

データ品質サブワーキンググループ（データ品質 SWG） 2025 年度活動報告書

令和 8 年 4 月 15 日

AISI Japan
AI Safety Institute

AI セーフティ・インスティテュート
事業実証ワーキンググループ
データ品質サブワーキンググループ

目次

1.	はじめに	1
1.1.	背景・目的	1
1.2.	本書の位置付け・概要	2
1.3.	対象読者	2
2.	データ品質に関する動向	3
2.1.	国内外の動向	3
2.1.1.	国内の動向	3
2.1.2.	国外の動向	6
2.2.	技術動向	11
2.2.1.	AI の性能とセーフティを支えるデータ品質要件	11
2.2.2.	データ品質管理の実装アプローチ	12
2.2.3.	データ品質向上のための AI 活用	14
3.	データ品質を取り巻く論点の整理	16
3.1.	データ品質マネジメントガイドブック	16
3.2.	2025 年度の活動で浮かび上がった論点	16
4.	データ品質 SWG における 2025 年度の活動	20
4.1.	ユースケースの整理	21
4.1.1.	ユースケースの整理アプローチ	21
4.1.2.	ユースケースの整理例	24
4.1.3.	得られた示唆・今後の改善策	29
4.2.	データ品質マネジメントチェックリストの作成・検証	30
4.2.1.	データ品質マネジメントチェックリストの概要	30
4.2.2.	データ品質マネジメントチェックリストの使い方	31
4.2.3.	データ品質マネジメントチェックリストの検証	33
4.2.4.	データ品質マネジメントチェックリストの今後の改善策	36
4.3.	データ品質レベルと AI 出力品質の検証	37
4.3.1.	検証の位置付け	38
4.3.2.	検証対象ユースケース	38
4.3.3.	検証の設計方針	39
4.3.4.	検証対象データ	39
4.3.5.	データ準備	41
4.3.6.	検証方法	43
4.3.7.	結果と考察	45
5.	まとめ	49
5.1.	2025 年度の到達点	49
5.2.	見えてきた課題の整理	50
5.3.	スコープ拡大（新たな利用形態への対応）	51
5.4.	継続的取組の方向性	51

1.1. 背景・目的

生成 AI をはじめとする AI 技術の急速な普及により、AI は研究開発段階を超えて、行政、産業、社会のさまざまな領域で実装・活用される段階に入っている。こうした中で、AI の性能や有用性、さらには信頼性や安全性は、モデルやアルゴリズムのみならず、学習や利用の前提となるデータの品質に大きく依存している。「Garbage in, Garbage out (ごみを入れればごみが出てくる)」と言われるように、データ品質は AI 利活用及び AI セーフティの成否を左右する基盤的要素であるとの認識が、国内外で共有されつつある。

一方で、AI 利活用が現場レベルで急速に進む中、データ品質に関する課題は多層的に顕在化している。現場では、データの鮮度や来歴が十分に把握されていないことにより、AI の出力結果をどの程度信頼してよいのか判断が難しい、あるいはデータに内在する偏りが AI の出力にそのまま反映されてしまうといった課題が生じている。他方、マネジメントレベルでは、データ品質に関する国際標準や体系化された知見が存在するにもかかわらず、それらを AI 時代の実務に即して活用・定着させるための枠組みや支援が十分とは言えないという課題がある。データ品質は、システムやプロジェクトにおいて明示的な要求として整理されにくく、結果として関心や投資が後回しにされがちである点も、こうした課題を構造的に難しくしている。

国際的にも、AI ガバナンスや AI セーフティに関する議論の中で、データ品質は重要な構成要素として位置付けられている。データ品質は、技術的な問題にとどまらず、組織運営やガバナンスの観点から横断的に取り組むべきテーマであり、その重要性は今後さらに高まると考えられる。

データ品質管理については、ISO 等を中心とした既存の標準類や体系化された知見が既に存在している。他方、これらの標準をどのように現場で使いこなすか、また学習データや評価データ、継続的改善といった AI 時代特有の観点をどのように取り込むかが必ずしも明確ではなかった。こうした課題認識のもと、実務目線での整理と共有を目的として、2025 年 3 月には AI セーフティ・インスティテュート (AISI) が「データ品質マネジメントガイドブック 1.0」を公表した。

以上の経緯から、データ品質 SWG は、AISI に設けられた事業実証 WG のうち分野横断 SWG の一つとして設置された。本 SWG では、データ品質を扱う必然性を感じている産学官の多様なメンバーが参画し、現場の知見や国際動向を持ち寄りながら、実証を通じて、AI 利活用のためのデータ品質管理のあるべき姿と実践的な枠組みを構築することを目的とした。また、既存の国際標準やガイドラインを前提としつつ、それらを補完する形で、実務で活用可能なガイドや評価ツールを整備することに重点を置いた。

これらの取組を通じて、AI 時代における信頼性の高いデータ利活用の実装へとつなげることが、本 SWG の目的である。併せて、日本が国際社会から期待される品質に関する強みを生かしながら、将来的には日本発のデータ品質評価フレームワークや指標定義を、国際的な議論へ反映していくことを目指していく。

1.2. 本書の位置付け・概要

本書は、データ品質 SWG が 2025 年度に実施した検討及び活動内容、その成果を整理・報告することを目的として取りまとめた活動報告書である。

前述の背景及び問題意識を踏まえ、本 SWG では、AI 利活用を前提とした実務の中でデータ品質をどのように捉え、管理し、評価していくべきか、検討を進めてきた。本書は、これらの活動を通じて得られた知見や成果を体系的に整理し、データ品質に関する共通理解の形成や、今後の実務及び議論につなげるための基礎資料として位置付けるものである。

