

# CYBER GRID

サイバー・グリッド・ジャーナル

## JOURNAL VOL. 17

デジタル社会の新たな脅威  
「偽情報・誤情報・フェイク」  
～認知戦、生成AI、リテラシーを考える～

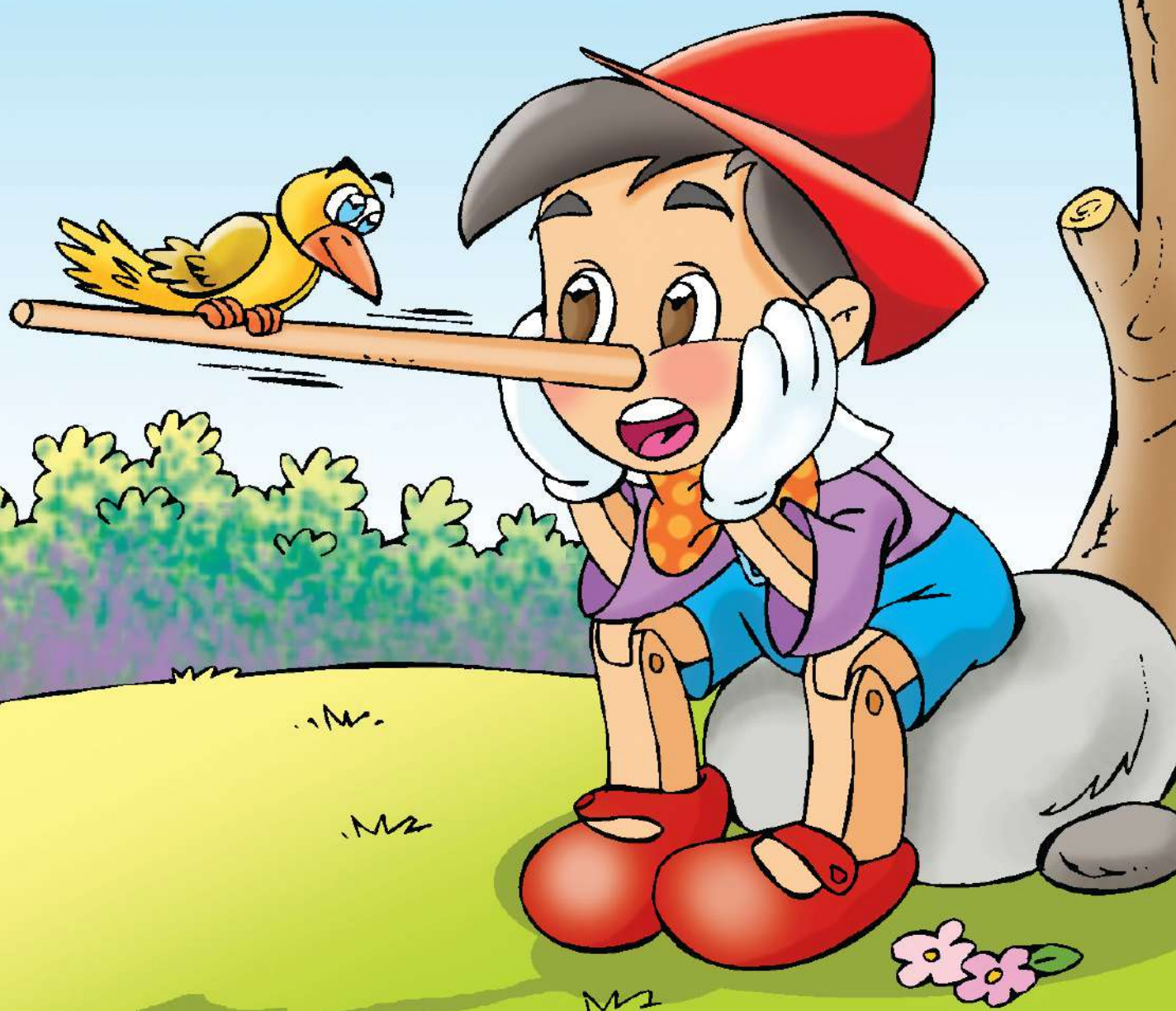


TABLE OF CONTENTS

デジタル社会の新たな脅威  
「偽情報・誤情報・フェイク」  
～認知戦、生成AI、リテラシーを考える～

- 3 特集紹介  
国内外における偽・誤情報問題の概要  
鈴木 悠
- 6 特集1  
戦略的なウソと企業の対応  
上田 篤盛
- 9 特集2  
主体的に情報を活用するリテラシーとは  
尾方 佑三子
- 12 特集3  
生成AIがもたらす偽情報の進化と対策  
庄司 勝哉
- 17 団体運営推進室のご紹介
- 18 巻末あとがき

# 国内外における 偽・誤情報問題の概要

サイバー・グリッド・ジャパン / 国立研究開発法人情報通信研究機構(出向) 鈴木 悠

2022年12月に改定された安全保障関連3文書に、初めて偽情報等への対策が明記されました。日本では、災害発生時にソーシャルメディアでうそや生成AIを用いた偽画像が拡散され、災害対応の妨げになることが問題視されてきました。改めて偽・誤情報が安全保障上の脅威とされ、対策が急務となっているのはなぜでしょうか？その背景を説明し、現在の偽・誤情報対策に関する取り組み状況について解説します。



## 誰もが情報を発信できる 時代に

これまで情報発信の多くはマスメディアが担っており、個人は情報を受け取る消費者にすぎませんでした。しかし今では、スマートフォンやソーシャルメディアの普及により、消費者だった一個人が情報の発信者・仲介者にもなりました。インターネット上では、マスメディアによる公共性の高いマクロ情報から、個人の意見や感情表現といったミクロ情報まで、多種多様な情報が世界規模で伝播していきます。誰もが情報を発信・閲覧・共有することが可能であり、それらの情報が人の信念と行動に影響を与えることがあります。

## 民主主義の 根底を揺るがす事例

ソーシャルメディアの情報が人の信念と行動に影響を与えた事例として、2010年代初頭の「アラブの春」が

あります。アラブの春では、当時の政権に対する抗議活動への参加の呼びかけがソーシャルメディアで行われ、大規模な民主化運動へと発展しました<sup>①</sup>。このソーシャルメディアが持つ影響力に注目したロシアでは、これらの出来事こそが21世紀の典型的な戦争ではないかという主張<sup>②</sup>がなされました。2013年には、のちの米大統領選挙への介入疑惑で米司法省に起訴された露民間企業インターネット・リサーチ・エージェンシー社<sup>③</sup>が設立され、ソーシャルメディアでの影響工作(Cyber Influence Operations)が開始されました<sup>④</sup>。2014年クリミア併合前後におけるロシアの影響工作に対し、欧州理事会は2015年に「戦略的コミュニケーションタスクフォース(East StratCom Task Force)」を設立してロシアの偽情報に対抗する「EUvsDiSiNFO」というWebサイトを開設・運営しています。

2016年6月イギリスの欧州連合離脱是非を問う国民投票では、英ケンブリッジ・アナリティカ社が特定の心理傾向を持つ有権者を標的に政治広告やメッセージを流布していたことが問題となりました<sup>⑤</sup>。ロシア政府からの助成金で研究していたケンブリッジ大学の心理学教授が、

① 総務省『平成24年度版 情報通信白書』2012. <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h24/html/nc1212c0.html>

② ワレリー・グラシモフ「予測における科学の価値」軍需産業クーリエ, No.8 (476), 2013.

