

AIに関する説明は利用者にどう受け止められているか（1） —意識調査の概要と課題の発見—

小川 隆一¹ 島 成佳^{1,2} 竹村 敏彦^{1,3} 渋谷 環¹

概要：AIの動作・判定に関する透明性は「説明可能性」とされ、信頼できるAIの重要な要素となっている。また誤用・悪用回避を含む「適切な利用に関する説明」も重要性が増し、AIシステム・サービス提供者の説明責任が問われるようになっている。筆者らは、AIの信頼に関する情報提供を「広義の説明」ととらえ、その品質管理の可否を検討しているが、実際のAI利用者がAIの説明可能性や広義の説明の意義をどう認識しているか、については実態がわかつていない。そこでAIを業務として利用している実務者を対象にアンケート調査を実施し、AIの信頼性に関してどのような説明を求めるかを質問した。本稿では、アンケート調査の概要と効果的な説明のあり方に関する課題抽出の結果について述べる。

キーワード：AIシステム、信頼、説明可能性、説明責任、説明品質、誤用・悪用回避

How explanation about AI is accepted by AI users - A fact finding survey -

Ogawa, Ryuichi¹ Shima, Shigeyoshi^{1,2}
Takemura, Toshihiko^{1,3} Shibuya, Tamaki¹

Abstract: Transparency regarding AI behavior and decisions is known as "explainability," and is an important element of trustworthy AI. "Explanation of appropriate use," including the avoidance of misuse and abuse, is also becoming increasingly important, and providers of AI systems and services are being held accountable. The authors consider the provision of information related to AI trust to be "explanation in a broad sense" and consider the feasibility of quality control for such information, but the actual state of AI users' perceptions of AI explainability and the significance of explanation in a broad sense are unknown. Therefore, we conducted a questionnaire survey of practitioners who use AI in their work, asking them what kind of explanations they would like to see regarding ensuring AI trust. This paper describes the outline of the questionnaire survey and the results of identifying issues regarding effective explanation methods.

Keywords: AI systems, trust, explainability, accountability, quality of explanation, avoidance of misuse and abuse

1. AIシステムの信頼性説明

AIシステム開発・サービスベンダーの責務として実装すべき性能の議論は主としてソフトウェア品質管理の観点で行われ、ガイドラインが公開されている[1][2][3]。また、AIシステムの信頼性の実装については、ベンダー向けのガバナンスのガイドラインが公開されている[4][5]。

一方、AIシステムの性能・機能とともに、信頼性に関する説明がベンダーの重要な責務となっている。例えば、ブラックボックスであるAIモデルが判定をどう導いたかを明らかにする説明、説明機能を備えたExplainable AIの検討が進んでいる([6]等)。判定結果の説明は、信頼性の要因としては「透明性(transparency)」「説明可能性(explainability)」等に含まれる。

* 本研究の意見は、著者たち個人に帰属し、所属機関の公式見解を示すものではない。

1 情報処理推進機構

Information-Technology Promotion Agency, Japan

一方、AIの信頼性に関して利用者が知るべき情報は判定理由に加え、安全性・セキュリティ・プライバシー・公平性・遵法性・誤用リスク等多岐にわたる。これらへの対応はベンダーが責務として開示すべきものであり、信頼性の要因としては「説明責任(accountability)」に包含されうる。しかしながら、「説明可能性」と「説明責任」の「説明」は上記のように必ずしも同じではなく、その整理やベンダー・利用者間の合意形成ができているとは言えない。また、信頼性の説明は最低限何が必要で、「説明可能性」や「説明責任」の品質をどう管理したらいいかの議論は十分行われていない。

筆者らはこれまでに、「AIシステムの品質として信頼性の説明を管理できるか」の検討を始めている[7]が、上記の状況を鑑みて、AI信頼性の説明の実態、あるいは説明に対

2 長崎県立大学
University of Nagasaki, Japan

3 城西大学
Josai University, Japan

する利用者の意識について基礎データを収集するアンケートを実施した。本稿ではその調査方式と結果を紹介する。

1.1 AI 信頼性説明の定義

筆者らは、AI の信頼性に関する説明の定義を「AI システムの品質の内容・程度を具体的に示す情報であり、利用者が安心して当該システムを使える、と判断するために有益なもの」と定義し、製品品質・利用時品質^aのいずれも含める。AI システムの品質特性は、前掲のガイドラインに含まれる安全性、セキュリティ、プライバシー、公平性、遵法性、動作・判定透明性に加え、AI システムの効果的な利用を支援するための特性を加え、以下のように整理する。

(1) 信頼特性

AI システムの信頼性に関する品質特性である。既存ガイドラインから、少なくとも以下が含まれる。

- ・ 動作・判定
AI の動作・判定結果の詳細。
- ・ 安全性・公平性
安全性・公平性に関する対処。
- ・ セキュリティ・プライバシー
セキュリティ・プライバシー保護の対処。