本書の構成は、以下のとおりである。なお、第 4 章で取り上げる成果物の一部については、本書中での概要説明に加え、別添資料として成果物そのものを併せて提示する。

- 第 1 章：はじめに
データ品質 SWG 設立の背景及び目的を整理するとともに、本書全体の位置付けや前提となる問題意識を明らかにする。
- 第 2 章：データ品質に関する動向
AI・データ利活用を取り巻く環境変化を踏まえ、データ品質に関する国内外の動向や技術的背景を概観する。
- 第 3 章：データ品質を取り巻く論点の整理
前章を受けて、データ品質を取り巻く論点の整理を行う。
- 第 4 章：データ品質 SWG における 2025 年度の活動
2025 年度にデータ品質 SWG が実施した具体的な活動内容とその成果について報告する。ユースケースの整理、データ品質マネジメントチェックリストの作成、官報データや分野別 SWG における検証等を取り上げ、それぞれの検討の経緯や得られた示唆を整理する。
- 第 5 章：まとめ
2025 年度の検討結果や成果を踏まえ、今後のデータ品質 SWG の取組や、継続的な検討に向けた方向性について整理する。

1.3. 対象読者

本書は、データ品質に関する共通理解の形成や、今後の議論・実務の検討に資することを意図している。

このような観点から、本書は、データ品質、AI 利活用、データガバナンスといったテーマに関心や課題意識を有する幅広い関係者に参照されることを想定している。

具体的には、企業や団体において AI システムやデータ基盤の設計・運用、データ品質管理に携わる実務者や技術者、行政機関において AI 政策やデータ利活用政策、標準化施策等に関与する担当者、さらにはデータ品質や AI、情報システム等を研究対象とする研究者・有識者などが、本書の内容と接点を持ち得ると考えられる。

本書が、これらの関係者にとって、データ品質を巡る現状の整理や課題認識の共有、今後の取組を検討する際の参考資料として活用されることを期待する。

2.

データ品質に関する動向

2.1. 国内外の動向

本章では、国内外のデータ品質に係る動向を紹介する。

2.1.1. 国内の動向

国内では、AI法を柱として、透明性確保を含む基本理念の下でAIの研究開発・利活用を推進する枠組みが整備されつつある状況である。EUのような規制主導の枠組みと比べ、国内はAI法で方向性を示しつつ、事業者向けガイドライン群で実務を具体化する構図である。なお、データ品質は、目的適合性を軸に、評価・改善を反復するマネジメントとして位置付けられ、共通言語化が進められている。さらに、相互運用やデータ連携基盤の整備が、組織・分野を跨ぐデータ流通における品質確保の基盤として位置付けられている。総じて、国内の動向は、制度の方向付け・実務ガイド・基盤整備を組み合わせ、信頼性の段階的な底上げを目指している。

主な動向を下表に示す。

表 2-1 データ品質に関する国内の動向

No	取組名	データ品質との関連	実施主体	発行年（最終更新年）
1	人工知能関連技術の研究開発及び活用の推進に関する法律（AI法）	信頼できるAIの原則の1つとしてデータ品質を提示	日本政府	2025年
2	AI事業者ガイドライン	目的適合・正確性/更新性・バイアス低減・記録/開示について提示	経済産業省・総務省	2025年
3	機械学習品質マネジメントガイドライン	妥当性/代表性・欠損/誤り・ドリフトの点検と評価方法について提示	産業技術総合研究所（産総研）	2020年（2023年）
4	生成AI品質マネジメントガイドライン	学習データ+プロンプト/参照データ/RAG/ログの管理・評価について提示	産業技術総合研究所（産総研）	2025年
5	政府相互運用性フレームワーク（GIF）	データモデル/語彙/形式の標準化（整合性・一貫性）について提示	デジタル庁	2022年
6	データ連携基盤を通して提供	役割分担、品質評価手	内閣府	2023年

	されるデータの品質管理ガイドブック	順、点検項目、簡易ツールについて提示		
7	データガバナンス・ガイドライン	品質 KPI の設定、監視、是正の統制について提示	デジタル庁	2025 年
8	AI プロダクト品質保証ガイドライン	データ適切性・偏り・更新管理・評価データ妥当性の確認（チェックリスト）について提示	AI プロダクト品質保証コンソーシアム	2019 年（2025 年）
9	センシングデータの品質評価に関するホワイトペーパー	品質測定量（計測条件/設置/通信/保守等）と評価手順について整理	データ社会推進協議会	2022 年（2024 年）

(1) 日本政府:人工知能関連技術の研究開発及び活用の推進に関する法律 (AI 法)

2025 年 6 月 4 日に公布・一部施行され、9 月 1 日に全面施行された AI 法¹は、人工知能関連技術の研究開発及び活用を総合的・計画的に推進するための基本法である。基本理念として、人間の尊厳の尊重、イノベーションの促進、リスクへの対応等を掲げ、関係主体の役割と取組の方向性を示している。

データ品質については、法律として個別の品質特性・評価手法（正確性や網羅性等）を直接定義するというより、透明性の確保等を通じた「信頼できる AI」の基盤整備を求め、具体的実装・管理は下位のガイドラインや実務指針に委ねている。

(2) 経済産業省・総務省：AI 事業者ガイドライン

AI 事業者ガイドラインは、AI の開発者・提供者・利用者がライフサイクル全体で配慮すべき事項を整理した統一ガイドラインであり、2024 年に第 1.0 版、2025 年に第 1.1 版が公表された²。リスクベースの考え方にに基づき、体制整備、説明責任、監視・改善等の実務的な行動指針を提示している。

データ品質については、目的適合性、正確性、更新性等の確保とともに、データに起因するバイアスの把握と低減を求める。文書化、記録、外部への情報提供等を通じた透明性確保を重視する点が特徴である。

