

# Trustworthy AI Foundation (TAF)

## AIのための「信頼」を築く。

AI技術が社会に浸透する今、その倫理的・社会的リスクに対応するため、金融界のS&Pや Moody'sのような「AI格付け機関」を目指します。

### 事業内容

「AI格付け制度」と「知の共有」を両輪に、信頼できるAIエコシステムを推進します。

#### AI格付け制度の策定・実施

- ✓ 公正性・透明性などを軸にした評価体系を設計。
- ✓ 自動判定を補助するAI評価ツールを開発。
- ✓ ホワイト/ブラックリストで開発者・利用者に安心を提供。

#### 情報提供・研究共有

- ✓ 最先端の論文や政策文書を集約した知識共有基盤を運営。
- ✓ 国際会議やワークショップを開催し、グローバルな知見を共有。

### 4つの基本理念

私たちの活動は、社会と産業に貢献するための明確な理念に基づいています。



#### 社会への開放と自由

標準ツールやデータセットを無償提供するオープンサイエンスを推進。



#### AI産業への全体的展望

産学連携を促し、AIライフサイクル全体をカバーするガイドラインを確立。



#### 国際標準とリーダーシップ

グローバルに通用するアーキテクチャを設計し、国際標準を主導。



#### 分散型第三者監査

透明で検証可能な監査プラットフォームを構築し、信頼性を保証。

### 初年度の計画

- ▶ グローバルな機関との連携 (10件以上)
- ▶ Trustworthy AI ホワイトペーパーの作成
- ▶ トレーニングプログラム&ワークショップの実施
- ▶ 産業領域への適応 (10社以上の企業)

# 未来へのロードマップ

短期・中期・長期の明確なステップで、自律的対話AIの実現を目指します。

## 中期：研究開発と展開

研究開発を本格化し、エコシステムを拡大。AIソリューションの地域適応やマルチモーダル具現化AIに取り組む。

## 短期：基盤構築

NPOインフラを確立し、産学パートナーシップを締結。パイロットプログラムを開始し、研究協力の強固な基盤を築く。

## 長期：グローバル統合

自律型対話AIを実現し、グローバルなエコシステムを統合。商業化を拡大し、完全な自律性とグローバル展開を達成する。

## Trustworthy AI Foundation (TAF) の概要

法人名	一般社団法人Trustworthy AI Foundation (TAF)
設立時期	2025年8月中旬 登記完了予定
所在地	東京都千代田区有楽町
代表理事	加藤隆哉（ミドクラジャパン株式会社 代表取締役）
理事	向山徳子（NTY Gridworks株式会社 代表取締役）、早川泰雄（日本千博グループ代表）

### 学術理事（予定）

- ランディー・グーベル（アルバータ大学、Explainable AIの国際的研究者）
- 馬雷（東京大学大学院 情報理工学系研究科 情報科学科 准教授）

### 学術顧問（予定）

- リチャード・サットン（Turing Award受賞、強化学習の世界的権威）
- ジョセフ・シファキス（Turing Award受賞、形式手法の第一人者）