

**経済安全保障重要技術育成プログラム／偽情報分析に係る技術の開発
／偽情報の検知・評価・システム化に関する研究開発
(中間評価) 2024年度～2025年度 2年間**

プロジェクト報告資料 (公開版)

2025年12月22日

偽情報分析に係る技術の開発

(60億円を超えない範囲／4年)

背景

- サイバー空間の「公共空間化」が進展する中、**事実と異なる加工**や情報の**組み合わせ等**によって、対象の印象やそれに基づく意思決定等を**発信者の意図する方向に導くような情報が存在**。
- このため、文書、画像、動画等の**情報の加工の有無の見極め**、情報発信時点の不自然性や**事実の歪曲等を、エビデンスを踏まえて、一体的、総合的に評価することが必要**。
- 加えて、SNS等のオープンなサイバー空間における偽情報への対処の判断には、社会的影響等、その情報の持つ脅威の評価が重要。
- 今後、**生成AI等の高度化に伴い、サイバー空間における偽情報の増加や複雑化が想定される中**、偽情報の検知からその脅威評価までを一貫して行う、偽情報への対処の判断を支援するシステムが求められている。

想定される利用ニーズ

- SNS等のインターネットにおける偽情報は国民生活に影響する可能性があり、マスメディアによる偽情報の打ち消し報道を行う等、対応を判断する必要がある。
- 例えば、特定企業（組織）や製品に関する偽情報は、その企業や製品のイメージダウンにつながり、企業の操業に関して、多大な影響を及ぼす可能性がある。また、災害情報、国内外の事件等についても、意図的に情報が作成された場合、国民生活や経済活動に大きな影響を与えるおそれがある。
- 本構想において開発する技術は、こうした偽情報への対応を迅速に行うための支援のツールとしての活用が考えられる。

研究開発の内容

(1) 要件定義

偽情報分析に係る最新の技術動向や、公的機関や民間企業によるユースケースを調査し、対象とするユースケースを特定するとともに、(2)~(4)の要件定義を行う。加えて、国際会議等において情報収集を行い、海外も含め可能な限り最新の動向を把握する。

(2) 偽情報検知技術

情報を構成する文章、画像、動画等における加工を総合的に分析する技術を開発する。加えて、情報に付随するメタ情報（時刻情報や地域情報等）に対する客観的な事実の紐づけや発信者の正当性等によるエンドースメントを作成し、情報の真偽性判定に資する要素技術を開発する。

(3) 偽情報評価技術

情報の受け取り手の反応や発信元の情報等を収集・分析することにより、当該情報に対する社会的影響等を定量化及び可視化し、脅威評価に資する指標・要素技術を開発する。

(4) 偽情報検知/評価システム化技術

SNS等のオープンソースを対象とした情報収集方式及び偽情報分析対象の情報の絞り込み手法、(2)及び(3)の技術との接続手法の検討を行い、偽情報検知/評価システムを構築する。

想定スケジュール

テーマ	2024年	2025年	2026年	2027年
(1)	ユースケース、技術動向の調査、要件定義の検討			
(2)	偽情報検知手法の確立	中間評価（ステージゲート）	開発・検証	事後評価
(3)	偽情報評価手法の確立	中間評価（ステージゲート）	開発・検証	事後評価
(4)	偽情報検知/評価システムの仕様確定	中間評価（ステージゲート）	システムの構築・検証	事後評価

本事業の目的・概要

- 偽情報がインターネット上に流布し、社会への影響が増加する中で、安定的で自律的な経済活動を維持するために偽情報対策の強化が求められている
- 本事業では、偽情報対策に資する要素技術を開発するとともに、それらをインテグレートし、個々の要素技術による分析結果やエビデンスに基づいて総合的に真偽を判定できる偽情報対策システムを構築する。さらに、このシステムが将来の社会実装を見据えた基盤の位置づけとなることを目指す。

