AI戦略会議（2023.05.26）

議論用資料

2023.05

山口 真一　博士（経済学）
国際大学GLOCOM准教授
syamaguchi@glocom.ac.jp

国際大学 INTERNATIONAL UNIVERSITY OF JAPAN

© Yamaguchi. 2023 All Rights Reserved.

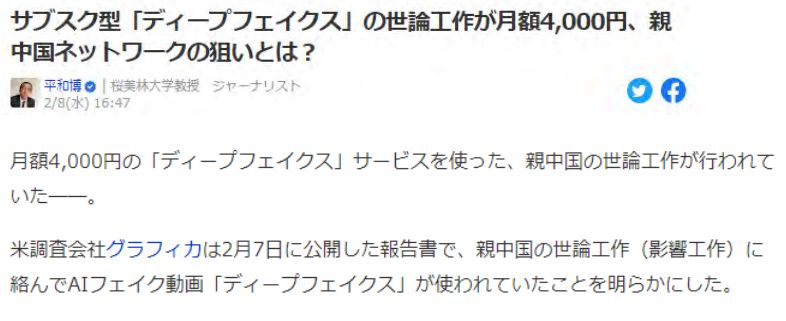
現状認識：優れた道具であり、ポジティブにもネガティブにも

- 生成AIは道具である。企業の生産性・創造性を高め、個人の人生を豊かにするポテンシャルを持っている一方で、様々なリスクもある。良い点を延ばし、悪い点を抑える施策が必要。活用を前提に、「正しく怖がる」。
- リスクの面では、既に政治的意図・ビジネス的意図をもって影響力工作（世論工作）に利用されている。また、ディープフェイクの民主化、影響力工作の民主化により、今後爆発的に偽・誤情報が増加する可能性がある。
- 知的財産権の侵害、情報漏洩、サイバー攻撃、戦争・プロパガンダへの利用、詐欺への活用、誤ったメディア記事の配信、など、様々な潜在的リスクが指摘されているため、社会全体で正しくリスク情報を共有しておくことが大切。

静岡県の水害に関連した実際の投稿（2022年）
https://www.itmedia.co.jp/news/articles/2209/26/news180.html

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上で拡散し、株価が一時、下落する騒動に発展（5/23）
- ブルームバーグを装った「ブルームバーグ・フィード」というアカウントも投稿したことで、株価にも影響。
- インドの主要テレビ局も誤って放送。

https://www3.nhk.or.jp/news/html/20230523/k10014075821000.html



ディープフェイク作成の安価なサブスクが世論工作に使われている
https://news.yahoo.co.jp/byline/kazuhirotaira/20230208-00336196

法律・政府のあり方：適切な利用を促すルール作り

- 強い法規制（禁止など）ありきで考えるべきではない。最小限の規制で、ネガティブポイントを抑えて社会的厚生を最大にすることが重要。強い法規制は適切な活用を阻害するだけでなく、結局利用を止めることはできず、問題がより見づらい場所で起こることにつながる。
- 一方、現行法で対応できる内容には、最大限厳格に対応すべきである。これは悪質な利用の抑止につながる。
- 災害時や選挙時に、社会の混乱を防ぐために何らか特別扱いすることは考えられる*が、慎重に検討すべきである。
- 技術の発展が著しい分野であるため、新たなリスクが顕在化した時に迅速に対応を協議できるような、会議体を常時持っておくことが求められる。
- 業界の自主規制・自主的取り組みを後押しするような役割が期待される。事業者との連携・コミュニケーションが必要。
- 透明性の確保を求めていく。その際には、「どういう社会を目指し、そのためにこのような透明性が必要」というビジョンを策定し、エビデンスベースで政策・方針を決定することに資する透明性を求めていくことが重要。
- メディア情報リテラシー教育、AIリテラシー教育を推進する。AIリテラシーには、ポジティブな活用に関するリテラシーと、リスクを正しく認識するためのリテラシーの両方向がある。
- 企業・個人の利用について、気を付けるべきリスクや、適切な活用方法などについて、参照できる情報を提供することは効果があると思われる。ただし、これには継続的なアップデートが求められる。
- 社会全体での効果的な対応を促進するため、ステークホルダー間連携を強化する必要がある。

* 例えば、個人データの第三者提供について、大規模災害等の緊急時に人の生命、身体又は財産の保護のために必要がある場合であって、本人の同意を得ることが困難であるときには、自治会等の個人情報取扱事業者が保有する個人データを本人の同意なく関係者等に提供することは可能とされている（https://www.ppc.go.jp/all_faq_index/faq1-q7-21/）。AI利用についても何らか取り決めをするということは考えられる。

技術による対抗：検証技術の民主化等

- 人のリテラシーで対応できる範囲を超えるため、技術的対抗が求められる。
- 例えば、生成AIによって作成されたか判断できる技術の開発が必要である。民間での研究開発の活性化を促すと同時に、研究支援を通して研究機関・大学での一体的な開発を促進する。重要なのは技術が民主化され、誰もが自由に使える状態になること。
- そのほか、生成物が既存著作権を侵害するものかどうか判断する技術、AIによって生成された有害コンテンツを弾く技術、誹謗中傷や名誉棄損に該当する生成物をブロックする技術など、需要のある技術は多岐にわたると考えられる。
- プラットフォーム事業者において、検証技術を開発し、コンテンツにラベル付けするなど、積極的な対策・実装が求められる。
- 情報環境が汚染されることを防ぐため、ファクトチェック組織やメディア企業は特にそのような対抗技術が利用できる環境にあることが望ましい。

偽動画 9割見破るAI

東大が開発、世界最高水準　米メタも封じ込め急ぐ

2022年6月7日 2:00 [有料会員限定]

保存

「ディープフェイク」と呼ばれる偽動画への対策が進化してきた。東京大学は人工知能（AI）を訓練し、9割前後と世界最高水準の精度で偽動画を見破る手法を開発した。米メタ（旧フェイスブック）や米マイクロソフトなどIT（情報技術）大手も検出ソフトなどを開発し、悪質な偽動画の排除を強化する。政治家などの偽動画がはびこれば社会に混乱をきたすため、封じ込めを急ぐ。

偽画像・動画を見破るにはAIを駆使するしかない

マイクロソフトは微妙な色あせなどをもとに偽動画かどうかを判定するソフトウエアを開発。偽情報対策を推進する活動を通じ影響力の大きな報道機関などに提供する。米アマゾン・ドット・コムも多様な偽動画を見分けられるAIを開発し、論文で公表した。

マイクロソフトは検証技術を開発し、報道機関へ提供
https://www.nikkei.com/article/DGKKZO61483110W2A600C2TEB000/