



第 959 号 ミニかわら版

令和 7 年 12 月 15 日

(この資料は全部お読みいただいても60秒です)

ヤマダ総合公認会計士事務所
代表 山田良平

〒124-0012

東京都葛飾区立石 1-12-11 ヤマダビル

TEL:03-3694-6091

FAX:03-3691-6680

総務省 2026年春をめどに評価AIの開発開始

総務相は先月18日の閣議後の記者会見で、総務省が生成AIの信頼性を評価するための基盤システムの開発に着手すると明らかにしました。この取り組みは、AIが差別的な表現や偏った回答をしていないかを確認し、評価結果を指標として公表することを目的としており、利用者や企業がAIを選択する際の判断材料を提供し、安全で公平なAIの社会実装を促進する狙いがあります。

開発は総務省所管の国立研究開発法人・情報通信研究機構(NICT)が担い、2026年度中に試作モデルの提供を目指すとしています。来春にも開発を開始し、複数の評価用AIが質問を自動生成し、対象AIの回答を分析する仕組みを導入します。生成AIの回答を評価する基準としては、①差別的な表現やプライバシー侵害の有無、②犯罪につながる情報の排除、③誤情報や根拠の欠如の防止、④バランスの取れた内容、⑤日本文化への忠実性、⑥人を欺かないこと、⑦未知のリスクへの対応力の7項目が想定されています。これらは、G7が合意した国際ルール「広島AIプロセス」を踏まえ、社会学者や法学者等専門家による会議で詳細が検討されます。

複数のAIが様々な質問を自動的に生成して評価対象となるAIに回答を求めてその内容を確認し、さらに人間による定期監査を行って評価用AIが正しく機能しているかを点検します。総務省は、日本の文化や価値観に沿った信頼できるAIの開発を促進し、外国製AIへの過度な依存を避けることを目指しています。また、評価の過程で日本企業のAIに課題が見つかった場合、NICTが補強データを生成し、開発元に提供する仕組みも検討されています。

総務相は、来年度中に関係企業が評価基盤の一部を試行的に利用できる環境整備に努めると述べています。世界でAI開発競争が激化する中、日本は国際的な信頼性評価の枠組みにおいて主導的役割を果たし、国内のAI産業の競争力を高めることを狙っています。将来的には、AIの安全性確保を担う政府系機関「AIセーフティ・インスティテュート」の評価システムとして活用される可能性もあり、国内外でのAI利用における安全性と透明性の確保に寄与することが期待されています。

