

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

**デジタル情報空間における多層的意味解析と拡散ダイナミクス解明
プラットフォームの開発・実証
成果報告書 概要版**

2026/3/19

技07_株式会社TDAI Lab

目次

1. 開発・実証における対策技術の開発

1. 開発技術によりアプローチする課題・目指す姿
2. 技術開発の取組・成果

2. 開発・実証における社会実装に向けた取組

1. 社会実装に係る取組・成果
2. 社会実装時のビジネスモデル等
3. 技術開発及び社会実装にあたっての課題・展望
4. 事業の拡大に向けた中長期的な計画

目次

1. 開発・実証における対策技術の開発
 1. 開発技術によりアプローチする課題・目指す姿
 2. 技術開発の取組・成果

2. 開発・実証における社会実装に向けた取組
 1. 社会実装に係る取組・成果
 2. 社会実装時のビジネスモデル等
 3. 技術開発及び社会実装にあたっての課題・展望
 4. 事業の拡大に向けた中長期的な計画

1-1. 開発技術によりアプローチする課題・目指す姿

開発技術によりアプローチする課題

- SNS言論動態の俯瞰的・定量的把握技術は未だ十分に発達しておらず、加えて論理的誤謬や印象操作を伴うため真贋判別が困難なミスリーディング情報は近年増加している。
- さらに今後は生成AIの発達等により、真実と虚偽が巧妙に混在する情報が一層増加することが想定され、従来型の真偽ラベル判定だけでは十分に対応できない。
- このため、偽・誤情報の構造特性や発生・拡散メカニズム、さらには訂正情報が届かない要因までを解明する深層分析アプローチの確立が急務であり、その成果が社会的混乱の抑止や健全な言論空間の維持に直結する。

上記課題を踏まえ目指す姿・ゴール

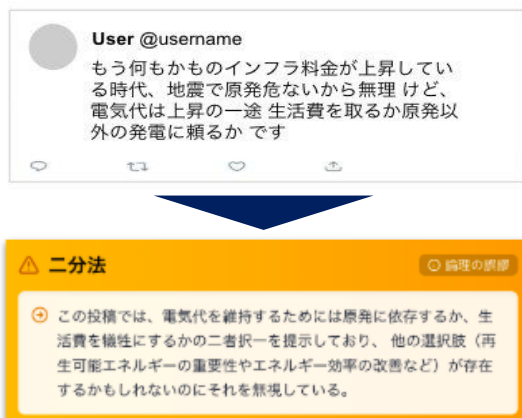
- 本開発・実証では、推論型LLMによる深層的意味的解釈（意見把握、プロパガンダ・論理的誤謬特定、反証構造分析等）と情報拡散分析を統合。これにより、X, YouTube上のミスリーディング情報の
 - ①構築論理・レトリック
 - ②拡散推移・アクター属性
 - ③影響度などを包括的に解明し、偽・誤情報の早期・効果的対策を支援する。
- さらに本事業を通じて、「ミスリーディング情報が可視化され、早期警戒と迅速是正が当たり前機能する予防型の情報空間」を目指す。具体的には、災害時や選挙時など社会的影響の大きい局面で、偽・誤情報の拡散を事前に察知・抑制できる仕組みを社会に根付かせることで、健全な言論空間と市民の批判的思考力を支えることをゴールとする。

1-2. 技術開発の取組・成果

取組及び成果

① 高度な意味的解釈エンジン

- 本エンジンは、SNS上の言説を単なるテキストとして扱うのではなく、言説がどのような論理構造・レトリック・役割構造に基づいて構築されているかを深層的に解釈する技術である。従来の真偽判定や感情分析では捉えきれない、論理の歪み、認知的誘導、正当化の構造を意味的に抽出することを目的として、大規模言語モデル（LLM）を用いた言説解析モジュールを実装した。
- ①二分法(下図)などの感情を煽る14種類の論理的誤謬・プロパガンダ技法、②事実を用いた反論など7種類の反証・論駁構造、③敵と被害者が対立する被害者ナラティブ構造の抽出が可能。