① 要件定義、ユースケース検討

偽情報分析に係る最新の技術動向や、ユースケースの調査と必要機能の抽出

② 偽情報検知技術

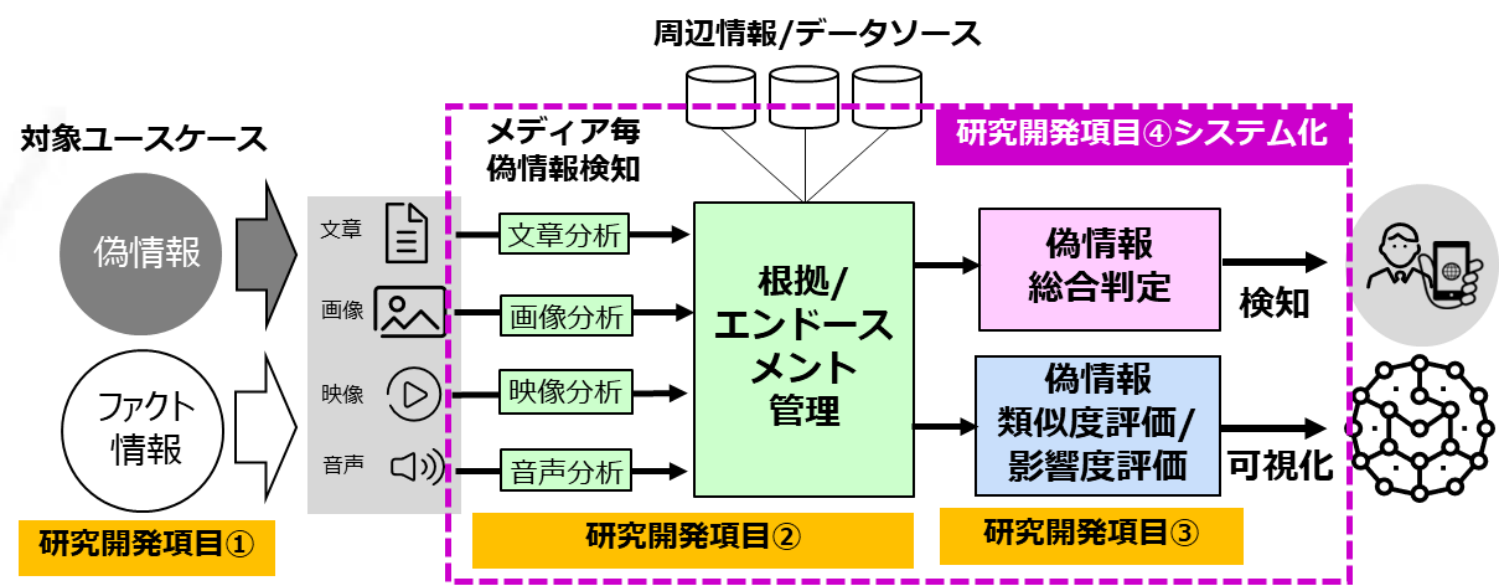
SNSの投稿内容などから文章、画像、映像、音声メディアをメディアごとに分解して抽出し、内容を分析し、その結果を根拠として利用するとともに、作為性の判定を行う技術を開発。また、インターネットの情報に対する様々な根拠を、グラフ構造化して管理するシステムを構築

③ 影響度評価技術

SNSデータからメッセージの情報源、情報内容、社会的文脈に着目し、LLMを拡張して偽情報評価用AIモデルを構築することにより、過去の偽情報との類似度や拡散速度などの偽情報の特徴を分析し、拡散規模や社会的な影響度などの指標を評価する技術を開発

④ 偽情報検知/評価システム化技術

②と③を統合するオーケストレーション機能とともに、根拠情報の分析によりLLM等のAIを活用して総合的に真偽判定を行う技術、およびユーザースタディに基づく検証を通じて、ユーザーがより正確に対象情報の真偽を判断できるインターフェースを開発



研究開発項目

①要件定義

1-1 民需・官需向けユースケースと要件定義(富士通)

最新の動向、想定ユースケースを分析し、開発システムに必要な機能を抽出・整理

②偽情報検知(メディア毎の情報分析)

2-1 ディープフェイク検知(NII)

画像/映像/音声メディアのディープフェイクに関する真贋判定に加え、ディープフェイクの改ざん箇所、ディープフェイク生成手法を判定の確信度を含める形で推定し、これらを根拠情報として出力する研究開発を実施

2-2 メディア理解(NEC)

画像/映像/音声メディアの内容を抽出する技術(どのような出来事が発生しているか/どのような会話がなされているか)を開発

2-3 位置検証(富士通)

