、リアルかフェイクだけではなく、説明可能な形式でユーザーに情報提供する方法や、FM 無毒化技術として、思考誘導や誤動作・誤判定が生じないように、例えば、敵対的サンプルといった、AI を誤判断させるノイズが画像に重畳されたとき、ノイズをきれいに除去してから学習データに使うとか、通常のメディアとしての視聴を可能にする方法が考えられる。こういったツールを使って、情報の信頼性を高める社会システムの原理と技術を確立したい。これまでの研究成果として、顔を対象とした真贋判定の手法、さらに、真贋判定と同時に改ざんされた領域についても推定を可能とした手法がある。そして、画像や音声などのフェイク生成手法によってつくられたか特定する説明可能な手法や、無毒化手法の究極の目的として、フェイクからリアルを復元するような手法を確立できないかを CREST プロジェクトでは提案している。これらにより、FM がもたらす潜在的な脅威に適切に対処すると同時に、多様なコミュニケーションと意思決定を支援するソーシャル情報基盤技術の確立を目指している。人間を中心に配置すると、悪玉菌を善玉菌に変えて視聴やモデル学習に使い、さらに意思決定支援や合意形成支援をすることで抗体が生成され、健全なサイバー社会を実現することを目指している。【越前構成員(3)】
- CREST プロジェクトの成果として、フェイクからリアルを復元する Cyber Vaccine という手法がある。具体的には、リアルの顔の中央部分の特徴を顔の周辺に分かりにくいように埋め込むことをワクチン接種と呼び、ワクチン接種済みの顔画像の見た目はオリジナルと分からない。そして、この画像が公開後、Face swap という顔の置き換えという攻撃を仕かけられたときに、この画像に復元モデルを用いると、この画像のみからオリジナルのものを復元できる。いろいろなユーザー企業と真贋判定について話し合っているが、どのようにフェイクが作られたかといった来歴情報が必要で、経緯がある程度分かることが重要という指摘もあり、例えば、このオリジナルの顔画像がどのような経緯を経て別の人に置き換えられたのかが分かるということで、このようなニーズがあって開発したという手法。【越前構成員(3)】
- CREST プロジェクトの成果として、Master Face と呼ばれる手法があり、これは顔識別システムに登録された複数の顔特徴と類似するような顔を、AI によっ

で生成しようという試みで、攻撃者側の手段として脅威となる。生体情報のデータセットは非常に多く公開されているので、そういった特徴を読み取って、複数の顔識別に登録されたものと類似するような顔を生成する。登録された複数の顔と類似した特徴を持つ Master Face と呼ばれている顔を作成可能だが、この Master Face は私たちの開発したフェイク顔映像の検出手法では検出可能で、こういった成果は我々のプロジェクトのウェブサイトで積極的に公開している。【越前構成員(3)】

- GREST プロジェクトで SYNTHETIQ VISION と呼ばれるフェイク顔映像の自動検出プログラムを開発した。これは、ユーザーから見ると問い合わせたい動画の投稿というアップロードと結果を取得するというダウンロードまでの全てのプロセスを Web API として利用可能になっているが、バックでは GPU サーバがうまくキューイングしながら推論を仕かけ、真贋判定の結果を保存していくような仕組みになっている。ユーザーから見て非常に簡便に使えるということで開発したところ、現状、多くの企業からリクエストを頂き、2022 年に、有償ソフトウェアライセンスとして、サイバーエージェントにおいて、デジタルツインレーベルに登録した著名人の deepfake 検知として利用され、国内初の実用化として、サイバーエージェントで事業利用開始のためのライセンスを 2023 年 1 月から開始している。また、2022 年に、有償ソフトウェアライセンスとして、国内企業において、eKYC 向け deepfake 検知として利用され、また、さらに用途が非常に広がっており、それ以外に複数の企業から問合せを頂き、現在複数の企業に有償のソフトウェアライセンスを実施中。潜在的なユースケースとしては、エンターテインメント業界だと著作権侵害のために真贋判定を使ったフェイクメディアの削除、報道機関においては報道前の真贋判定による誤報道の防止、ソーシャルメディアでは真贋判定結果（確信度）のユーザーへの提示による偽・誤情報の拡散防止、金融機関は eKYC のなりすまし検知、法執行機関では真贋判定による証拠などの真正性確認、コミュニケーションプラットフォームでは例えばオンライン面接や試験の不正があるので、真贋判定によるなりすまし検知や不正監視というのがある。【越前構成員(3)】
- 多種多様な偽・誤情報の生成手法が出現しつつある中、我々が開発した SYNTHETIQ VISION も定期的なデータセット更新やモデルの追加学習をしているが、既知の手法で生成された偽・誤情報の検知精度を確保しながらの追加学習はかなり難しく時間がかかる。さらに、極めて多種多様な生成手法を安定的に自動検知できるかというのは大きな問題。AI 製を示すような情報をコンテンツに不可分に埋め込む電子透かしの活用にも期待。【越前構成員(3)】
- 自動検知モデルやデータセット、ベンチマーク自体における課題もある。研究レベルでは様々な提案がされているが、ほとんどが現実の環境を反映していない。課題を解決するために、可能であれば国が主導して、産学連携による開発・実証が極めて重要。さらに、プラットフォーム事業者から、AI 関連事業者からでもいいが、研究者に対してデータを提供いただけると非常にありがたい。【越前構成員(1)(3)】



- 自動ファクトチェックの課題。自動検出と相互的な活用が期待されるが、重要なのは、クエリ情報と真贋を比較するために参照する信頼できる情報源について、誰がどのように収集して、新しい情報をどのようにメンテナンスしていくのかも今後大きな課題。【越前構成員(1)(3)】
- AIを活用したコンテンツ・モデレーションには、偽・誤情報に対するAIを用いた自動検知や自動ファクトチェックが提案されているが、課題もある。自動検知は、新たな生成AIで生成された偽・誤情報は高頻度で出現・拡散されているが、AIは未知の手法で生成された偽・誤情報の検出は苦手であり、学習する定期的なデータセットのアップデートや、モデル学習又は追加学習等が必要であるが、モデル学習とのこれまでの手法で生成された偽・誤情報の検知精度を確保する必要があり、この調整が最も難しい。【越前構成員(1)】
- 自動ファクトチェックについて、研究的にはテキストを対象に、テキストの文言がファクトなのか否かという研究がある。ただし、ドメインが限定で、学術論文のアブストラクトのデータセットがあったとき、テキストの文言が科学的に妥当か検証する問題がある。手法やアルゴリズム的には研究論文が出ているが、実装するとき問題となるのが、信頼できる情報源として我々が研究で使っているのが、たかだか論文のアブストラクトとかそういうレベルであること。現実で使うとなると、大規模な信頼できる情報をどうやって収集してメンテナンスしていくかも重要で、この課題をある程度解決できれば、研究開発が進むのではないかと。【越前構成員(3)】
- AIを用いた自動検知について、AIというのは100%真贋判定できるわけではない。場合によっては間違っただけで誤判定した結果、責任もアカウントビリティもあるが、技術者以外との連携が必要。技術的には、真贋判定はある程度の精度で出るが、エラーが出てきたときにどう対処するかが、技術的には対処できないという中で、場合によっては弁護士の先生方とかに相談しながら、コンテンツモデレーションを進めていければ。確率的には提示できても必ずエラーが存在する中で、技術だけでは解決できないところを他の分野の専門家と連携しながらモデレーションできれば、技術としては非常に進展がある。【越前構成員(3)】
- 技術が起こした問題はできる限り技術で解決する必要。しかし、技術は、実はそれ自体が開発の段階又は利用の段階で何らかの考え方や思想をはらんでいる可能性が十分ある。アメリカの憲法学者ローレンス・レッシングのアーキテクチャ論でも定義されているが、これを野放しにしてしまうと、いわゆる言論規制の道具、ないしは自由で公正な経済活動を阻害するリスクというものを技術自体が強くと内包してしまうという可能性がある。こういった問題をできるだけ回避していくということが最終的に世の中に普及させていく、イノベーションを広げていくときに重要。技術の使い方及び自律的な自由に基づく価値の最大化を目指して、どのように我々はこれを手なづけていけばいいのか。抑制したり、守ったりするような技術についても、どのようにそれを位置づけていけばいいのかを考えていくことが重要。【クロサカ構成員(1)】
- ディープフェイクみたいな議論はリテラシーの範囲を超えてくるので、やはり

技術的な開発とか進歩というものを欠かせない。だからこそ、事業者と連携して、継続的に問題を把握して対処を考えていくことが大事。【山口構成員(2)】

- オリジネーター・プロファイル、AI が作ったものかどうか判断する技術、偽・誤情報の検知技術、そういった対抗技術への研究開発の投資はもっともっとしていく必要。民間での研究開発の活性化を促すことも大事であるし、また、研究支援を通じて、研究機関や大学での一体的な開発を促進していくのが大事。【山口構成員(1)(2)】
- もう一つ大事なのが、技術が大衆化されるということ。研究開発を進めるというだけではなく、その技術を、例えば、メディア企業やファクトチェック組織が使える、さらに、一般の人々が手軽に使える、又は、分かる状況にしていくことが大事。例えば、SNS 上の投稿された画像が AI が作ったものかどうか自動的にラベリングするとか、AI が作った確率は何%か分かるなどの状況が望ましい。そういった検証技術は、ファクトチェック組織とかメディア企業に特に提供することが大事。【山口構成員(1)(2)】
- AI などの技術の文脈に応じた論点の整理と議論を行うこと。技術の問題は技術で対策するということが誤情報・偽情報のところでも議論されているが、レッシングが出している『CODE』みたいな話も含めて、いろいろな規律を使いながら議論していくことが重要。【江間構成員(1)(3)】
- 情報通信技術の進展の早さも考慮していく必要。【落合構成員(5)】

#### ◆ 研究機関等との連携・協力の在り方

- プラットフォーム事業者に求められることとして、研究者と連携した研究について、Innovation Nippon のような第三者を入れた研究とか、あるいは研究者へのデータ提供。X というのはその方向から逆行してしまったわけだが、そうではなくて、しっかり研究者にデータ提供するなど情報環境の研究を促進するということも大切。【山口構成員(2)】
- 研究者の立場としては、プラットフォーム事業者の取組、データ、コンテンツモデレーション、アルゴリズム等に関する透明性は高めていくことは、特に大事だというふう感じており、現在 X のアカデミック API は提供停止になっており、これまで有益だった偽情報流通の全体像の把握、特に大規模、そして過去に遡った解析というのが非常に困難。また、今後は X のみならずプラットフォームサービスとして SNS や掲示板、ニュースポータル、検索エンジン等々においても、研究者に対して、データやアルゴリズム等に関して分析可能な形で提供するような枠組みがあれば、もう少し全体像の把握というのが容易になるというふう考えている。例えばティックトックは、現在研究者向けの API を公開しているが、こちらは欧米のみに提供中で、国内のティックトックにおける偽誤情報のティックトックにおける全体図を見たいというふう考えているが、そこが今なかなか難しいという現状。【澁谷構成員(2)】
- プラットフォーム事業者から、AI 関連事業者からでもいいが、研究者にデータを提供いただけると非常にありがたい。研究者が使っているデータセットとい

うのはかなり特殊な環境で、例えば、テレビ番組のような背景がずっと同じような人物映像とか、人が1人しかいないとか、かなり理想的な環境で顔が照らされたような画像。これを実際の用途に使おうとすると、例えば、スマホのカメラのようなものでも光の環境は多様で、顔の向きも全然、斜めからとか違う向きのものとか、自然な環境とは、我々がやっている研究で使っているデータセットは大きく違う。実際に事業者から生のこういったデータを頂くことで、例えば、真贋判定における精度は確実に向上するし、さらにそういったものを想定してアーキテクチャ自身も開発できるというところで、実際にそういったデータ提供があるところの分野は極めて大きく進む。現状は、研究者は若干たこつぼ的で、与えられた研究用のデータセットに対して一番を取るというところに非常に熱心になるので、技術開発、本当に社会実装を目指すのであれば、事業者等から実際のデータを頂きながら開発するというのが非常に重要。【越前構成員(3)】

- アメリカの議論で日本で参照できそうな点は、なかなかこれ、難しいところがあり、私自身は、結構、アメリカを反面教師で見ているところがある。基本的に政府が規制を入れないのがアメリカでは前提になっているが、アメリカ版 DSA、プラットフォームの透明化法みたいなのがアメリカでも議論されていて、あそこでは研究者がデータにアクセスすることを結構重要視している。そこは DSA でもそういう規定が入っていますけれども、研究者がプラットフォーム側のデータにアクセスして、それこそ偽情報の影響とか、そういうものもきちんとリスク評価できる、そういう状況にするのが一つ、参照できる点。【水谷構成員(4)】
- プラットフォームと研究者との連携という中で、研究者へのデータ提供や、技術、研究開発の在り方もある。【落合構成員(5)】

#### ◆ 緊急事態（災害、サイバー攻撃など）への対応の在り方

- デジタル空間における悪意ある外部からの攻撃（偽情報・サイバー攻撃も含む）に対し安全保障も含めた対策をいかに構築できるか。【江間構成員(1)】
- 大規模災害や国家的危機等の緊急事態時におけるレジリエンスを検討しているか（特に人-技術システムの関係が複雑に関係しあっている場合、インシデント対応と復旧に向けた論点整理ができていくかどうか）。【江間構成員(1)】
- 情報流通については、いわゆる平時の場合と、自然災害や国際的紛争といった状況をはじめとする緊急事態、その両方がある程度区別しながら、両方を視野に入れた議論をしていくことが重要。例えば、集中的な偽情報・誤情報などの事象が生じた場合に、関係事業者に法に基づかないインフォーマルな対応のお願いをするのか、又は、法で枠組みをつくって、民主的なプロセスの中で対応の要請をしていくのか。そのどちらが望ましいのかは、国際的にも議論が分かれているが、そうした場合を含めて、多様なステークホルダーが関われる形でのレジリエンス確保の仕方というものが重要。【生貝構成員(1)】
- 緊急時の対応について、欧州の DSA は、イスラエル・ハマス紛争という国際事

案で、それに関わる情報の健全性対応として、デジタルサービスコーディネーターに関する指定が急遽、来年2月の期限から4ヶ月前倒しになり、EUがMetaやTikTokに対して指示を出したというニュースもある。悩ましい話ではあるが、時にはこういう有事の対応ということで制度的にも動いて頂く必要。【後藤構成員(1)】

- 健全性の確保は非常に重要。ただ、形は違っても、この問題は過去から将来にわたってずっとあるもので、永遠の課題に取り組んでいるという意識も必要。よく言うシルバーバレットがない、つまり、簡単に決着できるものではないということを実感する必要がある。健全性の確保のためには、社会の変化、つまり、サービスも技術も教育の中身も変わっていく変化にはじめから対応しながら、できたら将来の変化を先取りした取組の継続がすごく大事。1回の取組で解決できるものではなく、継続し続けなければならないという覚悟が必要。多角的かつスパイラル的に取組を継続する社会的な仕組みづくりが重要。法制度、教育・研究・メディア、政策、プラットフォーマー、それぞれが相互に少しずつ、タイミングに合った取組をして全体で連携・協調しながら解決に当たるという姿勢が一番に大事で、皆でうまく知恵を出し合う場が非常に大事。アジャイル・ガバナンスや、技術だけでは解決できないなど、この辺りはサイバーセキュリティ対策と同じ。【後藤構成員(1)】
- サイバーセキュリティ対策では、1つの技術や1つの法制度で解決できると思っている方はいない。長期戦覚悟で取り組み続けている。この中で、例えば、社会的な仕組み、いろいろな企業のCSIRTや、NICTやIPAの研究開発や情報の収集活動、幅広い人材育成、研究開発、また、それをビジネスにする取組、そういうものが、全体として必要ということを実感し、それに関して、政府・省庁も戦略面、対策面で政策を出しているという状況。そういう意味で情報の健全性確保とサイバーセキュリティ対策は似ている。【後藤構成員(1)】
- リスク評価と軽減をどうすればいいのかというと、なかなか包括的過ぎる、抽象的過ぎるというのもあるので、欧州委員会がある種、非常に強く後押しする形で形成するコードオブコンタクトや、特にディスインフォメーションは、どちらかというと平時というよりは、突然何かが起こったことが問題になったりする自然災害や武力紛争とか、そうした時には緊急で、これはかなりデジタルプラットフォーマーの裁量というのを重視した形にはなっている、緊急対応、危機対応メカニズムというのが置かれていて、こういう行動をこんな形で取ってくださいという要請を、まさに法に基づいて出すということになっている。その部分はまた別途、危機プロトコルという形で決められていたりするところ。これらがちゃんと守られているかどうかは、なかなか公開できる情報と公開できない情報もあるし、しっかり中に入ってデータを見ないと分からないということで、必ず独立した監査を徹底的に受けて、その修正要求にどう対応したかといったようなところもまさしく義務として入っている。ここの中でも、アメリカの文脈でのデータアクセス提供義務といったようなことも、外部の信頼できる研究者等に限った形で認めたりということもやっている。【生貝構成員

## (4)】

- 危機対応メカニズムというのがどうやって機能するかについては、すごく限られた状況というふうに限定はしているが、例えばコンテンツ・モデレーションのリソースはちゃんと増やしてくださいねとか、利用契約やアルゴリズム、システムを調整するようなことなんかも検討してくださいといったようなことが含まれている。【生貝構成員(4)】
- アメリカは修正1条と、それから特に巨大なプラットフォームやAIの会社さんが地元にいるということもあるので、非公式なお願いによってフェイクニュースを、いろいろと戦争のときだとか、消してもらったりすることが常態化しやすい。あるいは、することができる。しかしそれは、公法的な民主的なコントロールや透明性というのがなかなか及ばなくなってしまう。そうした時に、そのルールづくりというのは極めて様々論点あるが、非公式な要請と法に基づく要請というもののどちらが望ましいのかといったようなことは、まさに様々な観点から考えていく必要があるんだろうと思う。【生貝構成員(4)】
- サイバー攻撃がなされる中で、結局、攻撃がそういった偽・誤情報の拡散につながるということにもなる。これは発信者、全体に対して関わってくるであろう。【落合構成員(5)】
- 有事におけるレジリエンス確保、有事のほうが、より偽・誤情報の拡散が出てくる中で、自治体やメディアなどが特に、一定の立場を担うことはあろう。【落合構成員(5)】
- 伝送の部分においても、サイバー攻撃への対応は重要。【落合構成員(5)】

## ◆ 国際連携・協力の在り方

- 情報技術の進化に伴って重要な社会構造の変化が生じ、地理的にもグローバルというのと国内というので以前は分けることができていた部分もあったように思うが、国際的な影響というのも直ちに受けるような可能性が高まっている。【落合構成員(1)】
- 法制度と技術開発という観点において、技術進化のスピードはとても速く、それに対し、しっかり法制度をつくるのは大事だが、そこにどうしても対応の速度の差があり、その差を覚悟することが重要となる中、デジタルサービスの事業者にとっても法制度の裏づけは必須だが、法制度の対応には非常に時間がかかってしまう。せめて、国や地域の差が少なければという本音があり、企業からは、GDPR対応として、個人情報やプライバシーが国や地域によって異なり微妙な差が出て、この差への対応がすごく大変。巨大プラットフォームも個別対応しているが、日本のデジタルサービス事業者には個別対応のリソースがなく、グローバル競争上も不利になるという課題。【後藤構成員(1)】
- 偽・誤情報問題は国内で完結しない。だから、国際的な連携とか情報共有とか対策の実施、これは必須であるというふうに感じていて、例えばG7広島AIプロセスでも、国際機関と協力してプロジェクトベースの取組をこの偽・誤情報問題について実施していくということが明記されている。この間登壇した

IGF2023でも、EPaDというものを公開して、各国のこれまでにやった取組ということ、代表的な取組をまとめたわけだが、こういったことを参照にしながら、いよいよ一堂に会する会議とかをやって、それで情報共有しながら、こういうふうにしたらいんじゃないかということ議論できる、そういったことがもっともっと進んでいくといいんじゃないか。【山口構成員(2)】

- 国立情報学研究所に、シンセティックメディア国際研究センターを設置した。人間中心のAI社会を実現するために、多様なメディアの生成、メディアの信頼性確保、意思決定支援のための研究開発を、実世界の課題を取り上げながら、国際的な拠点として、新たな科学技術分野と研究潮流の創生、国内外の学術機関との連携、産学官連携を通じた実社会適用を推進している。【越前構成員(3)】
- プラットフォーマーは多くの国や地域を対象にしているのだから、基本的には何か全く違ったものを作ろうとするのはプラットフォームさんの対応としても大変だし、例えば、どんなことを透明化してほしいかみたいなことは、G7とかOECDなどでもいろいろ議論をしながら、しっかり標準化した形で我が国含めて各国、法制を考えていくことが現実的なんじゃないかなと考える。【生貝構成員(4)】
- 国際的な連携という部分も、実際には、かなり大きいというか、1つの国で閉じているわけではないということになってくるので、とても重要な点になってくる。【落合構成員(5)】

## 2. 主に情報発信段階の課題

### ◆ 発信情報の信頼性を得るためのコスト増への対応の在り方

- 巧妙な偽・誤情報の生成・拡散、メタバースの問題、情報流通・データ利用のグローバル化・広域化とスピード、それに伴う国際的な協調の必要性などのリスク増に伴って、情報やデータのトラストを毎回確認しなければいけない。このためのコスト増が本当に大きくなっていて、社会活動、これは個人でもコミュニティでも企業でも、全てにおいて質の劣化、効率の劣化が避けられない状況。【後藤構成員(1)】
- 大学のゼミでビデオジャーナリズムとファクトチェックを教えている。ビデオジャーナリズムとは、学生が自分で、カメラを持ち、相手を口説いて撮影し、ナレーションやテロップを入れる。どこの誰が来たというところから始めるため、公開を前提にした取材は、学生にとってはかなりの関門。あなたの行いや言葉が世の中にこういう役に立つから、あなたをさらし者にしに来たと言って相手を口説かなければいけない。このようなニュースが何のためにあるのかを学生が言語化するきっかけになるワークショップが、この会合でもヒントになることもある。【奥村構成員(1)】
- デジタル・ストーリーテリングとかファクトチェックの能力というのはメディアにとってプラスアルファで、ぜいたくと言われるもの（ニュースラグジュアリー）だが、そうやって手間暇かけたニュースは、ジャーナリスティック・エンタープライズと呼ばれる形で、説得力のある情報になる。表現のオプション

としてネットになると、テキスト・写真・映像・CG等に境界がなくなると、表現の豊かさが情報の説得力にもつながる。そうした情報をどのようにしてメディアが発することができるかというのは、これからの大きな課題になっていく。