(3) 産総研：機械学習品質マネジメントガイドライン

機械学習品質マネジメントガイドラインは、機械学習を用いる AI システムの品質をライフサイ

¹ e-Gov, 人工知能関連技術の研究開発及び活用の推進に関する法律

<https://laws.e-gov.go.jp/law/507AC0000000053>

² 経済産業省, AI 事業者ガイドライン

https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/20240419_report.html

クルで管理する枠組みを示すものであり、2020年に初版が公表され、その後改訂が重ねられている³。品質要求の整理、設計・学習・評価・運用の各段階の確認事項を体系化し、品質確保の実務を支援している。

データ品質は内部品質の主要構成要素として明示的に扱われ、例えば「データ設計の十分性」「データセットの被覆性」「データセットの均一性」等を品質管理の対象として位置付け、品質保証の要求事項として整理している。

(4) 産総研：生成 AI 品質マネジメントガイドライン

生成 AI 品質マネジメントガイドラインは、生成 AI 利用の品質確保の考え方を整理したものであり、2025年に第1版が公表された⁴。生成 AI 特有のリスクを踏まえ、企画・設計から運用・改善までの管理ポイントを提示している。

データ品質については、学習データに限らず、データ点／データセット／複数データセット組合せ時の品質観点を、正確性・完全性・一貫性・最新性、代表性、来歴性、文脈整合性等で整理している。位置付けとしては、(3)の機械学習品質マネジメントの知見を生成 AI に拡張した技術指針である。

(5) デジタル庁：政府相互運用性フレームワーク (GIF)

政府相互運用性フレームワーク (GIF) は、政府内外でのデータ連携を円滑にするための参照モデル及び実践ガイドであり、2022年に公表された⁵。コアデータモデル等を提示し、共通語彙基盤と連携してデータ項目・構造の整合を図っている。

データ品質については、正確性等の評価指標を直接規定するというより、データモデルや値の形式(書式)を共通化することで「定義のばらつき」や「形式不一致」を抑え、相互運用性の観点から品質の前提条件(整合性・一貫性)を底上げする点について記載されている。

(6) 内閣府：データ連携基盤を通して提供されるデータの品質管理ガイドブック

データ連携基盤を通して提供されるデータの品質管理ガイドブックは、データ連携基盤(都市OS等)を通じて提供されるデータの品質管理を推進するため、役割分担と評価手法の参考モデルを整理したガイドブックであり、2023年に公表された⁶。整備主体、データ提供者、データ利用者を想定し、品質確保の責任と実務を明確化している。

データ品質は本書の中心テーマとして扱われ、品質評価の手順、点検項目、簡易評価ツール等が提示されている。

³ 産業技術総合研究所, 機械学習品質マネジメントガイドライン

<https://www.digiarc.aist.go.jp/publication/aiqm/>

⁴ 産業技術総合研究所生成, AI 品質マネジメントガイドライン

<https://www.digiarc.aist.go.jp/publication/aiqm/genaiqm-guidelines-v1.html>

⁵ デジタル庁, 政府相互運用性フレームワーク (GIF)

https://www.digital.go.jp/policies/data_strategy_government_interoperability_framework

⁶ 内閣府, 「データ連携基盤を通して提供されるデータの品質管理ガイドブック」を公表します。(令和5年9月26日)

https://www.chisou.go.jp/tiiki/kokusentoc/supercity/supercity_230926_guidebook.html

(7) デジタル庁：データガバナンス・ガイドライン

データガバナンス・ガイドラインは、企業経営者向けにデータを戦略的資産として管理・活用するための要点を整理したものであり、2025年に公開された⁷。組織体制、方針・ルール、権限と責任、KPIによる運用管理等を通じて、全社的な統制を確立することを求めている。

データ品質は価値創出の前提として位置付けられ、品質指標の設定、監視、是正の仕組みを構築することが重要である。

(8) AIプロダクト品質保証コンソーシアム：AIプロダクト品質保証ガイドライン

AIプロダクト品質保証ガイドラインは、AIを用いる製品・サービスの品質保証の考え方と確認事項を整理したものであり、2019年に初版が公開され、以後改訂が継続されている⁸。品質要求の整理、評価観点、チェックリスト等を提供し、開発から運用までの品質確保を支援している。

データ品質に関しては、学習データの適切性や偏り、更新管理、評価データの妥当性等を品質要素として取り扱っている。

(9) データ社会推進協議会：センシングデータの品質評価に関するホワイトペーパー

センシングデータの品質評価に関するホワイトペーパーは、センシングデータの品質評価モデルと評価方法を整理したものであり、2022年に初版、2024年に実証結果を踏まえた改訂版が公開された⁹。

センシングデータ特有の品質観点として、計測原理や設置条件、通信安定性、運用・保守、セキュリティ対策等の要素を含めて整理し、品質測定量の定義と評価手順を示している。

2.1.2. 国外の動向

EUでは、AI Actを軸に、リスク区分に応じた義務と透明性要求を規則として整備し、市場全体の遵守を求める構図である。国内の基本法+ガイドライン型に対し、EUは法規制として要件を明確化し、データ品質を含むガバナンスをコンプライアンス事項として扱う点が特徴である。とりわけ高リスク領域では、データガバナンス要件が制度要件として位置付けられ、適合の説明責任が重視される。これを補完して、倫理・データ保護等の観点からの関連文書が、リスク管理や実装上の論点整理を担う構図である。総じて、規制主導で要求事項を統一し、関連文書で実務適用を厚くする動向である。

EUの動向を下表及び項目別に示す。

⁷ デジタル庁、「データガバナンス・ガイドライン」を策定しました

<https://www.digital.go.jp/news/71bf19c2-f804-488e-ab32-e7a044dcac58>

⁸ QA4AI, AIプロダクト品質保証ガイドライン(QA4AI Guidelines)

<https://www.qa4ai.jp/download/>

⁹ 一般社団法人データ社会推進協議会(DSA), センシングデータのデータ品質評価基準策定に向けた提案

<https://data-society-alliance.org/survey-research/data-quality-evaluation-standards/>