SNSコメントの論理的誤謬分析例

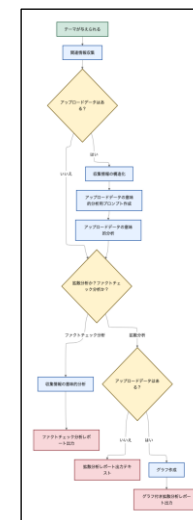
② 情報拡散ダイナミクス分析エンジン

意味的解釈エンジンの出力結果を統合し、分析目的に応じて3種類のレポートを自動作成するLLMエージェント（エージェントック・ワークフロー）を構築。

- A. 拡散分析(アップロードデータあり)：投稿間の関係性や時系列を解析し、拡散推移状況・主要アクター属性を特定。
- B. 拡散分析(アップロードデータなし)：関連収集情報に基づき、注意すべきミスリーディング情報の流通状況を特定。
- C. ファクトチェック分析：意味的解釈エンジンの分析結果を基に、主張と根拠の関係性を整理し、検証観点を提示。

意味的解釈エンジンの精度

意味的解釈エンジン	既存研究の精度	達成精度	目標
A. プロパガンダ技法検出	0.35~0.58	0.65	0.6
B. 反証・論駁構造分析	0.20~0.76	0.79	0.7
C. 被害者ナラティブ解析	0.34~0.38	0.56	-



LLMエージェントック・ワークフロー

1-2. 技術開発の取組・成果

取組及び成果

③ 情報空間分析プラットフォーム

- 本プラットフォームは、意味的解釈エンジンおよび情報拡散ダイナミクス分析エンジンを、専門知識を要さずに実行可能とするウェブアプリケーション型の統合分析基盤である。
- 災害・選挙・社会的関心事など、迅速な分析が求められる場面において、「テーマ入力 → 分析条件選択 → 自動分析 → レポート閲覧」までを一気通貫で提供することを目的として設計。
- 非専門家でも利用しやすいよう、分析観点をあらかじめ提示し選択式で設定可能。専門家向けに分析観点のカスタマイズ機能も実装。
- アップロードデータのコメント数が約4,000件規模の場合、60分以内で分析およびレポート生成が完了。

④ 有効性評価

- 5段階のヒアリング評価の結果、総合スコアは3.01で目標値の4には達しなかったものの、迅速性および理解容易性は比較的高く評価された。
- 一方、信頼性および業務継続性は課題として顕在化。「出力されるコメント例や図の根拠となる情報がわからない」といった信頼性に関するコメントや、データのラベル付けや可視化方法など完全なカスタマイズ性を求める声などが見受けられた。

有用性評価結果



情報空間分析プラットフォーム

評価軸	平均スコア
理解容易性	3.69
業務継続性	2.56
信頼性	2.13
迅速性	4.15
差別性	2.53
総合平均	3.01

目次

1. 開発・実証における対策技術の開発
 1. 開発技術によりアプローチする課題・目指す姿
 2. 技術開発の取組・成果
2. 開発・実証における社会実装に向けた取組
 1. 社会実装に係る取組・成果
 2. 社会実装時のビジネスモデル等
 3. 技術開発及び社会実装にあたっての課題・展望
 4. 事業の拡大に向けた中長期的な計画

2-1. 社会実装に係る取組・成果

取組及び成果

【取組】

- 偽情報対策が業務に直結する報道機関、行政機関、企業に焦点を当て、計9社へのヒアリングを実施。自治体向けサービスプロバイダもヒアリング先に含めることで全国の自治体のニーズを幅広く回収。これにより、実利用者が期待するユースケースや、導入時に懸念される点を把握し、得られたフィードバックをもとに、分析粒度の調整、ユーザーインターフェースの改善、結果提示の可視化方法の拡張といった機能改修を行い、懸念解消と利用現場に即した最適化を進めた。

【成果】

- 実務者ヒアリングを通じ、本技術の分析結果は最終判断を代替するものではなく、根拠情報や分類定義を明示した「判断補助情報」として提供することが社会実装上適切であることを整理し、今後の品質改善および説明性強化に向けた改善方針を得た。
- ユーザーニーズに基づきSNSデータのアップロードを必要としない分析機能（データなし情報拡散分析、ファクトチェック分析）や分析観点カスタマイズ機能等を追加。

ヒアリング結果

観点	実証結果
業務効率	数千件規模のSNSコメントを自動整理・可視化することで、初動調査・論点把握に要する時間を大幅に短縮。
有効な利用場面	リソースの限られる組織や、速報性が求められる初期調査フェーズで特に有効。
分析品質	ナラティブ分類・論点整理により、言説の傾向把握や構造理解が可能。
既存手段との比較	分析内容自体は人による分析で代替可能との認識がある一方、処理時間短縮・作業負荷軽減に明確な価値。
制約・課題	<ul style="list-style-type: none"> ・分類定義の明示、観点の調整など、分析手法の透明性とカスタマイズ性が重要。 ・分析結果の根拠となる情報源や代表言説（代表コメント）の選定方法について説明性が必要。 ・言説抽出・代表コメント選定ロジックを明示し、分析結果の解釈可能性を高める設計が重要。 ・社内規則により、生成AI出力をそのまま根拠・引用として使用することは困難。 ・「最終判断は人が行う」前提で、参考情報として利用する運用設計が必要。