【奥村構成員(1)】

- 「The Elements Of Journalism」のようにジャーナリズムのルール・目的・ミッションをニュースの消費者が分かるような言葉で表現し、日本はニュースメディアであるだけで信頼等された社会であったが、民主主義におけるニュースとは何かを考え直す必要。【奥村構成員(1)】
- 私たちが今扱っている映像というようなものも改めて考える必要が出てくる。例えば首から下しか映っていないインタビューというのがどういう意味を持つのかというようなことは、もう少しちゃんと考えられていいことだと思っている。【奥村構成員(1)】
- 世の中で流通する情報の中で重要な部分を出しているニュースというものを中心にして、情報の正確さというものがどのように担保されているのかや、世の中の人に信用してもらうためにどのような仕組みがあるのかについて、インプリケーションがある。【奥村構成員(1)】
- ジャーナリズム、ジャーナリストは Unlicensed Job と言われており、これさえ覚えていけばオーケーというような、コンクリートな基準があるものではなく、アートやサービスの側面もある。その中で最も重要な技術として、『ジャーナリズムの原則 (The Elements of Journalism)』(2001年刊行)という本が挙げる10の原則のうち3番目に、情報を検証する能力 (discipline) がある。かなり厳しく訓練をした上で情報を検証する能力を身に付けるというような意味合いと思われ、ジャーナリズムの原則の1番目にある truth (真実) に迫るためのものとして検証が必要であるという建付けになっている。一定の手続は共通のもので論理的なものだが、経験の積み重ねなどに依存する領域もあり、これが結構複雑。【奥村構成員(3)】
- 日航ジャンボ機事故を取材している北関東新聞という架空の新聞社のお話である『クライマーズ・ハイ』における事故調査委員会の委員長にエース記者を当て、締切りを1時間半延ばし、スクープを出すかどうか決めようというシーンや、「#Me Too」の原動力となったジャーナリストに関する『SHE SAID』における実名告発記事を出したニューヨーク・タイムズの記者のスキルなどは、おそらく情報の正確性、安心を考えるためのモデルになり得る部分がある。しかし、ニュースや読者が視聴者に信頼されるのに、すごいところがあるということでは通用しない。その手続が確実であること、そして誠実であることが保障されなければならない。そのためにメディアは倫理規範 (Code of ethics) が存在。この点、『ジャーナリズムの原則』という本は、その基盤を明確に表現している。【奥村構成員(3)】
- 日本のニュースメディアが倫理規定として公開しているものには、項目は網羅されているものがあるが、How がなく、目標の列挙になってしまっている。各メディアは内部文書としてもっと詳細なルールを設け、記者ハンドブックや、

イントラネットでスタッフが検索できるようなシステム等があり、自社や他社の事例等が書いてある。そのまま公開しなくてもいいが、自分たちがどのようにしてニュースの安心や安全や正確さを保つのかという、自分たちの行動原則というようなものはもう少し明らかにすると、もしかすると社会全体がこういう手続が真つ当なものというような認識も広がっていくのではないかと思うが、そこまで日本は行っていないというのが現状。この点、民放もNHKも、新聞社も、非常に問題意識が深い方はいっぱいいるが、組織として動けず、マネジメントが決断するかという問題と、例えば、倫理規程の中にも、複数の情報源を取るとということが書いてある社もあるわけだが、その部分だけでも公開するにあたり、それをより分ける作業が膨大で精緻にやらなければいけないとなるとそこまでのヒューマンリソースを割けないから後回しになってしまうという事情も非常に抱えていたりすることも理解している。また、欧米ではジャーナリズムスクールという大学院の教育システムがあり、それを経て一定のリテラシーと能力を得たものがメディアで仕事をするようになってきているが、日本のメインストリームのメディアは長らくインハウス・トレーニングをしてきた。かえって、メディア学があまりハッピーな発展状況ではなかったというのもあるが、そちらを専攻した人をあまりとらないというような傾向もあった。日本はメディアがメディアであるだけで信頼されていた時期がすごく長く続いていた、幸せな国だったのかもしれない。ただ、今はもうそうではないので、信頼をつくり直す社会的な営みに是非首を突っ込んでもらいたいと思う。【奥村構成員(3)】

- 信頼を取材の過程でどういうふうに得るかというシステムについて、これは匿名の情報源をどうするかという扱いに非常に影響するが、これについても扱いは様々で、匿名の情報源は一つだけでは絶対にニュースにしないというところまで明記している社もあるが、日本の政治ニュースを見ると、そういうようなものはほとんどなく、例えば、日本の政治欄を見ると、筋もののシングルソースの記事がどんどん乱発されており、普段のプラクティスからして匿名情報等に関してかなりカジュアルな面がある。これは取材元と取材先とも日々渡り合いながら、それをここまでは明かせるでしょうという形で、多分せめぎあいながら少しずつできることを増やしていかなければならないというようなことを怠ってきた結果なのではないか。匿名の情報源の取扱いといった手順も決められていないまま、何となくそういうプラクティスだけがあり、それも明文化されてなく、先輩から見よう見まねで取材をするというようなことになると、参照できる基準がなく、どう是正するかについての言及は甘いということになる。程度の差こそあれ、メディアには欧米でもそういう問題は多数発生しているが、参照する基準が非常に弱いと、そういうことは起きざるを得ない。実は参照する基準は持っていたりするわけで、それをなぜ公開して、議論にしないのかというのは、理解に苦しむ部分でもある。【奥村構成員(3)】
- 虚偽情報・誤情報に関する訂正、打ち消し報道が効果を持つためには、その発信元への信頼の獲得が重要。【安野構成員(1)】



- 社会調査をしてると、例えば、政府やメディアへの信頼が問題になってくることがあるところ、信頼し切っているというのは、これまた批判的に考えられていない可能性があるので、多少は疑うことも悪くないという指摘もあるが、例えば、リテラシー教育を出すような省庁、学校、メディアやプロバイダーへの信頼がまずないと、重要な情報、あるいは正しい情報であってもそれを信じてもらえないということは当然起きてくるので、提供側の信頼の向上・回復は本当に重要な課題。【安野構成員(2)】
- 信頼できる情報源、複数の対策を組み合わせで行わないと、最終的にはどれか一つだけで拾いきれないことは大前提だと思うが、信頼できる情報源の確保という点は非常に重要。【落合構成員(2)】
- もともとは電波の独占というものがあり、放送免許が一つの信頼性担保になっていたというところもありつつも、新聞に関してはそういったものもなく、他メディアにおいてもそういう中ではあるが、メディアと称されている方々に対する信頼感は一定程度あったところ、その中で独自の手法をそれぞれとられて、信頼性のあるスクリーニングをかけられていたと思っているが、インターネットの社会になってくる中で、いろいろな情報源であったり、どういう過程でこの情報がつくり出されているかが相対化されてきたり、個人であっても信頼性のある情報を発信できるような方も出てきたりするであろうという中では、信頼性担保に関する取組を行っていて、それをこういう検証をしてより向上させているという、ある種のディスクロージャーとそれに対する説明が重要になってくる。そういうものをしっかりやられているものは比較的信頼性が高いというか、受け手側の評価もされやすくなる。信頼性の高いものは何なのかといったときに、適切に情報公開を行って一定の手続を定めている人たちが信頼性の高い方々なんだというふうに見ていくのがよりよい。なかなか内容規制というのはできないと思うが、そういう適切な取組をしている方がなるべく報われるような形というのは比較的考えやすい。【落合構成員(3) ※一部再掲】
- オリジネーター・プロファイル(OP)はその情報を誰が発信したのかをユーザーの求めに応じて検証(ベリフィケーション)できる技術的仕組みであり、具体的には、記事を発信する主体及び記事そのものに対して、言わば電子透かしのような識別子を埋め込み、これをまた電子証明技術を用いて鍵をかけていく技術的方法を用いて、途中の改ざんが困難な存在証明を発信主体と記事そのものに付与していくことにより、発信元とその情報を発信したという事実、これを確認していくことを実現。いわゆる情報の真正性、確からしさを高めるということを目指した技術。当初 OP は、デジタル言論空間の情報流通の適正化を出発点として考えており、参画されている方々、例えば、メディア企業や通信企業が割と多いが、今般、生成 AI がかなり台頭した時代においては、教師データや生成 AI が生成したものについての取扱いについても、OP は一定の有効性があるのではないかと御指摘をいただいております、こういった可能性も踏まえて、今、開発中。直近ではオープンインターネット上での、非公開ではあるが、組合員をメンバーとした実証実験を既に開始しており、第1フェーズは

おおむね終了しつつあり、一応作って動いたというような段階。【クロサカ構成員(1)】

- 外形的な情報に対する評価と内部の規律、この2つをどのように接合させることによって、我々は信頼性や真正性を評価・検証し得る状態に至れるところ、ここが実はミッシングリンクになっているのではないか。例えば、エンジニアリング観点でいうと、ディープフェイクは明らかにオリジナルから改変されているという評価が、もちろんかなり巧妙にはなっているものの、一定程度可能。また、内部の規律として健全なジャーナリズムを目指す方々の自主規制、あるいは、欧州では法制化も進んでいるので、共同規制的なアプローチも含め、取組をしている。この2つがリンクしている状態であれば、情報を受け取る側は一定程度信頼することができる、ないしは何かエラーがあっても訂正してくれるというようなぎりぎりの信頼感の担保が可能になるわけだが、この2つがばらばらな状態だと、誰が何を信じればいいのか、どのような理由で信じればいいのかということがいま一つはっきりしなくなってしまう。ここをただユーザーリテラシーだけに帰するのではなく、ほかの方法も含めて、結びつきを持つことが、誰がどのように担えばいいのか、どのような結びつきがあるのかを検討してみるということが必要になってくるのではないか。誰が結びつければいいのかについて、拙速に政府や行政と言わなかったのは、本当にそれでいいのかという吟味が必要だが、もしかするとそこに何らかの制度的な担保を持つことによって、結果的に多くの利益が得られるとすれば、何らかの役割があるのかもしれない。情報そのものの性質と、その情報を生み出した主体における、中でどうやって生み出したかという、そういう話のつながりとして、外形的な情報の検証と、例えば、情報発信者の内部の規律のどちらかだけをやっている、もしかするとあまり意味がないということかもしれないと思いましたので、その辺りが論点の一つに今後なり得る。【クロサカ構成員(3)】
- 恐らく情報をつくる、あるいは取り扱う主体と、そしてその内部にはいろいろな規律が当然にある。その規律に服する主体が情報を発信することが結びついてきた時代は、一定のメディア環境といいますか、まさに発信した情報の評価がすなわちその主体の経済的な利益も含めた評価に関わっていて、なればこそ情報を発信する、あるいは取り扱う主体が内部規律をしっかりすることによって自動的に情報のレベルを確保することと、自動的な一致があった時代。それに対してデジタル空間、今現在の情報技術の現状がそうなのか、そうでないのだとすればどういうことをすればいいのかという話。【宍戸座長(3)】
- 情報元の信頼性について消費者が確認できる方策が必要。【増田構成員(1)】
- 取材等に基づく信頼できる情報・コンテンツを作成・流通させることへのインセンティブの設計・強化（その前提として、真正性が担保された情報・コンテンツなのかどうかをユーザーが認識できることが重要。こうした技術の1つとしてOP）。【山本構成員(1)】
- 信頼できる情報を発信するための経済的インセンティブ、つまりガバナンスをきちんとしながら生成した情報を提供している人たちに対して、経済的インセ

ンティブとして、より大きな収益を与える取組、例えば、信頼性のあるコンテンツをつくり出しているところの広告枠は高めに設定されるなど、ブランドを生み出していく方向性が必要。つまり、信頼できるものを出している人には助成をつけてあげるといようなところなど何か制度設計できればいいのではないか。【水谷構成員(2)】

- テレビでは昔、番組の視聴率ではなく視聴質の研究というものが行われてた。クオリティーが収益に結びつく仕組みというものが、今の状況下でなかなか難しいことは重々承知しているが、何か知恵を出し合って新しい評価指標を作って、アテンション・エコノミーの現状に少しでも楔を打ちこまなければならないのではないか。【水谷構成員(2)】
- 質の高い情報を提供しているところの広告収入を増やすというのは、仕組み的にはすごくアグリーするが、誰が信頼できる情報というふうに定義づけるか。政府はありえないので、第三者機関があるが、第三者機関の決めたことに全広告マーケットが従うというのは極めて強力な状況なので、法的枠組みでもかなり難しそうなので、なかなか実現しづらいのではないか。ただ、例えば、ヤフーニュースであれば、今ページビュー数以外の指標として、参考になった等のボタンが3つあって、そういったものに応じて記事の掲載料を変えるということをやっている。そうは言っても、そういうボタンは結局センセーショナルで過激なほうが押されそうな気がするので、もしかしたらあまり実効性がないかもしれないが、そういう試みが今始まっているところなので、より一層そういうふうにクオリティーに目を向ける、クオリティーがお金につながるような仕組みづくりというのを積極的に進めていくという方向性について非常に同意。【山口構成員(2)】

- もともと基幹放送制度というのは、質の高い情報を提供しているところの広告収入を増やす機能を、ある一定のときから意図せず持っていたということができるかもしれない。民間放送、地上波テレビジョン放送などが、非常に強いマス広告の仕組み、制度とそれから社会経済実態で裏返しになっていったというのが、まさにそうなのかもしれない。【宍戸座長(2)】

- 発信者側の論点としては、発信主体の真正性、信頼性の確保があり、これはもちろんメディアなのか、ファクトチェック機関なのか、それとも個人に近い方なのかで、それぞれ違う部分はあるかと思うが、どういう形で真正性、信頼性を確保していくかもあるかと思う。特に偽、誤情報を拡散された場合の効果的な発信手法、これをどのようにしていくのか。【落合構成員(5)】

#### ◆ 発信力強化のためのガバナンスの在り方

- 信頼できる社会の基本情報を取材・編集、ストーリーテリングできる事業者・人材の確保と育成、そしてそのために必要であると考えるのが効果的なオーディオビジュアル表現であるテレビ的手法、ノウハウの還元。【脇浜構成員(1)】
- AI等最新テクノロジーを地域コンテンツ制作にも活かしていけるのではないか。例えば、AI無人カメラでのスポーツの試合の中継など、コストを削減しつ

つ、地域コンテンツ流通を増やしていける可能性を感じた。ICT 活用のためのリテラシー向上に関するロードマップにあった言葉で言うと、デジタルコンテンツの作成・編集に関する能力。【脇浜構成員(1)】

- デジタル空間は決して健全オンリーのものではない。界隈性や猥雑性があるが、健全に過ごせる公園のような空間も分かりやすく確保することは必要で、その際、情報の受け手としての守りの強化として、ファクトチェックやオリジネーター・プロファイル等とともに健全な情報を生み出す人材というのをできるだけ多く育成する攻めの施策が重要。健全だけでは人は見てくれない。健全で面白い、魅力的な情報を生み出す人材を育てる、そこにリソースをつぎ込むべきではないか。【脇浜構成員(1)】
- **メディア側について、やはりデジタルでの発信強化を求めていくことは必要かと思う。【落合構成員(5)】**

#### ◆ 持続可能なファクトチェック推進のための仕組みの在り方

- 誤情報対策は主にプレバンクとデバンクに分けられることが多いが、誤情報が拡散した後の事後対応はデバンクと呼ばれており、ファクトチェックによる誤情報の訂正などはこのデバンクに該当。【田中構成員(1)】
- ファクトチェックの認知度が他国と比べて低いというような総務省の調査報告がある。また、別の調査ではアクティブなファクトチェックサイトの数が、例えば、アメリカは70件以上、インドは30件以上に対し、日本は5件という現状。これらの指標が即座に日本のファクトチェック行動そのものが弱いということを導くものではないかもしれないが、こういった違いが生じているのかを詳しく検討する必要。日本のデジタル空間が誤情報に対してどの程度脆弱性をはらんでいるのかについて、もう少し詳細かつ客観的に理解する必要。【田中構成員(1)】
- 認知科学の観点からは、デバンクにおけるユーザーの認知バイアスの影響が気になるところ、最近行った実験では、誤情報を信じているかどうかでファクトチェック記事のクリック行動が異なるという結果が出ている。具体的には、4割ほどの実験参加者は、誤りだと既に知っているような情報のファクトチェック記事はクリックするのに対して、誤情報を信じている場合はファクトチェック記事を選択的に避けるというような行動傾向も見られる。つまり、ファクトチェックの取組を充実させていくことが期待される一方で、訂正情報をデジタル空間で広くアクセス可能にするということと、誤情報が人々の認知に及ぼす悪影響に対処するということを区別することが重要。研究結果からは、これらの間に心理的なハードルがあるということが示唆。【田中構成員(1)】
- 訂正情報を提示した後、ユーザーがそれを読み、その訂正情報を記憶していても、なお誤情報を信じ続けるというような現象も確認され、心理学の分野では欧米を中心に「誤情報持続効果」という名称でメカニズムの研究が進められている。また、繰り返し誤情報に接触すると、情報処理の流動性が高まるがために、誤情報が正しく感じられるというような認知バイアス（真実錯覚効果）が

あり、もしこのメカニズムが訂正情報にも働けば、訂正情報も同じように繰り返せば誤情報の影響を緩和できるかと言ったら実はそうではなく、誤情報の3倍の頻度で訂正情報を出しても誤情報の影響が取り除けないというような非対称性というものも確認されている。こういったものは一度信じられた誤情報を事後的に訂正することの難しさの一因となっている。【田中構成員(1)】

- デジタル空間における情報流通の健全性確保と言ったときに、このようなユーザー側の認知的特徴（誤情報持続効果や真実錯覚効果等）が健全性の確保にどのように関わってくるのかについて、リテラシーの向上と併せて検討すべき。【田中構成員(1)】
- 既に誤情報が発生していたり広がっている場合は、事後的にその影響力を弱める介入方法が必要で、これがデバンクと呼ばれる手法。例えば COVID-19 に関する誤情報への対策として WHO や CDC が特設ページを開設して、それを訂正しようとした試みや、多様な誤情報を扱うファクトチェックサイト、マスメディアによる訂正情報を含むニュース、あるいは SNS ユーザーによる集合的な訂正など、様々な形式でデバンキングというものが日常的に行われている。このデバンクは一般的に用いられる手法だが、一方で社会科学研究を中心にその効果検証が行われていて、1度、誤情報の影響を受けると、それを事後的に緩和することが必ずしも容易ではないことも示されている。なぜ訂正情報を出しているのに、まだ誤情報が拡散し続けるのかとか、なぜ誤情報信じ続ける人がいるのか、こういった問いと密接に関わる心理現象が誤情報持続効果と呼ばれるもの。単に訂正情報を視聴していないとか、それを知らないから誤情報をうのみにしているパターンももちろんあるが、それに加えて、訂正情報を視聴していて、訂正情報が出されていることを知った後もなお誤情報信じ続けたり、その影響受け続ける心理現象を指す用語。これまでの心理学研究では、この訂正情報に視覚的な注意を払っていても、また訂正情報の内容を記憶していてもこの現象が生じることが確認されていて、認知的な処理のレベルの中でも高次認知レベルの観点からこの現象の解明が進められている。つまり情報を健全に伝達する時に、単に物理的に情報が遮断されることも重要な観点だが、ここでフォーカスが当てられているのは、物理的にはスマホのニュースアプリを開いていて、そこに訂正情報の記事が表示されていても、その後で心理的な障壁が生じることによって、その訂正の効果が制限されることに関する研究。【田中構成員(4)】
- 誤情報持続効果が生じる心理的な要因は、複数明らかにされている。その一つに、真実錯覚効果という人の認知的な特徴がある。これは、同じ情報に繰り返し接触するとその情報を正しく感じられるようになっていく現象。この背後には、人が同じ情報を繰り返し接触していると、親近性だとか情報処理の流暢性が高まることによって、それが正しさのシグナルとしてヒューリスティックに活用されていくというような認知的なメカニズムがある。接触頻度が1回の場合より3回繰り返し接触すると、その影響力が強くなり、これが真実錯覚効果の1例。もし、繰り返し接触することで真実錯覚が生じるのなら、訂正

情報も同じように繰り返し行えばよいのではないかと思われるが、実験ではそのようにうまくいかなかった。例えば、誤情報を1回提示した後で訂正情報を1回提示した場合と3回提示した場合は、有意な差は見られない。つまり、誤情報の3倍の頻度で訂正を行ったとしても、誤情報の影響は誤情報に触れなかった0回のレベルまでは戻らないことを示している。このような現象は真実錯覚効果の非対称性と呼ばれていて、誤情報の受け入れられやすさだとか信じられやすさに対して、それを事後的に訂正することの難しさの間にギャップがあることを示唆する。【田中構成員(4)】

- このような認知的な特徴に基づいて、できるだけ訂正の効果を上げるにはどのような具体策があるのかという問いに対して、2020年にデバンキングハンドブックという報告書が出され、心理学における学術的な知見に基づき、誤情報に関する心理的な特徴17点と実行可能な対策10点が整理されている。日本語版はないが、日本での対策を考える上で、先ほどのプレバンキングのガイドブックや、冒頭で述べたようなAPAの推奨事項と併せて、この辺りの知見はもう少し日本でも注目されても良いと思う。【田中構成員(4)】
- 私たちの研究チームが関心を持っているのは個人レベルとシステムレベルの相互作用について。人間の情報の認知処理や意思決定がデジタル環境と完全に独立して行われることが少なくなっており、デジタル環境と心理の相互作用を理解することの重要性が今後さらに高まっていくと予想している。そのような研究はヒューマンコンピューターインタラクション研究と呼ばれているが、例えば次のような研究がある。検索エンジンが返す情報のランキングが人の検索行動に影響を与えて、結果的に人の意思決定に影響を及ぼす現象として、ランキングが下がるごとにクリック率が顕著に下がっていく傾向が、この研究以外でも一般的に観察されている。デバンキングのための訂正情報に関してもデジタル環境を経由して伝達されることが多いため、このようなクリック行動は訂正情報へのアクセシビリティと関わってくるし、それが結果的に人の心理にも影響することが考えられるが、このクリック行動という重要だと思われる要因については、これまであまりよい測定指標が開発されてこなかったために、単純なクリック数だとか、クリック率だとか、そういった分析しか行われてこなかったような背景があった。そこで、クリック行動の個人差を測定する指標を新たに開発した。訂正情報が人をどの程度、訂正情報をクリックしているのかを客観的に示すための実験を行った。デバンクで重要になってくるのは多くの人々が誤情報を信じている時で。実験では誤情報に、誤情報というラベルが貼られていて、そこをクリックすると訂正情報が表示されるデザインになっている。クリック行動の特徴を分類したところ、信じている誤情報に対する赤色の訂正情報を積極的にクリックする集団と、選択的にそれを避ける集団がいることが分かった。例えば、10個の誤情報を信じている場合に、積極的にクリックする集団は平均すると4個以上の訂正情報をクリックしているのに対して、選択的にそれを避ける集団は1個にも満たない7%という数字。選択的にそれを避ける集団は、ラベルで避けているというよ