(2) 利用特性

利用時に信頼性を確保するため、利用者が留意・遵守すべき品質特性である。以下が含まれる。

- ・ 利用条件
AI システムが有効に利用できる応用分野・環境。
- ・ 誤用/悪用回避
誤用・悪用を回避するための対処。
- ・ 不適切な利用・不具合
AI システムが不適切な動作をした場合の対処。

1.2 説明の対象者

AI システムの利用者カテゴリとして、以下の 3 タイプを想定する。信頼性説明の内容はカテゴリごとに異なると考えられる。

(1) システム開発者(開発ベンダー)

AI ベンダーの開発した AI モデルを利用またはチューニングして AI システムを開発する。

(2) システム運用者(サービスベンダー)

AI システムを運用し、サービスを提供する。

(3) システム利用者(エンドユーザー)

サービスを利用して業務等を行う。

1.3 説明の流れ

ステークホルダ間の説明では以下のパターンを想定する。

(1) AI ベンダー → 開発ベンダー

AI ベンダーが AI モデルの信頼特性・利用特性を説明。

(2) AI ベンダー → サービスベンダー

サービスベンダーが AI ベンダーの API を提供するケースで、AI ベンダーが API の信頼特性・利用特性を説明。

(3) 開発ベンダー → サービスベンダー

開発ベンダーが AI システム全般に関する信頼特性・利用特性を説明。

(4) サービスベンダー → エンドユーザー

提供サービスに関する信頼特性・利用特性を説明。エンドユーザーには専門知識等を要求できないという制約がある。

以上から、AI システムの信頼性説明は AI ベンダーからエンドユーザーに至る連鎖が想定され、ステークホルダ間で適切な情報共有がなされるべきこと、またエンドユーザーへの説明では、非専門家であることへの配慮の必要性が確認できる。これらの課題を、筆者らは説明品質の管理という視点で考えていく。

2. 説明品質の評価の枠組み

筆者らは説明品質の評価の枠組みについて、以下のような具体化を試みている。

2.1 説明の品質分類と評価軸

説明品質を内容と外形の 2 面に大別して考える。

(1) 内容品質

Philips [6] は判定理由の説明性が満たすべき要件として、以下をあげている。

- ・ 説明性 (Explanation) 動作・結果の根拠・証拠。
- ・ 理解性 (Meaningful) 内容の理解の程度。
- ・ 正確性 (Explanation accuracy) 説明の正確性。
- ・ 知識制約 (Knowledge limits) 動作の環境条件。

これを参考に、筆者らは説明性の根拠・証拠に説明の目的との整合性を含めた「合目的性」を品質要件とし、知識制約は環境要件とみて説明品質には含めないこととした。

(2) 外形品質

説明の外形的・測定容易な品質として以下を想定した。

- ・ 形態
- ・ サイズ
- ・ 更新頻度
- ・ ユーザビリティ

これらの品質要件と、1.1(1)(2)の説明項目の対応を表 1 に

^a 製品品質は、製品の機能が想定顧客の要求を満たす度合。利用時品質は、特定の利用者が製品を実際に利用してニーズを満たせる度合。

まとめる。

説明項目	説明内容品質			説明外形品質			
	合目的	理解	正確	媒体	サイズ	更新	操作性
(A)	動作・判定						
	安全						
	セキュリティ						
	プライバシー保護						
	公平						
(B)	利用条件						
	誤用回避						
	不具合対応						

表1 説明項目と内容品質・外形品質のマトリクス
Table 1 Matrix of explanation items and quality components

2.2 内容品質の評価軸

上記で精査が必要なのは内容品質の評価項目である。

(1) 合目的性

「説明の目的に沿い、無関係な情報や冗長な情報がない」説明を行うため、誰に何を説明するのか、内容は説明対象が本当に求めているか、を評価する、合目的の定義や計測方法にまだあいまいさを残している。

(2) 理解性

説明の理解の程度は利用者のフィードバックが重要であり、試験利用やサービス中の利用者評価の適切な収集と反映が評価の指標となりうる。

(3) 正確性

正確性は重要であるが、理解性との競合調整が課題となる。正確性を重視すると長大・専門的な説明となり、理解性を損ねることが起こりうる。説明の目的・説明対象（専門家/非専門家等）・状況（法令順守が必要等）に応じた調整が必要となる。

以上の3項目はいずれも重要だが、これらが内容品質の評価軸として妥当かについて検証が必要である。このためには、AIシステム利用者が実際に信頼性説明をどう認識し、受容しているかを知ることが重要と考えられる。しかし、説明の実施状況や受容の実態、利用者がどんな説明を求めるのかはまだ自明でない。そこで筆者らは、AIの信頼性説明の実施状況とそれに対する利用者の意識を知るために調査を設計した。次節において、調査方法について概要を述べる。

3. 信頼性説明に関する実態・意識調査

筆者らの研究ゴールは、信頼性説明に関する品質管理の枠組みの構築にあるが、今回の調査は前段として、信頼性説明の実施の状況、およびそれに対する利用者の意識を問い合わせ、信頼性説明に関わる課題を抽出することとした。