根拠情報の候補となる位置情報の正当性を保証する技術。発信元デバイスの位置検証技術を開発

2-4 根拠/エンドースメント管理(慶應大、富士通)

インターネット情報に対する様々な分析結果を、構造化された情報として蓄積し、高次な分析を可能とする情報管理基盤を構築

③偽情報評価(影響度評価)(東京科学大、東京大、会津大)

偽情報特有の特徴量との類似性を評価し、拡散規模や影響度を評価。真偽検証された過去事例の多様な特徴量に基づくAIモデルを構築

④システム化

4-1-1総合真偽判定(富士通)

各研究開発項目における分析結果や根拠情報に基づき、大規模言語モデル(LLM)によって、真偽の根拠を分析し、判定結果とともに自然言語(文章)としてユーザに根拠を説明する技術を開発

4-1-2 システム化/オーケストレーション(富士通)

各研究項目の技術を管理基盤をベースに統合・体系化し、偽情報対策システムを構築

4-1-3 ユーザーインターフェース(富士通、名工大)

総合真偽判定や影響度評価の結果をユーザーに分り易く提示。真偽判定の根拠説明コンテンツをLLMで生成する手法、ユーザーの心理的要因を考慮した認知科学に基づくUI/情報提供手法を開発

4-2 日本語LLM(富士通)

真偽判定や影響度評価に高い精度を達成する技術、高速な推論技術、ハルシネーション(LLMの誤出力)の抑制技術、を開発し、これらをベースに透明性が高く検査可能であり、閉環境でのセキュア運用が可能なLLMを開発

4-3 IoT情報収集(大阪大)

根拠となるメディア情報(IoTカメラ映像等)の収集技術。IoTセンサーから網羅的に情報を取得することができない場合に、情報収集依頼に含まれるエリア情報や依頼内容にもとづいて真偽分析を合理化するために必要な情報源を選別し、能動的に取得する技術を開発

研究計画

• 2年目（2025年まで）研究計画

- 対象分野をフォーカスし、メディア毎の情報分析(画像/映像/音声)から総合真偽判定/影響度評価までの各技術を統合した偽情報対策プラットフォームの初版を構築し、組織ユーザーと連携してユーザ評価やシステム評価により有効性を検証

• 4年目（2027年まで）研究計画

- 各技術の改良や拡充を行うとともに、対象分野を拡大し、組織ユーザーが実業務で活用できるシステムの構築を目指す。偽・誤情報に関する状況等を考慮し研究開発を進める

＜評価項目 1＞ 研究開発ビジョン及び研究開発構想の実現に向けた研究開発課題の達成目標や内容の妥当性

- (1) 達成目標の妥当性
- (2) 知的財産・標準化戦略

達成目標の妥当性

- 本事業の研究開発構想では、各研究開発項目において以下の達成目標が設定
 - 研究開発項目 1.要件定義：ユースケースの特定、必要な機能の抽出・整理
 - 研究開発項目 2.偽情報検知技術：情報の作為性、真偽性の判定に資する要素技術を確立
 - 研究開発項目 3.偽情報評価技術：偽情報の社会的影響等の脅威評価に資する要素技術を確立
 - 研究開発項目 4.偽情報検知/評価システム化技術：システム仕様を確定
- 【設定目標】上記の達成目標に対して、様々な分析結果や収集した根拠情報を統合し、これらの整合性や矛盾を分析して真偽の判定を支援し、社会への影響度を評価する偽情報対策システムの詳細機能仕様を確定
- 【妥当性】システムの要件定義を実施、ユースケースを特定し、ユーザーニーズを仕様に反映
 - 要件定義では、偽情報対策を行う必要性が高い分野として、実事例をもとに発生度/拡散度/影響度の観点から偽情報を評価し、「自然災害」「政治」「医療・健康」を偽情報の影響が大きい分野として選定
 - 国内外の技術動向、市場動向、社会情勢等を調査・分析するとともに、公的機関や民間組織（ファクトチェック団体等）から要望・ニーズのヒアリングを実施し、これを踏まえてシステム開発機能を追加・更新
- 達成度の判定に関しては、機能要件として目標を明確化している

