



AIでフェイクを見破れ! 情報の信頼性を守る技術の最前線

AI

ファクトチェック

フェイクニュース対策

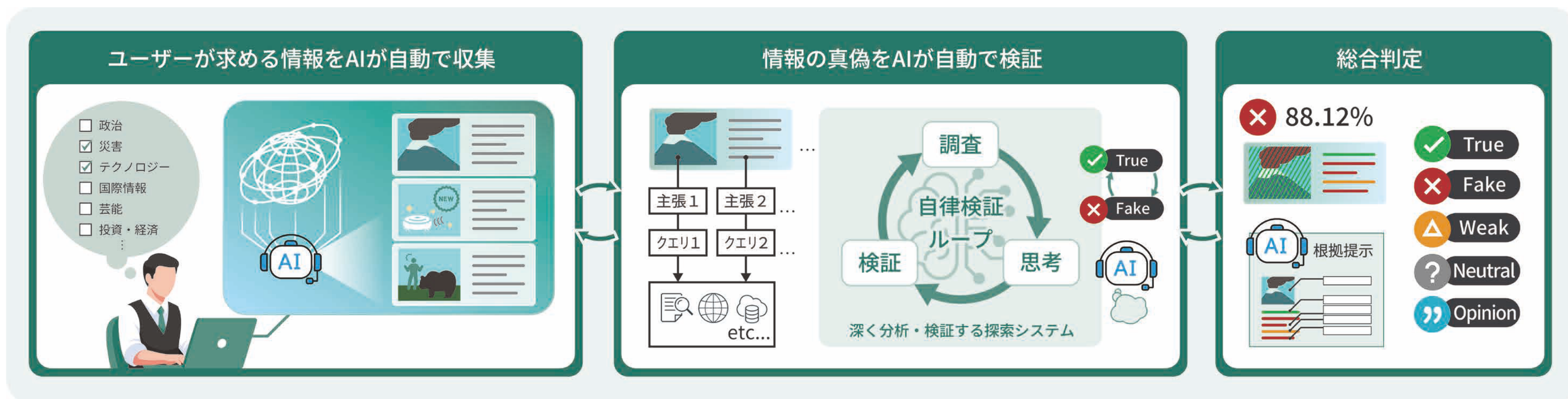
NABLAS (株) [東大発ベンチャー]

プロジェクト概要

- 生成AIの普及に伴い、ハルシネーション（AIによる誤情報）や偽情報が拡散するリスクが深刻化。
- 本事業では、情報の信頼性を客観的に評価する、ファクトチェックに特化した報酬モデルと自律AIエージェントを開発。
- 世界中のフェイクニュースに対応し、日本語環境にも利用できる高精度な検証技術の研究を実施。

展示物紹介

- 実証実験を経て進化した、AIエージェントによるファクトチェック体験デモを展示。
- ニュースやSNSの文章を入力すると、AIが信頼できる情報源を自動で集めて真偽を評価し、レポート化。
- 文章をパーツ分けし、パーツごとに主張をフェイク判定するだけでなく「根拠不十分」「発信者の意見・感想」なども識別して可視化。
- 専門知識がなくても、論理的な判定を元に正しい意思決定に繋がられるプロセスを体験可能。



AIエージェントによるファクトチェック工程

今後の展望

- 開発成果は 弊社プロダクト「KeiganAI」にて、**2026年度中にリリース予定**。
- 今後は、世界中の情報を自動収集/分析/検証するシステムへと開発を拡張。
- 事象ごとに起きうる偽情報のパターンを予測し、事前に対策・注意喚起できる未来の情報インフラを構築。メディアや官公庁、企業が偽情報に惑わされない安心・安全な社会の実現を目指す。



社会実装に向けたステップ

事業化フェーズ

フェーズ	基礎研究	応用・PoC	実証	事業化準備	商用化
TRL	TRL1-2	TRL3-4	TRL5-6	TRL7-8	TRL9

プロジェクト実施期間	2025年度(終了事業)
NEDOプロジェクト名	ポスト5G情報通信システム基盤強化研究開発事業／ポスト5G情報通信システムの開発／Factcheck RMとエージェントの開発研究
お問い合わせ先	NABLAS株式会社 HP : https://www.nablas.com/contact