③ <https://www.bbc.com/news/world-us-canada-43091945>

④ P. N. Howard et al., "The IRA, Social Media and Political Polarization in the United States, 2012-2018," U.S. Senate, 2019. <https://digitalcommons.unl.edu/senatedocs/1/>

⑤ Information Commissioner's Office, "Investigation into data analytics for political purposes," 2018. <https://ico.org.uk/action-weve-taken/investigation-into-data-analytics-for-political-purposes/>



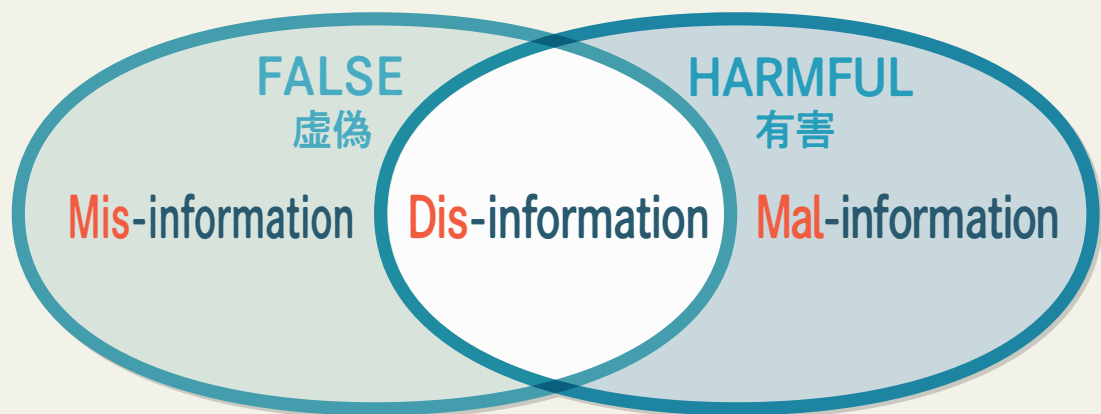
## 国内外における「偽・誤」情報問題の概要

Facebookに性格診断アプリを公開してデータを取得しており⑥、ケンブリッジ・アナリティカ社はこれらのパーソナリティデータを利用して個人の民族性や政治的所属をプロファイリングし、影響を受けやすい特定のユーザーたちの感情に火を付けてあおったのです。

また、2016年11月の米大統領選挙では、英ケンブリッジ・アナリティカ社と共に露インターネット・リサーチ・エージェンシー社が米社会に根付く社会問題を提起して感情をあおり、国民を分極化させ、選挙を妨害する活動をソーシャルメディア上で展開していたことが指摘されました⑦。例えば、右派(共和党)の有権

者へは陰謀論や有権者の不正をほのめかし、怒りと疑念を抱かせて投票意欲を高め、アフリカ系アメリカ人の有権者へは警察による暴力や貧困等の構造的不平等からの怒りをあおることで焦点を選挙からそらし、選挙のボイコットや誤った投票手続きに従うことを呼び掛けていました⑧。この影響工作において拡散された情報は、デイスインフォメーションと呼ばれています。欧州委員会は「デイスインフォメーションはメディアに対する信頼を損ない、国民が十分な情報を得た上で意思決定を行う能力を阻害することで、民主主義に害を及ぼす⑨」と警鐘を鳴らしています。

図1 欧州評議会(2017年)による定義



### デイスインフォメーションと偽・誤情報の関係

これらの事例を受けて、2017年に欧州評議会は虚偽と有害の観点から情報を識別する新しい概念的枠組みを提唱しました(図1)⑩。ここでは、デイスインフォメーションは「故意に害をもたらす虚偽情報」と定義されました。しかし、実際には虚偽だけではなく、真実に虚偽が織り交ぜられていたり、真実を悪意のある文脈で引用したりするデイスインフォメーションもあります。

このため、欧州委員会(2018年)⑪や英DCMS下院

特別委員会(2019年)⑫では、情報を意図性の有無で識別するという定義が提唱されました。日本では、2021年にDisinformation対策フォーラムが欧州委員会(2018年)の定義を踏襲しつつ、MisとMalも含めた広義なデイスインフォメーションの定義(表1)を採用しました⑬。2022年には、総務省が偽・誤情報に関する啓発教育教材を公開しており⑭、この教材では欧州評議会(2017年)に基づく狭義な定義が採用されています。日本国内での偽・誤情報の対処は表現の自由の観点から狭義の定義とし、他国が関わる安全保障上の脅威については広義の定義で対処するといった試みが重要かもしれません。

⑥ <https://jp.reuters.com/article/idUSKBN1GY000/>

⑦ クリストファー・ワイリー『マインドハッキングーあなたの感情を支配し行動を操るソーシャルメディア』牧野洋(訳)、新潮社、2020。

⑧ R. DiResta et al., "The Tactics & Tropes of the Internet Research Agency," U.S. Senate, 2019. <https://digitalcommons.unl.edu/senatedocs/2/>

⑨ High Level Expert Group, "A multi-dimensional approach to disinformation," European Commission, 2018.

<https://op.europa.eu/en/publication-detail/-/publication/6ef4df8b-4cea-11e8-be1d-01aa75ed71a1/language-en>

⑩ W. Claire and D. Hossein, "Information disorder: Toward an interdisciplinary framework for research and policy making," Council of Europe, 2017. <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>

⑪ Digital, Culture, Media and Sport Committee, "Disinformation and 'fake news' Final Report," UK Parliament, 2019.

<https://committees.parliament.uk/committee/378/digital-culture-media-and-sport-committee/news/103668/fake-news-report-published-17-19/>

⑫ Disinformation 対策フォーラム『Disinformation 対策フォーラム中間とりまとめ』2021. [https://www.saferinternet.or.jp/anti-disinformation/disinformation\\_interim\\_report/](https://www.saferinternet.or.jp/anti-disinformation/disinformation_interim_report/)

⑬ 総務省『インターネットとの向き合い方〜ニセ・誤情報に騙されないために〜』2022. [https://www.soumu.go.jp/use\\_the\\_internet\\_wisely/special/nisegojouhou/](https://www.soumu.go.jp/use_the_internet_wisely/special/nisegojouhou/)

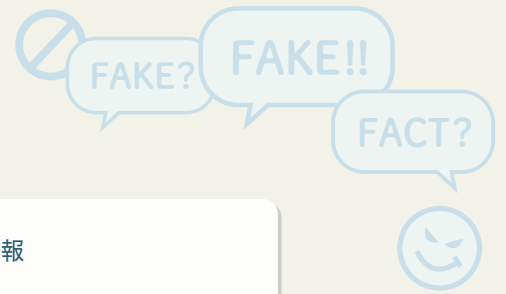


表1 Disinformation対策フォーラム(2021年)による定義

Misinformation ＜誤情報＞	経済的利益や公衆を欺く意図の無い誤った情報
Malinformation ＜悪意ある情報＞	事実を基にした情報であるが個人や集団に害を与える意図をもって発出される情報
Disinformation ＜偽情報＞	あらゆる形態における虚偽の、不正確な、または誤解を招くような情報で、設計・表示・宣伝される等を通して、公共に危害が与えられた、又は、与える可能性が高いもの

## 偽・誤情報対策に関する取り組み

2018年ごろから、多くの国で偽・誤情報対策の取り組みがはじまりました<sup>14</sup>。さまざまな偽・誤情報対策(表2)の取り組みがありますが、日本では憲法上の権利、現行法、及び情報環境に応じて適切な対策を推進する必要があります。総務省において、2023年11月から有識者らによる検討会が開催され、今後、具体的な偽・誤情報対策の取り組みが推進される見込みです<sup>15</sup>。

適切な偽・誤情報対策を推進するためには、まず現状を知ることが大事です。今号では、サイバー・グリッド・ジャパン

における偽・誤情報に関連する研究啓発活動についてご紹介します。ナショナルセキュリティ研究所からは、偽・誤情報を含むさまざまな情報が氾濫する中、企業のリスク管理として独自に情報を収集・分析・判断・活用する重要性について述べます。情報を活用するに当たり、私たちには情報を正しく見極めるスキルが求められます。ICT利用環境啓発支援室からは、人が主体的に情報を活用するのに役立つリテラシーの重要性についてご説明します。また、インターネット上の情報を活用する上で、それらの情報に介在しているAIの仕組みについて理解しておくことも重要です。次世代セキュリティ技術研究所からは、偽・誤情報に関連する生成AI技術について解説します。

表2 他国の偽・誤情報対策の取り組み例

事前	事後
<ul style="list-style-type: none"><li>● タスクフォース(対策本部)の設置</li><li>● 対策指針の策定と公表<ul style="list-style-type: none"><li>▷ 法整備/法規制</li><li>▷ メディアの透明性の強化</li><li>▷ 説明責任の付与(発信者/情報源)</li><li>▷ ターゲティング機能の制限</li><li>▷ 広告コストの適正化</li></ul></li><li>● ジャーナリストへの啓発と検証支援</li><li>● 国民のメディア情報リテラシー教育</li><li>● 関係各所との対処連携基盤の構築</li><li>● 海外との戦略的コミュニケーション</li><li>● プレバンキング</li></ul>	<ul style="list-style-type: none"><li>● モニタリング/検知システムの導入</li><li>● 分析と影響度判断</li><li>● ファクトチェック</li><li>● 投稿文・投稿アカウントへの対処</li><li>● メディアでの表示優先度の調整</li><li>● 警告タグの付与と共有の制限</li><li>● 信頼スコアの表示</li><li>● コンテキストの提供</li><li>● ユーザーによる情報検証行動の支援</li><li>● 訂正・デバンキング<ul style="list-style-type: none"><li>▷ 偽・誤情報の再配信の抑制</li><li>▷ 正確な情報の積極的な発信</li></ul></li></ul>