表 2-2 データ品質に関する EU 及び関連組織の動向

No	取組名	データ品質との関連	実施主体	発行年（最終更新年）
1	AI Act	高リスク AI のデータ要件（目的適合・代表性・誤り最小化・完全性等）について規定	EU	2024 年
2	Ethics Guidelines for Trustworthy AI	データ品質確保・バイアス低減・データ管理透明性（自己点検）について提示	欧州委員会	2019 年
3	General-Purpose AI Code of Practice	学習/評価データ・手法の文書化と情報提供（透明性）について提示	欧州委員会	2025 年
4	Guidance for Risk Management of AI systems	データ品質/偏りを含むリスク特定・低減・継続監視について提示	欧州データ保護監察機関	2025 年
5	AI Privacy Risks & Mitigations – LLMs	過剰収集・漏えい・再同定・偏り等のリスクと低減策について整理	欧州データ保護会議	2025 年
6	Bias Evaluation	代表性確保と評価用データ設計を中心とするバイアス評価手順について整理	欧州データ保護会議	2024 年
7	EU Data Quality Guidelines	FAIR に沿うメタデータ品質、標準語彙、文書化、更新・版管理について提示	欧州連合出版局	2021 年 (2022 年)

(1) EU : AI Act

EU AI Act は、AI をリスクに応じて規制する枠組みを定めた規則であり、2024 年 8 月に発効した¹⁰。禁止・高リスク等の類型に応じて要件を段階的に課し、高リスク AI にはリスク管理、技術文書、監視等を求めている。

データ品質については、高リスク AI の「データ及びデータガバナンス要件」として、訓練・検証・テスト用データセットに対し、目的適合性 (relevant)、代表性 (sufficiently representative)、誤りの最小化 (to the best extent possible, free of errors)、完全性 (complete) 等を明確に要求し、偏りや性能劣化の抑制、説明責任の確保を求めている。位置付けとしては、EU 域内で法的拘束力を持つ上位規則であり、下位の実務文書（例：(3)）や各種ガイダンス（(4)～(6)）は本規則や GDPR 等の実装を補完する関係にある。

(2) 欧州委員会：Ethics Guidelines for Trustworthy AI

Ethics Guidelines for Trustworthy AI は、信頼できる AI の実現に向けた規制を補完する指針であり、2019 年に公表された¹¹。合法性・倫理性・堅牢性の三要素を前提に、7 つの要件（人間の監

¹⁰ EUR-Lex, Regulation EU 2024/1689 EN
<https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>

¹¹ Shaping Europe's digital future, Ethics guidelines for trustworthy AI
<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

督、技術的堅牢性、プライバシーとデータガバナンス等)を提示している。

データ品質は「プライバシーとデータガバナンス」の要件の中で重要事項として扱われ、データ品質確保やバイアス低減、データ管理の透明性を求めている(自己点検用の評価リストも提供)。位置付けとしては、(1)のような法規制の「上位法」ではなく、制度規制を補完しつつ、組織の自主的な設計・運用原則の共通基盤(ソフトロー)として機能する文書である。

(3) 欧州委員会：General-Purpose AI Code of Practice

General-Purpose AI Code of Practice は、AI Act における汎用目的 AI (GPAI) モデル提供者の遵守を支援する自発的な実務文書として、2025 年に公表された¹²。透明性、著作権、AI 安全性等の観点から、取り組むべき管理策と情報提供の方法を整理し、モデルの学習・評価に用いるデータや手法の文書化、リスク評価、インシデント対応等を通じて説明責任の確保を狙う。

データ品質は、(1)のように品質特性を条文で直接定義するというより、透明性・説明可能性のための文書化と管理プロセス(何をどのように学習・評価したかの開示可能性)として位置付けられている。位置付けとしては、(1)の法的義務(特に GPAI 関連)に対する下位の「実装手段(任意)」であり、事業者の実務水準の底上げを目的としている。

(4) 欧州データ保護監察機関：Guidance for Risk Management of AI systems

Guidance for Risk Management of Artificial Intelligence systems は、個人データを処理する AI システムに関する技術的リスクの特定と低減を支援するガイダンスであり、2025 年 11 月に公表された¹³。EU 機関等を主対象とし、リスク特定、評価、対策、監視のプロセスを提示している。

データ品質は、バイアスや説明可能性、セキュリティ等と並ぶ共通リスクの一要素として整理され、運用段階での継続的監視や変更管理(再評価)が重要視されている。位置付けとしては、AI 規制体系というより、データ保護(EU 機関向けの個人データ保護枠組み)に軸足を置きつつ、AI ガバナンスを補完する下位の実務ガイダンスである。

(5) 欧州データ保護会議：AI Privacy Risks & Mitigations – LLMs

AI Privacy Risks & Mitigations Large Language Models (LLMs) は、大規模言語モデルに関するプライバシー及びデータ保護リスクの管理方法を整理した報告書であり、2025 年 4 月に公表された¹⁴。データ収集、学習、推論、運用の各段階で生じるリスクを体系化し、識別・評価・低減の手順を提示している。

データ品質は、過剰収集や再同定、機密情報漏えい、偏り等のリスクと結び付いた管理対象と

¹² Shaping Europe's digital future, The General-Purpose AI Code of Practice
<https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>

¹³ European Data Protection Supervisor, Guidance for Risk Management of Artificial Intelligence systems
https://www.edps.europa.eu/data-protection/our-work/publications/guidelines/2025-11-11-guidance-risk-management-artificial-intelligence-systems_en

¹⁴ European Data Protection Board, AI Privacy Risks & Mitigations Large Language Models (LLMs)
https://www.edpb.europa.eu/our-work-tools/our-documents/support-pool-experts-projects/ai-privacy-risks-mitigations-large_en