3.1 調査対象と調査内容

(1) 対象 AI システム

想定する AI システムの応用には制限を設けないこととし、二つのカテゴリを想定した。まず生成 AI 普及以前から利用されている判定・認識・予測・検知・監視等の AI 応用を「分類 AI」とカテゴリ化した。またコンテンツ生成・チェック・情報収集・要約・翻訳等の応用を含む「生成 AI」をもう一つのカテゴリとした。

(2) 調査対象者

信頼性説明は、AI システムの業務利用において多く求められると想定される。また、AI システムの信頼性やその説明に知見のある人に質問するのが妥当と考えられる。そこで調査対象者は、企業の従業員及び自ら事業を営む自営業者のうち、AI を業務利用している人を対象とし、個人的な利用のみの人は除外した。

全般傾向をみるため、1.2 節の利用者カテゴリ（AI ベンダー、開発ベンダー等）や調査対象者所属企業の業種は限定しなかった。一方で、AI 利用実態は企業規模・業種に依存する可能性が考えられ、所属企業規模を大企業・中小企業に2分し、自営業者は中小企業就業者とした。また業種は製造業・非製造業に2分し、企業規模（大・中小）と業種（製造・非製造）で決まる4カテゴリそれぞれで400人、合計 1600 人に対し、アンケート形式で質問することとした。

上記 4 カテゴリについては AI 利用実態データがないため、カテゴリごとの仮説は作らず、仮説の発見を主眼とした。ただし筆者らの調査研究の知見から、以下の調査課題を想定した。

(3) 調査課題

前提として以下の調査課題をおいた。

- ・調査課題 1：分類 AI カテゴリの AI システムは特に製造業で使われ、生成 AI カテゴリの AI システムは業種を問わず広く使われているのではないか？

- ・調査課題 2：AI システム利用者は動作経緯や判定理由だけでなく、信頼性に関する全般的な説明を求めているのではないか？

さらに、判定理由や処理の透明性は特に分類 AI をよく使う業種で重視され、それ以外の業種では安全・セキュリティ・プライバシー・誤用対応等の説明が重視されているのではないか、が想定された。ただしどの業種がどうである、等の仮説設定は事前調査が十分でなく行わなかった。この結果アンケートは、以下を目標として設計することとした。

- ・上記二つの課題設定が妥当かを検証する。
- ・検証過程で新たに顕在化する課題を特定する。

3.2 調査方法

アンケートの概要は以下のとおりである。

(1) 実施規模・形態・時期

- 回答者 会社員・自営業 1600 人
企業規模 (301 人以上, 300 人以下)・業種 (製造業, 非製造業) による 4 群各 400 人
- 形式 Web アンケート
- 実施時期 2025 年 7 月 18 日～7 月 27 日

(2) 質問内容

全質問数は 23 問で、以下の項目を設定した。

- AI 利用の役割・経験 2 問
職場における役割 (導入・開発・運用・普及教育を含む)・経験年数等をきいた。生成 AI 普及後 AI 利用を始めた層がどの程度か、が注目点だった。
- AI 学習経験 1 問
AI の学習経験をきいた。学習経験が AI 利用に影響するか、が注目点だった。
- AI 利用実態・効果/課題の意識 4 問
分類 AI・生成 AI の利用頻度と業務上の有効性、利用の課題意識をきいた。
- AI 情報の獲得・獲得の重要性/課題の意識 5 問
AI 利用に関する情報をどう得ているか、その重要性と課題意識をきいた。AI の情報を得て理解することを負担と感じているか、が注目点だった。
- AI 信頼性説明の実施状況・重要性の意識 4 問
職場で AI の信頼性説明を求める/求められることがあるか、その重要性の意識をきいた。どの信頼性要件の説明が重要とされるか、が注目点だった。自営業の回答者には顧客等への説明と読みかえてもらった。
- AI 利用を支援する説明情報提供のあり方 7 問
AI 利用を支援する説明の職場における共有、望ましい説明の形態、生成 AI や AI エージェントによる支援のあり方等をきいた。文書の事前参照ではない説明の形式が求められるのではないか、が注目点だった。

なお質問項目において、分類 AI・生成 AI を特定して回答するものがあるが、当該カテゴリの業務利用がない場合は仮に利用したら、として回答してもらった。これは特に分類 AI について経験者のみ回答とした場合、十分なサンプルが集められない可能性を考慮した結果である。次節で調査結果の概要を述べる。

4. 調査結果

単純集計を基本とした主要な調査結果を述べる。

4.1 4 群ごとの AI 利用状況

3.1(2)の企業規模・業種別 4 群ごとに、分類 AI・生成 AI の利用状況を質問したところ、分類 AI は各群ほぼ均等に 51～58%が利用した・利用したことがあると回答し。3.1 節調査課題 1 で想定した「分類 AI は特に製造業で使われている」はなり立たないことが確認された。一方、生成 AI の利用は各群すべて 95%以上で、調査課題 1 の「生成 AI は業種を問わず広く使われている」は確認された。