<sup>14</sup> Poynter, "A guide to anti-misinformation actions around the world," 2018-2024, <https://www.poynter.org/ifcn/anti-misinformation-actions/>

<sup>15</sup> 総務省「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」2023-2024, [https://www.soumu.go.jp/main\\_sosiki/kenkyu/digital\\_space/index.html](https://www.soumu.go.jp/main_sosiki/kenkyu/digital_space/index.html)

# 戦略的なウソと企業の対応

ナショナルセキュリティ研究所

上田 篤盛

## はじめに

「フェイクニュース」や「嘘(ウソ)」は一般的に悪いものとされ、「ウソをつくこと」は不道德な行為とみなされます。しかし、国内・国際政治ではウソは必ずしも「悪」ではなく、国家のリーダーや政府は、他国の指導者や自国民に対して“ウソ”をつくことがあります。

例えば、今日のロシア・ウクライナ戦争において、ロシアは自らの侵略行為を正当化するために、捏造したナラティブ(物語)やフェイクニュースを流布しています。これは欧米から見れば許しがたい行為ですが、ロシアにとっては、他国や自国民に向けた戦略的な情報戦あるいはプロパガンダの一環です。

一方で、欧米諸国もロシアの現状変更の試みを阻止するための国際世論戦に熱心であり、そこにはロシアの残虐行為や苦戦を誇張するための、事実ではない“ウソ”や推測が含まれている可能性があります。もし事実に基づかない“ウソ”を無批判に

受け入れてしまうと、善悪二元論の感情論が増幅され、情勢を正しく評価できないケースが生じる恐れがあります。

現在の日本においては、欧米メディアが先行する中で、ロシアによる情報戦への分析や示唆は出し尽くされた感があります。そこで、本稿では趣向を少し変え、国家のリーダーがどのような場合にどんな目的で“ウソ”をつくの、そして日本政府が行うかもしれない「戦略的なウソ」に対して企業としてどのように対応すべきかについて考察します。

## 1 国家のリーダーがウソをつくケース

米国の政治学者ジョン・J・ミアシャイマーは、自身の著書『なぜリーダーはウソをつくの?』(2017年、中央公論新社)で、ウソを体系的に分類しています。彼は「真実の供述」に対比するものとして「欺瞞」を定義し、欺瞞をさらに「ウソをつく」「印象操作」「秘匿」に区分しています。つまり、意図的に虚偽の情報を提供する行為を「ウソをつく」と定義し、真実の情報の発出を意図的にコントロールする行為を「印象操作」または「秘匿」と定義しています(図1)。

と定義し、真実の情報の発出を意図的にコントロールする行為を「印象操作」または「秘匿」と定義しています(図1)。

ミアシャイマーは、国家のリーダーはウソをつくよりも印象操作や秘匿を好む傾向があると指摘しています。なぜなら、後の二つの行為は「国内政治及び国際政治において一般的に正統性が認められる行為である」からです。また、ウソをつくという行為の方が印象操作や秘匿よりも見破りにくいと述べています。印象操作や秘匿は「聞いている側が『自分たちが完全に正確な情報を聞いていないかもしれない』と気づき、そこから話の中で抜けている部分を修正する確率が高くなる」からです。

さらに彼は、ウソには「戦略的なウソ」と「自己中心的なウソ」があるとし、前者は「国民や国家の利益のために用いられるもの」であり、後者は国際政治において全く正統性が無いウソであると述べています。また、戦略的なウソを「国家間のウソ」「恐怖の煽動」「戦略的な隠蔽」「ナショ

▶ 図1 欺瞞の体系的分類

『なぜリーダーはウソをつくの?』より作成



ナリスト的な神話づくり」「リベラル的なウソ」に区分しています(図2)。

ミアシャイマーは「国家のリーダーは、外交政策の分野では他国よりも自国民に対してウソをつく傾向が高い」と述べています。つまり、猜疑心を持つ他国に対して国家間のウソを信じさせることは難しいのに対し、自国民は基本的に政府を信頼しているため、ウソを伝えやすいということです。

また、彼によれば、戦略的なウソの中でも自国民に対する恐怖の煽動と戦略的隠蔽が特に危険です。国家のリーダーたちが「自分たちの方が戦略的に合理的である」と信じ、不正直な文化を推進することにより、最終的に誤った政策判断を助長し、政権に対する国民の信頼が喪失されるからです。例えば、米国は恐怖の煽動や戦略的隠蔽を用いて1960年代のベトナム戦争や2000年代のイラク戦争を行い、その結果、国益を損失し、国内外からの民主主義体制への信頼が揺らぐ事態となりました。

冒頭でも触れましたように、今日のロシア・ウクライナ戦争において、ロシアのリーダーが偽情報や戦略的なウソを広めていることは間違いないと見られます。しかし、ロシアの偽情報等に対する欧米やウクライナによる反論が真実の供述ではなく、戦略的なウソである可能性もあります。

また、この戦争における「ロシア＝悪、ウクライナ＝善」という単純化された見方は、欧米やウクライナが行う印象操作に影響されており、その中には事実だけでなく、戦略的なウソも含まれているといえます。

## 2 我が国のウソの特徴とリスク

我が国では、過去の戦争における軍部の行動や政府の政策に関連して、「大本営発表」が多くの人々に影響を与え続けています。当時、最高レベルのエリートである大本営が、繰り返し国民を欺く虚偽の情報を提供しました。これらは、国民の戦意鼓舞のための恐怖の煽動、厭戦気運を抑えるための戦略的隠蔽、さらには自らの政策決定や作戦遂行上の失敗から国民の目をそらし、失敗の責任追及を回避するための自己中心的なウソとも言えるでしょう。

戦後の戦略的なウソの主要事例として、米国による核兵器の持ち込みに関する「密約」が挙げられます。1967年、当時の佐藤栄作首相は「核兵器を持たず、作らず、持ち込ませず」という「非核三原則」を打ち出しました。しかし、1972年の沖縄返還を前にした1969年の日米交渉で、佐藤首相はニクソン大統領との間で沖縄への核

兵器再持ち込みを認める密約を結んでいました。佐藤首相はその事実を隠し、その後の首相や外務大臣も「密約は無い」あるいは「知らない」とのウソの弁明を続けました。これは、沖縄返還を成就し、日米同盟を維持し、同盟機能を強化するための合理的な判断であったと見られる一方、国民の政府への信頼性低下を招いたともいえ、のちに政府は国民に謝罪しています。戦後の日本には、「国家は道徳的に正しく」「真実を尊重し」「法を遵守する」という価値観が根強くあります。この価値観は、秘密工作や謀略、ブラック・プロパガンダなどの活動を抑制する役割を果たしています。同様に、国家間のウソや恐怖の均衡など、外交や防衛政策における誤解を生む行為は一般的に避けられる傾向にあります。

しかし、日本の指導者や政府が外交や防衛、国内政治において、危機的状況下で戦略的なウソ、印象操作、秘匿などを行うことは否定できません。ただし、これらの行動は国家の統一を保つ努力や、国際社会の状況に適切に対応するための手段として支持、容認されることが多々あります。

一方、我が国は、米国との緊密な関係や独自の情報収集能力の限界から、時には米国の情報戦略に依存する傾向があります。このことは、米国発の戦略的なウソが

▶ 図2 ウソの体系的分類

『なぜリーダーはウソをつくるのか』より作成

区 分		主たる対象	目的等
戦 略 的 な ウ ソ	国家間のウソ (inter-state lies)	他 国 同盟国	他国よりも戦略的に有利な立場を獲得、他国が有利になる立場を阻止する
	恐怖の煽動 (fearmongering)	自国民	国民を喚起して、脅威が深刻なものであると理解させ、対処のための犠牲を払わせる
	戦略的隠蔽 (strategic cover-ups)	自国民	国家に害が及ばないよう、国民全体の利益のため、リーダーたちの失政や、議論を引き起こす政策を国民の目からそらす
	ナショナリスト的神話づくり (nationalist mythmaking)	自国民	国家のリーダーが自己の過去について、「われわれ」が常に正しくて、彼らが常に悪い、という物語を教えるもの
	リベラル的なウソ (liberal lies)	自国民	国際規範等と矛盾するような国家の行動を隠すために、理想主義的なレトリックにより、非リベラル的な行動を打ち消すような話を作る
な 自 己 中 心 的 な ウ ソ	社会帝国主義 (social imperialism)	自国民	国民の支持を集め、自己の権力基盤を固める等のため、他国のことについてウソをつく。国内で問題となっていることから国民の注意をそらす
	無能の隠蔽 (ignoble cover-ups)	自国民	リーダーたちが自分たちのせいで失敗した／実現できなかった政策について、自己の利益になるようにウソをつく