研究開発成果の権利化／秘匿化／公表等の取り扱い方針

- 本事業では、研究開発成果の権利化／秘匿化／公表等の取扱い方針についても下記のように適切に定めている。
 - 論文を始めとして基本的には公開する方針とする。なお、論文レベルでの公開であり、それ以外のノウハウなどは非公開とする
 - 経済安全保障観点、他国との優位性・差別化の源泉となりうると判断する場合など、よりコアな部分がでてきた場合は権利化する方針とする

＜評価項目 2＞ 研究開発課題の達成目標に向けた進捗状況

- (1) 研究開発課題の達成目標に向けた進捗状況（国内外との比較を含む）
- (2) 今後の見通し（多様な分野における実現可能性含む）
- (3) 指定基金協議会において合意された内容の進捗状況

中間評価目標の達成状況

□ 研究開発項目 1.要件定義

項目	中間評価実施年度まで (2024年度) の目標	進捗・成果
1-1 要件定義	最新の技術動向や公的機関や民間企業によるユースケースを踏まえ、対象とするユースケースを特定するとともに、偽情報検知/評価システムに必要な機能の抽出および整理を行う。	偽情報による社会的影響が大きい重点分野を選定するとともに、対策技術・製品・規格の動向調査、分野毎のユースケース検討を実施。関係省庁/有識者との意見交換も参考に機能要件を抽出し、本事業で研究開発する偽情報対策システムの構成を導出。

□ 研究開発項目 2.偽情報検知技術

項目	中間評価実施年度 (2025年度) の目標	進捗・成果
2-1 ディープフェイク検知	本事業の想定ドメインにおいてSYNTHETIQ VISIONによるディープフェイク検知モデルの真贋性能を最適化する。顔映像に加えて、自然画像にも対応できるよう対象となるメディアを拡張する。音声メディアに対しては、ディープフェイク検知可能な基本手法を検討する。	大規模音声データで事前学習した基盤モデルを追加学習し、未知の生成手法に対して高精度な検知を実現。顔画像と災害画像等の一般画像の両方に対応。長時間音声ファイル内のディープフェイク音声領域を特定する技術も構築。
2-2 メディア理解	メディアデータからの情報抽出・結合技術を確認する。これら技術を実装したプロトタイプを開発する。開発したプロトタイプを用いて実現性を検証し、社会実装に向けた課題を調査する。	メディアデータからその要約やデータに含まれる人物・物体の属性、追跡情報及び音声・テロップ情報を抽出する技術の既存技術評価及び抽出機能の試作を完了。また抽出情報から偽情報判定に有効なグラフ情報を構築する技術の試作も完了。評価を通じ、検索内容に応じた動的な分析や、複数のモーダル間での情報の補完などの課題も抽出。
2-3 発信位置検証	位置検証機能 (Robust Localization機能、Image Geolocation機能) を偽情報対策システムと結合し、総合真偽判定に活用するとともに、有効性評価を行う。	Robust Localization技術でネットワーク遅延から発信位置を推定し、Agent評価で高信頼化を実現。Image Geolocation技術では、部分画像による整合性検証で画像位置推定の誤推定削減方式を設計。偽情報対策システムとの結合完了。
2-4 根拠・エンドースメント管理基盤	メディア分析基盤の設計、PoC実装および限定的検証を行う。国際標準化の調査を行う。	様々な分析結果やその過程を蓄積し、総合的な判断を支援するメディア分析基盤のアーキテクチャとして、主体(Subject)/述語(Predicate)/対象(Object)/時刻(Time)/位置情報(Location)の形式(SPOTL)で構造化する基本設計を実施。利用シナリオの検討と、部分的なプロトタイプ実装を元に課題を整理。関連する国際的な標準に向けた調査・活動を実施。

中間評価目標の達成状況

研究開発項目 3. 偽情報評価技術

項目	中間評価実施年度（2025年度）の目標	進捗・成果
3-1 影響度評価	偽情報評価システム全体の設計を行う。SNSの大規模ストリームから情報源、情報内容、社会的文脈に関わる情報抽出を可能とする技術を確認し、日本語全量規模の処理を毎日に実行可能とする技術を確認する。後段の分析に使えるようにするためのデータ前処理の技術を確認する。埋め込み表現の作成および可視化手法を開発する。	偽情報の拡散ネットワークに基づく影響度の計測とシステム構築を実施。tweet_idを入力として、影響度に関する指標を返すシステム・APIを実装。SNS(X)全日本語ポストから主要な情報拡散木をリアルタイムに抽出し、ポスト数、深さ、速度等の特徴量を用いてランキングするシステムを構築。Xの投稿内容に基づく埋め込みを可視化するシステムを開発。