りは、自分が信じているものに対する訂正を選択的に避けていることが分かる。このようなデジタル空間上で生じる心理的、かつ選択的な回避行動は、誤情報が拡散して社会問題となったときにどのように訂正情報を届けるのかという問いと密接に関わる。例えば、専門家の方々やジャーナリストの方々が事実と照らし合わせて丁寧にファクトチェックをしたとして、それをデジタル空間上に公表して誰でもアクセス可能な状態にしたとしても、この心理的な、この選択的回避という障壁によって誤情報が信じ続けられることが生じ得るということを示す実験結果。【田中構成員(4)】

- セレクティブアヴォイダンスについては、本当にこの研究で初めて明らかになった現象で、まだ複数のトピックで検証する段階までは行っていない。一応この研究ではワクチンを接種したかどうかや、周りにコロナに感染したかどうかみたいなことも聞いているが、それとの有意な関係は見られていないのが現状。【田中構成員(4)】
- ケンブリッジアナリティカで問題になったのは、パーソナルな情報を無断で取得していた点。それに基づいて、ビッグファイブのような指標を使って脆弱性のある人を、マイクロターゲティングのような感じで狙ったというところに問題がある。我々の研究は、そういったパーソナル情報はまず使わない前提で、あくまでも実験環境としてクリックコードを測定し、個人情報というものを使って狙っていくようなスタンスとは明確に線引きをしている。アヴォイダンス群がアヴォイダンスを少し緩和して、ファクトチェック見てもいいかなというような気持ちになるにはどういった介入が可能なのかというような方向性で研究を進めていく。【田中構成員(4)】
- 訂正のパラダイムとして、誤情報に触れたことがない人が最初に訂正から触れた時に誤情報を引用しながら訂正に触れることがある。そうすることによって、訂正の効果ではなくて誤情報がかえって信じられてしまうんじゃないかという、バックファイア効果と言われている現象があるが、結構研究が蓄積されていて、そういったバックファイア効果は発生する可能性は今のところ支持されていない。【田中構成員(4)】
- illusory truth の効果について、訂正情報がどうしても信じられないというのはすごく重要なこと。先に誤情報に接しているから訂正情報の効果がないのか、それとも誤情報が信念に合致してしまっているから訂正情報の効果がないのか。illusory truth は、とにかくたくさん接触すれば認知的に処理しやすいから真実だと見えてしまうということ。それは事実だと思うが、同時に信念に合致してしまっていることが問題の部分もあるのかもしれない。【安野構成員(4)】
- 真実錯覚効果について、訂正情報を先に出したときにどうなるのかというのは非常に重要な指摘で、今、記憶研究で研究が進められているが、そのメカニズムについてはまだ幾つか諸説ある。アップデートがうまくいっていないのか、訂正情報と誤情報が両方保持されて検索の時にうまくいかないのか、その辺りについては幾つかのモデルが提唱されているが、まだ一つに収束するほ

どまでは至っていない現状。【田中構成員(4)】

- 訂正をどういうふうにすれば効果があるのかについては、ハンドブックの中でも具体的な説明があり、訂正する情報ソースの影響についても研究が蓄積されていて、どういった情報ソースだとどういった訂正の効果があるのかとか、弱まるのか、そういった研究も蓄積されている。【田中構成員(4)】
- 人種だとかそういったものが関係するのではないかというのはそのとおりで、例えば受け手側と誤情報の人種が違う場合に、敵対的な人種的なステレオタイプを持っている場合に訂正情報が効きにくいですとか、訂正情報のほうが敵対的な人種、ステレオタイプを持っていると効きにくくなるとか、そういった研究もある。【田中構成員(4)】
- 情報元の信頼性について消費者は確認できる方策を知らないなので、やはりファクトチェック機関への支援や、偽サイトへの対応というのが必要。【増田構成員(1)】
- ファクトチェックというものは、情報空間における発信主体を塗り替える力があって、例えばファクトチェックが発信された以降はファクトチェックに関して言及する投稿がほとんどになったというような結果も出ている。それ以前は、偽情報を信じている人が多かったが、それが変わったと思う。ただ、偽・誤情報も膨大な量。AI でさらに膨大になっていくので、優先順位の高いものからファクトチェックするということが大事で、例えば災害とか選挙とか医療健康といったものが優先順位が非常に高いということがあり、こういったものを優先的に検証していくことが大事。【山口構成員(1)(2)】
- ファクトチェック組織の悩みとして、どれをチェックすればいいのか分からないということがある。そのため、AI による偽・誤情報スクリーニングとか、そういう技術をうまく活用した効率的なファクトチェックの推進が必要で、テキストファクトチェッカーなど偽・誤情報を検知する技術が既に出てきている。そういったものを上手く活用してファクトチェックを効率的にやっていくのが大事。【山口構成員(2)】
- ミス・ディスインフォメーション対策について、日本ではやっと国際ファクトチェックネットワーク(IFCN)のシグナトリーが誕生した。他の先進国並みに、情報を検証できるような総合的な社会的な仕組み、特にプラットフォームとメディア同士のコラボレーションがどのように進んでいくのか。ファクトチェッカーは非常に少ないという非常にバランスが悪い状況が起きている。今までは競争していたメディアがいかに協力できるのか、どこで一緒に仕事ができるのかという価値観の変換が必要になる。【奥村構成員(1)】
- ファクトチェックは信用されなければ成立しないので、IFCN の認証は定評のある大手メディアでも時間をかけて行われる。しかし、世界各国の国内事情も違うので、5つの原則を提示している。これらに合致した活動をしているかを、外部の人がアセサーが審査するが、この審査方法にも弱点があり、外部のアセサーの実力が一致しないということで、属人的な問題がある。それから、IFCN は、慢性的な人手不足と資金不足に悩んでおり、審査方法の検証が必ずしも十



分とは言えない、特にインド等ではかなりたくさん申請があるために、一旦取ったシグナトリーが翌年認められないというようなことがあったりもする。それから、言語の問題がある。日本は特に日本語というかなりユニークな言語を使っているために、アセサーの数が非常に限られるというような、アジアの国々特有の問題もある。【奥村構成員(3) ※一部再掲】

- 最近ではソーシャルメディア等だと、例えばXのバードウォッチみたいに、誤情報があった時に訂正に近い情報も同時に提示する場合は結構あり、最初はどうか分からないが、しかるべき後に同時に情報を出すケースが結構ある。【越前構成員(4)】
- 訂正情報、ファクトチェックのようなものを単純に届けばよいということではなくて、いかにそういったものを効果的に届けるのか、例えばプロメントのような話とか、やり方についても認知科学との連動というのは非常に重要。【山本構成員(4)】
- ファクトチェックについて見ると、資金、人材、効率性などファクトチェックがより効果的に行えるようにしていくにはどうするか、ということがある。【落合構成員(5)】

### 3. 主に情報伝送段階の課題

- ◆ アテンション・エコノミーが引き起こす課題（フィルターバブル、エコーチェンバーを含む）への対応の在り方
  - 情報流通の健全性を脅かしている「構造的要因」に着目すべき。構造的な要因として、プラットフォームのビジネスモデルといわれる「アテンション・エコノミー」のひろがりがある。アテンション・エコノミーの下では、どれだけユーザーのアテンション（エンゲージメント）を奪えるかが非常に重要（死活問題）となるため、ユーザーの選好、政治的信条、認知傾向等に合った情報が強く Recommend される傾向にある。それによって、フィルターバブル、エコーチェンバーといった問題が先鋭化・増幅する。また、この「構造」では、アテンションを得やすい、刺激的な偽情報や憎悪的表現などが優位性を獲得しやすい（逆に、刺激に乏しい「真実」は流通しにくい）。このような構造を無視して、「モグラたたき」をしてもあまり実効性がない。モグラが出やすい「土壌」（環境）を改善しない限り、モグラは次々と出てくる（認知戦についても、戦いやすい「構造」に対処していかないと有効なカウンターとならない）。【山本構成員(1)】
  - 現在のデジタル空間では、情報やサービスがユーザーの目から見ると非常に多種多様かつ大量にあり過ぎて、まさにジャングルの中にいるような状態。そこで、情報やサービスの一部を Recommend、優先的に表示してもらうことにより、我々は取捨選択ができるようになってきているが、アテンション・エコノミーが背景にあることにより、優先順位づけでアテンションを稼ぐものが表示されやすくなっている。【水谷構成員(1)】
  - 本検討会の中心的な課題は偽情報への対策だが、現に生じた現象への事後的な

対策だけでなく、偽情報の流通を容易にし、偽情報に社会を破壊しうる大きな力をもたせることになった要因、すなわち、フィルターバブルやエコーチェンバーのような状況を生み出すプラットフォーム事業者等のレコメンドシステムや、さらにはそれを可能にしている大量のデータ収集の問題など、プラットフォーム事業に起因する問題にも向き合う必要。【森構成員(1)】

- 事後の対応としてのコンテンツ・モデレーションだけでなく、偽情報に破壊力をもたせてしまう要因への対処も考えていく必要。例えば、DSA ではレコメンドの透明性確保や、特別なカテゴリー情報を用いたターゲティング広告の禁止、未成年者に対するターゲティング広告の禁止なども義務づけており、これらは我が国の制度設計においても参考になる。【森構成員(1)】
- 内部において発信者側でどう統制するかということは非常に重要で、今やプレーヤーは、そんな統制をどうするかなんて真っ当な人ばかりではなくなっており、欲望のままに発信をしている。一方で真面目な人たちが統制をどうするかという話があり、他方であまり真面目じゃない人たちが欲望に基づいて発信する、しかしながら欲望というのは一つの大きな形というか大きな流れを持っていて、それがアテンション・エコノミー。一方で真面目発信のときの統制の問題があり、他方でアテンション・エコノミーが個々の発信者と場合によってはプラットフォームさえも駆り立てるというところを、これからさらにこの検討会で詳細に見ていくのかなと思う。【森構成員(3)】
- アテンション・エコノミーのメカニズムの中でマスコミに対する批判のドライブがかかってしまうというか、まさにアテンション・エコノミーの中での議論ではマスコミが攻撃の対象になりやすい。そこが今の情報空間の健全性を変な形でゆがめている部分がある。【宍戸座長(3)】
- 消費者はフィルターバブルによって繰り返し同種のトラブルに遭っている。エコーチェンバーによる考え方の偏りが特に若年層には多いのではないか。【増田構成員(1)】
- エコシステムがどうなっているのかということで、最終的にはアテンション・エコノミーというのにも注目しながら議論する必要がある。その核心に至る部分というのは、むしろ、今後さらに掘り下げられるべきなのではないか。【落合構成員(5)】
- 全体としては、アテンション・エコノミーが、全体として、必ずしも質が高くないような情報を拡散させていくことについて、後押しをしている部分があるのではないか。これを見ていき、分析していくことが重要。【落合構成員(5)】
- アテンション・エコノミーについて、さらに分析をしていくことが必要ではないかと考えている。誤情報、偽情報の拡散だけではなくて、信頼性の高いコンテンツの拡散が進まないことについて、デジタル空間における関係者の行動インセンティブの設計に課題がある状況になっているのではないかとも思う。ただ、アテンション・エコノミーというのを見るときに、実際には様々なプラットフォームや、広告のエコシステムも関わって、データの移転、分析、コンテンツの利用がされていて、一口で言ってしまうと、間違ってしまう可能性があ

る。そういう中で、特に論点になるようなアドフラウド、ディープフェイク、その他の問題として指摘されるような幾つかのケースや、政治的な言論、サイバー攻撃などの場面が重要と思われるが、それぞれ場面によって、どういう当事者の関係性があるか、行動インセンティブが働いていて、実際どういう傾向で動いているのか、を分析していくことが必要。この点、デジタル市場競争本部で、経済的分析の関係で競争環境を見ていると思うが、これは、あくまで経済的規制の観点で見ている部分であり、情報の格差における問題点の分析はされていないと考えている。偽・誤情報への社会的な政策や対応を検討するという、この検討会での独自の分析意義は残るのではないかと。【落合構成員(5)】

- とりわけ、アテンション・エコノミーへの注目がやはり重要だというポイントだと思う。私なりに解釈すると、起きていて問題、あるいは起きることが容易に想像される問題について、個別の対応ということはもちろん必要であるし、できることについてはやっつけていくことが肝要であろう。しかしながら、問題は、もぐらたたきにどうしてもなってしまうところがやはりあるというところだと思う。そういう意味で言うと、視点を少し変えて、そもそも何でこういった偽情報、誤情報、とりわけ偽情報の悪意のある部分ですけれども、こういったものが発生し得るのかということ、つまりインセンティブに着目をするということの重要性を恐らく指摘されていると思う。アテンション・エコノミーの概念自体は、もう多くの方が御存じのとおり、ティム・ウーが既に整理をしてからかなり時間がたって久しい。ここの問題というのをさらに深掘りすると、アテンションを獲得するということは必ずしも悪ではない。あるいは、情報でエコノミーが発生するということが、これも当然悪ではない。しかしながら、このアテンションとエコノミーが密結合してしまうということが非常に大きな課題なんだということ。これが今日的な合意に近づいてきていると思う。つまり、耳目を集めることでマネタイズが可能になってしまう。それがさらに拍車がかかって、耳目を集めることが目的になってしまう。それでマネタイズができる。これがエコノミーという側面で、しかもこれは単純なエコノミー、つまりアドフラウド的なことも含めて金が発生して入ってくるということはもちろん、それだけではなくさらに違うより高次のエコノミー。つまり言ってしまうと経済紛争のような話。こういったところまでつながってきってしまう。つまりこの大きな構造の根っこがアテンション・エコノミーであるというような御指摘であったかなと思う。ですので、恐らくここでの検討もほかの様々な要素ということを当然念頭に置き、その取組を強化しながらではあるが、このアテンション・エコノミーに対してどのような対策をしていけるのかというところを一つ大きなポイントにしていくことが重要ではないかと。場合によっては、ここに着目の立脚点の一つを置くことによって、表面的といえますか現象に対して対処するというところで、言論の自由とどうしても対峙してしまうというような問題が出てきて限界が出てくる場合があるところ、いやそうではないんだと。ここは経済構造であるとか、そもそものインセンティブの構造に光を当てていくことによって、サニタイズというのはちょっと強い言い方かもしれま

せんが、ただそこで明らかに、既にもう経済犯罪になっているものもあるし、そういったことを意識しながら取組をするということは可能性であるとか、できることというのが広がるかなと思いましたが、そういった点をさらに検討を深めていけるといいのかなと思った。【クロサカ構成員(5)】

- アテンション・エコノミー自体は、今までのテレビの視聴率競争といったものもあった。それが今のデジタル空間では、アテンションとエコノミーの関係がより密接になってきているという点があるわけですが、その一方で現代において情報を発信・伝送するアクターが多様化したことにより、そうした発信・伝送を担うアクターの中で倫理面に格差が広がってきているのではないかと。つまり、今までだと、テレビはもちろん視聴率至上主義とか批判はされていたわけですが、放送法の規定もそうですし、あるいは自主規制、ジャーナリズム倫理といった観点から、一定の倫理観が求められてきた。これは放送のみならず、新聞もそうですし、雑誌には雑誌の倫理があったと思う。それが今では個人のクリエイターの方々が新規参入してきた一方で、そうした方々にはこれまでの新聞や民間放送の業界団体のような組織がきちっとあるわけではないし、特に生成系 AI の使い方ひとつとっても機能倫理である程度グリップを効かせていける仕組み、または構造そのものが欠如しているのではないかと。そうした点でいえば、今までは倫理的拘束から逃れにくい新聞やテレビが情報の伝送も担ってきたわけですが、他方、現在では主要な伝送アクターはプラットフォームに握られている。このようにコンテンツの生成だけでなく伝送を担うアクターの倫理面に関しても同様のことが言えるのではないかと。【水谷構成員(5)】
- 特に視聴率のような構造がテレビにもあるので、それとこのデジタル空間における違いがどこにあるのかはある。もちろんプラットフォームの中にも、モデレーションの基準や、倫理規範を定めている場合もある。一方で、実際に情報を発信、取捨選択するとき、メディアとプラットフォームでどのくらいお金が現実に関与が違ってくるのかなどは、ややまだ見えていない部分があるのではないかと。その辺りは、特に問題になり得るようなケースを参考にして分析をすることが必要なのではないかと。もちろん従来のメディアだけの場合に、今ほどフェイクニュースという話にならなかったということからすると、恐らく少なくとも10年前などのタイミングでは、メディアはしっかり対応されていたのではないかと推測。とはいえ、今回のこのインターネットの世界の場合ですと、プラットフォームが重要な位置づけを占める話もある。例えば広告のエコシステム、様々なデータ連携やインターネットの基礎レイヤー、そういった部分との関係で実際にはコントロールできるように見えていても、あまりできていないこともあるのではないかと。ただ全体として、やはりどうしてもそれぞれの方々がビューを取るといふこととお金を取るということがかなり密接な関係にあるといふか、もしかすると視聴率の場合以上に強い傾向になって現れていることによって、全体として質を確保するよりかは、まずは注目を高めるほうに流れる傾向になりやすい可能性があるのではないかとと思われる。このような点について、もう少し踏み込んで分析していくことで、本検討会での検

討の内容が真に迫ったものになるのではないか。【落合構成員(5)】

- やはりこうした議論で常に国際的にも議論に上ってくるものとしてショシャナ・ズボフのサーベイランス・キャピタリズムというものがある。それは恐らく1つはアテンション・エコノミーのような議論と個人データなどの議論を深く組み合わせたものというふうに考えられる側面があるかと思うのですけれども、まさにサーベイランスといったようなものが、どのように現代の情報空間において役割を果たしているのかといったようなことも、まさにこういった議論とオーバーレイして考えられていくとよい。【生貝構成員(5)】
- アテンション・エコノミーについて、若干補足。インターネットにおいて最も成功したビジネスモデルは広告であり、そこでの収益はアクティブユーザー数、ページビュー数、いいね！やシェアなどに基づいて計算される。どれくらい見てくれたか、いいね！してくれたか、シェアしてくれたか、コメントしてくれたかが経済的利益に直結するということ。このいいね！するとか、シェアするとか、コメントするとか、そういう行動をエンゲージメントというふうに言うそう。御紹介したいのは、フェイスブック（当時）で、フランシス・ホーゲンという人の内部告発が2021年にあった。内部告発のきっかけが非常に共感できるというか本当に響く。「フェイスブックは20年秋の米大統領選後、私が当時いた（偽情報やヘイトスピーチに対処する）部署を解散しました」——2020年秋の米大統領選だから、バイデンが当選した大統領選挙。「その時、フェイスブックは、大統領選での（SNSが社会を分断した）教訓を忘れたのだと感じました」——こちらの「大統領選での教訓」というのは2016年の大統領選。ケンプリッジアナリティカの舞台となったトランプが当選した大統領選。そのときの教訓を忘れたんだ。だから偽情報やヘイトスピーチに対処する部署を解散したんだ。それは私にとっては耐え難いことであつたと。そんな内部統制の重要な要素を失っていいのかというふうに彼女は思ったんだろう。続きですが「路上で血が流れるようなことは起きず、フェイスブックは成功したんだと。リスクの大きさを無視したのです。でもその後、翌年1月に米議会の襲撃事件が起きました」ということ。ホーゲンさんの発言を紹介しているのが下のところですけれども、『FBは極端な反応を引き出すコンテンツが、より多くのクリックがコメント、シェアを得られると知っている』と指摘。同社のアルゴリズムの変更で、怒りや分断をあおる内容が拡散しやすくなり、子どもたちの安全を脅かしているとの見方を示した」ということ。【森構成員(5)】
- もう一つ御紹介したいのは、アテンション・エコノミーはしばしば偽情報の原因であると言われるが、それだけじゃなくて誹謗中傷の原因にもなっているよという徳力さんの記事。「羽生結弦さんの離婚報道で考えるべき、過剰報道と誹謗中傷の『負のスパイラル』」というタイトル。「しかも更に悪いことに、現在のネット広告の仕組みでは、そうした過剰報道の記事がたくさんの人に読まれれば読まれるほど、より広告収入が上がる仕組みになっています」。「こうした便乗ネットメディアの記事や便乗 YouTuber の動画の多くは、過激なタイトルやネガティブな内容のものが多いた特徴です。メディアの記事にしても

YouTube の動画にしても、基本的にはアクセスが多ければ多いほど広告収入が増える仕組みです。仮にネガティブな記事や動画に対して、ファンが怒って批判してきたとしても、多くのファンが注目してくれて記事のアクセスが増えれば、収入が上がる仕組みなので批判も怖くないわけです。その結果、ファンが軽く検索しただけで、ネットや YouTube 上に、目をおおいたくなるような内容の記事や動画が出てきてしまう状況が生まれてしまうわけです」ということ。

【森構成員(5)】

- 権利侵害情報との関係では、今日このアテンション・エコノミーと権利侵害情報という交錯点で徳力さんのお話を御紹介させていただきましたけれども、私はこの3つの問題（権利侵害情報、偽情報、利用者データ）というのはいかにこれを中心に検討するということはいかに大きく変わってはいかないのではないかと思いますけれども、やはりアテンション・エコノミーということからホーゲンが言うような様々な問題というのが生じていっている。特に彼女の言うところの怒りや分断をあおるとか、それは権利侵害情報とか偽情報とまたちょっと違うものなのかもしれない。もしかしたら偽情報に分類されるのかもしれないけれども、そういう縦の軸にももしかしたらインパクトを与えるのではないかと。もともと「アテンション・エコノミー、すなわち偽情報」みたいな感じでしたけれども、実はアテンション・エコノミーは権利侵害情報にも関係する。もしかしたら何かそれ以外違うものも出てくるかもしれませんねという、そのような意味でサイバー空間の情報流通の健全性に全体的に影響してくるんじゃないかなと思っている。【森構成員(5)】
- 森先生の御指摘の中で DPF 透明化法とはちょっと射程が違うのではないかというお話があった。我が意を得たりというか、先生に整理いただいて腹落ちした状態ですが、やはり競争政策の観点に基本的に軸足を置いて、そこから派生してきているということを DPF 透明化法は基本的に志向しているものであろう。そういう意味で言うと、競争環境がいびつであるということがアテンション・エコノミーの遠因であるという説明の仕方もあるかもしれないけれども、アテンション・エコノミーそのものというものは恐らくより大きかったり、より小さかったり、いずれにしてもエンティティによって変わるかもしれませんが、様々なところに既に発生してしまっている。だから例えば、個人の誹謗中傷の一つの大きな要因にもなってしまっているということもあろうかと。つまり産業構造の話だけではなくて、こういう金の回し方、インセンティブの発生のさせ方があるんだということが大きく広がってしまっているということなんだろうということを今日御指摘いただいたんじゃないか。ですので、アテンション・エコノミーに当然、光を当てながらも最後のほうの話になるかと。モニタリング／ヒアリングの留意点という辺りの補完と思いますけれども、まさにこのコンテンツモデレーションと利用者データの保護に関する透明性やアカウントビリティの確保ということ、この辺りが背中合わせになっている状態で議論をしていく。その中で一元的に検討することが適正であることとそうではないことということが恐らくあると思うんですが、さらに区分けをしていくとい