4.2 AI 利用の役割・経験

図 1 に、業務における AI 利用経験の役割別分布 (4 群ではない全データ) を示す。生成 AI 普及以前 (3 年以上前) から業務で AI を利用、とする回答はどの役割でも 15%未満であり、回答者は経験の少ない AI 利用者が大半を占めている。

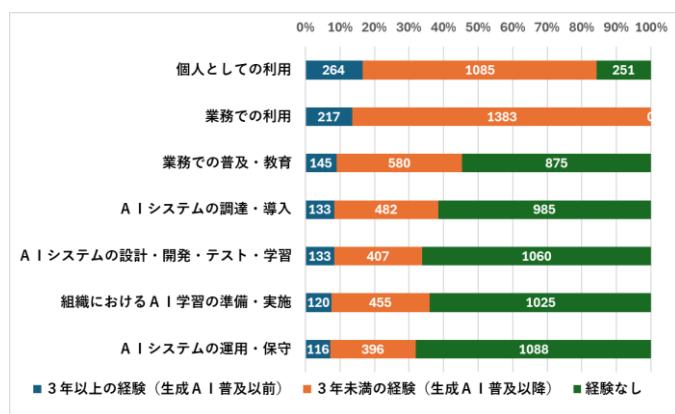


図 1 AI 利用の役割別の経験値

Figure 1 Experiences of AI usage by role

4.3 AI 学習経験

図 2 に、利用者自身の AI 学習の状況を示す。AI の仕組み・動作に関しては授業・研修、独学ともに 30%程度で、過去 3 年以内に開始した回答者はその 6 割を占める。また 40%近い回答者が学習経験なしで業務用 AI を利用している。

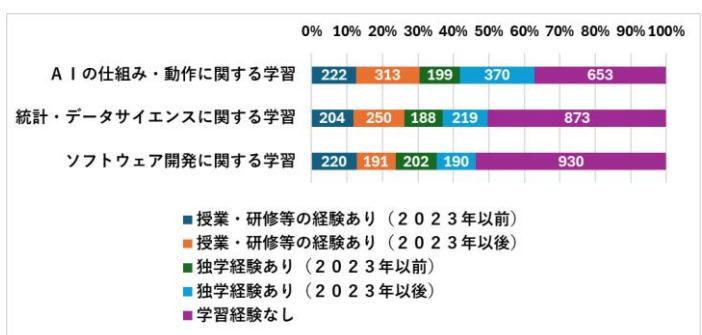


図 2 利用者の AI 学習経験

Figure 2 Experiences of AI learning

4.4 AI 利用実態・効果/課題の意識

図3に、AIシステム利用分野ごとの利用頻度を示す。図の縦軸(利用分野)は分類AI・生成AIと紐づけしていない。例えば業務自動化はどちらの場合もありうる。高頻度の利用分野は業務・処理自動化、情報収集・要約・分析、文書作成・修正であり、生成AI利用の影響が大きいことが伺われる。なお分類AIの分野である制御・認識・検知・予測等を利用する、という回答が30%前後あるが、図1では生成AI以前からの利用は15%に留まつた。分類AIも生成AI普及後に利用が進んだのか、は精査が必要である。