研究開発項目 4. 偽情報検知/評価システム化技術

項目	中間評価実施年度（2025年度）の目標	進捗・成果
4-1-1 総合真偽判定	研究開発項目2と3.を統合し、偽情報検知システムの総合真偽判定アルゴリズムへ組み込み、複数の分析モジュールを活用することで、特定ユースケースにおける判定精度を評価する。	SNS投稿の文脈を抽出し、テキストと画像の文脈の一致を判定するとともに、テキスト真偽判定やテキストと画像の関係性分類技術と組み合わせることで総合的な真偽を判定するアルゴリズムを考案・実装。テキスト真偽判定で最高精度88%達成(Knowledge-based QAタスク)。
4-1-2 オークストレーション機能	各研究開発項目の技術をインテグレートした偽情報対策システムの第一版を開発し、複数のユースケースにおける評価を行うとともに、複数ユーザーによるトライアル評価を行う。	ワークフロー形式で各分析技術を柔軟に組み合わせ、分析を実行するためのオーケストレーション機能の基本設計を実施。マイクロサービスアーキテクチャのような疎結合な形で分析モジュールを実装し、Web API越しにデータ連携を行う構成をもつ偽情報対策システムの第一版を開発。民間組織・公的機関の複数ユーザーによるトライアル評価を実施。
4-1-3 ユーザーインターフェース	真偽判定結果の提示の仕方とユーザーの判断の相互作用を検証し、判定結果を効果的にユーザーに伝達するインターフェース・パイロット版を設計・実装する。結果概要を示すアイコンから根拠説明へのスムーズな接続を実現するアイコンの考案とその効果検証を行う。	従来のファクトチェック説明の課題を分析し、偽情報の判定結果を根拠の種類ごとにモジュール化・階層化することで、ユーザーが必要な根拠に最短距離でアクセスできる「モジュール型ファクトチェック説明」UIの基本設計を構築。アイコンのデザイン、効果検証、再デザイン、再効果検証を実施し、説明カテゴリーの認知に対応したアイコンセットを開発。
4-2 偽情報対策特化の日本語LLM	各種データ（記事、SNS等）を用いて偽情報対策特化型LLMを強化する。量子化・蒸留技術による推論高速化を実現する。LLM内部状態の解析によるハルシネーション検知・緩和技術を確認する。	報道領域とSNS領域のコーパスを利用した継続事前学習によりLLMの知識を強化、さらに学習データ合成により偽情報対策タスクの処理能力を強化し、汎用LLMの精度を越える特化型LLMを開発。自己整合性に基づくハルシネーション検知技術を開発LLMに適用、強化学習により中間思考の妥当性を高めることによるハルシネーション緩和技術を開発。
4-3 情報収集手法	IoTセンサー情報を対象とした能動推論型情報収集技術を偽情報対策システムに組み込み、総合真偽判定との連携が可能であることを確認する。	真偽分析の合理化に必要な情報源の選別・能動的な取得手法を考案。数値評価によって、より少ない情報収集回数で真偽判別が可能であることを示した。広域マップへの対応のための計算量削減を図り、複雑な氾濫シナリオに対応出来ることを確認。

アウトカム達成の見通し

研究開発構想におけるアウトカム目標

- 本事業で開発した偽情報検知技術、偽情報評価技術及びこれらの技術を統合した偽情報検知/評価システム化技術を通じて、様々な社会実装に向けたシステム構築が可能。本研究開発の成果は、多種多様な企業/業種を始めとする社会経済基盤に対して、過度な混乱及び影響を及ぼす可能性がある偽情報発信に対して、迅速な意思決定及び対処の支援が可能
 - 民生分野：SNS等のインターネットにおける偽情報は国民生活への影響等が懸念され、マスメディアによる偽情報の打ち消し報道等の対応・判断が必要であり、また、特定企業（組織）や製品に関する偽情報においては、特定企業（組織）や製品のイメージダウンにつながり、企業の操業に関して、多大な影響を及ぼす可能性があるため、速やかな対応が必要。本事業で開発した技術は、こうした対応を迅速に行うための支援のツールとしての活用が可能
 - 公的分野：例えば災害情報、国内外の事件等について、意図的に作成され、国民生活や経済活動に大きな影響を与えるおそれのある情報（偽情報）の迅速な発見、対処等の判断の支援が可能