うことになるのかなど。その観点からも、恐らくモニタリングをさらに進めていくということ、ここを軸足にこの検討会として置いていくということなんだろうというふうに理解。【クロサカ構成員(5)】

- 森先生、それから落合先生の御提言で重要だと私も思うところは、収益のシステムがもしかするとやはり重要なポイントなのではないかということ。以前、私もアクセス数が経済的な利益を生むシステムを変えたほうがいいんじゃないかというようなことをどこかで申し上げて、ちょっとそれは難しいみたいな感じになったようなことがあった。どうしても利用者の側のリテラシー教育というのは必要ではありますけれども、限度があるというふうに私個人は思っておりまして、そう考えますと誤情報をばらまくことが必ずしも利益を生まない形にしていくということが一つ、規制をせずにより健全な情報空間をつくるポイントだというふうに感じた。ではどうすればいいかということは、ちょっと私もよい案がなかなかないのですけれども、その意味で森先生、それから落合先生に今回のようにまとめていただいたことは、私は非常に重要な御提言だと思った。【安野構成員(5)】
- 今、この権利侵害情報と偽情報の関係で、コンテンツモデレーションということをやっているわけですが、そこで我々が考えているのは悪い情報を削除したり、そういうこと発信するアカウントをBANしたりということコンテンツモデレーションと言っているわけですが、やっぱり、それはもぐらたたき。なので、今、コンテンツモデレーションというよりは、アテンション・エコノミーのモデレーション。アテンション・エコノミーに対してめちゃくちゃビビッドに全てのプレイヤーが反応しているという状況のモデレーションをしないとイケないというふうに本当に私も思っている。【森構成員(5)】

#### ◆ 偽・誤情報の拡散への対応の在り方

- 情報環境において偽・誤情報問題は深刻化している。2016年は偽・誤情報元年と呼ばれ、米国大統領選挙では様々な偽・誤情報が拡散されて、選挙前の3か月間で、トランプ氏に有利な偽・誤情報は3,000万回、クリントン氏に有利な偽・誤情報は760万回もシェアされたということが指摘。このフェイスブック上のシェアの回数は、事実のニュースのシェアの回数を超えることがわかっている。2020年以降は新型コロナウイルスやコロナワクチン関連の偽・誤情報が広く拡散されて、WHOがinfodemicと警鐘。5G電波がコロナウイルスをまき散らしたというような偽・誤情報を信じて電波基地局に対する破壊活動を行った例も。【山口構成員(1)(2)】
- 最近では教育コンテンツの体を取った偽・誤情報というものが動画サービスで見られていて、こういったものが青少年に及ぼす影響が懸念。【山口構成員(2)】
- 日本の事例で、安倍元首相が被災地を訪れた際の写真が、スタジオで撮影されたかのような偽・誤情報投稿があった。この偽・誤情報は海を渡って台湾で拡散されたために台湾のファクトチェック組織によるファクトチェックがなされた。このように、日本でも災害時のデマ投稿、新型コロナ関連、政治的なも

の、多くの偽・誤情報が時にグローバルに展開してしまっている。【山口構成員(2)】

- コロナワクチン関連の偽・誤情報、政治関連の偽・誤情報、陰謀論の各 6 件、合計 18 件を使い、実際の偽・誤情報の調査と人々の行動を分析。コロナワクチン関連は 37.1%、政治関連は 26.4%、陰謀論は 19.1%の人が、たった 6 件の事例にもかかわらず、少なくとも 1 つ以上を見聞きしていたという結果が出た。少なくない人が偽・誤情報や陰謀論に接していることが言える。さらに見聞きした上で誤っていると気づいている人、これが政治関連では平均して 13%しかない。コロナワクチンと陰謀論ではそれよりは多いが、それでも 43.4%と 41.7%にとどまっている。大半の人は、偽・誤情報に出会った後にその情報が誤っていると気づけていないという結果が出た。また、年齢別に見ると、50 代、60 代の中高年の方が若い世代よりも誤っていると気づきにくい傾向も見られた。とりわけ政治関連の偽・誤情報と陰謀論。その結果を見るに、偽・誤情報が、SNS をよく使っている若者だけの問題ではなく、老若男女問わず身近にある問題であることが言える。【山口構成員(1)(2)】
- 保守の政治家に不利な偽・誤情報と、リベラルの政治家に不利な偽・誤情報両方用意し、それぞれの政治家への指示がどのように変化するかを調査。1 つ分かったのが、偽・誤情報を見て、その政治家への支持を下げの人が少なからずいたということ。よって、ネガティブな影響がある。もう一つ特徴的だったのが、特に弱い支持をしていた人ほど偽・誤情報によって支持を下げやすいという傾向が見られたこと。この弱い支持層というのは、人数でいうと、むしろ多い人たち。つまり、選挙ですごく重要な層。そういう人たちの考えを偽・誤情報は変えやすいということで、偽・誤情報は、民主主義の根幹をなす選挙結果に影響を与えているのではないか。【山口構成員(1)(2)】
- 偽情報の生まれる背景は、主に経済と政治、この 2 つの理由。経済的理由については、アテンション・エコノミーが広まる中で、広告収入目当てに偽情報を作ったり、あるいは世論工作をビジネスにすることも。また、政治的理由としては、政治選挙に介入して荒らすということがある。3 つ目の動機として、注目を集めたいとか、ふざけてやるというものもある。【山口構成員(2)】
- 昔はフェイスブックファンページと言われてたフェイスブックページについて、こちらの投稿を対象に 2020 年 1 月から 2021 年 5 月までの投稿を収集し、その中から、最も影響力のあるアカウントとして、コロナ関係の投稿が、この収集期間中 100 回以上あって、なおかつ獲得反応数として、いいねなどのリアクション、シェア、コメントの総数が 10 万回以上ある比較的大きな影響力のあるアカウントだけを抽出すると、75 あった。このうち 2 つが、ファクトチェックイニシアティブで虚偽もしくは誤りとされる内容を半数以上占める投稿を繰り返している偽・誤情報系のアカウントだった。内容としては、陰謀論、ワクチンに関する偽・誤情報の定期的な発信のほか、偽情報とまでは言えないものの不安をあおるような投稿、例えば、日本で何々が全く報道されないといったような発言、偽情報の拡散を促すような投稿、あるいは誰かに行動を促す



ような投稿、具体例としては、まずできることとしては不買運動であるといったような、少し不安感あおるような投稿というのを拡散していた。この偽・誤情報系のアカウント、コロナ関連のみの投稿だが、それだけで見ると、主流のメディアアカウントよりも非常に大きな反応を獲得しており、我々の調査の中では一番大きな反応数を獲得していたアカウントもあった。【澁谷構成員(2)】

- フェイスブックグループについては、影響力が大きいグループのうちの3つが偽・誤情報系アカウントだった。政治的なイデオロギーに基づくもの、自然療法・代替療法に関する話題に関するグループ、陰謀論などに特化した情報をシェアするグループが主に見られた。こちらでも、不確かな情報やほかのユーザーに共感を求める投稿も多く見られた。共感を求めるというのは、具体的には、例えば、ワクチンに対して非常に消極的な、反対の方の発言だが、「両親がいよいよワクチンを打つことになってどうしたらいいのか」などと共感や意見を求め、そこで非常に会話が盛り上がるといったような投稿群もあった。【澁谷構成員(2)】
- フェイスブック投稿におけるコロナの話題について、偽・誤情報系のアカウントとそれ以外でトピック分析により、どんなトピックがそれぞれ話されていたのかを見ると、偽・誤情報系のアカウントでは、コロナ関連の話題の中でも特に解説をしたり、又は、自らの主張を展開するような投稿群が多く、それ以外では、コロナの症状、コロナの起源、又は国内外の感染状況等に関する発信等の占める割合が、その他の影響力が大きいアカウントに比べて多かったことが特徴。【澁谷構成員(2)】
- 偽情報系アカウントについて、内容だけでなく、どんな情報源を用いて発信をしていたのかも見るために、外部 URL の共有状況を確認したところ、ユーチューブへの言及がコロナ禍に関しては最も多いという結果になっている。また、その他の、例えば、営利団体サイト、市民団体サイト、又は、ブログ、一部ツイッターやヤフーニュースもあるが、こういった大きなソーシャルメディア以外の URL を含んだ投稿も多く、特定のセミナーの紹介や商品の販売といったところに誘導するような内容も見られたということが特徴として挙げられている。【澁谷構成員(2)】
- 外部 URL の共有状況について、ネットワークでどのようなクラスターがあるかを分析した結果、アカウントごとにクラスターで分けてみると、例えば、地方公共団体の感染状況に関する情報等を共有しているものや、ニュース記事の共有を中心としたアカウント群などがある。偽情報系アカウントを含むアカウント群には、非常に多い反応数を獲得している主な偽情報系アカウントがネットワーク図のクラスターの真ん中付近にあるが、これらの周りの大小様々なノードも具体的に見ると、影響力はそこまで大きくないが、非常に似たような発信を繰り返していたりしており、特に影響力が大きいアカウントの周りで大小様々なアカウントがネットワークを形成しているということが明らかになっている。【澁谷構成員(2)】
- コロナ禍でのフェイスブックにおけるアカウント運営に関して特徴が2つあ

り、1つ目が組織的なアカウント運営が確認できた。プラットフォーム事業者によるアカウント削除というのが当時も行われていたが、削除への対応策として、複数アカウントを運営したり、又は削除対象とならないように、コロナとかワクチンといった特定のキーワードを避けて、いわゆる隠語のようなものを繰り返し使った投稿というのも見られた。2つ目の特徴としては、多様なグループが偽情報系アカウントの周辺で非常に緩やかなネットワークを形成し、そこで協調し合っているということ。具体的には、偽情報系アカウントが、別のグループやアカウントのコメント欄で、情報不足や、それから不満不安を訴える投稿を繰り返したり、又は、特定の外部 URL の投稿を、一部はボットのようなものを使って繰り返すというのも見られた。また、偽情報系アカウントは、政治イデオロギーに基づくグループ、一部の自然療法、代替療法、子育てに関する情報共有のグループ、教育に関するコミュニティと非常にグループのつながりがあるということが見られ、内容を少し確認すると、例えば、一般の方が切に、より健康になりたい、より自分の子供に健康に育てほしいとか、よりよい教育を受けさせたいという情報に関して、切に探している方に対し、ある種狙っているような形でそこにコメントの投稿をするというようなところも多少見られた。また、特にフェイスブックグループを中心に、感染拡大に伴う不安に乗じた関連商品の販売、ビジネス勉強会、研究会等への勧誘などを行うといったところも見られた。投稿例としては、講演情報とか、希少な情報を得られますといったものをうたったものが非常に多く見られた。【澁谷構成員(2)】

- フェイスブック上、コロナ禍において少数のアカウントが非常に中心的な役割を担っていた。コロナ関連だけに注目すると、主流メディアのアカウントと同等もしくはそれ以上の反応を獲得。情報発信の動機としては、金銭的なインセンティブ、それから政治やその他のイデオロギーに基づくものと考えられるものが、日本のコロナ禍でも見られたということが確認。またユーチューブその他のウェブサイトへの外部の情報源に言及した偽情報の流通があった。何よりその中心的なアカウントは存在するが、周囲に非常に大小様々なグループやアカウント群が存在し、協調的に偽情報の生成流通に寄与している全体像が少し見えてきた。【澁谷構成員(2)】
- プラットフォーム事業者の単一的なアカウント削除は効果があるが限定的である可能性。組織的に対応策を取っていたり、緩やかにつながっているので、1つのノードを削除したところで、ネットワークとしての偽・誤情報系アカウント群の復活は非常に簡単であることが理由。偽情報は単に誤った情報の集まりというよりも、ソーシャルメディアのネットワーク性を生かしながら、戦略的な情報発信やネガティブ形成が行われ、一見無関係な集団との協調が取られていたり、不安定化する論理の要素を利用したような流通が行われている。【澁谷構成員(2)】
- 単一のアカウントやユーザーと捉えるのではなく、ネットワーク性を持つアカウント群として、情報源や情報流通を理解し対応する必要もある。その他に、

もうちょっとマクロな形で、現在の例えばユーザーを取り巻く社会経済的な状況、あるいはそのプラットフォーム以外の他のプラットフォームでどんな情報が流通しているのかとか、プラットフォームを超えた流通というところのこのメカニズムというのを把握していく必要がある。【澁谷構成員(2)】

- フェイクという言葉について、欧米でミスインフォメーション、ディスインフォメーションという言い方をしてフェイクニュースという言葉を使わないのは、フェイクという言葉に政治的な意味合いが帯びているからである。日本語にはそのようなニュアンスが共有されていないが、それをそのまま英語にすると問題が出てくる言語にもなっている。これはトランプ元米大統領の存在が大きく、2019年にニューヨーク・タイムズが記事にしたが、彼がひと月に40回ぐらいメディアをフェイクニュースと言って気に入らないメディアを罵倒することから、世界で四十か国ほどの、特に独裁者の国々の政府の高官、政治家等がメディアをフェイクニュースと呼び、政治的な攻撃を行っている。ファクトチェックコミュニティでは、ほとんどフェイクという言葉はうそぐらいにしかならなくて、ほかは全部、ミスインフォメーション、ディスインフォメーションという言葉を使う。ただし、日本語ではすごく使いにくい言葉であるため、何かほかのいい言い方を編み出して頂けないか。【奥村構成員(3)】
- プラットフォームサービス研究会で議論し始めたときに、一番最初はフェイクニュースと、2018年当時は言っていたが、ディスインフォメーション、ミスインフォメーションの訳語として偽情報、誤情報を充てたが、ディスインフォメーションに偽という言葉を使っているのが、このままいいのかどうかは問題のような気がする。【宍戸座長(3)】
- (デジタルサービス法において) 特に直接的に偽・誤情報対策に関わるところでは、VLOP、VLOSEの偽・誤情報を含むシステミックリスクへの対応がある。一つ一つのミクロなレベルでの対応の透明性とそれを集計した透明性レポートというのも重要だが、まさにシステミックなリスクが問題。それはしかし、ディスインフォメーションだけに限らないが、しかし当然、市民言説や選挙への影響というのがデジタルサービス法の立法経緯としては非常に大きくある。その他いろいろなリスクを自ら評価し、そして、それをちゃんと緩和する義務を取らないといけないという、非常に重要な包括的な規律が置かれている。【生貝構成員(4)】
- プラットフォーム通信事業者と利用者との関係の中で、偽、誤情報が流通、拡散する中で、情報のトラストを得るためのコスト増がある。ここはもしかすると、プラットフォームも対策を行うことによって、プラスになる面もあるかもしれないし、一方で、こういったプラットフォームの上のレコメンドですとか、コンテンツモデレーションとも関わってくる部分もあるかと思うが、フィルターバブル、エコーチェンバーといった課題が、アテンション・エコノミーをより推進するという構造があるのではないかということもあるし、データ保護の在り方もあろう。【落合構成員(5)】
- ミスインフォメーションというのが発生する過程は様々ではあるが、非常に時

間がかかる。一瞬にしてぼっと何かクリエーションが起きて世の中にぼんと流通するわけではなくて、ファーストドラフトというファクトチェッカーの供給機関で、今ちょっと資金問題で頓挫してしまって活動が全くされていませんけれども、彼らが沈黙のトランペットというものを発表している。それはコンスピラシーセオリーがどのようにして生成していくかというような過程を考えたときに、最初はクローズドな、最初はダークウェブのようなところで地中から湧いてくるように議論が起きたものが少しずつ練られていって、信憑性を帯びるような形でいろいろな意見が加えられていって、それで内輪の陰謀論サイトのようところで紹介されたものを陰謀論や何かに多少影響を受けた芸能人などのインフルエンサーがソーシャルメディアに出てきたところでぼんと世の中に広がるというような過程を取るということになって、その直前で何とか抑えたい。ただこのような陰謀論があるから注意せよというのをあまり早めに出すと、今度はその陰謀論の存在をたくさんの人が知ってしまうというデメリットのほうが強いと言われていて、そのタイミングを計るというのがファクトチェッカーの中では非常に重要な大きな課題になっている。というか、ファクトチェッカーははっきり言ってそれに対応できていないのが世界的な現状。そうすると、そのようなところまでどうやって分け入っていくかということになると、例えばBBCや何かには手だれがおりまして、全然コンスピラシーセオリーの人ではないんですけれども、そのようなところに深く深く入り込んでいけるようなスキルのある人間が世界に数人はおりますけれども、そのようにはなっていないということになるとそれをどうやって発見していくかというのは、こういう大きな国や何かで考えるような枠組みとしてどういうふうに扱えるかどうかは分かりませんが、リアリティーとしてはそういうこと。【奥村構成員(5)】

#### ◆ 広告を巡る課題への対応の在り方

- 情報流通の健全化に際しては、アテンション・エコノミーとの関連でも、偽情報やそのほかの有害情報を発信する情報発信者の背景にある経済的インセンティブ（主として広告収入）にどのように対応していくかについても議論した方がよい。その場合、PF側によるデマネタイゼーション、デプラットフォームのみならず、PFに広告を出稿する広告主・代理店の「ブランドセーフティ」を高める必要もある。【水谷構成員(1) ※第1回会合終了後の追加意見】
- 偽情報を使ってもうけるとか、あるいはアテンションエコノミーの中で広告収入を得るために過激なタイトルつけたりとか、それも偽情報、流したりとかということが起こっているが、こういった広告収入がそういう悪意を持った集団に流れてしまうことが結構、問題の根底の一つにあり、それをどうやって対策すればいいのかという話で、プラットフォーム事業者が頑張っても半分も減らないようなデータも出てたと思うので、例えばGoogleが対策したら、まだ60%ぐらい残っているみたい話がある。【山口構成員(4)】
- デマネタイゼーションの話はなかなかいい案が浮かばないが、一つ単純な方法

としては、デマネタイゼーションをきちんと DPF 事業者側に法的に義務づける方法があるが、DPF 事業者のデマネタイゼーションもモデレーションの一種なので、それには限界がある。また基本的にコンテンツ自体を直接制限するわけではなくてお金を止めるというだけだから間接的にはだけれども、直接的な表現規制とは違うという議論はあり得るが、果たして法規制として国家がそこまで介入していいのかというのが、まだ憲法的に検討しきれてないところがあるので、今後の宿題としてそういう規制が憲法上可能なのかというのを考えていきたい。【水谷構成員(4)】

- 削除までいかななくても、例えば収益を奪うという行為もすごく有効なんだろうなと思いつつ、例えば動画共有サービスとか、そういったところで収益化をいきなり止められて、それで生活できなくなるみたいなことというのは十分起こり得る。そうすると、かなり巨大なプラットフォーム事業者が1人の生命を結構握っているみたいな状態にもなるので、その辺りをすごく最近考えている。

【山口構成員(4)】

- 一定の広告について、それを非収益化する、さらには禁止すべきなんではないかと考えている。一つは表現の自由と言っても、広く皆さんの前にパブリックに示される表現でなければ、それはその個人に対するメッセージとして見せられた場合には、検証の不可能性によって個人がうっかり信じてしまう、「それは間違っているよ」と誰も教えてくれないというようなこともあるし、また他方で、広告なのでお金があればどんどん行けてしまうと、幾らでも出せてしまうこともあるので、そういうことを考えると政治的なメッセージをマイクロターゲティングで出すようなことは禁止されても憲法上、問題がないんじゃないか。

【森構成員(4)】

- 山本先生は公選法の戸別訪問の禁止とのアナロジーを使っておられたが、何かここまで弊害が大きくなってしまうと、それでも表現の自由だから、そのかつての表現とは随分表現の形が変わってきているかなとは思っている。【森構成員(4)】
- デマネタイゼーションの話とマイクロターゲティングの話に分けて考えなくてはいけないのではないかと。デマネタイゼーションの議論については、参考にしようと考えているものとして、アメリカでサムの子息法という法律が、表現の自由との関連で議論されたことがある。これは、犯罪告白本で収益を上げることへの規制だったが、これが修正一条上の議論になる。そのあたりをアナロジーで捉えて今後議論していかなくてはいけない。【水谷構成員(4)】
- マイクロターゲティングに関しては弊害についての議論ももちろん存じており、まさにあれはより広告を見させることが背景にあり、広告に関しては表現の自由としても規制が可能なので、そこをベースに規制を考えるということはあると思う一方で、マイクロターゲティングが、政治的なマッチングとしての側面をもつという点も無視できないと思う。つまり自分の考えと近い政治候補者とか、そういうところとどうやってつながっていくかという民主政治の活性化の問題とも、ターゲティングの話は関連してくると思うので、弊害とメリットの部分と両方見据えながら規制の枠組みを議論していったほうが良い

のではないか。【水谷構成員(4)】

- ◆ 事業者の取組（コンテンツモデレーションの方針・体制・実施状況、プロミネンスなど）の透明性・アカウントビリティ確保の在り方
  - EU のデジタルサービス法（DSA）のように、コンテンツ・モデレーションのポリシーの公表や、モデレータに実施している訓練内容や、AI による自動処理のエラー率などの記載を求めていくことも一案。また、削除やアカウント停止などの対象になったユーザーに具体的に理由を説明することや、判断が間違っていた場合の対応など苦情処理体制の整備も求めていく必要。【森構成員(1)】
  - 課題として思っているのが、具体的に何をどういうふうに透明性を確保して、それをどういうふうに活用するかというところの具体を詰められていないんじゃないか、並びに、それを外資系の企業も含めてどのように実行していくか。さらに、日本ローカルの透明性をどのように持たせていくか。あるいはユーザーに日本語で対応できる体制をつくる、こういったことを求めていくことが大事。また、TikTok のような、今すぐく伸びているサービスも対象にする必要。【山口構成員(1)(2)】
  - デジタル空間における情報流通の健全性について、偽・誤情報の拡散に対する技術的対策として、AI を用いたコンテンツモデレーションは、効果と効率性の観点から必須。しかしながら、透明性・アカウントビリティの確保が極めて重要であり、課題もある。AI による推論は原則ブラックボックスであり、たとえソースコードを公開したとしても、どのように説明責任を確保するのか、難しい状況。また、ここでセキュリティ的には AI の学習データやベンチマークを公開すると、それを逆手にとって AI の自動検知を迂回するような偽・誤情報の生成手法が出現する可能性があるため、そういうところに注意しながら透明性、アカウントビリティを確保するのが肝要。【越前構成員(1)(3)】
  - 構造上の問題と認知バイアスの問題が絡み合う状況下で、社会的なステータスであるとか社会的な評価であるとかいうことを使って緩和させていかざるを得ないのだと仮にしたら、そういったところをちゃんとプロミネンスさせていくということが、いわば認知バイアスに引っかかってしまったとしても最悪の状態にならないみたいなことになるのか。【クロサカ構成員(2)】
  - DPF 事業者の情報管理には、コンテンツ・モデレーションについても AI が非常に重要であり、コンテンツの種類によっては9割がた、AI によってフラグ立てが行われている現状がある。これは1日当たりで膨大な量のコンテンツをチェックしなきゃいけないので、人力ではもう到底手が回らないということで AI が活用されている側面ももちろんある。【水谷構成員(4)】
  - コンテンツ・モデレーションの世界というのは昔ながらの法学がイメージしてくるような、法廷でこの言論は虚偽かどうかとか、あるいはこの言論は権利侵害がされているかどうかというのを緻密に時間かけて議論して判断するというような仕組みというよりは、どちらかというところパブリック・ヘルスの世界、例えば誤情報でいうと誤情報にユーザーが接触する確率を減らすといった世