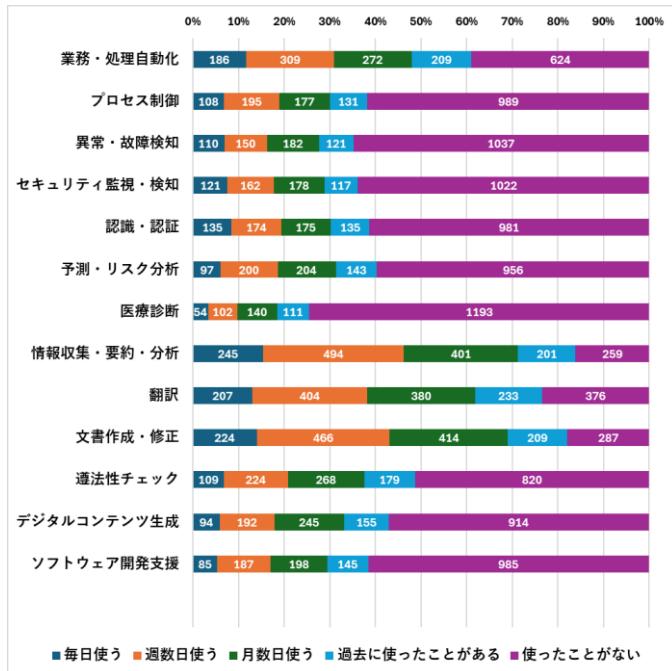


図3 AIシステム利用分野ごとの利用頻度

Figure 3 Usage frequency of AI system by application field

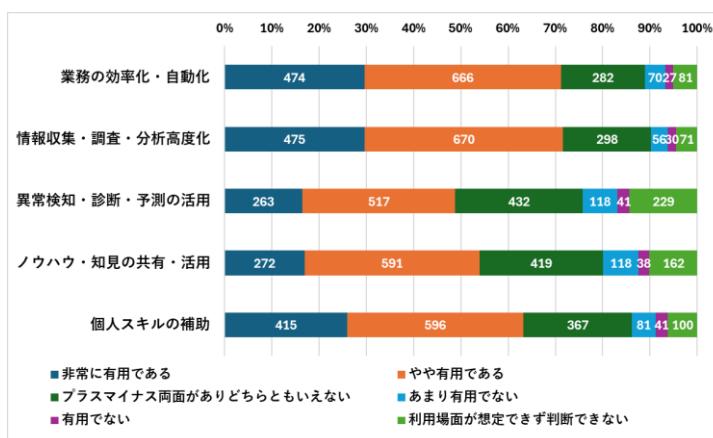


図4 AI利用の有用性の意識

Figure 4 Perception of AI usage effectiveness

図4に、職場におけるAIの有用性に対する意識を示す。「やや有用」まで含めると、業務効率化・自動化、情報収

集・調査分析について70%超が有用としている。

図5に、職場におけるAI利用の懸念に対する意識を示す。組織における知識・スキルの不足への懸念が際立っており、AIの精度・誤回答の不安、誤用・セキュリティへ等への懸念が続いている。また別質問において、具体的にリスクと思うものをきいたところ、情報漏えい(セキュリティ)、判定理由の不透明性(精度・誤用等の不安)、が50%を占めて突出した。

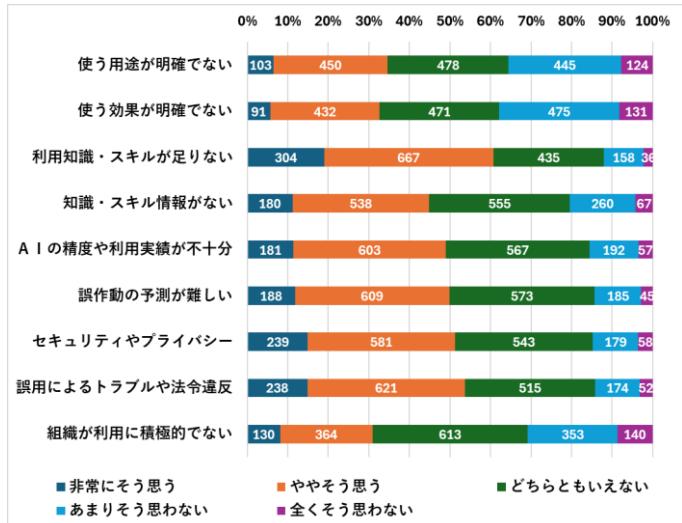


図5 AI利用の懸念点の意識

Figure 5 Risk awareness of AI usage

4.5 AI情報の獲得・獲得の重要性/課題の意識

図6に、AI利用のためにどんな情報を参照しているか、を示す。

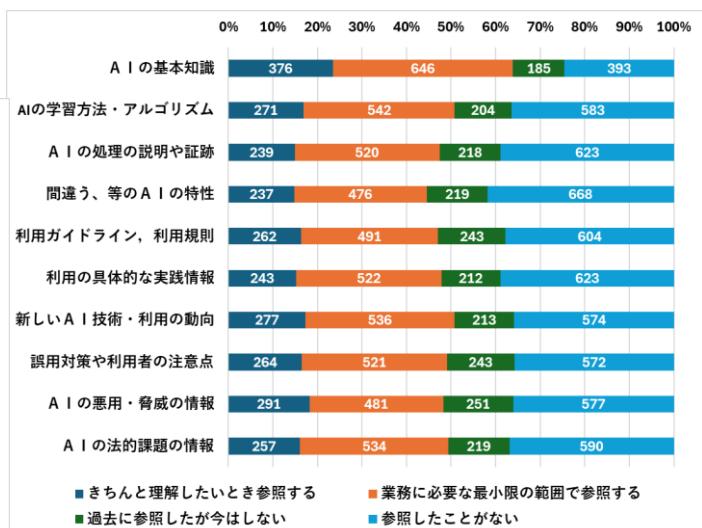


図6 AI利用のために参照する情報

Figure 6 Information referring for AI usage

また図7に、当該参照情報がAIの利用にどの程度重要

か、の意識を示す。図6で「AIにできること、利用場面等の基本知識」を参照するとした回答が60%以上、図7で「どちらかといえば」までを含めて重要とした回答が70%以上となったことは注目される。基本知識の重視傾向は別質問でも現れており、図5で確認された「AI利用知識・スキル不足」の懸念の裏返しである可能性がある。

他の注目点は、別質問で「判定理由の不透明性」がリスクとして抽出されたが、AI処理説明・証跡の参照（図6）や重要性意識（図7）の程度は必ずしも大きくないことである。透明性に対して漠然と不安はあるが、個々の証跡を見るまではいかない、という利用者の意識があるのかもしれない。