• アウトカム達成の見通しの状況

- それぞれの研究開発項目における技術開発とともに、それらを統合した偽情報対策システムの初期評価版を構築
- 指定基金協議会に参画する公的機関、ならびに民間組織（ファクトチェック団体、メディア関連団体）からの要望・ニーズの抽出を実施し、偽情報対策システムの開発機能を見直すとともに、初期評価版システムに対するトライアル・実証実験を開始。民生分野と公的分野における実ユースケースに基づき、技術・システムの深化を加速させていく

• 多様な分野における研究成果活用の状況

- 民生分野：ファクトチェック機関、メディア領域、など
- 公的分野：指定基金協議会に参画する公的機関等

指定基金協議会で合意された内容

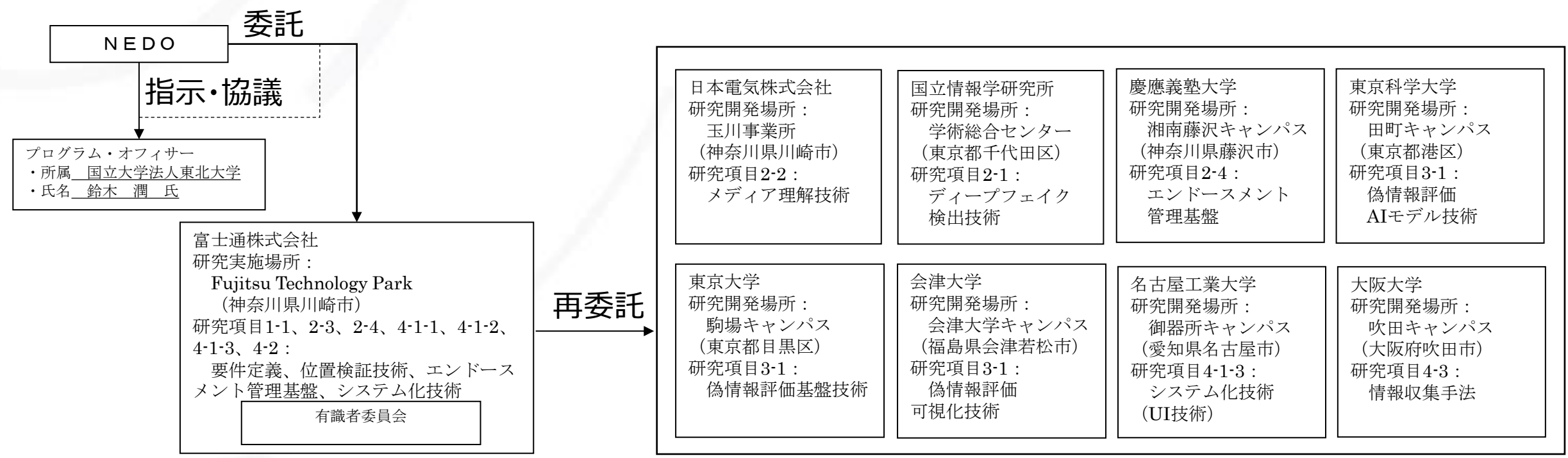
第1回・第2回の指定基金協議会において、活動に反映すべき特段の合意事項等はない。

＜評価項目 3＞ マネジメント

- (1) 実施体制
- (2) 研究資金の効果的、効率的な活用
- (3) 国民との科学・技術対話に関する取組

実施体制

- 本事業内容である偽情報の検知から根拠収集、分析、評価までを統合的に行う偽情報対策システムの開発において、実施者は国内屈指のアカデミアや企業による9組織から構成されるオールジャパン体制を構築しており、事業遂行に十分な技術力及び事業能力を有している。
- 偽情報対策システムの構築に向けて、各組織の担当領域を明確化している。9者による全体会議を定期的を開催することに加えて、組織間の個別会議も適宜実施することで密な連携を図り、全体の方向性をアラインさせて進捗のスピード化を行っている。また、各組織から文書による月例進捗報告を行うことで全体管理を実施。