して扱われ、技術的・組織的対策の例示により「適法・適切なデータ取扱い」を運用に落とし込んでいる。位置付けとしては、AI Act の下位文書というより、GDPR 系の実装（プライバシー・データ保護）を補完し、LLM 実務に適用するための参照資料である。

(6) 欧州データ保護会議：Bias Evaluation

Bias Evaluation は、AI のバイアスの発生源と評価方法を整理した技術文書であり、欧州データ保護会議の Support Pool of Experts の成果物として 2025 年 1 月に公開された（プロジェクト自体は 2024 年 3 月に完了）¹⁵。学習データの偏り、アルゴリズム上の偏り、評価過程の偏りを区別し、代表性の確保や評価用データセット設計等の重要論点を整理している。

データ品質は、バイアス評価の前提として「データの代表性」「文脈外利用」等の観点から扱われ、評価の設計（テストデータの妥当性）まで含めて実務者の参照枠を提供している。位置付けとしては、(5) 同様にデータ保護当局の監督・評価のための実務資料であり、(1) (3) の AI 規制・実装文書を「公平性／データ保護」側から補完する下位の技術文書である。

(7) 欧州連合出版局：EU Data Quality Guidelines

Data.europa.eu Data Quality Guidelines は、オープンデータの提供者向けに高品質データの作成・公開の推奨事項を整理したガイドラインであり、最新版は 2022 年に公開された¹⁶。FAIR 原則に沿って、データ及びメタデータの品質改善策を提示し、標準語彙の利用、文書化、更新・版管理等の実務項目が含まれている。品質指標・メトリクスの概観やチェックリストを備え、公開データの再利用性向上を狙う。位置付けとしては、(1) ～ (6) の AI 規制・データ保護の系統とは別に、公開データ提供を焦点としてデータ品質実務ガイドである。

また、米国、カナダ、英国、豪州では、政府情報・行政データの品質確保を軸に、データを「用途適合性 (fit for purpose)」で評価する枠組みが存在する。例えば、米国連邦総務庁 (GSA) の Data Quality Guidelines は、公衆配布情報を公開前レビューで審査し、有用性・信頼性・完全性に加え透明性・再現性も考慮して品質を担保する手続である。また、カナダの Guidance on Data Quality は、アクセス性等 9 次元で品質概念を統一し、データライフサイクル全体での実践として政策立案や自動化意思決定 (AI を含む) の信頼性を支えるものである。さらに、英国の Government Data Quality Framework は、政府横断の指針として品質次元とライフサイクル指針を提示し、具体的なツールも併せて継続的改善の運用を重視する枠組みである。

主な動向を下表に示す。

¹⁵ European Data Protection Board, AI: Complex Algorithms and effective Data Protection Supervision
https://www.edpb.europa.eu/our-work-tools/our-documents/support-pool-experts-projects/ai-complex-algorithms-and-effective-data_en

¹⁶ Publications Office of the EU, Data.europa.eu data quality guidelines
<https://op.europa.eu/en/publication-detail/-/publication/b601d9cc-b3c0-11ec-9d96-01aa75ed71a1/language-en>

表 2-3 米国、カナダ、英国、豪州のデータ品質に関する動向

No	取組名	データ品質との関連	実施主体	発行年（最終更新年）
1	Data Quality Guidelines	公開前レビュー基準（有用性/客観性/完全性）と透明性・再現性、訂正手続について規定	米国連邦総務庁	2002 年（2024 年）
2	Guidance on Data Quality	9 次元の品質定義とメタデータ/来歴/検証を含むライフサイクル実践について提示	カナダ政府	2024 年
3	Government Data Quality Framework	fit for purpose に基づく把握→記録→改善と継続モニタリングについて提示	英国政府	2020 年
4	Data Quality Guideline	品質次元の評価方法、改善計画、品質声明（テンプレ）について提示	オーストラリア・ビクトリア州政府	2019 年
5	Data Management Body of Knowledge	データ品質を含む知識体系（概念・役割・プロセス）について整理	DAMA International	2009 年（2017 年）

(8) 米国連邦総務庁：Data Quality Guidelines

米国連邦総務庁（GSA）の Data Quality Guidelines は、情報品質法（Section 515）に基づき、公開情報の品質を確保するための手続きを定めたガイドラインであり、2002 年に制定され、2024 年に更新されている¹⁷。

公開前審査により、有用性、客観性、完全性等の基準を満たすことを求められている。影響力の大きい情報については、透明性や再現性に配慮した説明と文書化を重視している。誤りが判明した場合の訂正手続や、利用者からの訂正要請への対応手順も規定されている。位置付けとしては、共通指針を各機関が実装するためのガイドラインであり、AI 個別規制というより政府情報の品質統制として上流で横断的に機能する文書である。

(9) カナダ政府：Guidance on Data Quality

カナダ政府の Guidance on Data Quality は、政府内で共通に用いるデータ品質の観点と実践策を整理したガイダンスであり、2024 年に公表された¹⁸。

データ品質を 9 次元（access, accuracy, coherence, completeness, consistency, interpretability, relevance, reliability, timeliness）で定義し、目的適合性（fit for purpose）として捉えつつ、データライフサイクル全体で考慮すべきことを明確化している。メタデータや標準の活用、データの制約・限界について利用者に周知すること等の推奨プラクティスを通じて、解釈可能性と再利用

¹⁷ GSA, Data Quality Guidelines

<https://www.gsa.gov/reference/reports/information-quality-guidelines/data-quality-guidelines>

¹⁸ Government of Canada, Guidance on Data Quality

<https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/information-management/guidance-data-quality.html>

性の向上を狙う。

(10) 英国政府：Government Data Quality Framework

英国政府の Government Data Quality Framework は、政府横断でデータ品質を理解・記録・改善するための枠組みであり、2020年に公表された¹⁹。