また図7で、誤用・悪用・法的課題の情報を「重要」「どちらかといえば重要」とする回答は、処理説明・証跡への回答をこえて60%後半となった。調査課題2を支持する結果といえる。

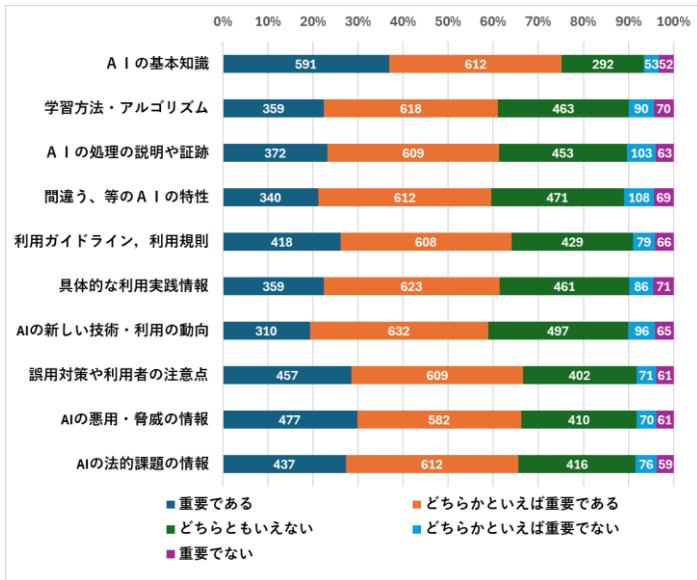


図7 参照情報の重要性の意識

Figure 7 Perception of importance of information referring

4.6 AI信頼性説明の実施状況・重要性の意識

図8に、AIの動作透明性・精度・安全・セキュリティ・プライバシー・誤用等の信頼性項目について職場で説明を求めるか/求められるか、の実施状況を示す。グラフの左側4区分、ほぼ55%が「実施」と回答した。また別質問で、職場における説明の重要性をきいたところ、実施状況に近いほぼ50%が「どちらかといえば」を含め重要と回答した。なお、これらの質問では分類AI・生成AIを明示的に分離した信頼性項目に回答してもらったが、カテゴリ間で大きな差はなかったb。

b 分類AI利用者は回答者の約半数と判明しており[8]、図8からは分類AI利用回答者の多くが説明実施に関与していることが推定される。

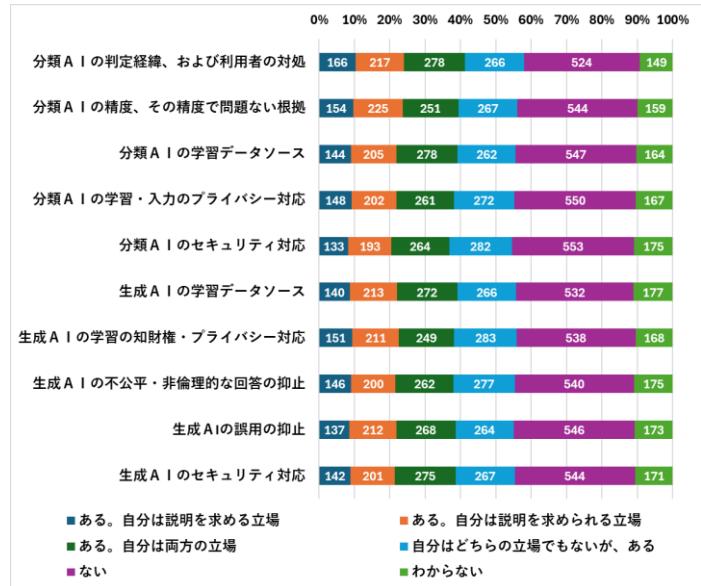


図8 信頼性に関する説明実施の状況

Figure 8 Status of explaining AI system trustworthiness

なお重要性の意識については、企業規模・業種別4群のうち、中小企業・製造業群に特徴が見られた。特に生成AI系の信頼性項目の説明重要性が40%に落ち、全データの説明重要性の回答率に比べ低かったc。重視されそうな分類AI系信頼性項目、例えばセキュリティ説明も40%だった。これらの値は他の3群のそれと比べ特異的に低く、精査が必要である。

4.7 AI利用を支援する説明情報提供のあり方

AIの信頼性説明の提供をどこに求めたいか、の質問に対しては、項目別に以下の回答が多かった。全体としてAIベンダー・サービスベンダーへの期待が大きい。

- ・基本知識：専門家（書籍・Web）
- ・学習・アルゴリズム：AIベンダー/サービスベンダー
- ・透明性・証跡：AIベンダー/サービスベンダー
- ・利用例（誤用含む）：AIベンダー/サービスベンダー
- ・学習データ：サービスベンダー/AIベンダー
- ・セキュリティ・インシデント：AIベンダー/サービスベンダー/公的機関
- ・法制・制度：公的機関