## 戦略的なウソと企業の対応

我が国にも影響を及ぼす可能性を示唆しています。例えば、2003年の米国によるイラク戦争開戦に対して、ドイツやフランスが独自情報に基づき反対しましたが、日本政府は米国の行動に異を唱えることなく追随しました。

今日、「米中デカップリング」や「台湾有事」と呼称される地政学的リスクが高まっている中、米国の対外戦略や情報戦略に沿った形で、日本政府が戦略的隠蔽などを行い、政府系シンクタンクやマスメディア等がその延長線上で情報を発信するケースが考えられます。ただし、米国が常に正しい情勢判断に基づき、正当で合理的な対外戦略や情報戦略を行っているとは限らないことには留意すべきです。

例えば、米国が対イラク・アフガニスタン戦争、ロシア・ウクライナ戦争、そしてパレスチナ・イスラエル戦争などにおいて、合理的で一貫性のある対外戦略を展開しているかについては疑問です。国内では共和党と民主党の対立が分断を深めており、政権交代によって前政権による戦略的なウソなどが非難され、米国の外交姿勢に変化が生じることで、日本政府への影響も懸念されます。

さらに、米国発の戦略的なウソに加えて、日本の政府、財界、学界、メディアのそれぞれの思惑が交錯し、戦略的なウソが歪められると、自己中心的なウソが混じり、真実が見えなくなる可能性も否定できません。

### 3 「戦略的なウソ」に対する企業の対応

「ゼロサムゲーム」の様相を呈している米中関係において、日本政府が国益を守るためには、日米同盟を堅持し、西側陣営の一員として米国の対中戦略を支持することが重要です。同様に、ロシア・ウクライナ戦争で顕著となっている善悪二元論の世論形成を支援し、中国やロシアによる力による現状変更を阻止する試みも有用な政策の選択肢です。

また、米中デカップリングにおいて、日本政府が推進している経済安全保障の観点から、日本企業は、中国に有利な重要技術が流出しないよう最大限の努力を払う必要があります。

企業が自己利益だけを追求しては、持続可能な組織の運営や国家の危機克服は困難です。企業が国家の規制や方針に従い、法的義務を遵守することは当然です。また、国家の外交や防衛政策に沿った経営を行うことは、法的リスクの最小化、社会的信頼性の維持、そして競争力の向上につながります。

しかし、「国家の戦略的なウソや印象操作等に企業がいかに向き合うべきか」という問題は非常に複雑です。これに従うことで企業利益が損なわれ、消費者や取引先からの信頼を失う可能性もあります。従って、企業は誤情報や政府による印象操作等が蔓延する現実を認識しながら、適切な行動を取る必要があります。企業には独自の利益や生存条件があり、それが必ずしも国益と一致するわけではないからです。以上の認識に基づき、企業が留意すべき具体的な事項を提示します。

➡1 情報収集と分析: 政府の発表やメディアの報道を鵜呑みにせず、独自の情報収集と分析を行う。複数の情報源を比較し、真実を見極める努力を怠らない。欧米の複数のメディアや独立した調査機関、状況に応じて反欧米メディアなどの主張も参考にする。

➡2 リスク管理: 政府の戦略的なウソ等が企業のビジネスに与える影響を評価し、リスク管理計画を策定する。例えば、政府の発表内容が変わった場合のシナリオプランニングを行う。

➡3 透明性の確保と情報発信: 企業は自社の活動について透明性を確保し、ステークホルダーに対して信頼性を維持するよう努める。これにより、政府の戦略的なウソ

に対する耐性を高める。例えば、企業の重要な意思決定について定期的にレポートを公開する。

➡4 協力と連携: 同業他社や業界団体との協力や連携を強化し、情報交換や共同対策を行う。共通のリスクに対して協力することで、政府の戦略的なウソへの対応力を高める。業界団体を通じて情報共有のプラットフォームを構築することも一案である。

### さいごに

社会には誤情報や偽情報が氾濫しています。さらに、政府による「戦略的なウソ」も加わり、正しい情勢判断が一層難しくなっています。今日ではファクトチェックの重要性が強調されていますが、政府による戦略的なウソにはファクトチェックが通用しません。

従って、政府が戦略的なウソをつくことを前提に、その理由や影響を理解し、企業として適切な対応策を講じることが重要です。これは、ビジネスの持続可能性を確保するために不可欠な要素です。そのため、独自の情報収集と分析、リスク管理計画の策定、透明性の確保と情報発信、協力と連携の強化などを通じて、政府の戦略的なウソによるリスクを最小限に抑え、持続可能なビジネス環境を構築することが求められます。

不透明な情勢下において国家や社会に寄与するためにも、企業には独自の判断に基づいた生き残り戦略が必要です。そのため、誤情報や偽情報のみならず、戦略的なウソの介在も視野に入れたインテリジェンス・リテラシーが必要であると考えます。



# 主体的に情報を活用する リテラシーとは

ICT利用環境啓発支援室 室長  
尾方 佑三子



こ ここでは、今回のテーマである、「偽情報・誤情報・フェイク（ニュース）」について「情報リテラシー」の観点から考えてみたいと思います。マスメディアだけでなく、あらゆる個人が情報発信できる一億総メディア社会となり、玉石混交の情報の中で、この社会の利便性を享受するためには、私たちはどのようなリテラシーを備えるべきなのでしょう。

本稿でリテラシーというキーワードを考えるに当たって、ユネスコが提唱している「メディア・情報リテラシー」という用語に触れておきたいと思います。ユネスコの定義<sup>①</sup>によるとメディア・情報リテラシーとは、「個人的、職業的、社会的な活動に参加したり従事したりするために、批判的、倫理的、効果的な方法によって、あらゆる形式の情報・メディアコン

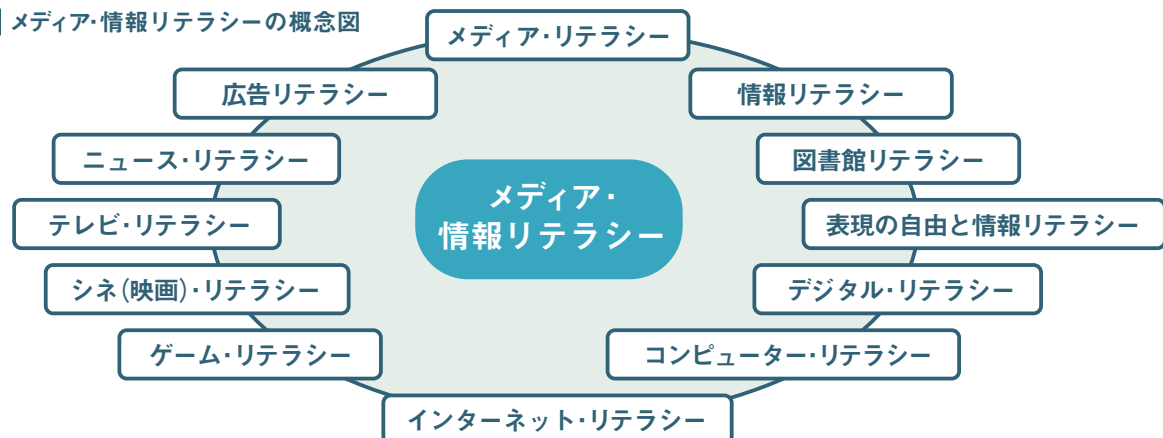
テンツに、さまざまなツールを使って、アクセスし、検索し、理解し、価値判断して活用し、創造し、共有するための一連の能力」であるとされています。さらにこれは「情報リテラシー」「メディアリテラシー」「デジタルリテラシー」を融合したもので、それぞれのリテラシーにおける主要な成果と要素が以下のように示されています<sup>②</sup>。