データ品質を目的適合性として捉え、データ品質活動は予防的（proactive）かつ証拠に基づき（evidence-based）、狙いを定めて（targeted）進めるべきことを示し、品質の把握・文書化・改善を一体として扱っている。組織文化の醸成、課題の源流対策、継続的モニタリングと報告を重視し、実務適用を支援している。

(11) オーストラリア・ビクトリア州政府：Data Quality Guideline

ビクトリア州政府の Data Quality Guideline は、共有・重要データ資産の品質を評価・維持・改善するための指針であり、2019年に公開された²⁰。データ品質標準と併せて位置付けられ、品質次元、評価方法、改善計画の作成手順を示す。データ品質ステートメント及び品質マネジメント計画のテンプレートを提供し、品質の可視化と説明責任をサポートしている。所有者、管理者等の役割分担を明確化し、ガバナンスの実装を促す。行政横断で一貫した品質管理を可能にする実務文書である。

2.2. 技術動向

2.2.1. AI の性能とセーフティを支えるデータ品質要件

生成 AI の登場により、AI システムの形態は大きく変化した。さらに RAG（検索拡張生成）や AI エージェントの普及に伴い、AI が利用するデータは学習用データセットに留まらず、推論時に参照するインターネット上の多様な情報、組織内のナレッジベース、外部ツール/API から得られる情報、ユーザ入力などへと大幅に拡張している。対象とするデータが膨大かつ多様化した結果、どの時点・どの経路で品質上のばらつきが生じるかを把握し、安定的に運用することが重要になっている。

このような環境では、データ品質は AI の性能やハルシネーション低減に直結する。併せて、誤情報の抑制、偏りの緩和、機密情報の混入防止など、AI を安全かつ安心に活用するための基盤としても機能する。運用の観点では、問題発生時に迅速に原因を切り分け、影響範囲を把握するために、データがいつ・誰により・どのように作成・更新されたかという来歴を追跡できる状態が望ましい。さらに、対外説明や監査対応を見据えると、品質判断の根拠を来歴と結びつけて再現可能な形で保存・提示できる状態が重要になる。これにより、性能・安全性に関する主張を「説

¹⁹ Gov.uk, The Government Data Quality Framework

<https://www.gov.uk/government/publications/the-government-data-quality-framework>

²⁰ vic.gov.au, Data Quality Guideline

<https://www.vic.gov.au/sites/default/files/2019-07/IM-GUIDE-09-Data-Quality-Guideline.pdf>

明可能な事実」]として提示しやすくなる。

さらに、AIによる自動処理・検索・統合を前提としたデータ設計の重要性も増している。人間が読むことを想定した文書形式やデータ構造は、AIによる自動処理に最適とは限らない。構造化の度合い、メタデータの充実度、API経由でのアクセス可能性など、機械可読性を考慮したデータ整備が、AIの出力の品質を底上げする。加えて、学習データや評価データとして合成データ（AIが生成したデータ）を用いるケースも増えており、現実のデータ分布との乖離や意図しないパターンの混入、生成元モデルのバイアスの継承等を踏まえた取り扱いが求められる。

また、医療診断支援、採用判断、信用スコアリングなど、説明や検証が重視される高リスク領域では、データ品質・来歴追跡・説明可能性をあらかじめ整備しておくことで、運用上の納得性や対外説明のしやすさが高まる。

総じて、データ品質はAIの出力品質を高め、運用を安定させるための主要因である。加えて、問題が起きた際の原因究明や説明、必要に応じた監査対応を支える基盤にもなる。

データ品質の概念自体も拡張している。従来のデータ品質は正確性や完全性が中心だったが、AIによって組織横断利用・再利用や自動処理が前提となる環境では、多次元的な品質要件として捉える必要がある。これらを満たすデータをAI-Ready Dataと呼ぶ。上述した来歴や機械可読性に加え、以下のような特性がある。

- 相互運用性：フォーマットや用語の意味が統一されているか、交換可能か。
- ラベル整合性：正解ラベルが付与されているか。その基準は一貫しているか。
- 公平性：特定の属性や集団に対するバイアスがなく、偏りのない判断を支えられるか。
- 規制準拠：機微な情報が適切に扱われているか。法規制や利用規約に従っているか。

なお、本資料では上記の観点も含めて広くデータ品質と捉える。

例えば、相互運用の意味的な一貫性が高いほど、RAGシステムの検索精度や引用の整合性が向上する。同じ概念が部署や自治体で異なる語で表現されている場合でも、用語やメタデータを整備することで関連情報を適切に検索でき、根拠に基づく回答生成を支えやすくなる。さらに、公平性に配慮したデータ整備は、AIによる判断の偏りを抑え、特定の集団に対する不利益が生じにくい運用につながる。

こうした観点を踏まえたAI活用には、技術的対策だけでなく、データのライフサイクル全体を通じた品質管理とトレーサビリティの確保が有効である。データ品質は、個別の技術課題に留まらず、組織的なガバナンス体制と一体として設計・運用されるべき基盤となっている。

2.2.2. データ品質管理の実装アプローチ

ここでは、データ品質管理を網羅的に論じるのではなく、AI活用の実務で参照されやすい実装パターンに着目し、関連する技術動向を概観する。具体的には、来歴や利用条件を含むメタデータを機械可読に整備し、検証・監視に接続する流れである。以下では、メタデータ記述、自動評価・観測という2つの軸から整理する。

(1) メタデータ記述

AI セーフティに資するデータ品質対策は、一般にデータそのものだけでなく、データの周辺情報（メタデータ）が整備対象になりやすい。特に重要となりやすいのは、実務上、次のような観点でメタデータを整理しておくことである。