界、むしろ感染症対策に似ているかなと思うが、そういう世界観になってきている。当然、そうした接触をゼロにすることはできないので、コンテンツの削除においても、日々、膨大な量のエラーが出る。ここでいうエラーは過剰にコンテンツを消してしまったり、あるいは本来消されるべきルールに反するコンテンツが残っていたりというような状況を指すが、これらが膨大な量のエラーが毎日出るということを前提に考えなければいけない。【水谷構成員(4)】

- ケイト・クロニックという、DPF 事業者をニューガバナーと位置づけた研究者が、自身のブログでここ数年のビッグテック改革、まさに透明性の確保とか、そういう部分については、直感に反するかもしれないけど実は黄金時代だったと指摘している。確かに DPF 事業者は透明性レポートも自主的によく出しているし、それこそ、フェイスブックは監督委員会という最高裁みたいな外部組織を、非常に多額のお金を出して作ったりしていて、実はそうした改革が自主的に進んできた時代でもあったが、昨今の経済不況でビッグテックも経済的にその力が衰えてくると、社会に対する影響力は弱まっていないけれども、トラスト&セーフティだとか、公共政策の部門などから人とか資金が引き上げられてしまい、自主的な取組の後退が見られるようになる。【水谷構成員(4)】
- 国連の特別報告者のデビット・ケイさんが、DPF 事業者のコンテンツ管理の仕組みを、プラットフォーム法と評していて、非常に不明瞭な仕組みをしていると指摘している。デジタル空間の秩序形成の主導権が公的機関から私的機関であるプラットフォームに委ねられていると、これが憲法が大切にしてきた法の支配という仕組みを後退させるのではないかというような懸念があったりとか、あるいは、もうちょっと怖い話でコンテンツ・カルテルというのが形成されているというのをアメリカ研究者が指摘をしたりしている。DPF 事業者間や DPF 事業者と政府の間で、有害な情報への対策等についての協調体制がとられ、その中で、不透明な中で適切なモニタリングもなくコンテンツやアクターをサービスから排除するようなことを、カルテルのように一致協力してプラットフォーム間が内々で取決めをしてしまい、さらにそこに政府関係者の関与を強めると、DPF 事業者が検閲代理人化するんじゃないかという懸念が指摘をされている。【水谷構成員(4)】
- (AI を用いたコンテンツ・モデレーションに関し) 透明性と一口にしているけれども、どこまで開示すれば十分といえるのかというのは、論文とか読んでいてもいろんなレベルで議論されているところがある。あくまで私個人が考えているところだが、特に AI がどの程度、プラットフォームのモデレーションに食い込んでいるのかというのをきちんともっと明示していただく必要。恐らくコンテンツの種類によって AI が適切に使える、使えないといった差があり、例えば、難しいのはヘイトスピーチだと思う。ヘイトスピーチは国によって文脈が全然違い、ある国ではヘイトスピーチになるものが、ある国では全然ヘイトスピーチとして通じないというようなことがあるので、その問題がおそらくあるため、まずは、どこまで AI が入っているのかというところ。もう一つは、絶対エラーは一定程度出ているはずなので、このエラー率、どれぐらいエ

ラーが出てしまっているのか。AIを入れたけど、特定の分野ではむしろエラー率が高いといったことが分かれば、どこに課題があるかわかるので、偽陰性、偽陽性がどの程度、出てきているのかというのをきちんと開示していただくことが、ポイントではないか。【水谷構成員(4)】

- 透明性というと全ての人に公開しろということが前提になっているようにも思ってしまうが、アメリカの PATA は資格を持った研究者にのみアクセスさせるみたいな、そういうことを前提にもしているのだから、透明性で出す情報でアクセスできる人たちというの範囲を限定しながらやる方法もあると思いき、その辺を複合的に考えていくべき。【水谷構成員(4)】
- コンテンツ・モデレーションの仕組みは、先ほどもお話しした通り、だいぶシステム的な仕組みになっているので、個別のエラーについて事後的に責任追求するということが、非常に難しい。そのうえで、DPF 事業者の場合は、コンテンツ・モデレーションの部門と経営部門というのが基本的には報道機関のように分離していない、ピラミッド型でモデレーションの部門、トラスト&セーフティ部門等がある。つまり報道機関で言われるような編集と経営の分離といったエシカルな議論を基本的には採っていない。しかも、モデレーション実務の多くを外注しているといった側面もあり、これがモデレーションの責任の所在を外からますますわかりにくくさせているので、責任の所在を議論するためにもコンテンツ・モデレーションのルール形成とか、コンテンツ管理のオペレーションに関する政策といった点に関する年次計画みたいなものをまずは示してもらった必要があるのではないかと。そこにどういう人たちが関わっていて、どのようなプロセスでできているのかという、ここも結局、透明性の話になるが、責任所在をはっきりするためにも透明性の確保が重要になってくる。【水谷構成員(4)】
- AI の透明性の観点では、全て公開してしまうと、それを迂回して攻撃する攻撃者も考えられるので、そのバランスをどう取るかというのは結構、重要な課題かと思う。【越前構成員(4)】
- (デジタルサービス法は) コンテンツ・モデレーション、端的には少なくとも違法コンテンツではないものの、有害情報やディスインフォメーション等に対して、少なくともこの部分では積極的に削除せよということはない。何を言っているかという、透明性と、そしてそれによって影響を受ける利用者等の救済という形。DSA は、コンテンツ・モデレーションというものに対して恐らく初めて明確な法的定義を与えていて、まずは違法情報の削除、あるいは利用規約に適合しない情報、有害情報や、ディスインフォメーションが広く含まれる。しかし、その定義はあくまでデジタルプラットフォームの側がする形になっている。識別したり、削除のほかにも非常にいろんな手段がある。降格、収益不能化・デマネタイゼーション、アクセス不能化、可視性に影響を与える措置など、非常に広くとっており、例えばしばしば話題になるようなシャドウバンのようなものも含まれてはくる。それを徹底的に透明化する、利用規約にこういうコンテンツ・モデレーションをどうやっているのか、どんな人員体制



でどんなモデレーションをアルゴリズムを使ってやっているのかといったことを記載する。特に大きいこれらの方々は各加盟国、少なくとも 27 プラス、EEA の言語で提供しなければならない。そして、それをどのように実施したかということについて、少なくとも年に 1 回、VLOP は年に 2 回、透明性レポートというのを提示する。政府の透明性という意味だと各国政府からどんなコンテンツ削除の要請があったり、情報請求の要請があったのか、あるいは違法・規約違反別の対応件数、それから大体対応 1 件あたりに中央値としてどのぐらいかかったのか。それからコンテンツ・モデレーションの人員では、大体何語ができる人がどのくらいいて、どんなトレーニングをしているのか。または、自動処理のエラー率指標と、それからそのエラーがあった時にどういうセーフガードをとっているのかといったようなことを事細かに出さないといけない。

#### 【生員構成員(4)】

- ディスインフォメーション対策等で重要なのは、何語ができるモデレーターをどのくらい抱えているか。X は、言語の偏りが大きいといったようなことも言われているところだが、例えばオランダ語ができる人は 1 人確保しているだとか、イタリア語ができる人は 2 人確保しているだとか、そういったことがしっかりと透明に公開をされることになっている。【生員構成員(4)】
- 透明性の側面として面白いのは、コンテンツが削除されたりする。それはしばしば、大体自動的にやるので間違えることもあるといった時に、コンテンツのモデレーションをやったら、その影響を受けた情報発信者等にちゃんと一々、1 件 1 件理由の通知をしないとイケないことになっている。そしてその通知を受ける、そうすると初めて反論の確保、機会が確保される。GDPR のプロファイリングのアナロジーで言えば、まさに機械の決定だけに服さずに人間の関与を求める権利と言っても良い。こういった苦情があった時は、これには別のルートで個人で削除してくれと言ったのに、例えば誹謗中傷等、削除してくれないといったようなものの苦情処理も、ちゃんと理由を通知した上でこういうシステムの対象にしないとイケないが、ちゃんと公平な判断をしないとイケない。それでも納得いかなかったら、例えば ADR を使うことができるといったようなことを定めていたりする。【生員構成員(4)】
- データベースもすごく面白く、例えば今、既に 1 か月ぐらいで 8 億件ぐらい理由通知データベース集まっているが、この中で種類ごとに生データが当然匿名化された形で出てくる。例えばミスインフォメーションというタグで検索をすると、タグがまだ整理されてないので 100 万件という比較的、数は少ないが出てくる。例えば、TikTok とピンタレストの中身を見てみると、例えばこういう理由で利用規約のこれに違反するから削除されました、利用規約、Terms of Service の根拠はミスインフォメーションであると。コンテンツタイプはビデオである。そしてカテゴリーは市民の議論や選挙への影響だといったようなこと。誰かから言われたんじゃないじゃなくて自分で探して消した。そして意思決定は人間がやったのか、それともアルゴリズムがやったのかというと、これはフリー・オートメテッドだといったような、こういうデータが非常に事細かく出

てきて、こういったデータ一つ一つ確認できるだけではなくて、今、まさにこういうビッグデータを使ったコンテンツ・モデレーションの全体的な研究みたいなのもだんだん進んでいたりする。【生貝構成員(4)】

- 利用規約が大事であるがコロコロ変わる、フォローできなければ後で確認可能性がないということで、利用規約が変わっていくのをちゃんと時系列でもしっかりと把握した上で、変遷も後でチェックできるようなデータベースも作っている。こういうテクノロジーに基づく、ある種の透明性の確保というののもいろんな形で重要なんだろうなと思う。【生貝構成員(4)】
- モデレーションの透明性と救済というのは果たしてどのような情報空間が健全であり、望ましいのかということ和社会で全体で議論する、それに対して個人が異議を申し立てる、そういうのを大前提と考えるべきであろう。【生貝構成員(4)】
- プラットフォーム側と受信者との関係で言いますと、レコメンドシステムの在り方、プロファイリング、コンテンツモデレーションの在り方もあるし、プロミネンスの在り方ということで、メディア等をできるだけ前に表示させるかどうかといったこともあろう。【落合構成員(5)】
- また、デジタルプラットフォームやデジタルサービスの伝送場面においても、コンテンツモデレーションの実施状況や、エコシステムの状況についても分析を深めることが有益。【落合構成員(5)】

◆ 事業者の取組（レコメンド、データ取扱い、広告収入の発信者への分配など）の透明性・アカウントビリティ確保の在り方

- モデレーション、レコメンデーションといったものは、個人データを処理したプロファイリングなどの活動と密接不可分。全ての事業者にも適用される個人情報保護法などの規範は重要であり、デジタル空間における情報流通という文脈の中での情報保護、データ保護の在り方もよく検討していく必要がある。【生貝構成員(1)】
- 国際的にはGDPRの議論に即したものが非常に多く、最近の情報発信としては、世界プライバシー会議（GPA）で採択された生成AIに関する声明がある。これによると、開発段階、運用段階、導入段階のそれぞれに分け、いわゆるGDPR的な諸原則がうたわれているが、中でもAIシステムが個人に関する意思決定を行ったり、意思決定を支援したりするような場合には、生成AIを使うことに慎重な姿勢が示されていることや、学習データに個人情報を利用する場合には、事前にプライバシー影響評価にかけよう求めている。これはプライバシー・バイ・デザインが強くうたわれる傾向にある中で、どう実装するかというときにプライバシー影響評価を行いなさいとなっている。また、不正確なデータを取り除くためのガバナンスの手続きを取ること、児童のデータ処理を行うときの権利保障に注力することが挙げられており、日本における議論でも留意しておいた方がよい事項が最近の国際的な動向から見て取れる。生成AIの問題は、今、議論すべきインターネット上のデジタル空間の健全な情報流通に係る論

- 点を浮き彫りにするという面でも注目しておく課題。【石井構成員(1)】
- 偽情報が、その影響を受けやすい人物のもとにピンポイントで届けられ、人心操作（マインドハッキング）が行われる可能性があることは、ケンブリッジアナリティカ事件によって広く知られるようになった。フェイスブック利用者の大量のデータを入手したケンブリッジアナリティカ社が同データを詳細に分析し、「陰謀論に傾きやすい」などとプロファイリングされたグループに集中的に政治広告などを送り、2016年米大統領選や英国のブレグジットにおける世論操作を試みたとされる事件。現在、プラットフォームによって大量に収集された利用者のデータを使って操作対象者を選び、プラットフォーム事業者がターゲティング広告の配信のために作り出したシステムを使って操作対象者に偽情報を含むメッセージを容易に届けることが可能になっている。偽情報はプラットフォーム事業者のビジネスモデルを利用して、その破壊力を強めており、プラットフォームのビジネスモデルによって生じる各種の問題は、偽情報の悪影響の要因分析として重要。【森構成員(1)】
  - レコメンドに使われるのはオンライン上で集められたユーザーのデータであり、データ収集の方法・態様とプロファイリングに代表されるその利用方法についても留意する必要。【森構成員(1)】
  - 食品表示法は、特定の食品の摂取を禁止するのではなく、食品に関する基本情報を表示させることで、食品の安全性を確保し、食品摂取に関する消費者の自主的かつ合理的な選択の機会を確保するもの。情報の摂取についても同様の試みが必要。「情報的健康」のためには、どのような情報を「食べているのか」、「食べさせられているのか」をユーザー自身が知ることが重要。何を「食べている」のかを理解するためには、情報・コンテンツの作成者等のわかりやすい表示や、レコメンダーシステムの透明性が重要。【山本構成員(1)】
  - EUではGDPRでプロファイリング規制を置いているのに加えて、オンラインプラットフォームなどはターゲティング広告のパラメータ、レコメンダーもパラメータを出さないといけない。正確に言うと、コンテンツ・モデレーションの定義の中にこの2つは入らない。同じに扱うのではなくて、ユーザー投稿コンテンツへの対応とは違うため。しかし、ここまで視野に入れた形での規律というものを考えていくことが包括的な対応にはなるということなのだろう。センシティブデータを使ったプロファイリング広告禁止、青少年保護と未成年個人データのプロファイリング広告禁止ということで、青少年の脆弱性に特に配慮したような規定なんかも置かれている。【生貝構成員(4)】
  - データ保護、プロファイリングはモデレーションと同じか、場合によってはそれ以上に重要なルールというものにもなってくるだろう。そしてVLOP、VLOSEに関しては日本の立法技術上、こういったような仕組みがどの程度可能かというのは難しいが、しかし個別のディスプレイフォメーションを削除せよというのではなく、まさにシステムリスクを管理してもらい、軽減してもらい、そして、その具体的な形は行動規範等を通じて具体化していくようなことは、今の時代のテクノロジー分野のルールで不可欠なのかなと感じているところ。同

じようなことが多分、汎用・生成AI、特に大きなものについては必要だという議論も出てきている。【生貝構成員(4)】

- (欧州では) 政治ターゲティング広告規則というのもそろそろ終盤に差しかかっているところで、政治と言っても全然選挙だけではなく、特定の政策 이슈に関わるような活動なんかも含めた規律や透明化といったようなものも同時並行で作られていて、ホリゾンタルなDSAと、それから政治・民主主義といったようなところのバーチカルな問題というのも併せて出てくるのかなと思う。【生貝構成員(4)】
- より重要になってくる部分として、プラットフォーム事業者、サービスと発信者側とのトラフィックですとか、広告収入の分配、コンテンツ枠の設定といった辺りがどういう形で行われているのか。ここでのどのような動きによって、今の情報格差の仕組みというのにインセンティブ付けがされているのかを見ていくことは、より重要ではないか。【落合構成員(5)】
- 発信、受信に関するデータの授受と、拡散される情報そのものの伝送は必ずしも一致しないような場合もある。データと拡散される情報とを切り分けて分析をしていく部分も必要ではないか。また、個別要素としては、データの利用や生成AI、サイバー攻撃等については、技術の進展、また、実務、ルールの変化によってどういう変化が生じているかを確認することは重要ではないか。【落合構成員(5)】

#### 4. 主に情報受信段階の課題

##### ◆ 認知的・社会的バイアスを前提としたリテラシー向上策の在り方

- 誤情報対策は主にプレバンクとデバンクに分けられることが多いが、プレバンクは誤情報が拡散する前の備えのことを指して、リテラシー向上などはこれに該当。【田中構成員(1)】
- デジタル空間における情報流通の健全性確保という、検討会のテーマに関連して一つ、心理学では最近大きな動きがあった。11月にアメリカの心理学会が誤情報対策に対する共同声明というものを出し、この共同声明とともに、心理学を誤情報の理解と対策に活用するというタイトルの報告書が出されている。この報告書では、なぜ人が誤情報を信じるのかや、誤情報に対抗するにはどのような介入が効果的か、などの問いに対して人の要因に焦点を当てながら論じられており、専門的なところで言えば、記憶や推論のメカニズムや介入の長期的効果検証などは、まだまだこれから解明が必要な部分もあるが、この報告書の最後に数百本の実験論文が引用されているように、学術的な知見が急速に蓄積されていることも事実で、現時点で明らかになっている心理学的知見を社会問題の対策に活用する観点から、この報告書では制作者や実務家の方々に向けて誤情報による脅威に対処するための推奨事項というものが整理されている。【田中構成員(4)】
- 誤情報に対する介入の捉え方は、いろんな捉え方があるが、個人レベルとシステムレベルに分けて捉える考え方があり、下のシステムレベルにはアルゴリズム

ムによる介入や法的介入が含まれている。心理学が対象とするのは個人レベルによる介入で、Boosting や Nudging、デバンキングなどが含まれている。これらは、目的・背景にある理論・期間・タイミングが異なるもので、例えば、ブースティングのところにあるプレバンキングは、誤情報に対する事前の備えという位置づけ。一方で、デバンキングは誤情報が既に拡散した後の事後対応という位置づけ。基本的な考え方は not a one-size-fits-all solution という考え方で、単一で万能な手法があると仮定するよりは、まずどのような介入手法があり得るのか、この 10 年ほどで急速に研究が発展する中で複数提唱されてきた流れがある。同時に効果の検証が行われていまして、ようやく図で表されているようなカテゴリーに整理されてきた。【田中構成員(4)】

- プレバンクは、事前に誤情報に対する心理的な耐性を築くことを目的とした介入手法。もともとは 1960 年代に社会心理学者が提唱した接種理論というものを土台としたものだが、現在はケンブリッジ大学の研究チームを中心に心理的予防接種という介入方法に発展していて、WHO や Google などとの共同研究による大規模な実証研究が行われている。プレバンクの主な内容は、誤情報に近い将来、出くわす可能性のある人々を対象にその可能性を警告すること、また、誤情報に対抗するためのあらかじめ反論材料を提供する構成。【田中構成員(4)】
- プレバンクは、具体的な手法から、能動的なプレバンキングと受動的なプレバンキングに分けられている。能動的というのはインタラクティブだとか対話型とか、ユーザーの積極的関与が前提とされている手法。例えば、この黄色の BAD NEWS というゲームでは、ユーザーはあえてフェイクニュースを拡散するお題が与えられ、その目的のモットーは、よりフォロワーを獲得するためのメッセージを選択するというような構成。その中で、誤情報を広めるためによく使われるテクニックがどのようなものであるのかを学んでいく。ゲームではこのテクニックを理解して使うと、ポイントの代わりにフォロワーが増えていたり、各テクニックを理解するごとにバッチを与えられたりとか、ゲームの中でも言葉遣いが平易で会話調にするなど工夫されていて、比較的ユーザーが飽きにくく使い続けるような工夫がされているのが特徴。BAD NEWS 以外にも政治やコロナなど特定のトピックに関連したゲームが後続で開発されていて、WHO と共同開発された GO VIRAL は、できるだけ利用しやすいように 5 分程度という短めに設定されている。BAD NEWS ゲームも Go Viral も残念ながら日本語版はないが、ヨーロッパを中心に複数の言語に翻訳されている。また、効果検証もメタ分析が可能になる程度までは蓄積されていて、受動的なプレバンキングと比べると、比較的效果が持続する報告がされている。能動的なプレバンクは効果が得られやすい一方で、時間が長かったり、ユーザーの積極的な関与が求められるために広くリーチしにくいというような特徴がある。【田中構成員(4)】
- 受動的プレバンキングという手法では比較的短い動画や、ポップアップ形式のテキストメッセージを提示する形で誤情報で使われるテクニックに対抗するための情報を提供する方法をとる。例えば、ケンブリッジ大学のチームが Google と共同で実施したフィールド調査では、各テクニックごとに 30 秒から