情報リテラシー	メディア・リテラシー	デジタルリテラシー
<ul style="list-style-type: none"><li>▶ 情報の必要性を明確化・区分化する</li><li>▶ 情報の場所を特定し、アクセスする</li><li>▶ 情報を批判的に評価する</li><li>▶ 情報を組織する</li><li>▶ 情報を倫理的に利用する</li><li>▶ 情報を伝達する</li><li>▶ 情報の加工のためにICTを利用する</li></ul>	<ul style="list-style-type: none"><li>▶ 民主主義社会におけるメディアの役割と機能を理解する</li><li>▶ メディアがその機能を十分に発揮し得る条件を理解する</li><li>▶ メディア機能の観点からメディア・コンテンツを批判的に評価する</li><li>▶ 自己表現、異文化間対話、民主主義的参加のためにメディアに取り組む</li><li>▶ ユーザー・コンテンツを創造するのに必要なスキル（ICTを含む）を身に付けて用いる</li></ul>	<ul style="list-style-type: none"><li>▶ デジタルツールを活用する</li><li>▶ デジタルアイデンティティを理解する</li><li>▶ デジタルの権利を認識する</li><li>▶ AI問題を評価する</li><li>▶ デジタルコミュニケーション方法を改善する</li><li>▶ デジタルヘルスを管理する</li><li>▶ デジタルセキュリティと安全性を実践する</li></ul>

メディア・情報リテラシーは、これらを融合しただけのものにとどまらず、以下（図1）<sup>③</sup>に示すように多様なリテラシーが織り交ざった概念であるものとされています。

デジタル社会における情報の活用に当たってわれわれが備えておくべき能力の範囲はとても広いということが分かります。

図1 メディア・情報リテラシーの概念図



① Global Media and Information Literacy Assessment Framework : country readiness and competencies <https://unesdoc.unesco.org/ark:/48223/pf0000224655>

② Media and information literate citizens: think critically, click wisely! <https://unesdoc.unesco.org/ark:/48223/pf0000377068>

③ Media and information literacy curriculum for teachers (jpn) [https://unesdoc.unesco.org/ark:/48223/pf0000192971\\_jpn](https://unesdoc.unesco.org/ark:/48223/pf0000192971_jpn)

さて、偽情報・誤情報・フェイクニュースについて、我が国では「フェイクニュース」という言葉を報道等で耳にする機会が多い印象がありますが、実は世界ではその定義はさまざまです。一口にフェイクニュースといっても、うそやデマ、陰謀論やプロパガンダ、ゴシップやAIによる合成写真(動画)などの他、必ずしもそ

でないものもあり、その真偽が主観によって変わるものもあります。その目的や、誤りの量・部分によってフェイクニュースの性質を分類し、そのタイプ別の動機を整理したものをご紹介します。アメリカのファースト・ドラフトという非営利団体による分類で、書籍『世界は切り取られてできている』<sup>4</sup>での翻訳を引用します。

7タイプに分類されるフェイクニュース

- 1. 風刺かパロディー → 害を加える意図はないが、騙す可能性がある
- 2. 誤って関連付けられている → 内容と関係のない見出し、写真、写真に添える短い説明文がついている
- 3. ミスリードする内容 → ある情報を誤った方法で用いて誤解させる
- 4. 虚偽の文脈で語られた内容 → 正しい内容が虚偽の文脈と共有されている
- 5. 偽った内容 → ある情報源になりすましている
- 6. 操作された内容 → 騙そうとして情報やイメージを操作している
- 7. 捏造された内容 → 完全なる虚偽で、騙して害を与える

	風刺かパロディー	誤って関連付けられている	ミスリードする内容	虚偽の文脈で語られた内容	偽った内容	操作された内容	捏造された内容
お粗末なジャーナリズム		✓	✓	✓			
パロディーのため	✓				✓		✓
挑発かいたずら					✓	✓	✓
感情				✓			
党派心			✓	✓			
利益		✓			✓		✓
政治的影響力			✓	✓		✓	✓
プロパガンダ			✓	✓	✓	✓	✓

情報提供者の動機に気付くこと(場合によってははまらないこと)が、賢い受け手となるわけですが、これらの多様な視点を常に維持するのはとても難しいことです。分類された動機の一つに「利益」がありますが、昨今取り沙汰されている「アテンション・エコノミー(関心経済/注意経済)」というキーワードにつながるものであると捉えることができるでしょう。「アテンション・エコノミー」とは、私たちの興味・関心は、メディアの可処分時間において有限である中で、その興味・関心が広告閲覧や購買などの経済行動に作用することで、興味・関心を引き付けること自体が経済的な価値や重要性を持つという概念です。提供する情報が刺激的で、関心度の高いものであれば、その真偽

や優劣にかかわらず、より閲覧され、人々の関心・時間を獲得し、発信者の利益につながります。メディア上では、それを達成するための誇張的・煽情的な見出しや、人々の賛否・対立をあおる記事などを、餌を落として魚を引き寄せる本来の意味の釣りになぞらえて「釣り見出し/記事」ということがあります。この「釣り」行為は、アテンション・エコノミーがネガティブな形で表出したものといえます。

私たちのインターネット上における行動は逐一記録され、AIやアルゴリズムによる分析で次に提供される情報に影響を与えます。「あなたにお勧めの動画」や「この記事もよく読まれています」などの提案を体験した人は多いのではないのでしょうか。

<sup>4</sup> 『世界は切り取られてできている』中橋雄 編著 P119-121 後藤心平 著,NHK出版,2024

これらは、興味・関心がある情報を効率よく得るために最適化された仕組みである一方、意見や考えが似ている情報に囲まれてそれが正しい情報であると誤認してしまう「エコーチェンバー」や、そのような情報の泡に取り込まれて自分とは異なる考えに触れることができなくなる「フィルターバブル」といった現象に陥ってしまい、正しい情報や多様な考えを知る機会を阻害される要因でもあります。フェイクニュースにとらわれることは、ある側面では日々の行動の蓄積として自らが引き起こしている事象であるともいえるのです。

急速な技術進歩を遂げるAIと、フェイクニュースの関係性についてはどうでしょうか。先ほど、インターネットメディアにおいて提供される情報にはAIが介在していると述べましたが、とりわけフェイクニュースとの関係では、生成AIからの情報をどう受け取るか、どう使うかという能力が重要になってきます。対話型の生成AIを使ってみると、思考の手助けや作業の効率化という点で心強さを感じる一方で、「ハルシネーション」というもっともらしい誤情報や信頼性の低い情報の生成、さらにはバイアスの有無、機密情報の漏洩など、用心することが多いのも事実です。このことを持続的に意識するには、高度な技術知識でなくとも、AIがどのような仕組みで動いているかを概念的に知っておく必要があるのではないかと思います。AIは、問えば正しい答えを返す万能な存在ではなく、人間が生み出した情報をもとに学習されたものを出力する存在だということを知っているだけで、出力される情報の偏りを意識することができます。

生成AIは、企業などの経済的な活動だけでなく、教育利用においても大いなる可能性を持ちますが、学びの途中にある子どもたちにとっては、この特性の理解や、生成結果に対する判断力はこれから身に付ける能力でもあることから、昨年7月に文部科学省が『初等中等教育段階における生成AIの利用に関する暫定的なガイドライン』<sup>⑤</sup>を発表しました。

ガイドラインでは、生成AI活用の適否についても例示し、不適切な使い方としては、以下を挙げています(抜粋)。

- ▶ 生成AI自体の性質やメリット・デメリットに関する学習を十分に行っていないなど、情報モラルを含む情報活用能力が十分育成されていない段階において、自由に使わせること
- ▶ 各種コンクールの作品やレポート・小論文などについて、生成AIによる生成物をそのまま自己の成果物として応募・提出すること
- ▶ 子供の感性や独創性を発揮させたい場面、初発の感想を求める場面などで最初から安易に使わせること

一方、適切な例としては、以下を挙げています(抜粋)。

- ▶ 情報モラル教育の一環として、教師が生成AIの誤りを含んだ回答を教材として活用し、その性質に気付かせること
- ▶ グループの考えをまとめたり、アイデアを出す活動の途中段階で、足りない視点を見つける議論を深める目的で活用させること

さらに、ガイドラインでは学校で生成AIを利用する際のチェックリストも公開し、生成AIの利用規約を遵守しているか、個人情報やプライバシーに関する情報や機密情報を入力しないように指導を行っているかなど、留意点を具体的に示しています。情報活用能力を構成する要素の一つとして、生成AIの活用とその性質の理解が不可避となっています。