図9に、説明をどのような媒体で提供してほしいか、を示す。PPTスライド、PDF文書、動画を单一選択で選んでもらったところ、個々の説明項目を問わずほぼ均等に選択された。文書で説明、では必ずしもないことがわかる。

また、図9と同じ説明項目について、どのような状況で参照したいか質問したところ、「AI利用中に、自分から要

c 本稿に記載していないが、全データの結果では、生成AI系の当該項目を「重要」「どちらかといえば重要」とする回答は45~50%だった。

求」「AI 利用と異なる環境で自分から要求」「AI 利用時に AI エージェント等が判断して提供」の選択肢の中で、「利用時に自分から要求」が項目を問わず突出して選択された。利用時の説明は、例えば動作説明や適切な利用（誤用回避等）の確認でニーズが大きいと思われる。

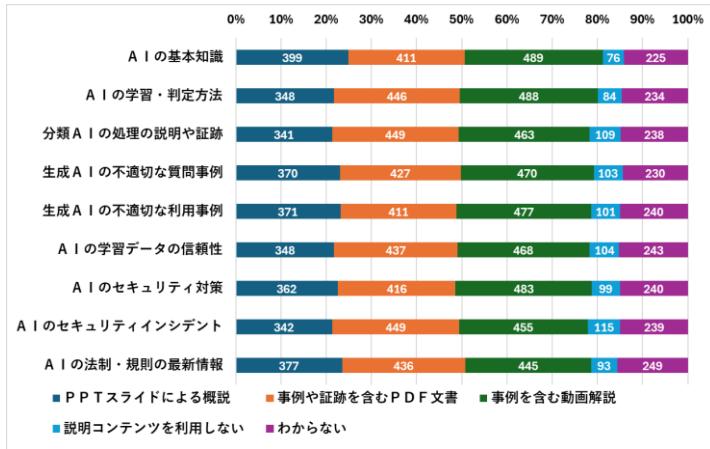


図 9 信頼性説明の望ましい媒体

Figure 9 Desirable medium of trustworthiness explanation

さらに、今後の説明提示のあり方として、生成 AI による説明を利用したいか質問したところ、「利用したい」が各説明項目で 50% をこえた。一方で、「どちらともいえない」「利用したくない」が合わせて 30~40% となった。利用したい理由では、「短時間でわかる」「最新情報が得られるはず」「AIについて大量に学習している」「対話的」等が選ばれた。一方、利用を躊躇する理由は「説明が間違ってもわからない」「生成 AI の知識レベルがわからない」「質問で回答の質が変わる」「体系的に理解できない」等が選ばれた。利用したくない理由の回答結果を図 10 に示す。

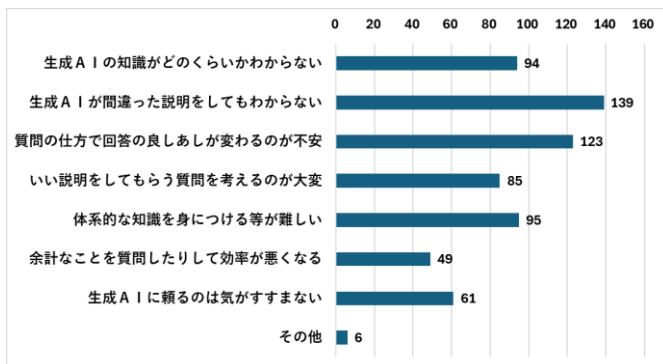


図 10 生成 AI による説明を利用したくない理由

Figure 10 Reasons of unwillingness of generative AI for explanation

5. 分析

5.1 調査課題検証のまとめ

調査課題 1 については 4 章記載のとおり、製造業以外の業種においても分類 AI は利用され、説明も一定程度実施されている。生成 AI の利用は調査課題 1 のとおりであり、全業種の AI 業務利用の 9 割以上が生成 AI を含むと思われる。ただし、全般的な AI 利用の有用性は主に業務効率化と認識される傾向がみられ、生成 AI 固有の個人スキル支援等は目立たない。

調査課題 2 の信頼性全般に関する説明の要請は、参照する情報、重要性の意識、説明実施の状況等で繰り返し確認された。特に、実際に説明を求める/求められる内容が AI の動作追跡・理由説明に留まらず、安全・セキュリティ・プライバシー・誤用対応等に広がっていることはステークホルダ間の説明責任の意識を合わせるうえで重要な示唆と考える。

なお、説明を実践しているとする回答は全体の約 50% に留まった。例えば職員が業務文書作成に生成 AI を利用するとき、結果の妥当性を組織がチェックする、等はまだ行われていない可能性がある。個人作業における生成 AI 利用の説明は精査が必要と思われる。