（例）

- 来歴：データの出所、加工履歴、更新の経緯
- 品質：欠損率、整合性、更新頻度などの評価結果
- 利用条件：ライセンス、制約事項、想定される利用目的
- 責任主体：管理者、作成者、問い合わせ先

従来の説明文に加え、AI 活用のためには、これらを構造化して記述し、自動検証や運用監視につなげやすい形にすることが重要になる。

データカタログ標準である DCAT (Data Catalog Vocabulary) を基盤にしつつ、来歴記述の PROV-O (Provenance Ontology) と品質記述の DQV (Data Quality Vocabulary) を組み合わせることで、メタデータ自体を検証対象として扱うアプローチがある。さらに、人間可読のドキュメント整備だけでは、ソースデータに問題があった場合の影響範囲の把握が難しくなり得るため、機械可読で相互運用可能な形でメタデータを整備する方法も研究されている²¹。これには、品質評価結果そのものや、評価に用いた根拠と評価手順を、後から比較・再検証できるよう機械可読に保持する設計の議論も含まれる。

(2) 自動評価とオブザーバビリティ

データ品質は「良いはず」と主張するだけでなく、評価してはじめて管理できる。例えば前項で述べた来歴・品質・利用条件・責任主体といったメタデータは、整備状況そのものが AI 活用の土台になるため、機械的に評価できる状態にしておく取組が始まっている。こうした自動評価は、運用者の主観に依存しない客観性を確保し、第三者に対しても「なぜそう言えるか」を示す説明力を高める点で重要である。一例として、欧州では、Metadata Quality Assessment²²というツールが提供され、メタデータの DCAT Application Profile (DCAT-AP) への準拠性を SHACL (Shapes Constraint Language) によって検証している。

そもそも AI が参照・学習・統合するデータは、静的な資産ではなく継続的に変化することを前提に扱う必要がある。一方で、品質は悪くなってから気づく形では手遅れになりやすい。RAG や AI エージェントのように、外部情報やナレッジを継続的に参照する仕組みでは、参照先の更新停止、リンク切れ、語彙や定義の変更、内容の陳腐化などが、そのまま回答品質の揺れや誤引用につながり得る。そこで重要になるのが、(1) 設計段階から品質要件を織り込むことと、(2) 運用中も継続的に観測して変化を捉えることである。

この「常に観測し、異常の兆候を早期に捉え、是正につなげる」考え方は、近年、可観測性

²¹ 松澤 有三, 高橋 陽一, RAI のためのメタデータ整備手法の検討～機械学習用データセットに関する記述の機械可読性向上を目指して

https://www.jstage.jst.go.jp/article/jsaisigtwo/2025/SWO-066/2025_01/_article/-char/ja/

²² EU, Metadata quality

<https://data.europa.eu/mqa/methodology>

(Observability) として注目されている。単発の点検ではなく、実データとメタデータの状態を継続的に観測し、品質劣化の兆候（例：更新頻度の低下、必須項目の欠落、参照関係の破綻、語彙不整合の増加）を検知できる状態を作ることによって、AI の性能と信頼性を安定して引き上げやすくなる。

2.2.3. データ品質向上のための AI 活用

ここまでは、AI を活用するためにデータ品質を高める観点から技術動向を整理した。一方で近年は、データ品質を高めるために AI を活用するアプローチも現実的な選択肢になりつつある。

一般に、Text2SQL のように自然言語からデータベースへの問い合わせを支援する技術は、データへのアクセス性や探索性を高め、品質課題の発見・分析を支援し得る。また、従来のデータベース手法では扱いにくかった非構造化データ（文書、FAQ、議事録、ウェブページ等）についても、AI により構造化や要約、分類付与、矛盾・欠落の検知といった処理を支援できるようになってきた。これにより、品質管理の対象を「構造化データ中心」から「非構造化データを含む運用」へ広げ、改善サイクルを回しやすくする余地が生まれている。

ただし、AI は標準や品質管理の枠組みを置き換えるものではない。むしろ、標準（スキーマや語彙）があることで AI の出力を揃えやすくなり、実務に組み込みやすくなる。一方で現場には、標準や DCAT 自体の理解が難しいこと、仕組みがあっても運用できる人材が不足しがちなことというボトルネックがある。そこで本項では、AI が品質改善に寄与し得る使いどころと注意点を整理する。

(1) メタデータ整備支援

これまでに、来歴・品質・利用条件・責任主体といったメタデータを構造化し、評価・観測につなげる重要性を述べた。実務上の課題は、ここで必要になる標準（例：DCAT のスキーマや語彙）を理解し、記述を継続的に整備・更新できる人材に限られる点にある。その結果、データカタログ等の仕組みを導入しても、入力負荷の高さから記述が薄くなったり、表記ゆれや粒度の不統一が残ったりして、運用が形骸化しやすい。

このボトルネックに対して、LLM は、自然言語で存在する情報を、標準記述の下書きとして変換する役割を担わせやすい。例えば、仕様書、README、既存の説明文、データ定義書、運用手順書などから、DCAT 等のスキーマに沿った項目の候補を抽出・補完し、人が最終確認して確定する、という使い方である。これにより、標準の学習コストを下げつつ、メタデータ整備を継続運用に乗せやすくなる。

また、語彙付与や用語揺れの検出、異なる語彙間のマッピング候補提示など、意味整合の作業も支援対象となる。組織横断・分野横断でデータを活用するほど、同じ概念が異なる言葉で記述される状況が増えるため、AI により揺れの検出や候補提示を行い、統一語彙への寄せ又はマッピングを人が決める形が実務的である。

さらに、メタデータの不足や矛盾（例：利用条件の欠落、更新頻度の記載と実更新の不整合、参照先のリンク切れ）を検知し、レビュー観点を提示するといった形で、品質管理の見落としを減らす使い方も考えられる。評価・観測と組み合わせることで、整備→評価→是正のループを回

