

令和7年度 インターネット上の偽・誤情報等への対策技術の開発・実証事業

**グローバル・メタアナリシスと国内実証による
対策技術の有効性の研究・調査
成果報告書 簡易版**

2026/3/19

研01_株式会社新領域安全保障研究所

グローバル・メタアナリシスと国内実証による対策技術の有効性の研究・調査

アプローチ
する課題
・目指す姿

- 英語圏偏重の既存研究を相対化し、日本固有の文脈に即した対策を立案するため世界的な誤情報研究動向を明らかにする。その上で、日本の文脈を再現可能なLLMエージェントを用いた「シリコンサンプリング」による安全な検証環境の構築を実施し、コストの低い再現実験の手法を提示する。

研究・
調査区分

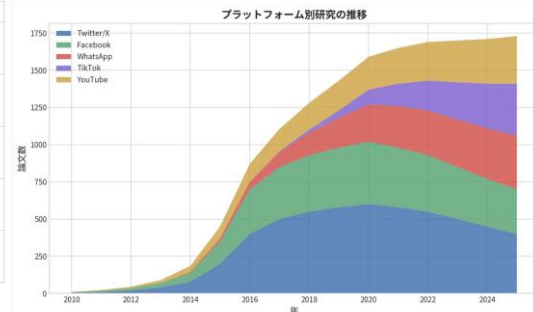
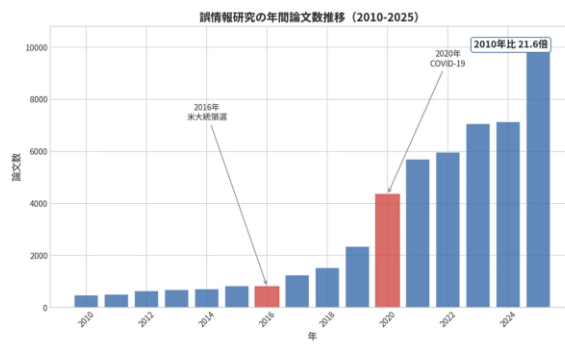
偽・誤情報対策技術に係る研究

実施体制
(下線：研究・
調査主体)

株式会社新領域安全保障研究所

研究および有効性等に関する検証の取組・成果

- 誤情報研究動向の分析：OpenAlexのAPIを用い、誤情報研究が過去15年で爆発的に拡大し、選挙・パンデミック・生成AI等の社会的危機が研究を牽引する「イベント駆動型」であることを整理。また、研究対象の地理・言語・プラットフォーム偏りを可視化し、誤情報研究における欧米の「政治」への関心に対しグローバルサウスでは「健康」への関心が高いことなどを示した。
- 国内実証（シリコンサンプリングの提示・比較検証）：WVS（世界価値観調査）日本データを用い、LLMエージェントによる社会シミュレーションを構築。個別ペルソナ法／集約統計法、GPT系／Gemini系、複数介入条件を組み合わせることで比較した。モデル・手法・入力情報によって推定が揺れること、介入によっては平均では効いて見えても特定層で逆効果を取り得ることを示唆。



研究・調査にあたっての課題・展望

- 研究動向分析はOpenAlex + タイトル中心の抽出のため、非英語圏・ローカル語／領域固有語の取りこぼしがあり得る。→多言語・多概念での研究ウォッチへ拡張。
- シリコンサンプリングは有望だが、現時点では妥当性（現実をどれだけ当てているか）検証のベンチマークが不足。→比較用の人間データ整備と評価枠組みの確立が必要。
- ペルソナ設計（どの情報を与えるか）や、ネットワーク効果・反復曝露・画像／動画など現実環境要因の取り込みが今後の課題。→より実環境に近いシミュレーションへ改善し、日本での誤情報対策に資するエビデンス創出へつなげる。

代表者コメント



新領域安全保障研究所
代表取締役 齋藤孝道
(明治大学
理工学部教授)

世界的な研究動向を理解することは重要ですが、一方海外で成功したとされる対策手法を無批判に日本へ導入することは、予期せぬ副作用を生むリスクを伴います。今回は世界の研究動向を追いかけながらも、海外の知見についてどのように日本に適用するか低コストで検証する手法の提示を目指しました。