90 秒程度の短い動画を作成し、YouTube で視聴させたところ、誤情報の共有意図を低減させた結果も報告されている。【田中構成員(4)】

- 能動的プレバンキングも受動的プレバンキングも誤情報に対する心理的耐性を向上させる効果が見られてはいるが、その効果は時間経過とともに減少する傾向があり、この効果検証そのものが北米やヨーロッパ中心で行われている現状で、日本を含めて他の文化圏でどの程度効果があるのかについては、研究が不足しているのが現状。【田中構成員(4)】
- 認知プロセスがあって、信じ続ける人がいて、そこから行動に移るかどうかに関する知見については様々な別のファクターが絡んでくる。例えば、ワクチンに関する誤情報を信じ続けている人がワクチンを打ちに行くとか、行かないとか、逆に海外ですと妨害しに行くみたいなパターンがあると思うが、そういったものも近くにあるかどうかや、どれぐらいの数があるかどうか、どのようにWHO とかが介入しているかによっても変わってくるので、複雑な要因が絡んできている。【田中構成員(4)】
- アメリカで連邦議事堂の襲撃事件があり、その時に各ニュース機関がいろいろ選挙結果についての報道をして一生懸命やったが、結局それが大規模な人たちの行動を招いてしまった。政治的ないろんな働きかけも要因としてあり、どこまでニュースメディアができたことは分からない。【奥村構成員(4)】
- 現状の情報空間というのを健全化することの難しさについて、これは非常に個別の現象面というよりも構造的な問題だとか、これは例えばアテンションエコノミーというビジネス構造にも関わるが、さらに個人のレベルでいくと認知システムのような、我々がふだん意識しないような、そういう非常に認知のレベルで一つ原因・根源的な問題があり、その難しさを感じた。そう考えると、単純にディスインフォメーションをモデレーションしていけばよいとか、あるいはファクトチェックを増やしていけばよいということ、これも非常に重要だが、リテラシーでもこういった認知科学の知見を積極的に参照していくことが求められると感じた。【山本構成員(4)】
- 情報環境における様々なバイアス（偏り）について、社会心理学的にまとめるとすると、認知的バイアス、情報環境のバイアス、社会的バイアス、大きく分けて3つに分けることができる。【安野構成員(2)】
- 情報の処理における認知的バイアスについて、様々な段階で私たちの認知システムの中に入っているバイアスが発動する。なぜこのようなバイアスが起きるかといえば、私たちが認知的に、様々な情報を合理的に、また完全情報で判断することはできないという脳、それから情報環境の制約によるというもの。特に私たちの認知処理のバイアスのことを認知的バイアスというふうと呼ぶ。虚偽情報、誤情報に関係して、まず情報収集の段階、それから推論や判断の段階、それから思い出す記憶の段階の3つ、それぞれにおいてバイアスが影響する。【安野構成員(1)(2)】
- 情報収集の段階では、いわゆる確証バイアスと言われるもの。ウェイソンの4枚カード問題（カードの片側にはアルファベット、裏には数字が書いてある。

母音の裏側には必ず偶数があるというルールが成り立っているかどうかを確かめるためには、どのカードをめくらなくてはならないだろうかという問題)では、偶数の裏に子音があってもそれは別にルール違反ではないが、奇数の裏に母音があったらルール違反なので、それを見なきゃいけないが、多くの人があくまでも予想を確かめる情報検索をしてしまう。自分が持っている前提で情報を評価することは、メディア報道の評価でも起きる。あくまでも自分の予想を確かめる方向で情報を集めようとしてしまう。変形4枚カード問題(居酒屋で学生たちが飲物を飲んでいるところ、そのうち2人の年齢は分かっているが、何を飲んでいるか分からない。別の2人は何を飲んでいるか分かるが年齢は分からない。お酒は20歳になってからというルールが守られているかどうかを確認するには、どの人に年齢のあるものを確認すればよいか)による日常的な規則違反探しの場合は、ウェイソンの問題と構造的には同じ問題でも正解率が高くなる。どうしてこういうことが起きるかということ、最適化ではなくて満足化を私たちはしているからであり、サイモンの限定合理性という概念で、意思決定において必ずしも最大限の努力は払われない。目立つ刺激への注目をしやすく、これが因果推論にも用いられやすい。【安野構成員(1)(2)】

- 推論の段階のバイアスについては、私たちは特に確率について合理的に判断しにくい。参照点に近いほど僅かな差も大きく感じる。これはプロスペクト理論(あなたがある4択クイズの番組で、賞金1万円まで勝ち進んでいるところ、今やめたら1万円持ち帰ることができる。先に進んだら4分の1の確率で80万円がもらえるが、4分の3の確率で何ももらえない。期待価値が1万円と20万円のところ、あなたならどうしますかという質問をすると、80万円にチャレンジするという回答がかなり圧倒的な多数になる。また、あなたがある4択クイズの番組で、賞金20万円まで勝ち進んでいるところ、今やめたら20万円持ち帰ることができるが、先に進んだら4分の1の確率で100万円がもらえるが、4分の3の確率で何ももらえない。期待価値は20万円と25万円のところ、あなたならどうしますかという質問をすると、20万円もらって帰るという回答が結構多くなる。期待価値を比較した場合、どちらも次の質問に進むのが合理的だが、そうはならない。なぜかということ、1つは、参照点、いわゆる原点的なものだが、これに近いほど僅かな差も大きく感じる。例えば、1万円と2万円の差が非常に大きく感じるけれども、100万円と101万円の差はあまり大きく感じないといったようなことがある。また、利得よりも損失に敏感であって、1万円の価値は、もらえる1万円よりも払う1万円の価値の方が重く感じる。より重要なこととしては、参照点が文脈によって移動する(フレーミング効果)。例えば、20万円もらえるはずだったのに、先に進んだことでもらえなかったと思ってしまうと、実際には何も損はしていないが、損したというふうに思ってしまう。)と言われるもの。例えば、ワクチンの評価も、助かるフレームと副作用フレームで異なる可能性があり、虚偽情報や誤情報もどのようなフレームで提示されているかに注意する必要がある。また、認知的反射(熟慮性)テスト(バットとボールは合わせて1ドル10セントで、バットはボールより1ドル

高いところ、ボールはいくらか。)において、アメリカの名門大学の学生でも10セントという回答が多くなるということをカーネマンが報告しており、冷静に考えれば小学生の算数で解ける問題でも、熟慮せずに直感的に答えてしまう。また同様に、因果関係を判断できない事例として、共変性の誤認知がある。例えば、あるダイエットサプリが売れているところ、このサプリで痩せたというレビューが5000件集まっていて、これらの感想はうそではなさそうな中、このレビューだけでこのサプリに効果があると言えない理由を考えるという問題において、よく学生は5000件じゃ少ないという数に着目した回答が多かったりするが、そうではなく比較がいる。つまり、薬を飲んで痩せた人痩せない人、薬を飲まないで痩せた人痩せない人、ここを比較しないといけないが、そこまで比較せずに判断してしまうと、因果関係を推定するために必要な情報が何かということに私たちは気づきにくい。【安野構成員(1)(2)】

- 記憶の段階については、虚偽の記憶というのは後から提示された情報でゆがむという Loftus による研究結果がある。例えば、車がぶつかったときに出していたスピード、車が衝突したときに出していたスピードという聞き方によって、見ていたものは同じ映像であっても、その出していたスピードの推定が変わる。また、ムードや感情によっても思い出す情報が違う。私たちの記憶はなかなか当てにならない。ネット上で見た記憶、あるいはネット上で接した情報を判断する際に私たちが参照する記憶も正しいとは限らない。【安野構成員(1)(2)】
- そもそも情報をきちんと見ていない人の存在がある。ウェブ調査をしていると、いかげんな回答が出てくることがあるが、ウェブ調査の設問について読み飛ばし傾向がある人たちがどのような情報接触をしたかということを見ると、意思決定と情報探索の実験において質問項目の読み飛ばしをする回答者は、しない回答者よりも情報探索数が少なく、全く探索しないという人も4割に達していた。正しい情報が提供されていても、そしてそれが簡単にクリックできる場面でも、あまり関心がなければ情報を確認しない。全く確認しない人もそれなりにいる。【安野構成員(2)】
- 2つ目の情報環境における様々なバイアスとしては、アルゴリズムによる「フィルターバブル」、選択的情報接触や等質な対人的ネットワークによる「エコーチェンバー(echo chambers)」、関心を集めることが経済的利益を生む「アテンションエコノミー」、SNSなどによる大量の拡散情報をもたらす「幻想の真実(illusory truth)」効果による認知的流暢性の影響、伝統的マスメディアへの信頼の低下による「世論」「公共の問題」の見えにくさ等がある。【安野構成員(2)】
- 3つ目の社会的バイアスについて、認知的バイアスと情報収集の偏り、この2つを合わせると、正しい知識が乏しいことが考えられる。また、関心のないトピックについては、わざわざ情報収集のコストを支払わない。面倒なことはしないことから予想されることは、特に関心が薄い政治といったことについては、あまり知識がない人が多い可能性がある。実際に政治的知識の研究では、有権者の政治的知識は総じて低水準にあるという指摘もある。これは1960年代か



らずっと報告されてきており、例えば、有権者の政治的知識や態度の安定性は最小限であるため、ふらふら変わってしまうし、民主政治の担い手であるはずの市民が政治的知識や関心に乏しいとか、また、所得や学歴、性別などによって政治的知識に格差があるという指摘もある。アメリカにおいて、特に、豊かな白人男性と貧しいアフリカ系女性との政治的知識の量が、知識の量に大きな差があったが、現代でも、こういった社会的な属性と政治的知識の格差については報告されており、また、日本でも報告されている。なぜこのような差が出てくるかという、例えば、学歴の差とか若年期の政治的な社会化の影響というのが考えられるが、ウクライナ、国交、海外の政治家に関する知識を質問し、この正解、不正解を従属変数として、あえて社会的属性だけを独立変数として分析してみると、男性のほうが女性よりも正解率は高く、年齢が高い人のほうが正解率が高く、また資産、貯蓄がある人が正解率が高い。また、大卒ダミーはおそらく学校教育の効果と思うが、これが効いている。自家用車はマイナスだが、これはおそらく住んでいるところの属性の可能性もある、貯蓄がまたプラスに効いている。社会的に恵まれた層のほうが政治的知識があるという傾向は、現在の日本でも見出されている。【安野構成員(1)(2)】

- 認知的バイアスと社会的バイアス（政治的知識の乏しさ、社会的属性による知識の格差）、この2点を前提とする必要があり、デジタル空間において正しい情報を提供できてさえいればよいとは言い切れない。正しい情報を提供する努力をしても届かない層が存在する可能性を前提とする必要。正しい情報を提供するということはもちろん大切であるが、それを提供する努力をするだけでは虚偽情報や誤情報の影響を読めることは難しい可能性がある。【安野構成員(1)(2)】
- 私たちは必ずしも積極的な情報接触を行っているとは限らない。情報環境には正確とは言えない情報が溢れていても、目新しさや興味で情報に接触してしまいがち。また、情報の真偽を判断するときには、自らの既存の態度や対人環境の影響を受け、分かりやすいものは事実だと思われやすい、認知的に処理しやすいものは事実だと思われやすいということに注意する必要がある。そして、私たちにとって政治は最重要事項ではないことが多いので、一般に有権者の政治的知識が乏しく、また、社会的に不利な立場にある人ほど政治的知識が少ないという指摘もあり、これは民主主義社会を維持していく上で非常に問題になり得る。【安野構成員(2)】
- 虚偽情報・誤情報の拡散が人々の認識に影響を与えてしまうという懸念。繰り返し接触すると正しい知識を持っている人も illusory truth の影響を受け、誤った情報も正しく見えてしまう。認知的な処理が簡単になることで正しく見えてしまい、しかも、知識があってもだまされてしまうことがあるということも重要な問題。一方、正しさを意識すると、興味深さを意識した場合よりも影響を受けにくいという指摘があるが、生成 AI を用いたサービスなど、作成又は拡散の責任が曖昧な虚偽情報・誤情報が拡散する中で、そのような情報にたくさん触れてしまった場合、正誤の判断が私たちにはますます難しくなってい

- くことが今後の課題。【安野構成員(1)(2)】
- まずは私たち自身がいつでも間違えるということ認識すること。私たち自身が、また虚偽情報や誤情報が世の中にはたくさんあって、それによって利益を得る人がいるということに気が付く。また、面白さよりも正しさにちょっと注意を払うということで、少しは防げるかもしれない。まずは私たちが、全員が虚偽情報や誤情報に対して決して強いわけではない、必ず見抜けるわけではないということを広く知っていただくということが重要。私たち人間は基本的に認知的に怠け者であり、自分視点でしか判断していないということを全員が前提とする必要。【安野構成員(1)(2)】
  - 認知バイアスはリテラシー教育で改善するのかについて、どこに着目するかとか、情報の提供の仕方で認知バイアスの出方が少し変わってくるという研究がある。完全にはおそらくなならないと思うが、リテラシー教育で少し弱めることはできるのではないか。虚偽情報、誤情報などについても、知識がある人も誤情報にはだまされてしまうが、面白さよりも正しさに着目すると、要するにその情報が正しいと思うかどうかということに着目すると、ちょっとだまされる確率というか比率が減るという研究もあるため、どういうところを見るか、どういうところに気をつけたらいいかというリテラシー教育で、恐らく改善はする。問題は、それをどのようにできるだけ多くの人に届けるかということになる。【安野構成員(2)】
  - だまされやすさや知識の吸収能力に関して年齢が持つ効果について、政治的知識に関して言うと、年齢はむしろプラスの効果を持つことが多い。比較的新しい制度的な変化などは、学校教育を最近受けた若い世代のほうが正解率が高いこともあり、学校の効果が出やすいのは若い層。また、新しい技術などに関係したことであれば若い世代が有利で、古い知識が必要なことであれば年齢が高い方が有利。おそらく年齢の効果というのは、何の知識かによって出方が変わってくるのではないか。【安野構成員(2)】
  - ウェブサイトの認証技術でSSL/TLSという、昔http://だったのがhttps://になって鍵付きになったというのがあるが、残念なことに今、フィッシングサイトの多くが既に鍵付きになっているという状態が発生していて、あまり意味がないという状態になってきている。このように、多くの人間はURLの窓の中身なんて確認もしなければ、鍵が掛かっているのかもほとんど気にせず、引っかかる時は引っかかる。つまり、構造上の問題と認知バイアスの問題が絡み合ってしまう、いたちごっこ化している状況。データプライバシーの研究者によって、実はどの企業を信用するのかというのは結局ブランドでしかないという、身も蓋もない実証研究をされている方がいて、こういった社会的なステータスであるとか社会的な評価であるとかということ、やはり我々は使って緩和させていかざるを得ないのだと仮にしたら、そういったところをちゃんとプロミネンスさせていくということが、言わば認知バイアスに引っかかってしまったとしても最悪の状態にならないみたいなことになるのではないか。【クロサカ構成員(2) ※一部再掲】

- ダークパターン等による詐欺サイト・定期購入・デジタル勧誘等に関する消費生活相談からのまとめとしては、消費者は、情報の真偽を確かめることができず、ネットの情報を信用する傾向があり、真偽を確かめる方法を知らない。インターネット上の細かい規約を確認できていない。デジタル契約について法律が追いついていないなどの問題があり、情報流通の健全性確保は喫緊の課題。  
【増田構成員(1)】
- ICT リテラシーの向上、インターネットの基本的な仕組みについての教育や啓発（フィルターバブルによって、繰り返し同種のトラブルに遭う、エコーチェンバーによる考え方の偏りなど）というのも非常に重要。【増田構成員(1)】
- 情報選択等において個人が合理的に判断するということには限界が実はあるということがかねてから指摘をされており、私たち個人には刺激に弱い脆弱性がある。こういったことを前提に議論を進める必要。【水谷構成員(1)】
- 1つの例として、お子さんを持って不安に思っている方、又は、お子さんへとにかくいい教育を受けさせたいと思っている方といった個人のライフステージや社会経済的、文化的なバックグラントかに応じて、情報に対する考え方が変わってくるというところも感じている。あるいは、治療法を求めて情報を求めている方も、ある種ライフステージの中で、病気になったタイミングとかで一生懸命情報を求める中で行き着いたのが偽情報だということもあり得るということで、受容者側の動的な変化を捉える必要性も分析の中で感じている。  
【澁谷構成員(2)】
- ライフステージなりライフイベントなりに由来して、ある種の情報に対する渴望というか、そういうものが発生して、そこに乗せやすいところもある。【森構成員(2)】
- メディア情報リテラシー教育の拡充が大事。AIによってリテラシーの枠を超えるものが生じているが、そうでない偽・誤情報も大量にあり、AIという意味でも、画像や映像にも誤ったものがあることを知ることも重要。メディアや情報の環境、特性も含めたメディア情報リテラシー教育を老若男女に実施していくことが何よりも重要。情報社会において、リテラシーを高めることは、教育を受けた人が生きる上で欠かせないだけでなく、リテラシーが高まるのは社会全体にとって非常によい結果を生み出すので、算数や国語の科目に匹敵するぐらい重要。実際、欧米ではメディアリテラシー教育は進んでおり、義務教育に入っているケースも少なくない。ただ、リテラシーと一口で言っても様々なリテラシーがあり、だからこそ、こういったリテラシーが重要なのかということの研究によって特定して、そのエビデンスを踏まえた教育啓発を行うことが重要。最近 RISTEX の研究プロジェクトで、リテラシー研究グループ長をやっており、今後の実証研究で重要なリテラシーを特定する予定なので、皆さんにシェアする予定。【山口構成員(2)】
- 教材を作るだけでは圧倒的に不足している。例えば、これまでのベストプラクティスという意味では、ユーチューブクリエイター、インフルエンサーと組んでキャンペーンをやったところ、760万回以上の再生回数になり、ものすごく

効果があったこともあり、そういったキャンペーンと、教材、講座、あるいは教育課程の導入、こういったほかの手段も含めていろいろなことを検討して、多様な、多角的な手段で啓発していくのが大事。【山口構成員(1)(2)】

- 音声詐欺の被害拡大が懸念される状況を踏まえ、積極的に中高年以上に啓発していく必要がある。1つは、ものすごい草の根の活動でいうと、自治体での講座とか、中高年以上の方へ地道に啓発していくというのは1つある。一步踏み込んでいくと、例えば高齢者にスマホとかを購入するときにそういったチラシを渡すこと、ほかにはマスメディアの活用。NHKが以前、私はだまされないという特殊詐欺のことをタイトルを変えてやっているが、唯一非難があった点は、だまされないと言ってしまうと、むしろだまされるようになるという研究結果も出ている。だから、みんなだまされるかもしれない事例を、マスメディアを活用して啓発していくこと。自分事として捉えてもらえるように、私もだまされないという話で啓発していくのが、1つ。【山口構成員(2)】
- AIの音声詐欺は、海外でものすごい問題になっていて、日本ではまだ少ないが、米国では14%以上、インドでは20%程度、要は個人を狙った、特にシニア層を狙ったAIの音声の詐欺は、脅威になっている。さらに、だまされやすいという特性もあり、特にシニア層にリテラシー教育をうまく周知させるための工夫が大事。【越前構成員(2) ※一部再掲】
- リテラシー教育や啓発活動は、サイバーセキュリティ対策としても取り組んでいるが、簡単な話ではなく、いつも若年層や高齢者層、ビジネス面、いろいろなところでそれぞれの啓発活動が要ると悩んでいるところ、情報の流通の健全性においても同じ。【後藤構成員(1)】
- 高齢者層については、今から教育と言っても困るため、実被害を避ける取組や、被害を受けてしまったときの救済策が要る。【後藤構成員(1)】
- 山口先生のお話では、いわゆるシニア世代はだまされやすい、また安野先生のお話では、資産があるシニア世代はちゃんと情報を見ているところもある。年代によって、大分情報を見る特性が決まっている。【後藤構成員(2)】
- 昨今のアンサーカルチャーというものの影響はすごく大きくて、分からなくて不安だが分からないもの分からないままに置いておくことができない人が増えている印象は非常に強く持っている。白黒すぐつけてリツイートしたりシェアしたりすることをすぐしたくなってしまうところを、リツイートやめましようではなく、分からないもの分からないまま放っておく強さをリテラシーの一部として取り入れていかないといけない。リテラシーも万能ではないので、その部分を補うような少し涵養するような議論をしていかないといけない。【奥村構成員(2)】
- 個人の観点から見れば、リテラシー教育、啓発、普及の在り方があるし、ただ一方で、消費者における対応に委ねることがどうしても難しいことはあろう。また、認知的バイアス、社会的バイアスや、誤情報の持続効果、真実の錯覚、選択的回避といったようなこれまで検討会で議論されてきたような論点も、対策の、特に個人側の対策の難しさに関わってくる部分もあろう。【落合構成員

## (5)】

- リテラシーについて従来は、誹謗中傷の加害、被害を防ぐとか、福祉犯を防止するとか、そういう観点からの情報共有がメインでありまして、「ユーザーリテラシー」とか「賢いサービスの使い方」とかそういうことにとどまっていた。近時では、ケンブリッジアナリティカや外部送信の問題への関心からサービスの背景にあるインターネット全体の仕組みとかビジネスモデル、そしてそれらがもたらす弊害についての知識の共有に関心が移ってきているのではないかと。また同時に心理学的アプローチ、認知科学的アプローチが重要になってきている。リテラシー向上検討会の成果物が公表されており、世代共通の課題の洗い出しみたいなことをしておりまして、(1) インターネット上で自分の好みの情報や自分と似た意見に触れやすくなる特性・仕組みの理解。フィルターバブル、エコーチェンバー、まさにこういうところを課題としてくり出していく。(3) インターネット上の情報を熟慮する機会の確保ということで、アテンションが重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ということ。【森構成員(5)】

## ◆ 発信・拡散主体となり得る受信者側のガバナンス・リテラシー向上策の在り方

- 偽・誤情報に出会った後に拡散している人というのが15~35%おり、拡散手段として最も多いのが直接の対話であり、インターネットだけの問題ではない。【山口構成員(1)】
- 偽・誤情報を信じている人は、誤っていると気づいている人に比べて、はるかに高い確率で拡散する。また、メディアリテラシーや情報リテラシーが低い人ほど偽・誤情報を拡散する結果も出た。こういった背景から、米国の研究では、偽・誤情報は事実のニュースよりも約6倍も早く拡散することがわかっている。つまり、我々が接している情報空間は、偽・誤情報を信じている人や、メディアリテラシー、情報リテラシーが低い人が偽・誤情報を拡散しやすい空間にある。だからこそ、訂正情報はなかなか拡散しにくいと言える。【山口構成員(1)(2)】
- 偽・誤情報は、消費者側・供給者側の行動パターンがネットワークの中で複雑に動いていて、消費者・供給者の明確な線引きができないというのも1つの問題。あるとき消費者が供給者側に変わり、受動的に受け取っていた方があるとき主体的に関わるようになったり、どのように関わっているのかというところがアクター間でいろいろと多様であって、そこをしっかりと分析することに意味がある。【澁谷構成員(1)】
- (AI) を使う側のリテラシー向上だったり、セーフティーネットをつくっていくというようなことも非常に重要。【江間構成員(3)】
- 高齢者層については、今から教育と言っても困るため、実被害を避ける取組や、被害を受けてしまったときの救済策が要る。一方、若年層に向けては、静止画と音声の録音さえあれば、3次元のおしゃべりするビデオが作れてしまい、中高生のスキルで使えるレベルまで、使いやすくなってしまっている中、最新の

ツールの利用を止めても仕方ないので、積極的に情報合成ツール等を活用してリテラシー向上させ、偽画像が簡単にできてしまうことを肌身で感じてもらったほうが早いのではないかと、という議論もある。今後、5年後、10年後には、個人差の拡大や、年齢層がだんだんシフトすることもあるので、それにどう対応していくのかということが大事。【後藤構成員(1) ※一部再掲】