2-1. 社会実装に係る取組・成果

取組及び成果

【成果】

- 社会的インパクトの大きい社会問題など4テーマについて本開発技術を利用した分析記事をメディアへ寄稿。
 - ① JICAがホームタウン撤回 SNS炎上招いた「私は被害者」(2025年9月29日公開)
 - ② 高市氏、SNSでは敵無し「愛国」と「感情」の支持投稿(2025年10月15日公開)
 - ③ 国も動いたクマ騒動 AI動画が描く「もうひとつの世界」(2025年11月18日公開)
 - ④ SNS、自民党嫌いのサナエ推し ファンが支える高市政権(2026年2月16日公開)

- サービスプロバイダーへのAPI連携を通じた650自治体へのサービス提供を実現予定(2026年4月予定)。

2月16日

SNS、自民党嫌いのサナエ推し ファンが支える高市政権

...2025年の参院選から今回の衆院選に至るまでのデータを、東大・鳥海不二夫研究室発AIスタートアップ、TDAI Lab(東京・中央)の最高技術責任者(CTO)、野田昂希氏が分析する。...

SNS



2025年11月18日

国も動いたクマ騒動 AI動画が描く「もうひとつの世界」

...デジタル空間分析の視点から、東大鳥海研究室発AIスタートアップ、TDAI Lab(東京・中央)の最高技術責任者(CTO)野田昂希氏がクマ騒動を読み解く。今年ほどクマが話題になったことはないだ...

政治・渉外 偽情報 SNS



2025年9月29日

JICAがホームタウン撤回 SNS炎上招いた「私は被害者」

...SNSの約1万件のコメントを基に、東大鳥海研究室発AIスタートアップ、TDAI Lab(東京・中央)で最高技術責任者(CTO)をつとめる野田昂希氏と同社データサイエンティストの樋口海斗氏が分析す...

SNS 偽情報

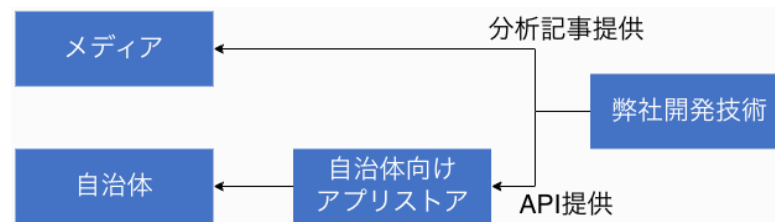


2025年10月15日

高市氏、SNSでは敵無し「愛国」と「感情」の支持投稿

...東大鳥海不二夫研究室発のAIスタートアップTDAI Lab(東京・中央)のCTO野田昂希氏が自民党総裁選をデータから分析する。...