事業遂行の技術力・事業能力

・ 富士通株式会社

- ・ SNS投稿をはじめとするインターネット上の情報の真偽を判定する技術や、エンタープライズ向け日本語特化型LLM「Takane」の研究開発実績、およびシステムインテグレーションの実績を有する。

・ 国立情報学研究所（ディープフェイク検知）

- ・ ディープフェイク検知、特にディープフェイク音声と顔映像の検知に関する研究に取り組んでいる。ディープフェイク音声の検知性能のベンチマークを行うデータベース公開、コンペティションも定期的に開催する等の実績を有する。

・ 日本電気株式会社(メディア理解)

- ・ 映像・画像・音声を横断的に解析できる高度なメディア解析技術を有し、オブジェクト認識、関係抽出、シーン理解、関係・イベントのグラフ構造化までを実現している。

・ 慶應義塾大学（根拠/エンドースメント管理）

- ・ データをどのように扱うかという視点で「データアーキテクチャ」研究を推進。データの検証可能性を様々な形で検討。Trusted Webを含め、デジタル署名技術活用についての実績を有する。

・ 東京科学大学（偽情報評価モデル）

- ・ 偽情報が拡散するネットワークに関して構造的特徴や拡散速度を推定し、偽情報の影響度を評価するためのDBシステムの構築、及び大規模データ分析を行なった実績を有する。

・ 東京大学（偽情報評価基盤）

- ・ 大規模Webデータの収集・蓄積・分析の実績（1999～ Web 600億URL規模のアーカイブ構築）や、ソーシャルメディアデータの収集・蓄積・分析の実績（2006～ Blog 30億記事規模、2011～ Twitter 500億ツイート規模、2020～ Twitter, X 日本語全量分析）を有する。

・ 会津大学（偽情報評価可視化）

- ・ 大規模ウェブデータの可視化システムの開発や、ソーシャルメディアにおける大規模ネットワーク構造のクラスター分析・時系列分析の実績を有する。

・ 名古屋工業大学（ユーザーインターフェース）

- ・ 訂正の効果を妨げる心理バイアスなどの偽誤情報の訂正に関する心理学的知見（関連論文複数）や、心理学的知見にもとづくユーザーインターフェースの設計、ユーザー実験によるインターフェースの効果検証のためのスキルを有する。

・ 大阪大学（情報収集手法）

- ・ ベイズ手法をベースとする状態推論技術の研究開発実績、マイクロサービスのAPI連携／プロトコル接続による情報通信システムの運用実績を有する。

研究資金の効果的、効率的な活用

- スケジュールや研究開発費に関して、想定成果をみすえた費用対効果の観点から再委託先も含めて精査したうえで研究開発費を活用している。
- ソフトウェア製作においてAI活用を積極的に進めることにより、研究開発の加速と研究資金の効率化を両立させることに成功している。
- GPUリソース価格が高騰する中で、オンプレ・クラウド等の様々な提供形態から最適な選択を行い、購入費用を圧縮している。

成果発表実績

評価基準：展示会出展、学会発表、論文等の対外的な成果発表を適切に行っているか。（又は行う予定であるか。）

・ 展示会出展，講演等

※2025.12.09現在

イベント名	タイトル	実施日
外国特派員協会 Press Conference	Confronting Deepfakes: Japan's Next Steps in the Age of AI	2025/6/24
LLMの安全性シンポジウム2025	フェイクニュースの社会課題を解決する偽情報対策技術	2025/7/23
Kプロシンポジウム	技術シンポジウム「フェイクニュースの社会課題を解決する偽情報対策技術の最前線」	2025/7/31
AI・人工知能EXPO2025秋	偽情報分析技術～フェイクニュースの社会課題を解決する偽情報対策技術・プラットフォーム～	2025/10/8-10
NHK 「明日をまもるナビ」	(177) 災害時 SNSに広がるフェイク情報	2025/11/23