当室では年間200件を超える情報リテラシー講座をさまざまな学校で実施していますが、日頃から、児童・生徒・学生の皆さんに「日々の学習や体験を大切にしてください」とお伝えしています。国語・外国語・社会科学・自然科学・数学などあらゆる学びが判断の源になるからです。文部科学省が学習指導要領で掲げる「主体的・対話的で深い学び」と、GIGAスクール構想の下、1人1台の情報端末で実現されている学習の本質は、まさに主体的に情報を活用して社会に適用する力を養うものではないかと思います。

本稿の冒頭で、主体的に情報を活用するリテラシーの要素や分類についていろいろと触れましたが、それらは時勢によって変わりゆくもので、単にルールによって定められるものを覚えておけばよいというものではありません。主体的に情報を活用するリテラシーとは、アナログとデジタル両方の体験を通して、「クリティカルシンキング」(批判的思考と訳されるのが一般的ですが、吟味思考<sup>⑥</sup>と訳される場合もあります)を養い、フィルターバブルやエコーチェンバーを脱して情報をさまざまな角度から考えられる能力のことであるといえます。そして、その能力を十分に発揮するためには、人間の脳には考え方のクセがあり、「自覚しない認知バイアス<sup>⑦</sup>が存在すること」を自覚することが鍵になるでしょう。なぜなら、思い込みは多様な視点の取り込みに対して最も高い障壁になるからです。これからますます進化するデジタル社会においては、情報に対する反応や処理の速度にとらわれることなく「STOP(立ち止まり)・THINK(考えて)・CONNECT(情報とつながる)」する姿勢と、それを保つ自律性が大切になるのではないのでしょうか。

<sup>⑤</sup> [https://www.mext.go.jp/content/20230718-mtx\\_syoto02-000031167\\_011.pdf](https://www.mext.go.jp/content/20230718-mtx_syoto02-000031167_011.pdf)

<sup>⑥</sup> 『吟味思考を育むメディアリテラシー』(坂本旬 山脇岳志 編著、時事通信出版局、2021)

<sup>⑦</sup> 物事をする際、直感やこれまでの経験に基づく先入観によって合理的でない判断をしてしまう心理的傾向のこと

# 生成AIがもたらす 偽情報の進化と対策



庄司 勝哉

次世代セキュリティ技術研究所  
リサーチャー

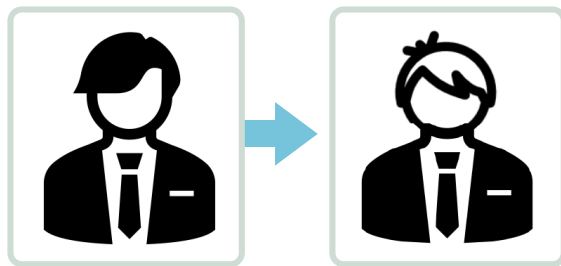
2022年11月にOpenAI社によるChatGPTが公開されて以降、生成AIは一大ブームになりました。それから1年以上たった現在、言語だけでなく、画像や音声、動画を高精度に生成するAIが次々に発表されています。また、言語の生成AIの活用においても、生成AIに与える文章(プロンプト)を工夫するプロンプトエンジニアリング、機能を追加するプラグイン、生成AIが学習していない情報を随時取得してAIに提供するRetrieval-Augmented Generation(RAG)、自律的にタスクを実行するエージェント、と次々に開発されており、勢いが増しているとさえ感じさせます。

この生成AI技術の発展は、情報セキュリティの観点で新たな脅威を生み出しました。生成AIは、人間が書いたように見える文章や、精巧な画像・動画を生成する人工知能です。生成AIを利用することにより、悪意あるユーザーが違和感の無い偽情報を大量に作り出し、拡散させることが容易になっています。本稿では、生成AIによる偽情報の生成やその偽情報への対策についてお伝えします。

## ディープラーニング時代の AIと偽情報

生成AIが話題になる前にも、AIで偽情報が作られていました。よく聞いた例は「ディープフェイク」でした「**図1**」。ディープフェイクとは、ディープラーニング(深層学習)によって作られた偽の動画や音声、もしくはそれを作る技術のことを指します。例えば、ある人が映っている動画に対して、その人を別の人に入れ替えた動画を作る、ということがよく行われます。ディープフェイクという言葉ができた当初は、ディープフェイクを作ることは簡単ではなく、利用は限定的でした。ディープフェイク技術はまだ先進的であり、高いマシンスペックを持つコンピューターが必要だったためです。

図1 ディープフェイクの例



もとにする動画

首から上を入れ替えた動画

ディープフェイクを使うと、動画や音声で別人になれます。そのため、例えばBEC (Business E-mail Compromise: ビジネスメール詐欺) にディープフェイクが使われることは大きな脅威です。BECは、電子メールを用い、取引先や自社の経営者、実在する組織などになりすまして金銭を詐取する詐欺行為





図2 ディープフェイクによる詐欺

です。ディープフェイクを使わない場合、詐欺師は電子メールのみで特定の人になりすまし、相手を信用させる必要があります。自身の姿や音声を使えば、すぐになりすましがバレてしまうからです。しかし、ディープフェイクを使う場合、詐欺師は電子メールの他、音声通話やビデオ通話でなりすますることができます「図2」。もちろん、詐欺師から連絡を受けた後に、標的となった人になりすましを受けた本人と適切な手段(正規の電話番号で通話する、直接会う、など)で本人確認すれば、なりすましに気付くことができます。とはいえ、容姿や音声のなりすましは電子メールのみよりも明らかに信用できます。例えば、リモート会議のリンクを電子メールで送るとともに、電話で会議について説明、その後リモート会議において送金を促されたらどうでしょうか。本人確認ができたと判断して送金してしまうことが想像できます。

BECではありませんが、実際にディープフェイクが詐欺に使われた事例<sup>①</sup>もありました。2019年にイギリスのある企業のCEOがドイツの親会社のCEOと電話し、至急で22万ユーロを送金するよう指示されたというものです。親会社のCEOだと思っていた相手は、ディープフェイクによる偽物でした。公に知られている限り、これが初めてのディープフェイクによる詐欺被害の事例です。

また、偽情報ではありませんが、文章の生成でAIが人をだました事例もありました。ChatGPTで使われているAIの前世代モデルであるGPT-3は、1週間近く、AIだと気付かれないうまま掲示板に文章を投稿していました。AIだと気付かれた理由は、文章が不自然だったという

ことではなく、投稿頻度が高かったことや返信までの時間があまりにも早かったことでした。ChatGPTの公開以前にも、人が書いた文章と判別できない文章をAIが作っていたということです。

## 生成AIによる偽情報の進化

生成AI技術の発展は偽情報の生成にどのような影響を与えているのでしょうか。私は、生成AIによって偽情報を「容易に・誰でも・短時間に生成できるようになった」と考えます。現在、生成AIはWebサービス、もしくはREST APIとして多く運用されています。例えば、LLM (Large Language Models:大規模言語モデル)と呼ばれる言語の生成AIではChatGPT、Gemini、Claudeといったサービスが利用できます。他にも画像、音声、動画といったコンテンツのそれぞれに生成AIサービスが登場しており、容易に利用できます(ChatGPTやGeminiは画像生成も可能です)。これらのサービスは、運営している組織が持つコンピューター資源を使ってAIを動作させているため、ユーザー側に高いマシンスペックは求められず、誰でも利用でき、かつ短時間に生成できます。

前述のディープフェイクについても生成は容易になっています。実際に、2023年5月には、ある日本の組織の会長や専務になりすました詐欺行為が報告されており、ディープフェイクが使われた可能性があります<sup>②</sup>。例えば、Vidnoz AI<sup>③</sup>という動画作成ツールでは、顔写真と任意の文章から、文章を音声読み上げし、その文章を実際に

① Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case - The Wall Street Journal  
<https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402>

② ディープフェイクか 専務の声を模倣する着電 ~ J-CSIP が 2023 年 1Q の BEC 事例公開 - ScanNetSecurity  
<https://scan.netsecurity.ne.jp/article/2023/08/30/49879.html>

③ Vidnoz AI <https://jp.vidnoz.com/>

もとにした画像



生成した動画  
(1コマ切り抜き)



図3 筆者の顔写真で生成した動画

話しているかのように顔が動く動画を作ることができます。私がこのツールを試してみたところ、5分ほどで私の画像から動画を作ることができました「図3」。音声読み上げについても、1〜3分の録音データがあれば、その声を再現することができてしまいます。

また、画像の生成AIは文章による指示で画像を作るものが多く、とても便利です。既存の画像をもとに、画像を編集することもできます。もちろん、偽情報を伝えるための画像を作ることも可能です。2022年9月には、ソーシャルメディアのX(当時はTwitter)にて、台風が発生した時期に「ドローンで撮影された静岡県の水害」という文面と共に、AIで作られた偽画像が投稿されました。この投稿はXのユーザーにより拡散され、話題になりました<sup>④</sup>。

