

令和6年度 生成AI に起因するインターネット上の偽・誤情報等への対策技術に係る調査の請負 (実証事業)

個人の署名によるコンテンツの真偽表明データベース 成果報告書 概要版

2025年3月3日

株式会社DataSign

実証番号04 株式会社DataSign: R5補正 成果概要

主たる実証の成果

オンライン上のコンテンツの真偽について、著名人やジャーナリストが表明を投稿するシステムであるBoolcheckを開発した。

投稿者の所属等の属性情報を証明する証明書（Verifiable Credential）を組み合わせることで、身元を証明しながらコンテンツの真偽に言及することを実現している。なお、この証明書を管理するためのアプリケーションであるDIW（デジタルアイデンティティウォレット）についても併せて開発を行なった。

生活者は、本実証で開発したブラウザエクステンションを予めインストールしておくことで、オンライン上のコンテンツを閲覧している際にその真偽に関する投稿がBoolcheckに存在するか否かリアルタイムに確認することができ、偽誤情報の見極めに役立てることが可能となる。真偽に関する投稿が存在する場合は、閲覧するコンテンツの一部を遮るようにポップアップを表示することで、生活者に注意を促すことが可能である。このブラウザエクステンションは公開済みであり、容易に利用を開始することができる状態にある。

また、Boolcheckからデータを取得するためのAPIは一般に公開していることから、各種プラットフォームと連携し広告審査等に活用する可能性を備えている。

開発システムに対する実証実験では、前澤友作氏をはじめとする著名人やメディア関係者が参加し、実際にDIWやBoolcheckを利用して真偽情報の表明・確認を行うことで本システムの効果をより高めるための意見を得た。

今後の課題・展望等

＜技術的観点＞

- コンテンツの永続保存とプライバシー保護の両立
オンラインコンテンツはURLで表現されるため、投稿時にスナップショットを取得し改ざん防止の上で永続保存が必要である。しかし、センシティブな内容を含む場合を考慮し、匿名化やアクセス制御、保存期間設定などのプライバシー対策も必須である。

＜非技術的観点＞

- 一般利用者の拡大と認知度向上
対応ブラウザ・デバイスの拡充、利用ガイドの整備、広報活動強化により認知度向上を図る。利用者のフィードバックを反映し、ユーザビリティ向上と社会実装の促進を行う。
- マネタイズの検討と連携強化
公開するBoolcheckのAPIを活用し、広告・情報伝送事業者と連携することで新たなビジネスモデルの構築が考えられる。関連組織との連携を通じた具体的なマネタイズ戦略で、持続可能な運用基盤の確立を目指す。
- 投稿者・認証組織の信頼性向上
現状、投稿者の属性を証明するのみでありその投稿自体が正しいか否かに本システムは関与していない。最終的な偽誤情報の判断は生活者に委ねるモデルであることから、その判断材料として有益な検証可能な情報の拡大を進める必要がある。具体的には投稿者が投稿を行うだけの資格を有しているかや、その審査を行う主体の信頼性の表現についてである。