・ 学会発表（発表済：21件，発表予定：3件）

※査読付き国際会議案件. 2025.12.09現在

学会名	タイトル	発表年月
The 63rd Annual Meeting of the Association for Computational Linguistics	Neuron Empirical Gradient: Discovering and Quantifying Neurons' Global Linear Controllability	2025/06
International Conference on Web Services (ICWS) 2025	A Hybrid Approach Combining LLMs and Web-Based Information for Automated Fact-Checking	2025/07
IEEE International Conference on Control, Decision and Information Technologies (CoDIT)	Evaluation of Person Search System Using Multidirectional Cameras and a Drone Based on Active Inference	2025/07
IEEE International Joint Conference on Biometrics (IJCB) 2025	LENS-DF: Deepfake Detection and Temporal Localization for Long-Form Noisy Speech	2025/09
IEEE ASRU Workshop 2025	Post-training for Deepfake Speech Detection	2025/12

付録：学会発表リスト（2025/12/09現在）

評価基準：展示会出展、学会発表、論文等の対外的な成果発表を適切に行っているか。（又は行う予定であるか。）

※太字は査読付き

学会名	タイトル	発表者所属	発表年月
第4回計算社会科学会大会(CSSJ2025)	keyATMを用いた2024年東京都知事選の主要候補者に関するXのデータの分析	東大	2025/02
言語処理学会第31回年次大会(NLP2025)	フレーム意味論に基づく言及先情報を含むSNS投稿の事実忠実度のアノテーション	東大	2025/03
第17回データ工学と情報マネジメントに関するフォーラム (DEIM2025)	部分画像を用いた地理的位置推定の精度向上手法	富士通	2025/02
第17回データ工学と情報マネジメントに関するフォーラム (DEIM2025)	画像を再利用した偽情報の検出に対する情報補完を用いた精度向上手法	富士通	2025/02
第31回 画像センシングシンポジウム(SSII2025)	Refining Image Geolocation with Coarse-to-Fine Multi-Task Learning	富士通	2025/05
人工知能学会全国大会 JSAI 2025投稿	2024 年東京都知事選におけるXのデータを用いた主要候補者の比較分析	東大	2025/05
International Conference on Web Services (ICWS) 2025	A Hybrid Approach Combining LLMs and Web-Based Information for Automated Fact-Checking	富士通	2025/07
The 63rd Annual Meeting of the Association for Computational Linguistics	Neuron Empirical Gradient: Discovering and Quantifying Neurons' Global Linear Controllability	東大	2025/07
第16回ソーシャルコンピューティングシンポジウム (SoC2025)	マイクロブログにおける情報共有構造の変遷に基づく分極化現象の経時分析	東大	2025/06
電子情報通信学会ネットワークシステム研究会	偽情報判定のための能動推論を用いたIoTセンサ情報収集手法の提案と評価	阪大	2025/07
IEEE International Conference on Control, Decision and Information Technologies(CoDIT)	Evaluation of Person Search System Using Multidirectional Cameras and a Drone Based on Active Inference	阪大	2025/07
電子情報通信学会サービスコンピューティング研究会	マルチモーダルファクトチェックのための包括的なデータセットの提案と評価	富士通	2025/07
IEEE ASRU Workshop 2025	Post-training for Deepfake Speech Detection	NII	2025/12
2025 MIT Polarization Workshop	Cascades as a Lens: A Systemic Approach to Misinformation on X	東京科学大	2025/12
情報処理学会 データベースとデータサイエンス研究会(第181回DBS・第160回IFAT合同研究発表会)	モダリティごとの類似度判定を用いた偽情報検出手法の検討	富士通	2025/09
情報処理学会 データベースとデータサイエンス研究会(第181回DBS・第160回IFAT合同研究発表会)	短文投稿の真偽判定に向けたコンテキスト制御型情報補完手法	富士通	2025/09
IEEE International Joint Conference on Biometrics (IJCB) 2025	LENS-DF: Deepfake Detection and Temporal Localization for Long-Form Noisy Speech	NII	2025/09
情報処理学会 データベースとデータサイエンス研究会(第181回DBS・第160回IFAT合同研究発表会)	SNS投稿内容に関連する災害情報を特定するための5W1H抽出手法	富士通	2025/09
第28回情報論的学習理論ワークショップ(IBIS2025)	LLM製知識グラフに対するEntity Alignmentの有効性評価	NEC	2025/11
IEEE ASRU Multimedia 2025, 1M-Deepfakes Detection Challenge	Detecting the Undetectable: Robust Defense Strategies Against Audio Deepfakes	NII	2025/10
CYDEF 2025	Measuring Influence in Information Diffusion on X Using a Cascade System	東京科学大	2025/12