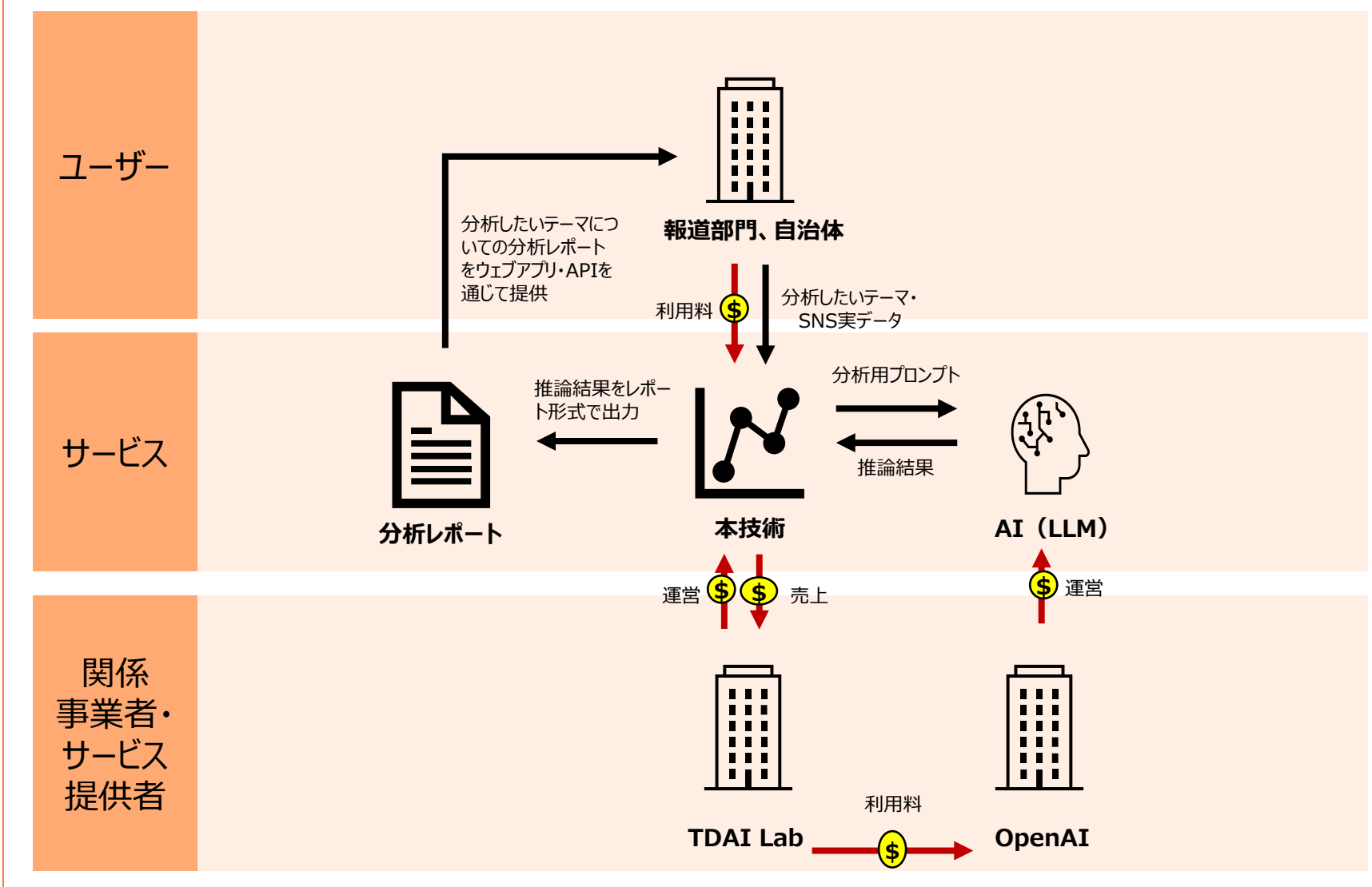
政治・渉外 SNS



メディア・自治体への提供方法

2-2. 社会実装時のビジネスモデル等

社会実装時のビジネスモデル



2-2. 社会実装時のビジネスモデル等

ユーザー・導入先の詳細とそのペインポイント

- 本技術はメディアおよび、SNS監視・分析者を想定ユーザーとして提供される。
- ユーザーごとのペインポイントと本技術による解決方法は以下のとおり。

ユーザーごとのペインポイント

ユーザー	主なペインポイント	本技術による解決方法	ユーザー視点での効用
メディア(新聞社、テレビ局)	<ul style="list-style-type: none"> • SNSの膨大なテキストから論点抽出が困難 • 速報性と分析の深さの両立が難しい • SNS分析を行える専門人材の不足 	<ul style="list-style-type: none"> • 意味的解釈エンジンによりナラティブ・論点を自動整理・可視化 	<ul style="list-style-type: none"> • 初動調査の省力化 • 注目論点を迅速把握し記事の質と速度を両立
SNS監視・分析者(省庁、自治体、サービスプロバイダー)	<ul style="list-style-type: none"> • 選挙時の情報操作や災害デマへの優先順位付けが困難 • SNSはデータ数が多く、確認する人的コストが高い • 従来感情分析などでは偽・誤情報など意味的理解が必要な事象への対応が難しい 	<ul style="list-style-type: none"> • 拡散分析・ファクトチェック補助により重要言説を抽出 • どの情報・言説に優先的に対応すべきか判断を支援 	<ul style="list-style-type: none"> • 対応優先度の明確化 • 初動判断の迅速化・社会不安の抑制