- デジタル空間の中で利用者が発信者に回ってくることもかなり多くなってきている。利用者に伝達された情報が、また拡大再生産されるような、そういう構造も中にはあると考えている。【落合構成員(5)】
- もう一つ、重要と思っているのが、先ほどの図面が1対1か、1対nの形式で記載をしているが、受信者側のほうにグラデーションがついていたように、実際には、n対nの形で拡散が進んでいるように思う。このため、実際にはなかなか正確に一枚の図面では書きにくいところがあるが、その状況も見ていくことは重要。【落合構成員(5)】

## 5. その他課題に関するご意見

- 偽情報・誤情報は、様々な課題が絡み合っている複合問題。対症療法と根治療法の両方を並行して考える必要があり、こういった観点でできることを最大化するというところを目指すということが望まれていることではないか。【クロサカ構成員(1)】
- 情報流通の健全性を脅かしている「構造」を揺さぶるために有効な施策を検討すべき。【山本構成員(1)】
- この検討会は、タイトルが、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となっており、情報流通の健全性確保という比較的広いスコープを持っている印象を受け、偽情報には限られておらず、アテンション・エコノミーや、それとの関係で利用者データの取扱いもスコープに入ってくると思うところ、あまりセグメントを限定して対策とか原因を考えるのは、むしろ困難であるという御指摘もあったので、なかなか偽情報だけで話をしようとするとかえって無理がある。【森構成員(2)】
- 健全性検討会の開催の趣旨、設置の趣旨を見ますと、キーワード的には生成AIとかメタバースとか、デジタル空間、ステークホルダーとありますけれども、偽情報に限定するというのではなくて、デジタル空間における情報流通の健全性に関するこの3つの問題（権利侵害情報、偽情報、利用者データ）を中心に広く見ていくということ。このように本格的なクロスオーバーを前提に検討が進むということになっているのかなと思う。このようなクロスオーバーということは、やはりこれは全体に共通するというか各問題の相関関係があるというか、そういうことによるものだろうと思う。クロスオーバーの傾向は3つの問題に共通点があることの認識の深まりによるものではないか。特に対応の観点、どう対応していくかという観点に共通点があることが強く意識されるようになったのではないかと。どなたも割と思いつくところとして、共通の対応の観点として、リテラシー、これには心理学、認知科学を含みますし、AI それか

らプラットフォーム、アテンション・エコノミー、そういうところが考えられるのではないか。例えば、リテラシー、AI、プラットフォーム、これについてはその在り方が3つの問題の状況を直接左右すると言える。例えばリテラシーは高ければ権利侵害情報の被害というのは軽減されますし、低ければ権利侵害情報が増える。シンプルな話だと思いますし、AIに関しても山口先生、越前先生、江間先生から非常に分かりやすい御発表をいただきましたけれども、攻撃側でも使われるし守る側でも使われる、検知側でも使われるということ。それぞれの問題を深刻化させるほうにも、軽減させるほうにも働くということ。プラットフォームについては言うまでもないのではないか。ちゃんとやればよくなるし、やらなければ悪くなるということ。アテンション・エコノミーは、言ってみれば3つの問題の共通の元凶と言うことはできるんじゃないか。それ自体の解消は困難だが、やはりここに直接メスを入れていただくというのは非常に重要なこと。私の現状の認識ではそれ自体の解消は難しいということにとどまっていますけども、ただ一つ言えるのはこれを踏まえてこの3つの問題に対応するということは必須ではないか。【森構成員(5)】

- 国際動向、ジャーナリズム、人権、サイバーセキュリティ、これは重要性においてはこの上の4行（リテラシー、AI、プラットフォーム、アテンション・エコノミー）に決して劣るものではない。国際的動向なんていうのは明らかでございまして、当然これを参考にして対応していかなければいけない。ジャーナリズムについても奥村先生から御発表もありましたし、やはりジャーナリズムという固有の空間において情報流通の健全性をどう考えてきて、どう対応してきたかということが明らかにされている。これをサイバー空間一般の情報流通の健全性の問題に当ててみるということは非常に重要なこと。【森構成員(5)】
- ここで検討した上で、最終的にスコープが違うということであれば、ほかの検討会に報告するなどして対策を連携していくことになるのだろう。打ち手の整理も、技術的な対策も含めて各国において実施している官民の打ち手を課題ごとに整理しておいて、我が国でどういう選択肢があり得るかを準備運動として、まず、可能性を検証していくための整理も必要ではないか。【落合構成員(5)】

### ③ 各ステークホルダーが果たすべき役割・責務

#### 1. 総論：各ステークホルダーに期待される役割・責務の在り方

- プロファイリングであっても、偽・誤情報であっても、ダークパターンであっても、いろいろなアプローチから情報がゆがんだ状態で流通することによって人の意思決定が操作されてしまう。これに対する対策をどうするか。これらの課題はプライバシー侵害の側面もあるが、社会的な利益や国の利益が侵害されるリスクがある。いろいろな権利や利益が複層的に関わり合うのがゆがんだ情報の流通の問題。こうした状況を踏まえ、政府として新しい技術がもたらす課題に対して制度上の手当が十分であるのかを検証する必要があり、事業者にとってはバイアスのかからない情報をいかに担保するのかを検討しなければいけない。利用者においては、社会が複雑化している中で新しいサービスが出て

きて、それが自己にいかなるリスクをもたらすかを自覚する機会を持つ必要がある。【石井構成員(1)】

- 広くデジタル空間の情報流通の全体像を見てみたい。これまでは情報単位で、何々情報、そして何々情報をもたらしている弊害、リスクに対して何々対策みたいな形で、個別の対策が取られてきた部分があるが、今回、デジタル空間における情報流通ということで切り出したというのは、その全体像、取りわけアクターに着目して見ることができるといいんじゃないか。【穴戸座長(2)】
- 今回の検討において非常に広い範囲で議論をしていく中で、どういう打ち手を組み合わせていくとよいのかという話は、別に AI を規制すればいいとか、メディアに発信をさせればいいとか、そういう単純な話ではない。情報の流れの全体像において、人の認知バイアスや、AI においても誤った処理をするということもあるし、こういった点にどういう要素があるのかをマッピングしていく中で、実際にどういうプレイヤーがいて、何らか協力や期待を持って対応してもらうことの可能性があるのかもはっきりマッピングをまずしていくことが重要。例えば、プラットフォームなどが何かを行うと、こういう効果がある可能性があるということについて、まずはそこをある程度できる限り一覧性があるような形でまとめていくことについて、意見だけではなく、研究結果、エビデンスというふうに言っていいたいようなものも多く示していただいていたところ、そういうものを見ながら、有効であるからといって最終的にその手を打てるかどうかは、権利制限や何かに関する倫理観の調整や、もしくは過度な負担になるかどうかもあるもあるので、打ち手については、まずは全体像を見えるようにした上で考えていくことが重要。まずは全体像を示していきつつ、マップとして足りない部分を埋めていく作業を前半で行っていきけるといい。【落合構成員(2)】
- アカウンタビリティ、説明責任や、あるいは日本語でもう少しきちんと言うならば、説明をする責任ではなくて誰が何か問題があったときに責任をとるのかということを一応考えなければいけないということで、その点に関してはそれぞれの関係者が果たすべき役割や責任というのも、サプライチェーンの中で考えいくことが重要。個人や法人では全てが責任を吸収できない場合は、公的機関が事件や事故の原因究明や被害者救済の仕組みなどをつくっていくということも重要になってくるし、ある程度のリテラシー向上、使う側のリテラシー向上でしたり、セーフティーネットをつくっていくというようなことも非常に重要になってくる。【江間構成員(3) ※一部再掲】
- 次回以降のこの検討会における短期的対応について御提案をしたい。まずはこの3つの問題、権利侵害情報、偽情報、利用者データの保護ということを検討してきたプラ研の第3次取りまとめから見てみたいと思うが、偽情報への対策について「第2章 今後の方向性(結論)」というのがありまして、プラットフォーム事業者に対するモニタリングですけれども、「投稿の削除やアカウントの停止等の措置の運用の改善につながることを期待される。したがって、こうしたモニタリングの取組については、継続的に実施していくことが適当である」



としている。そして「2023年11月より『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』（座長：宍戸先生）を開催し、検討を継続している。今後、インターネット上の偽情報の生成・拡散やプラットフォーム利用者の情報に対する選択的接触の問題については、以上述べた観点を踏まえ、当該検討会において議論を深化させていくことが期待される」というふうにされている。このように、プラ研の第3次取りまとめ案は偽情報のモデレーションに関するモニタリングを継続的にまずやるべきであると言っていて、かつそれを健全性検討会、ここにおいて議論を深化させていくことが期待されるというふうに提言している。健全性検討会ではこれまで偽情報に限らず、権利侵害情報も利用者データも、デジタル空間における情報流通の健全性をめぐる各問題について検討してきたわけですし、具体的にはAI、リテラシー、認知科学・心理学、ジャーナリズム、国際的動向等の様々な観点も含めて検討してきた。このようにしてまいりましたので、来年以降、本検討会において情報流通の健全性に関する各問題に関連するオブザーバー団体や対応の観点に関わるプラットフォーム事業者、生成AI関連事業者からもヒアリングを実施すべきではないかということをご提案させていただきたい。【森構成員(5)】

- ヒアリングを続けていくという点、私のほうも先ほど発表させていただいた中で、プラットフォーム等のコンテンツモデレーション等については課題があるということであったりですとか、もうちょっと広くアテンション・エコノミー全般に関する情報収集という意味も含めて調査の内容自体は改めて設計していく必要があるのかなど。ヒアリング内容自体はしていく必要があるのかなどは思いますが実施していくのが必要なことなのではないかと感じた。【落合構成員(5)】

## 2. 政府の役割・責務

- アメリカでも今問題になっているが、政府が権力を持って表現空間に介入してくるというのがどこまで許容されるべきか。政府と事業者の間の透明性とかアカウンタビリティの向上も同時に確保されていくべき。【水谷構成員(1)】
- AIに関して強い法規制を入れたほうがいいんじゃないかということ聞かれることもあるが、最小限の規制でネガティブポイントを抑えるということが一番重要と感じている。強い法規制というのは、適切な活用とか技術革新を阻害するだけでなく、こういう便利なサービスは結局利用を止められないので、問題がより見づらい場所で起こるだけなんじゃないかという懸念がある。また、偽・誤情報についても、アンケートを取ると、74%の人が偽・誤情報には法規制が必要だと考えているという結果が私の調査で出ているが、一方で、表現の自由という観点から非常に危険性もはらむということが言える。例えば、ロシアやマレーシアでは、偽ニュースを対策するといった名目の法律ができて、その結果として、政府が、お前らはフェイクニュースを言っているというふうに言ったジャーナリストがもう逮捕されている。あるいは、対立候補が捜査の対象なったりということが起こっている。日本で今その法律を定めて、急にそう

いった問題が発生するかといったら、ない気がするが、ただ、スリッパリースロープという議論がある。最初は適切に運用されていても、どんどんどんどん拡大解釈されていく中で、やがて許容できないものを引き起こす。例えば、何が誤って何が正しいということは、人によって言っていることは結構違う。そういう線引が曖昧なものについて強い法規制を入れてしまうと、それがだんだん拡大解釈されていって、例えば数十年後に強い政権が誕生したときに、それを基に言論封殺するということも考えられる。そういう懸念があるので、法律を作るというのは慎重であるべき。【山口構成員(1)(2)】

- 現行法で犯罪に当たる行為は常にある。AI を使って名誉毀損的な動画を上げるとか、そういったものに対しては厳正に法的対処をしていくのが大事。【山口構成員(2)】
- 表現の自由とかいろいろ絡んでくるので、偽・誤情報を直接規制するとか AI を規制するとかはないと思う一方、みんなが分かる状態にするのはとても大事。例えば AI が作ったものであるということの開示の義務化。その辺りを厳しく法律で決めるのか、あるいはプラットフォーム事業者とコミュニケーションしながら、そういう規約はどうですかみたいな話になるのか、促していくのかというのがいろいろあると思う。【山口構成員(2)】
- アジャイル・ガバナンスの実践という意味では、もちろん根拠法の整備が必要になるような場合はあると思うが、詳細な対応については柔軟に調整ができるように設計をしていくことが重要。官民共同規制という言葉を使うだけでは、作ったものの、あまりうまくワークしていないと言われるような場合もあるので、どうすると皆さんにリソースを割いていただけるようになるのかが重要。これは国もそうですし、国も十分にそれで定員を確保できないという場合もあるでしょうし、民間側も十分に協力してくれないという場合もある。これをどう考えていくのか。【落合構成員(1)】
- 個人とか法人とかでは全てが責任を吸収できないというような場合は、公的機関が事件や事故の原因究明とか被害者救済の仕組みなどがある種つくっていくことも重要になってくるだろう。【江間構成員(3) ※再掲】
- アメリカのデジタル空間の規制は、基本的には政府は介入しないし、介入しにくいといったような状況。特にソーシャルメディアに関しては、アメリカには憲法修正1条、これは表現の自由を定めた条文だが、この保護が日本なんかよりもはるかに強力な保護を私的アクターに与えるため、端的に言えば、例えば偽情報の発信者に制裁を加える法規制なんていうのはアメリカでは即違憲になると思うし、DPF 事業者に対する法規制を行うことも原則として非常に困難である。【水谷構成員(4)】
- アメリカでの検閲代理人化の観点で今、議論が沸騰しているものとして、jawboning「口先介入」と呼ばれるものがある。要するに、DPF 上で事業者が合法的な言論、つまり法には反していないけどプラットフォームのポリシーには違反しているというようなものもモデレーションしているが、そういった本来合法的に許されるはずの言論の取締りを行うために、政府関係者が DPF 事業者

対して非公式の圧力をかけるというようなことがアメリカで問題になっている。【水谷構成員(4)】

- 例えば、バイデン政権は COVID-19 対策の一環で偽情報対策をしなければならなかったが、アメリカでは偽情報の規制はかなり困難なわけで、法規制でなかなかできないから、それこそ FBI とか政府、ホワイトハウスの関係者が DPF 事業者と綿密に協力して、プラットフォーム上から例えばワクチンに対する偽情報とかウイルスに対する偽情報をモデレーションするように圧力をかけていたんだと指摘されている。これが一部、表に出てきたことによって、モデレーションされた側の人たちが裁判を起こしていて、DPF 事業者が政府関係者と接触をすること、これを表現の自由の観点から差止めを求めるといった訴訟が現在係争中。これも連邦最高裁に上がっているが、恐らく今年も終わるので来年ぐらいに判断が出るかなという感じ。ただ難しいのは、政府によって事業者にとって適切な説得をする行為は、これ自体は本来許されてしかるべき行為。憲法上もこれは問題ないはずとアメリカも考えているが、そこに強制が入るとまずい。例えば有害な言論、違法じゃない言論を DPF 事業者に消させるために、事業者の担当者に政府機関の警察が、これを消さないんだったらお前も逮捕されるぞ、とか言って脅迫をする場合、こういったものはまずいだろうというのは何となく分かるが、ではどこで線引きができるのかというのは非常に難しいということで、アメリカでも議論になっている。そのため、むしろここにも透明性確保のための規制を入れろという提案が在野からあがってきていて、ソーシャルメディア検閲透明化法というものを作ったらどうかというのが在野の研究者からあがっているところ。これは単なる在野の提案のため、今のところ連邦議会にはあがっていないと思う。【水谷構成員(4)】
- アメリカの議論で参考になるのは、DPF 事業者を検閲の代理人化させてはいけないので、事業者への政府機関のコンテンツ削除要請なんかに関しては、政府機関側の、政府機関サイドの透明性をいかに確保していくかが今後、非常に重要になってくる。【水谷構成員(4)】
- 政府側の透明性を確保する方法としては、もちろん色々方法はあると思うが、一つは法律で、こういう部分についてはきちんと透明性レポートを政府が逆に出すというのを定めるということが一つ。例えば、アメリカで提案があったソーシャルメディア検閲透明化法を参考にすると、政府の職員が表現の自由とか、そうしたことに関係しそうな案件で DPF 事業者と接触した場合は、それをきちんと記録として残せと。それを事後的にチェックできるようにしろ、というようなことを規律で定める。そこまでできるかどうかは別にして今、DPF 事業者側が政府からの要請についての記録を透明性レポートで出してくれている。だから、あれに対応するものを逆に政府側が、例えば白書とかできちんと出すだけでもすごく変わってくるので、個々のコンテンツの削除について、どういうものを行っているのかということをしちゃんと明示していただくための仕組みというのを、何かしら作るというのが重要。【水谷構成員(4)】
- アメリカの議論の拝見をしていて、いろいろな主体がそれぞれ透明性を持って

いくことについて、分け隔てなく、何らかのアクションをするものについては、それぞれ透明性を持っていくというのが一つ、意味がある対応なのではないかという示唆もあると受け止めた。日本の中でもそういう意味では情報発信者とか、情報の媒介者、または、それに対して影響を与え得る政府と、それぞれの主体が透明性を高めていくことも一つ大事なのではないかとも思う。【落合構成員(4)】

- 透明性の確保について DPF 事業者が主軸になっているが、政府側も当然必要だし、そういうコンテンツを生成する側がどういうプロセスで自分たちがコンテンツを作っているのかと、こういうところにも透明性、もっと言うと説明責任、自分たちのコンテンツの生成プロセスの説明責任が、まさに生成系 AI 使い始めると重要になってくるかなと思う。【水谷構成員(4)】
- 先ほどの裁判で、プラットフォームはコンテンツ・モデレーションしてはいけないんだという判決というものだが、何かアメリカは表現の自由を非常に重視する、そしてそれに対する規制を恐れる、いろんな規制を恐れることによって反面教師というか、危険なところにいるのではないかと私も感じている。【森構成員(4)】
- 国、自治体や、場合によっては外国政府、国際機関も発信者としては存在する。【落合構成員(5)】
- 事業者への要請、自治体、国が関係する事業者に、どういう形で情報のやり取りをさせているのか、開示させたりしているのか、といった透明性確保の部分も関わってくる。【落合構成員(5)】
- モニタリング／ヒアリング、特にプラットフォームのモニタリングということをやするわけだが、これが必ずしもこれまでうまくいってなかったということは教訓として持つておくべきではないか。これまでのプラ研における権利侵害情報、偽情報、このモデレーションに関するモニタリングは法的根拠を持たないお願いベースのものであり、プラットフォームの事業者から十分な情報開示が行われたとは言い難いのではないか。特に偽情報についての X の対応について、第 3 次取りまとめ案は、「再三の求めにもかかわらずヒアリングシート及び発表資料が提出されなかった。任意とはいえ、資料が提出されなかったことは遺憾である」と、厳しい言葉で記述されている。委員のほうもフラストレーションがたまっていたようでして、単なる「遺憾である」ではなくて、「極めて遺憾である」にしてくれみたいな、そういうことを言い出す人もいた。利用者データに関するモニタリングもありまして、これは経産省の取引透明化法に係るモニタリングと連携してやることになっておりますので、法的根拠がないわけじゃない。ところが 2023 年の 11 月 3 日に実施されたモニタリングについては、読売新聞の報道でこのモニタリングにおいても十分な情報開示が得られなかったとされているところ。これからモニタリング／ヒアリングをやっていくわけですけども、それについての留意点ということだが、プラットフォームは本来、最も法制化の歴史が古い 2001 年にはプロ責法ができています。ただそのプロ責法は私法上のプラットフォームの責任に関する法律。3 つの大きな問題

とプラットフォーム横軸の交錯点には、投稿型プラットフォームとしての SNS が存在。プラットフォームに関する公法規制として、取引透明化法と取引 DPF 消費者保護法が存在するが、これはいずれもこの 3 つの問題、権利侵害情報、偽情報、利用者データの保護の関係で SNS における一般利用者の保護を図るものではない。そういう事情がありますので、短期的提案として先ほど御提案しましたヒアリングを進めるに当たっては、特に一般利用者の権利保護を直接の目的として、つまりこれはビジネス利用者ではなくて、一般利用者の権利保護を直接の目的とすべき。「ビジネス目的の利用者を保護しなきゃいけないから一般利用者にも配慮する」とかではなく、一般利用者、ビジネスではない利用者の権利保護を直接の目的として、投稿型プラットフォームにおける①コンテンツモデレーションと、②利用者データの保護に関する透明性やアカウントビリティの確保の観点から行うべきではないか。コンテンツモデレーションというのは、権利侵害情報と偽情報のコンテンツモデレーションをプラットフォームにやっていただくということ。【森構成員(5)】

- 一番最後に、第 3 次取りまとめでなく第 2 次取りまとめにありました透明性アカウントビリティ確保の留意点について一言御紹介しておきたい。これは権利侵害情報に関すること。「総務省は、透明性・アカウントビリティを確保する観点から、行動規範の策定及び遵守の求めや法的枠組みの導入等の行政からの一定の関与について、速やかに具体化することが必要である」。第 2 次取りまとめ、偽情報についても以下のとおり述べております。「総務省は、偽情報への対応に関する透明性・アカウントビリティの確保に向けて、行動規範の策定及び遵守の求めや法的枠組みの導入等の行政からの一定の関与を具体的に検討することが必要である」となっておりますので、今後は行動規範の策定とか法的枠組みの導入といったことも検討対象として入ってくるのではないかと。【森構成員(5)】

### 3. プラットフォーム事業者の役割・責務

- 情報流通管理プロセスの透明性・アカウントビリティの向上、信頼性の高い情報の発見可能性の向上、メタ情報の可視化。【水谷構成員(1)】
- 偽情報、誤情報の発信において経済的インセンティブに問題があるが、その対応策として、まずはプラットフォーム側がデマネタイゼーションをする、つまり広告収益を止めるといったハードな方向での対応があるところ、プラットフォームもすでにやっているし、継続していただくということが重要。【水谷構成員(2)】
- ゲートキーパー機能を誰にどういうふうを持たせるのか。ゲートキーパーとしては、電気通信事業法の規制事業者、プラットフォームの役割がある。その他、メディアに対するプロミネンスも放送の分野では放送事業者を主に議論されているが、別に放送事業者だけに限ることではない。欧州 DSA 等を参考にできる部分も踏まえ、これをどう考えていくのか。【落合構成員(1)】
- 発信者側の情報、特に AI を使ったかが一番大きいと思うが、そういう部分が、

なるべく分かりやすく情報開示されている状態になった方が多分判別率が上がる可能性があるかとも思う。プラットフォームの側も、例えば個人が情報発信するに当たっても、できる範囲でスクリーニングして、そういう注記を出していくことを対応として求めていくことが、偽・誤情報の拡散防止という意味では比較的効果がありそう。権利との調整は最終的に必要だが、必要性、有効性という意味では意味がありそう。【落合構成員(2)】