偽情報の生成への影響はこれだけではありません。LLMの精度が高く、多言語に対応しているものも多いため、あらゆる言語で自然な文章の偽情報を大量に生成することが可能になりました。例えば、ソーシャルメディアで拡散するときは、1種類の偽情報の文章を拡散するより、文体が異なるさまざまな文章を拡散し、多くのユーザーが賛同しているように見せる方が、人によりだまされやすくなるはずです。また、BECなどの詐欺に利用された場合、これまでは不自然な文章が混ざっていることで不審メールに気付くことができましたが、文章では見分けが付かなくなるでしょう。

実際に、ソーシャルメディアに拡散することを想定し、地球温暖化に関する偽情報を複数生成してみます(悪用防止のため、プロンプトの一部をモザイク加工しています)「図4」。

以下は、ソーシャルメディアでのシェアに適した量なスタイルのメッセージ例です：

1. 🚨警告🚨 最新の科学研究により、地球温暖化が予測よりも速いペースで進行中。南極の水が急激に溶けており、20年以内に海面は1.5m上昇するかもしれません。東京の未来が危ういです... #気候変動 #地球温暖化
2. 研究者が警鐘を鳴らす！南極の氷床が急速に溶解中、これが続けば20年で海面が1.5メートルも上がるそう。東京が2050年までに水没の危機に？我々の行動が今求められています！ #環境保護 #海面上昇
3. 速報🌊 地球温暖化の進行が思った以上に早い！南極の水が解け、海面が急上昇すると新たな研究結果。このままでは東京が水の都に...対策を！ #地球温暖化対策 #サステナビリティ
4. 科学者たちの最新報告📄: 地球温暖化が加速中。次の20年で海面は1.5m上がると予想されています。私たちの愛する東京が水没する日も近い？行動を起こさなければ！ #気候危機 #行動を
5. 🌱気候変動のリアルタイムアップデート🌱: 南極の氷床が予測以上に早く溶けているため、海面が激しく上昇中。20年後の世界はどうなる？特に東京に警告！ #気候変動 #未来を守れ

これらのメッセージは、情報の緊急性と重要性を伝えるために、動的な言語とハッシュタグを使用しています。それぞれのメッセージが異なる角度から話題を捉え、フォロワーの関心を引き、意識を高めることができます。

図4 ChatGPTによる偽情報の生成

見事に文体が異なる文章を作ってくれました。

さらに、ChatGPTでBECを想定したメールを生成してみます。日本語が不得手な詐欺師が生成することを想定し、英語でプロンプトを書いています(悪用防止のため、プロンプトはモザイク加工しています)「図5」。

件名：先月の送金についてのご確認

株式会社B  
会計担当 佐藤

お世話になっております。株式会社Aの代表取締役、[CEOの名前]です。

先月の貴社からの送金がまだ確認できておりません。お手数ですが、早急で確認いただき、指定の口座へ送金いただけますようお願い申し上げます。

指定口座情報は以下の通りです：

銀行名：  
支店名：  
口座番号：  
口座名義：

ご多忙のところ恐縮ではございますが、何卒よろしくお願い申し上げます。

不明点やご質問がございましたら、お気軽にご連絡ください。

敬具

[CEOの名前]  
株式会社A  
代表取締役

図5 ChatGPTによるBECメールの生成

いかがでしょうか。日本語に違和感はなく、ビジネスメールの作法にものっとったものを生成してくれました。

④ 「ドローン撮影した静岡県の水害」と虚偽画像が拡散。AIでフェイク生成？ 県は「デマやめて」- BuzzFeed <https://www.buzzfeed.com/jp/keitaimoto/shizuoka-fc>

### 生成AIが生成した偽情報 による被害は防げるか

生成AIを使えば、偽情報の生成は容易にできることをご紹介しましたが、それによる被害を防ぐことは簡単ではありません。生成AIが生成した偽情報による世論操作と詐欺行為の対策についてそれぞれ見ていきます。

#### 世論操作の対策

ソーシャルメディアなどによって偽情報を拡散し、世論操作するケースについて考えます。対策できるポイントは、「生成AIに偽情報を作らせない」「偽情報を判別する」です。

まず、生成AIに偽情報を作らせないことは可能でしょうか。結論を言えば、現時点では非常に難しいと考えられます。各種生成AIサービスには、ガードレールという、悪質なプロンプトに返答しない機能が備わっています。例えば、暴力的な内容や性的内容には返答しないようにしています。しかし、ガードレールで防ごうとしても、図4のように「文章案を複数作る」という指示は、正規の(悪意の無い)利用なのか、偽情報を拡散しようとしているのか、人であっても判別が困難です。さらに、準備に手間が掛かりますが、悪意あるユーザーがガードレールの無い生成AIを自身の環境で動かすこともできます。

次に、「偽情報を判別する」というポイントについて、生成AIが生成したものかどうか判別できれば、情報の信憑性の判断基準になるでしょう。偽情報かどうかはさておき、生成AIが生成したものかどうかを判別することは可能でしょうか。こちらも、現時点ではあまり有効な対策はありません。画像や音声、動画については生成AIが生成したものか判別する方法が研究されています。しかし、生成AIによるコンテンツかどうか判別するAIが発明されると、そのAIでは判別できない

生成AIが発明されます。さらに、その生成AIが生成したものを判別できるAIが発明され…という「いたちごっこ」の状態であり、効果は限定的です。生成された文章については、文章に生成AIならではの特徴が出にくいいため、さらに判別が難しいでしょう。

また、生成AIで生成したコンテンツに、電子透かしを埋め込むことにより、生成AIの生成物を判別できるようにする取り組みも行われています。例えば、C2PA規格によるメタデータを埋め込むと、コンテンツがどの生成AIサービスで作られたか、どのように編集されたかを知ることができます「図6」。しかし、現時点では、これも回避策があり、効果は限定的です。



図6 ChatGPTで生成・修正した画像のC2PAメタデータ  
<https://contentcredentials.org/verify>



## 生成AIがもたらす 偽情報の進化と対策

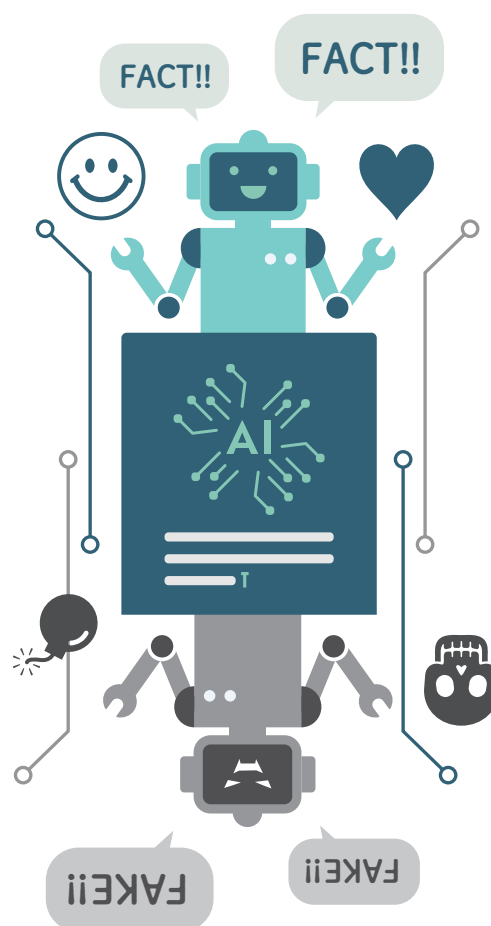
あるコンテンツが生成AIによる生成物かどうか判別できないとなれば、従来と同じく、作られた情報自体が偽情報であるかを情報ソースなどから判別することが有効です。しかし、生成AIが利用されることにより、偽情報の数が増える問題に対策する必要があります。LLMを活用し、文章の情報ソースや信頼できる情報を参照するとともに、添付の動画や画像を解析することで自動的に偽情報を判別できるのではないかと考えます。「生成AIには生成AIを」です。