2-3. 技術開発及び社会実装にあたっての課題・展望

技術開発及び社会実装にあたっての今後の課題

- ① **根拠提示と分析の透明性**：本技術は生成AIを用いた分析結果を含むため、メディア・自治体・官公庁においては「分析結果をどこまで信頼し、どのように説明できるか」が導入判断の重要要素となる。特に、生成AI出力をそのまま根拠・引用として使用できない運用制約が存在する。
- ② **実務に即した分析カスタマイズ**：実証を通じ、メディア、自治体、ファクトチェック機関等で分析観点・求められる粒度・最終判断プロセスが大きく異なることが確認された。汎用的な分析設定のみでは、各組織の実務フローに十分適合しないケースが存在する。事業拡大には、一定の標準化と同時に、業務特性に応じた柔軟なカスタマイズ提供が課題となる。
- ③ **導入先組織の理解・リテラシー醸成**：分析内容自体は人手や内製ツールでも代替可能との認識が一定数存在するため、「何が新しいか」ではなく「どの業務負荷をどれだけ下げるか」を明確に示す必要がある。事業拡大においては、省力化・初動速度・再現性といった定量・定性的価値の整理が不可欠である。

上記課題を踏まえた今後の展望

- ① **ユースケース別にパッケージ化した提供モデルの確立**：用途別（調査初期の論点整理／SNS大量データ整理による省力化／ファクトチェック初動支援など）に提供モデルをパッケージ化し、導入時の理解負荷と設定負荷を下げる。その上で、利用者属性（報道／行政／企業）に応じた観点カスタマイズを段階的に提供し、標準化と個別最適の両立を図る。
- ② **「最終判断を代替しない補助ツール」から始める段階的導入設計**：初期段階は「最終判断を代替しない補助ツール」として位置づけ、判断補助情報としての活用に限定した導入設計を行う。利用実績の蓄積に応じて、説明可能性を段階的に高度化することで、信頼性への懸念を抑えつつ継続利用・横展開を促進する。
- ③ **省力化・初動速度・再現性を中心とするKPIに基づく価値の定量化を軸として、国内定着から周辺市場拡張**：ユースケース単位でKPIを定義し、PoC→本導入の評価枠組みとして提示する。加えて、リソース不足だが速報性・網羅性が求められる組織へ市場を拡張し、代替手段が乏しい領域で価値を先に実証することで訴求力を高める。
- 上記①～③の事業拡大を進めつつ、テキスト分析に加え画像・動画を含むマルチモーダル分析へ拡張し、情報流通全体を俯瞰可能な分析基盤としての位置づけを確立する。

2-4. 事業の拡大に向けた中長期的な計画

事業の拡大に向けた中長期的な計画

【基本方針】

実証で得られた知見を基に、(1) ユースケース別にパッケージ化した提供モデルの確立、(2) 「最終判断を代替しない補助ツール」から始める段階的導入設計、(3) 省力化・初動速度・再現性を中心とするKPIに基づく価値の定量化を軸として、国内定着から周辺市場拡張、さらに中長期の機能拡張（マルチモーダル化等）へ段階的に推進する。

【フェーズごとの方針】

フェーズ1（2026年度）：国内市場への定着

- ・商用運用の安定化と、実利用事例・運用知見の体系的蓄積を進める。
- ・「補助ツール」前提の運用ガイド、根拠提示・出力の前提／限界の明示等により、信頼性・説明責任のハードルを低減。

フェーズ2（2027～2028年度）：周辺市場への拡張

- ・利用者属性に応じて、観点・粒度・レポート形式を調整可能なカスタマイズ機構を拡充し、業務フローへの組み込みを容易化。
- ・API連携や既存サービスとの統合を進め、利用チャネルを拡大。

フェーズ3（2029年度～）：機能拡張と横断領域化による持続的成長

- ・テキストに加え画像・動画を含むマルチモーダル分析へ拡張し、情報流通全体を俯瞰可能な分析基盤としての位置づけを確立。
- ・国内での運用実績を踏まえ、海外機関との連携やルール・標準化議論への参画を進める（必要性が高い領域から段階的に検討）。

フェーズ1（2026年度）

- ・国内市場への定着。
- ・商用運用の安定化と、実利用事例・運用知見の体系的蓄積を進める。

フェーズ2（2027-2028年度）

- ・周辺市場への拡張。
- ・業務特性に合わせたカスタマイズとAPI／業務ツール連携を拡充し、周辺領域へ横展開しながら処理性能と説明可能性を強化する。

フェーズ3（2029年度～）

- ・機能拡張と横断領域化による持続的成長。
- ・マルチモーダル分析へ拡張して横断的な分析基盤として位置づけを確立し、国内実績を土台に国際連携・標準化も段階的に進める。