- デジタルプラットフォーム事業者に関しては、透明化法、取引 DPF 消費者保護法の対象範囲を拡大していただく。偽情報・誤情報の削除を行う。取引 DPF 提供事業者による消費者への啓発を行っていただきたい。【増田構成員(1)】
- 偽情報の流通に利用されるプラットフォーム事業者は、コンテンツ・モデレーション等の偽情報対策を実施することについて、社会からの強い期待を受けている。【森構成員(1)】
- プラットフォーム事業者には求められることがいっぱいある。透明性の確保、特に日本ローカルの透明性。また、既に有効と考えられている施策を積極的に導入したり、ファクトチェック効果を効率的に配信したり、生成 AI を見破る技術の開発・実装とか、実施できることは様々にあるし、例えば AI が作った画像や映像を投稿するときそれを明示する必要がある、そういった規約を作るというのもありと感じている。【山口構成員(2)】
- メディア企業やファクトチェック組織などとの連携の継続・発展も大事。【山口構成員(2)】
- 広告収入が偽情報を流しているサイトとかに流れないように取組、これは EU もそういう要請をしているが、そういった取組を進めていくことが大事。例えば、Google は結構頑張っ対策しても、現実的には 60%以上残っているみたいなことが指摘されている中で、多分技術的な対抗は限界がある。そうすると、広告収入には流れないようにするという対策だけでは多分駄目。【山口構成員(2)】
- デジタルプラットフォーム事業者が私たちユーザーの情報の発信や閲覧や需要に対して、環境のデザインをもって非常に大きく作用しているのを憲法やメディア法の観点からどのように考えていくべきか、こうした作用は、アーキテクチャとアルゴリズム、コンテンツ・モデレーションの大きく分けて3つの諸要素から構成されている。こうした各種機能を有している DPF 事業者は、ニューガバナー、新たな統治者と呼ばれたり、あるいは、これは山本龍彦先生が指摘をしている国家をリヴァイアサンと例えた場合の対抗存在としてのビヒモスというような存在として語られる、つまり単なる大企業・社会的権力を超えるものとして語られることが多くなっているが、これは DPF 事業者が国家と同一視される存在であるということを示しているわけでは必ずしもなくて、国家に類似する機能を持ち始めて匹敵する存在になりつつあるんだけど、ただやはり、国家とは違う側面も持つ、そういう特殊なアクターとして位置づけられている。【水谷構成員(4)】
- アメリカ憲法の特徴は、ヨーロッパ圏と違って憲法というのはあくまで国家権力を縛るものであり、そして DPF 事業者は、あくまで国家ではなく私的アクタ

一にすぎない。私的アクターに憲法の統制を効かせる場合は、アメリカではステイトアクションの法理というのがあるが、その非常に厳しい条件を満たさないと憲法の縛りというのは基本的に私的アクターに及ばないという側面がある。また、表現の自由が世界的に見ても強力に保護されている側面がアメリカにはあり、一部の論者がこれを修正1条の「武器化」だというような批判をしたりしているが、連邦最高裁の判例や多くの修正一条理論では、今でもそうした強力な保護を後押ししているところがある。【水谷構成員(4)】

- 通信品位法 230 条という非常に強力な法律の存在もあり、アメリカの DPF 事業者の成長を支えた一側面として、この CDA、通信品位法 230 条というのがよく注目される。これは、ある論者に言わせると憲法の修正1条のステロイド強化版だと評されたりするが、日本のプロバイダ責任制限法3条に相当するような条文が書いてある。ただ条文では、DPF 事業者も含め、双方向型コンピューターサービスのプロバイダまたはユーザーに対して、発信者や権利侵害情報の被害者からの責任追及に対して非常に広範な免責規定を設けている。もともと CDA230 条ができる以前、インターネット草創期のオンライン企業に対してどの程度、情報媒介者として法的責任を負わすべきかというのが下級審の判例で議論されたが、モデレーションをしっかりとやった人たちは責任を負わされ、逆にモデレーションをやらないという人たちは免責される可能性があるという、いびつな状況が生じそうになり、その課題を解消するために CDA230 条は、96 年に制定された。この頃は今の DPF 事業者のような巨大な存在はなく、むしろネット企業は新興産業だったので、ここに強力な法的責任を負わせるような状況が生じると企業の発展可能性を阻害してしまうことが懸念されたため、逆に広範な免責を与えて、むしろ事業者にコンテンツの管理は自主的にやらせる、そういう自主的なモデレーションを促進するようなことを目的にこの条文というのは作られたと言われている。【水谷構成員(4)】
- 特にトランプ政権前後からアメリカでは DPF 事業者に対する規制の議論があがっている。しかしそれらの議論の中には、EU とは方向性の違うものが散見され、特にトランプさんがツイッターから Ban されたように、保守派は、DPF 事業者が保守派を狙い撃ちしてモデレーションしていると勘ぐっているように、保守派の人たちが反発をして、むしろ DPF 事業者はコンテンツ・モデレーションをすべきでない、コンテンツに手を触れるなどというような規制、いわゆるコモンキャリア的な規制をしようとしている州などが見られる。その典型例がフロリダとテキサスの州法で、これら2つがどちらも保守派の知事のもとで出来たが、基本的にソーシャルメディアプラットフォームがユーザーに対してアカウント凍結とか、コンテンツの削除とか、あるいはシャドーバンといったことを行うのを制限しようという法律を作っている。ただ、これは DPF 事業者がコンテンツをモデレーションする自由も表現の自由としてアメリカでは考えられている節があるため、こうした州法はそれを規制することになるとして、差止めを巡って訴訟で争っている最中。そろそろ連邦最高裁で判断が出ると思うが、そういう状況がある。【水谷構成員(4)】

- プラットフォームのコンテンツ・モデレーションの規制に関する裁判も含めて、この手の議論というのは本当につい最近で、ケンブリッジ・アナリティカの事件よりも後。むしろケンブリッジ・アナリティカ事件以降、黄金時代と言ったが、まさに DPF 事業者はフェイクニュースを含めてコンテンツ・モデレーションをかなり強化してきた。それに対する反発として出てきているのではないかというのが私の見立て。つまり、保守派がヘイトスピーチとかフェイク、偽情報で消されることが多く、トランプさんのツイッター・アカウントの凍結、プラットフォームがかなり保守派の間では緊張感を高めることになったというのは報道レベルとかでも見るので、その流れではないか。【水谷構成員(4)】
- デジタルサービス法は、もともと、日本のプロ責的な規定を中心に置いていた電子商取引指令をもとに、違法・有害情報に対するプラットフォームの受動的な責任と、そしてそれからその社会的影響力等を鑑みた、よりポジティブなレスポンスビリティ、そして透明性等の在り方を全面的にアップデートするもの。非常に様々規律あるわけだが、わけてもディスインフォメーション等の舞台になってしまいやすい、この特に巨大なデジタルプラットフォーム及びオンライン検索エンジン、これを VLOP と VLOSE と略すが、第3章でそれらに対して偽・誤情報対策を含む、かなり他の事業者たちよりも追加的な義務を課するような作りになっている。既に 17VLOP と 2VLOSE が指定されており、既に特に巨大な部分に関わるものは8月から適用開始されていて、来年の2月には全面適用されるようなスケジュールで、これ自体が非常に巨大な法律である。【生貝構成員(4)】
- デジタルサービス法は、まず、媒介サービスを導管、キャッシング、ホスティングと3つに電子商取引指令と同じように分けた上で、その中でコンテンツを預かるサービスに追加的な義務が課されて、その中で特にコンテンツを配布するようなサービスは OP という形でまた義務が増えてきて、一番大きいものは4,500万人以上のユーザー数、EU域内の10%以上の人口といったところで、御紹介するシステミックリスクの評価と軽減といったようなことの義務等々がかかってくる。他にも取引 OP に関わる特別の規律なんかもあったりするが、VLOSE は、ほぼ VLOP と同様の義務がかかる。【生貝構成員(4)】
- AI 法案とデジタルサービス法の関連性について、まず AI 法の中でディープフェイクの場合はラベルづけをしなないといけないような義務が含まれる予定であるが、正式に成立して施行されるまで大分時間がかかる。だが、デジタルサービス法ができる前から、適用開始される前から、このデジタルサービス法の共同規制における行動規範になることを前提にして偽情報行動規範というものを作ってきた。去年6月の段階で AI が生成したり、操作、マニピレートしたコンテンツに対しては、これは注意喚起や、そういうものに対抗するための措置というのを様々とらないといけないということディープフェイクや偽情報は結局、この文脈だと生成する AI があって流通するプラットフォームがあり、両方から考えていけないといけない。まず、法律がある後者からやったということだと思う。【生貝構成員(4)】



- DSA はかなり強い法律だと思うが、経済学的に考えると、市場規模が大きいからできるというようなこともあるのかなと思う。つまり、5 億人市場だからプラットフォーム事業者もそれに従うわけで、一方、例えば日本というのは1 億何千万人だし、また、さらに人口が少ない、あるいは経済規模の小さい発展途上国がいっぱいある中で、プラットフォーム事業者に対してこういう法律を施行していくのは、そもそも難しい国のほうが多いんじゃないかなと思っている。  
【山口構成員(4)】
- 全体的に特に透明性、救済について、10 億人、20 億人使うようなサービスについてリスク評価して緩和せよというのは、社会的なレスポンスビリティとして、それがデジタルプラットフォーマーであっても General purpose AI であっても、あるべきなんだろうなとは思っているので、我が国でもバランスのとれた形で前向きに考えていく価値があるのかなとは感じている。また他方で比較的本当にホリゾンタルな部分を御紹介したが、バーチカルには例えばテロの情報に関して、当局の要請後 1 時間以内でちゃんと対応しなきゃいけないとか、また CSAM 系ですと非常に難しい問題であるだけに、強い措置というのも EU レベル及び各国レベルで行われている。まさにバーチカルにどうしていくかといったような問題は一つ一つ、その問題の深刻さとの兼ね合いでしっかり考えていかなければいけない部分が多いなと思う。【生貝構成員(4)】
- もちろん伝統メディアは伝統メディアを通じての発信というのを続けているが、一方で、電気通信事業者の通信を使ったサービス、さらに言うならば、プラットフォーム事業者、SNS、ブログ、掲示板検索、動画共有、ニュースポータルなど様々あるかと思うが、プラットフォームと称されるものや、デジタルサービスを通じて、情報拡散していく構造があるように思う。【落合構成員(5)】
- プラットフォームサービス事業者と受信者との関係では、コンテンツサービス、アテンションや、個人データが戻っている部分があるのではないかな。もちろんここは利用料のようなものを、特に有料コンテンツも徐々に多くなってきている部分もあるので、もちろん直接的な対価が支払われている場合もあるかと思うが、テンション、個人データということで返されている部分もあるだろうということ。【落合構成員(5)】

#### 4. 生成 AI ・メタバース関連事業者の役割・責務

- 生成 AI の特に広範な提供者について、何らかの協力を求めていくことがどうだろうか。【落合構成員(1)】
- 生成 AI、AI 利用者の事業者については、基本的理念の遵守をして頂くこと、メタバース、仮想空間関係事業者においても基本的な理念、ルール の 制 定 を し て 頂 き、それを消費者にしっかりと広報していただく必要。【増田構成員(1)】
- AI ガバナンスに関する原則を方法論に落とし込んでいく際、ガイドラインのようなノンバイディングなもの（日本の新事業者ガイドライン、米国大統領令など）や、国による法令というバイディングなもの（欧州 AI Act、欧州評議会 AI 条約など）が入り交じっているが、最近では Voluntary commitments とい

うところで、様々な組織に対して自主的にガイドラインとかをつくって、安全性・公平性といったところに対応していこうということを促すような層が新たに出てきた。自主ガイドラインを自らつくっていったり、業界で考えていくということが非常に大事になってきている。【江間構成員(3) ※再掲】

- いろいろなところで報じられているとおり、AI 法案、最初ハイリスクと4分類といったような、2021年の最初の段階では生成AIを意識しない形で法案を作っていたが、去年末から急遽、生成AIが重要になってきたということで、それとはまた別に生成AIなどの、彼らは最終的に汎用AI、general purpose AI、GPAIと呼ぶようにしたわけだが、その内容が大きく含まれることになった。テキスト自体はまだ公開されていないが、トリローグ合意の直後に欧州委員会が詳しいQ&Aを出していて、まず、全てのそういった汎用AIモデルに課せられるような透明性の義務というものに加えて、わけても10の25乗フロップ以上の計算能力を持って計算されるような非常に大きなモデルは、取りあえずはシステムリスクがある。そのシステムリスクを持つモデルのプロバイダはリスクの評価と軽減、まさにシステムリスクの評価と軽減、その他、最近AI規制について言われるようなことをやらないといけないということで、デジタルサービス法のVLOP、VLOSEに課せられている義務というのをスライドさせてきた。まさにAI議論のプラットフォーム化というところの一つの大きな象徴かなと思う。【生貝構成員(4)】
- AIの開発、提供、利用者について言いますと、発信、伝送の様々なプレーヤーが実際には利用しているような状況になっているであろうと認識。【落合構成員(5)】

## 5. 通信・放送事業者の役割・責務

- 信頼性の高い情報の発見可能性を向上させるというのは流通を担っている事業者にはしかできないが、一方で、信頼性の高い情報を生成する人たちがいなければ、それも絵に描いた餅になる。特に、例えば、報道が典型例だと思うが、職業倫理などによって内部的にも、あるいは外部的にもガバナンスを利かせて情報の信頼性を確保してきた事業者の持続可能性をどう考えるか、あるいはこの生成プロセス自体のガバナンス、応答性、持続可能性もアテンション・エコノミーが加速する中では考えなければいけないのではないかという点が今後重要。【水谷構成員(1)】
- 信頼できる情報源の確保という観点から、メディアの貢献については、放送法の関係でも情報空間の健全性確保ということで議論してきた部分があった。しかし、放送法については、一部の論点を除いて、あくまで放送事業者の話だけになってしまうが、これをもう少し広くメディア全般を捉えて、多元性を踏まえて考えることが重要。【落合構成員(1)】
- 民間放送事業者などのメディアをどういう形で、よりインターネットでの発信を増やしていただくのかが、また、それを目立ちやすいような形にしていけるのかがあり、プロミネンスという名前もあるが、改めて重要かと思うし、フ

アクトチェックの機能も、もちろん既存のメディアだけに限られるものではないが、既存のメディアの方々にもやはりそういった役割を担っていただくことは、1つ期待していくべき。【落合構成員(2)】

- 災害報道について、今までテレビが競っていたところ、スマホに自分の地域の情報が入ってくるようになったという、伝達経路が一変している中で、メディアの役割というものは再定義されて良い。【奥村構成員(1)】
- 日本はメディアがメディアであるだけで信頼されていた時期がすごく長く続いていた、幸せな国だったのかもしれない。ただ、今はもうそうではないので、信頼をつくり直す社会的な営みに、ぜひとも首を突っ込んでいてもらいたい。そういう意味でオブザーバーに民放連が加わっていらっしやらないことは非常に残念。【奥村構成員(3) ※一部再掲】
- 公共の電波を長年にわたって寡占することで蓄積されてきた映像で伝えるというノウハウを情報発信が民主化された時代に広く市民に還元していくと情報空間の健全性や豊かさに貢献できるのではないか。それがインフォメーションヘルスの担い手としての放送事業者の責務の1つ。【脇浜構成員(1)】
- 地上民間放送事業者に期待したい。今日の時点でオブザーバーとして入っていないようなことをお見受けしたが、民放局にはより積極的な公共性へのコミットを求めたいし、特に、ローカル局はそこに存在意義があるのではないか。実際に、地域の課題解決に放送局のノウハウを生かす事業(効果的なオーディオ・ビジュアル表現、一貫性と継続性のある情報発信として、例えば、KBS九州朝日放送のGlocal KやLivePark)も散見されるようになったし、より明確なソーシャルビジネス的なアプローチとして、千葉テレビがチバテレミライチャンネルを地デジのサブチャンネルで設定(地デジのマルチチャンネル活用)等している。【脇浜構成員(1)】
- 偽情報・誤情報についての通知を受けた場合に確認し削除等の対応をすることの検討をしっかりと頂く。【増田構成員(1)】
- 偽・誤情報が増え、情報環境が悪化しているからこそ、信頼できる情報の存在は非常に重要。例えばメディアの存在、マスメディア、これが非常に重要になってくる。【山口構成員(2)】
- 発信側について見ると、もともとは伝統メディア、デジタル空間が広がる前は伝統メディア、公共放送、民間放送、新聞等のメディアのいわゆる放送波ですとか、紙媒体等の伝統メディアの通信が大きかったところではあったと思う。一方で、クリエイター、フリー記者、ネットメディア、個人、こういった主体による情報発信も、さらに強まってきている。【落合構成員(5)】
- 伝統メディアについては、受信料、購読料を直接的に利用者から受領する部分があってコンテンツが提供されているということがある。もちろん伝統メディアでも、近時だと、例えばユーチューブやニュースポータルなどに記事を出すこともあるので、もちろんこれだけでもないとは思いますが、次第に業務の幅を広げてきている部分がある。【落合構成員(5)】
- 電気通信事業者自体は、通信サービスそのものを利用してもらうということで、

利用者に対して、通信料でお金のやり取りをしている部分がある。【落合構成員(5)】

- こういう情報流通全体を支える、さらに下のほうにある基盤というものが多分、インターネットのアーキテクチャの中でいろいろある。それは例えば DNS だったり、CDN だったり、あるいは場合によってはカードなんかを含めた決済システムといったような、様々なコントロールポイントと言われるものが、実はかなり様々低いレイヤーのところにもありますよねといったようなものを、果たしてこの射程のかなり端のほうにはなると思うんですけど、どのように考えていくかということも重要。【生貝構成員(5)】

## 6. 広告主・広告関連事業者の役割・責務

- DPF 事業者に対する規律は、基本的にアメリカなどは自由放任でやってきたが、では DPF 事業者は野放しになっていて無秩序なのかということ全然そんなことはなくて、まず DPF 事業者は国家と違って企業なので、ユーザーや、特にソーシャルメディアに関しては広告主の存在が非常に大きいと考えられている。ユーザーや広告主からそっぽ向かれるような状況は彼らにとって好ましくないの、規制がされていなくても彼らは自主的にコンテンツ・モデレーションを行う。このように DPF 事業者は、背景に非常に複雑なステークホルダーを抱えており、それによって DPF 事業者のガバナンスが決められている側面がある。【水谷構成員(4)】
- 伝統メディアのほう見てみると、広告料、広告枠でのやり取りがもともと重要な部分ではあったと思うが、さらにこの部分については、プラットフォームサービスが入ってくることによって、プラットフォームとの関係でも、広告主とプラットフォームとの関係性がどうなっているかという部分がさらに論点になってくる部分がある。【落合構成員(5)】
- やはり特にアクターとしての広告というところが重要なのだなということを感じているところ。それはやはりこれまで様々この分野で議論になってきた政治的イシューに関する広告といったようなもの、あるいはもっとステルスマーケティングといったようなところを含めて、非常に情報流通とまさにアテンションエコノミーというところの大きな接点になってきている。そうしたときに恐らく場合によっては、発信というところにも広告の位置づけというのは、あるいは広告主のような位置づけというのはどこかに探してもよいのかなということを感じたところ。そしてこのことというのは、またマクロにはアテンションをエコノミーに変える手段としてのそういったビジネスモデルの在り方をどうするべきなのかということにも深く関わってくるので、よく考えてまいたい。【生貝構成員(5)】

## 7. 教育・研究機関の役割・責務

- 大学にも果たせる役割がたくさんあるのではないかと。欧米にあるようなジャーナリズムスクール、メディアスクールなど、国内では体系的・実践的にそうい

った情報発信を学べる場が見られない。こういったオーディオビジュアル表現を有効に使う、それを教えられる人材、教えられる場というところも重要。【脇浜構成員(1)】

- 欧米ではジャーナリズムスクールという大学院の教育システムがあって、それを経て一定のリテラシーと能力を得たものがメディアで仕事をするようになってはいるわけだが、日本のメインストリームのメディアは長らくインハウス・トレーニングをしてきた。かえって、メディア学があまりハッピーな発展状況じゃなかったというのもあるがそちらを専攻した人をあまりとらないというような傾向もあった。【奥村構成員(3) ※再掲】
- 政治に対する知識のディバイドがあるとすると、公教育、すなわち親の収入とか身分とかそういう属性に関係なく、みんなが公平に受けられる教育の現場でこういう知識に触れてもらうというのはますます重要なのではないか。【水谷構成員(2)】
- 政治的知識の分析をすると、特に制度面とか手続面の知識については、学校教育の効果が大抵有意に出てくる。政治家の名前や、今問題になっている事柄などはメディア接触がききやすいが、政治のシステムに関する知識は、かなり学校教育の効果が出ているので、学校現場で虚偽情報、誤情報のほか、市民教育といったものを進めていくことは非常に有意義。【安野構成員(2)】
- 心理学はインディビジュアルレベルの中での研究が主になりがちだが、誤情報という社会問題を捉えていく上では、この2つのレベルがどうやって連携していくのかというのが今後の課題になっていく。【田中構成員(4)】

## 8. 専門家の役割・責務

- 特に医療従事者などの専門家がどんな役割を果たしたのか、果たせるのかという点に関心を持って研究されたものとして、ツイッターのデータを集め、そこからユーザーのプロファイル情報に基づき、医療従事者関連のアカウントを抽出して調査した。医療従事者関連のリツイートのカラスタのネットワークにおいて、大きなノードは医療従事者関連のノードで大きさはフォロワー数になっているところ、この医療従事者関連のユーザーが他のユーザーと比べてどうだったのかを見ると、医療従事者アカウントはそれ以外のアカウントに比べ、返信をより受け取っており、報道でも少し目にしたが、誹謗中傷なんか受け取るという事例なんかも報告されている。それから、より多いフォロワーを持っているという傾向があるということが分かった。一方、医療従事者関連の投稿は、それ以外のアカウントに比べ、投稿がシェアされにくい傾向があるということも分かっている。投稿内容の情報源に関して、医療従事者は、海外メディアとか医療系メディア、それから政府のサイト、論文誌サイト等々、一般の方とはまた違う情報源を使って発信しているということも分かっている。このように、ツイッター、現在のXにおける医療従事者関連のアカウントの非常にユニークな立ち位置、役割というのが少し見えてきており、この辺りもう少し今後詳しく見ていくことが大切。【澁谷構成員(2)】

- 専門家等によるカウンターナラティブの刷新戦略と、あるいはネットワークの形成といったところの検討や効果検証というのにも必要。【澁谷構成員(2)】
- 例えばWHOのような専門機関、ファクトチェック機関、それに関連するような団体なども情報を様々な方式で発信をし続けている状況がある。【落合構成員(5)】

### 9. 利用者の役割・責務

- プラットフォーム利用者、メディアの視聴者のICTリテラシー向上も非常に重要。【越前構成員(1)】
- 利用者については、やはり教育だが、若年者、親の教育や、情報が届きにくい高齢者への啓発などが必要。【増田構成員(1)】

以上