5.2 新たな抽出課題

(1) 抽出された利用者特性

本調査の質問回答を整理すると、以下のような業務 AI 利用者の特性が導かれる。

- ・業務 AI 利用者は生成 AI 普及以降急増しており、利用形態は生成 AI による個人業務効率化が多くを占めると推定される。
- ・急増した利用者の多くは AI 学習経験や AI 利用の知識・スキルが少ない。AI 理解のためにほしい情報として「基本知識」が最上位を占めることは、これを裏書きすると思われる。
- ・信頼性説明、あるいは説明の実践が重要であると思う利用者は約 6 割であり、知識・スキル不足はあるものの、信頼性説明の重要性について一定の理解が得られている。
- ・AI の判定理由説明（動作説明）の重要性の意識は他の信頼性項目と同等以上であるが、実際に証跡を見るまでではいかない。重要性はわかるが、実経験がない利用者が多いと思われる。
- ・従来の「事前に読む説明」だけでなく、「利用中必要に応じた説明」を求める利用者が多い。また媒体はプレゼン・文書・動画等の多様な形式を求めている。
- ・以上の傾向に企業規模・業種の差は顕著でない。

(2) 抽出された新たな課題

信頼性説明は技術・利用ルール・法制等多様な要素を含み、急増中のAI利用者がその基本となる知識・利用リテラシーを獲得するのは容易でない。一方で、生成AI普及をトリガとして「知識・経験値が少ないAIビギナー」である利用者は今後も増え、この「ビギナーグループ」の特性を明らかにして適切な説明を提供することが重要と考えられる。そこで新たな検討課題として、例えば以下の項目が導かれる。

- ① 生成AI利用を主体とし、AI知識・経験等が少ない利用者のAI利用特性は具体的にどういうものか
- ② 彼らに対して適切な信頼性説明とはどのような内容・形態か
- ③ 信頼性説明が利用者特性に合っていることの評価をどのようにするか

上記①については、アンケート・ケース作成等により特性を抽出する必要がある。例えばどのような利用ケースで知識・スキル不足が問題となるか、の深掘りが考えられる。なお、知識・スキル不足はアンケート質問で明らかにできただが、組織の教育や説明体制等については問うていない。組織体制のあり方というサブ課題もありうる。

②については、第一義的には「利用者のAI利用リテラシーのレベルを想定した説明」を作成する必要があるが、一方で、「説明自体に、基本情報への遡及などリテラシー向上に資する情報を入れる」が有効となる可能性がある。これらはフィールド調査に加え、説明生成の試行が必要と思われる。また4.7節に記載したとおり、「AI利用中の説明提供」の必要性が示唆されたが、AIエージェントや生成AIによる説明提供には効果と副作用の両面が考えられる。これも新たなサブ課題になりうる。

さらに③については、それらの説明のあり方を2章で述べた説明品質の枠組みでどう評価するか、が新たな課題となる。例えば表1の内容品質の特性項目「合目的性」「理解性」「正確性」が評価に適合するものか、どんな拡張や改変が必要かを検討する必要がある。

6. まとめと今後の課題

本稿では、AIの信頼性項目（精度・安全・セキュリティ・誤用対応等）に関する情報提供を広く「信頼性説明」と定義し、信頼性説明の品質を評価する枠組み案を提示したのち、信頼性説明の企業での実践状況や同説明に対する業務用AI利用者の意識調査の概要を述べた。回答者は業務AI利用を開始したばかりの人が多く、AIの知見は少ないものの、信頼性説明について一定の理解と重要性の意識が見られ、業種や分類AI・生成AIによる差は顕著ではなかった。これらを総合して、筆者らは「AI知識・知見の少ないAI利

用者の利用特性の探索」「彼らへの信頼性説明のあり方」「その評価の方式」という新たな課題を抽出した。

ただし、これらの抽出課題は検討時間が短く、さらなる精査が必要である。今後は本稿のアンケート調査結果を深掘りするとともに、抽出課題の検討を進めていく。

参考文献

- [1] デジタルアーキテクチャ研究センター, “機械学習品質マネジメントガイドライン 第4版,” 産業技術総合研究所, 2023年12月22日.
- [2] ISO/IEC JTC1/SC42, “ISO/IEC42001:2023 Information technology - Artificial intelligence - Management system,” the International Organization for Standardization, December 2023.
- [3] ISO/IEC JTC1/SC7, “ISO/IEC 25019:2023 Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) –” Quality-in-use model, the International Organization for Standardization, November 2023.
- [4] NIST, “Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, NIST AI 600-1,” National Institute of Standards and Technology, July 2024.
- [5] 総務省、経済産業省, “AI事業者ガイドライン 第1.01版,” 総務省、経済産業省, 2024年11月22日.
- [6] Philips et al., “Four Principles of Explainable Artificial Intelligence, NISTIR 8312,” National Institute of Standards and Technology, September 2021.
- [7] 小川隆一、島成佳, “AIシステムの「説明」を品質として管理できるか,” 2025年度人工知能学会全国大会, 4L3-OS-38-02, 2025年5月29日.
- [8] 竹村敏彦、小川隆一、島成佳、渋谷環, “AIに関する説明は利用者にどう受け止められているか (2) 一生成AI利用者の属性分析—,” コンピュータセキュリティシンポジウム2025講演予稿集, 2025年10月27日.