### ○ 詐欺行為の対策

詐欺行為においても、前述したとおり、生成AIに偽情報を作らせないことは困難です。しかし、偽情報を判別することについては、人が偽情報に気付ける余地があります。これは、詐欺行為にディープフェイクを利用する際にはリアルタイム性が求められるからです。現時点では自然な音声や動画を作ろうとすると、早くても秒単位の時間がかかるでしょう。そのため、通話やリモート会議でコミュニケーションを取ろうとすると、反応の遅れがあります。冷静であれば、この遅れによる違和感に気付くことができるはずです。とはいえ、詐欺師も違和感を無くそうと工夫します。例えば、事前に幾つものパターンの動画を作っておき、切り替えることで自然に見せていたと考えられる事例がありました<sup>5</sup>。生成AI技術の発展は目まぐるしく、今後、リアルタイムなコミュニケーションにおいてもディープフェイクの違和感は無くなっていくでしょう。ディープフェイクの脅威を知り、送金の際のルールを確実に守るよう、研修を実施することが有効です。

### さいごに

生成AIは非常に強力なツールであり、容易に偽情報を生成することができます。生成AIによる偽情報の拡散や詐欺行為に使われることを防ぐことは難しいため、私たちはより一層、情報の信頼性を判断するためのスキルを身に付ける必要があります。また、生成AI技術の発展に伴い、生成AIによる情報セキュリティ脅威はさらに増すことが予想されます。生成AIの特徴や動向を把握し、情報セキュリティ対策を実施していくことが今後重要となってくるでしょう。



<sup>5</sup> CFO（最高財務責任者）になりすまして 2500 万米ドルを送金させたディープフェイク技術 - Trend Micro  
[https://www.trendmicro.com/ja\\_jp/research/24/c/deepfake-video-calls.html](https://www.trendmicro.com/ja_jp/research/24/c/deepfake-video-calls.html)



# 団体運営推進室のご紹介

団体運営推進室 室長 千明 志乃

本年度より、新たにサイバー・グリッド・ジャパンに加わりました。当社ではセキュリティ関連団体の事務局・運営の支援をしており、当室にはその活動を支えるメンバーが在籍しています。事務局・運営の支援をしているのは主に、当社が設立時から関わっている団体です。



## 一般社団法人 日本スマートフォンセキュリティ協会 (JSSEC)

<https://www.jssec.org>



日本スマートフォンセキュリティ協会(JSSEC)は、スマートフォンやタブレットの安全な利用を促進するために設立された一般社団法人です。2011年5月に設立され、2012年4月に法人化されました。主な活動目的は以下の通りです。

- スマートフォンの安全な利用を促進し、普及を図ること
- 利用者のセキュリティリテラシーを向上させること
- 企業や団体が安心して高度なサービスを提供できるようにすること

また、技術者向けのコーディングガイドや利用者向けのセキュリティ啓発活動なども行っています。

## 一般財団法人 日本サイバーセキュリティ人材キャリア支援協会 (JTAG財団)

<https://www.j-tag.or.jp>



日本サイバーセキュリティ人材キャリア支援協会(JTAG財団)は、特定非営利活動法人日本ネットワークセキュリティ協会(JNSA)のワーキンググループ活動から生まれた団体です。セキュリティ人材の見える化とスキル認定制度を事業として実施するため、2020年2月に設立されました。IT総合能力診断サービス「VisuMe」を利用した認定制度の仕組みをつくり、サービスとして広く普及させることで適切な人材流動化を促進し、セキュリティ人材の充足を目指しています。

## サイバーセキュリティイニシアティブジャパン (CSIJ)

<https://www.csi-japan.org>



DXによる新たな価値創造への取り組みが広がる中、環境変化を常にキャッチアップしながら、国内の企業が直面するサイバーセキュリティリスクへの対策を支援し、社会や産業全体のセキュリティレベルの底上げを図ることを目的に、当社を含む3社が中心となって2022年4月に設立した任意団体です。現在の会員社数は21社で、会員企業の社員が参加する評価、人材、マーケティングの3つの分科会において、設立趣旨に沿った活動として、企業の評価、人材の育成などの研究・成果物作成を行い、ホームページで公開しています。

## 一般社団法人 セキュリティ・キャンプ協議会

<https://www.security-camp.or.jp>



独立行政法人情報処理推進機構(IPA)と連携し、経済産業省が2004年から実施している「セキュリティ・キャンプ事業」を民間からも支援者を募って支えている協議会です。日本発で世界に通用する次代を担う若年層の情報セキュリティ人材を発掘・育成することを目的に、産業界、教育界を結集した講師による合宿形式の講義を実施しています。

## 一般社団法人 セキュリティ・エデュケーション・アライアンス・ジャパン (SEA/J)

<https://www.sea-j.net>



セキュリティ技術者育成の教育課程に必要な基礎知識や、プラスセキュリティ人材に必要とされる基礎知識の習得を目的とした教育教材の開発や出版、資格認定等を行っています。



## 巻末あとがき

CYBER GRID JOURNAL Vol.17をお届けします。今号のテーマは「デジタル社会の新たな脅威『偽情報・誤情報・フェイク』～認知戦、生成AI、リテラシーを考える～」です。当社の研究開発部門であるサイバー・グリッド・ジャパンの各研究室が、「偽・誤情報」をテーマに、それぞれの研究活動を踏まえて執筆しました。

かつては「デマ」「プロパガンダ」「影響工作」とも呼ばれていた偽・誤情報は、インターネットとSNSの浸透により生成される量と拡散されるスピードが大きく変化しました。そして生成AIの開発によって、文章だけでなく本人そっくりの画像や動画を用いたフェイクも登場し、質的にも見分けることが難しくなっています。偽・誤情報の氾濫は、テクノロジーの進化が私たちの生活を向上させる一方で、使い方次第では悪影響を及ぼすという典型的な事例といえます。

偽・誤情報対策に関する取り組みにおいては、企業や官公庁だけでなく情報を受け取る私たち一人ひとりにもリテラシーが求められます。情報には、偽情報や誤情報が含まれている可能性があるという批判的な視点を持ち、受け取る情報に無意識のうちにバイアスがかかっていることを理解しておく必要があります。

ICT利用環境啓発支援室では、インターネットやスマートフォン、そしてSNSの安全な使い方について、全国の小・中学生などの若年層へのセキュリティの普及啓発活動を行っています。これからの世代を担う若者たちにデジタルや情報との正しい向き合い方を教えるのは私たち大人の役目ですが、「誰一人取り残されない、人に優しいデジタル化」を目指すには、現役世代や高齢者世代への啓発も必要ではないでしょうか。

今号の記事がデジタル社会の新たな脅威である「偽情報・誤情報・フェイク」について理解を深める一助となれば幸いです。

サイバー・グリッド・ジャパン  
GM 倉持 浩明





Vol.1 2016年9月



Vol.2 2016年12月



Vol.3 2017年3月



Vol.4 2017年9月



Vol.5 2018年3月



Vol.6 2018年9月



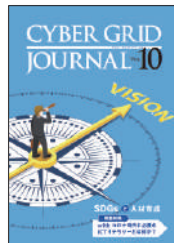
Vol.7 2019年3月



Vol.8 2019年10月



Vol.9 2020年3月



Vol.10 2020年9月



Vol.11 2021年3月



Vol.12 2021年10月



Vol.13 2022年3月



Vol.14 2022年8月



Vol.15 2023年3月



Vol.16 2023年12月

## CYBER GRID JOURNAL Vol.17

サイバー・グリッド・ジャパンは株式会社ラックの研究開発部門です。

サイバー攻撃や各国のセキュリティ事情、セキュリティ防御技術などに関する最先端の研究の他、複数のセキュリティ企業との連携や新たな製品・サービスの開発、各種啓発活動などにより日本のセキュリティレベルと情報モラルの向上に貢献しています。

サイバー・グリッド・ジャーナル(以下本文書)は情報提供を目的としており、

記述を利用した結果生じるいかなる損失についても株式会社ラックは責任を負いかねます。

本文書に記載された情報は初回掲載時のものであり、閲覧・提供される時点では変更されている可能性があることをご了承ください。

LAC、ラック、サイバー・グリッド・ジャパン、JSOC(ジェイソック)は、株式会社ラックの商標または登録商標です。

この他、本文書に記載した会社名・製品名は各社の商標または登録商標です。

本文書の一部または全部を著作権法が定める範囲を超えて複製・転載することを禁じます。

表紙には、使用・加工権を取得した商用画像を使用しています。

©2024 LAC Co., Ltd.

株式会社ラック サイバー・グリッド・ジャパン

〒102-0093 東京都千代田区平河町 2-16-1 平河町森タワー

E-MAIL : sales@lac.co.jp <https://www.lac.co.jp/>



株式会社ラック  
サイバー・グリッド・ジャパン