しやすくなる。

(2) 注意点：Human-in-the-loop と責任分界

AI をデータ品質向上に用いる際には、提案生成と最終判断の分担が論点となる。LLM は候補生成や差分検知、レビュー観点の提示を通じて整備作業を加速し得る一方、誤りや偏り、根拠不明な生成を含み得るためである。そのため近年は、重要な更新や公開に関わる判断では、人の確認・承認を挟む Human-in-the-loop (HITL) を前提に位置付ける整理が一般的になりつつある。加えて、誰が何を見てどの理由で承認したかを記録し、後から追跡できることが説明責任の中核となる。例えば EU AI Act²³では、運用中の出来事や人間による監視の内容を記録できることが規制・ガバナンス上の要請として位置付けられている。

併せて、LLM の出力を確定値ではなく下書き／候補として扱い、確定内容・判断者・参照根拠を記録して後から検証可能にする、といった運用設計も注目される。標準や評価ルールが曖昧な状態では出力のばらつきが増えやすいため、前節で述べた標準化・評価の枠組みと組み合わせて活用することが現実的とされる。

²³ EU Artificial Intelligence Act, Article 12: Record-Keeping
<https://artificialintelligenceact.eu/article/12/>

3.

データ品質を取り巻く論点の整理

本章では、AISI のデータ品質の取組である「データ品質マネジメントガイドブック」について紹介するとともに、本ガイドブックの内容も踏まえて、2025 年度のデータ品質 SWG 内で浮かび上がった論点とその論点を踏まえた 2025 年度の活動の概要を紹介する。

3.1. データ品質マネジメントガイドブック

国内の取組では、AI 法や AI 事業者ガイドライン等において、AI 活用を安全に促進させるために必要な観点は取りまとめられている。一方、その中でも、AI セーフティのためにはデータが基盤であることに注目し、データ品質の観点を解説した文書や、各観点を具体的に検討するためのガイドラインやフレームワークが少ない。特に、AI のためのデータ品質の観点までを解説した文書は少ないことが課題であった。そこで、様々なデータ品質に関する国際標準を参照してデータ品質の取組に関する考え方を網羅的に整理した『データ品質マネジメントガイドブック』（データ品質ガイド）が AISI より公表された。

データ品質ガイドは、AI とデータの価値最大化に必要なデータ品質を継続的に確保するための考え方と実践事項を整理したものである。具体的には、データ品質を AI の卓越性と信頼性の基盤として位置付け、プロセス、ガバナンス、品質特性の観点から管理体系を示している。

データ品質 SWG では、データ品質ガイドの Ver1.0 のレビューを行い、そのレビュー結果を踏まえて Ver1.01 が 2025 年 12 月に AISI より公表された。今後も、データ品質 SWG のメンバーやデータ品質ガイドの利用者のコメントを踏まえて、データ品質ガイドは継続的に更新されていく想定である。

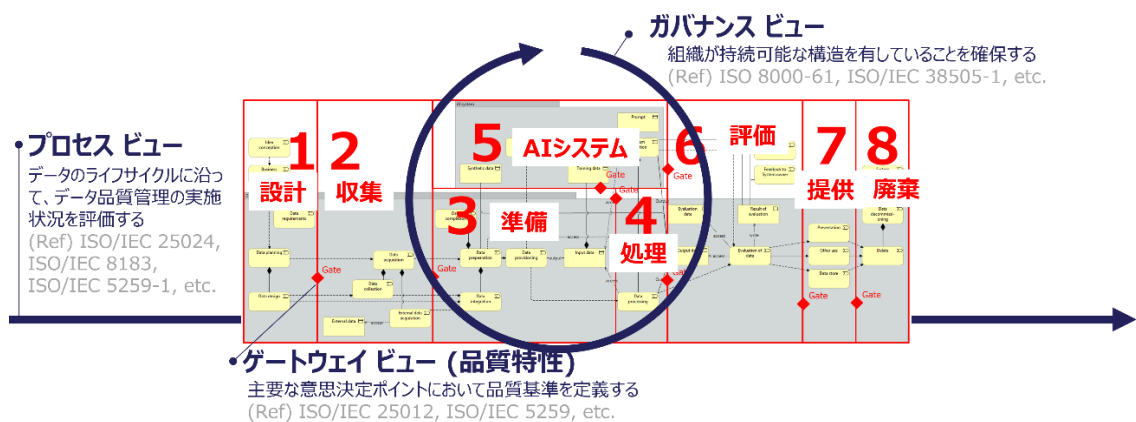


図 3-1 データ品質マネジメントガイドブックの概要イメージ

3.2. 2025 年度の活動で浮かび上がった論点

2025 年度のデータ品質 SWG では、SWG メンバーによるデータ品質ガイド Ver1.0 のレビュー

をはじめとした様々な議論を SWG 内で実施した。レビューにおいては、データ品質ガイドが以下の観点で評価された。

- ISO 等の国際標準をベースにした包括的フレームワークであり説得力と網羅性がある。
- 「プロセス／ガバナンス／品質特性」の3視点構成やライフサイクルの網羅性は有用性が高い。
- ハイレベルな記述粒度でのガイドブックと見れば非常に整理されたドキュメントである。
- 意識すべき観点がおよそ全方位的に語られている。
- 「Garbage in, Garbage out」を前提としたコンセプトの明確さから、AI/データ活用の成否がデータ品質にかかっているという本質的な重要性を示している。
- AI の最適な活用に向けて教師データ等の外部調達を含め品質管理するというコンセプトに共感できる。
- 体制やリソースの必要性について言及されている点が現場に即している。
- 教育資料や監査などの網羅性の高い利用ケースにおいて活用可能だと考える。

また、レビューを含めた議論を通して、評価された点以外にデータ品質の取組促進に関して 5 つの論点が浮かび上がった。各論点の概要を以下に示す。

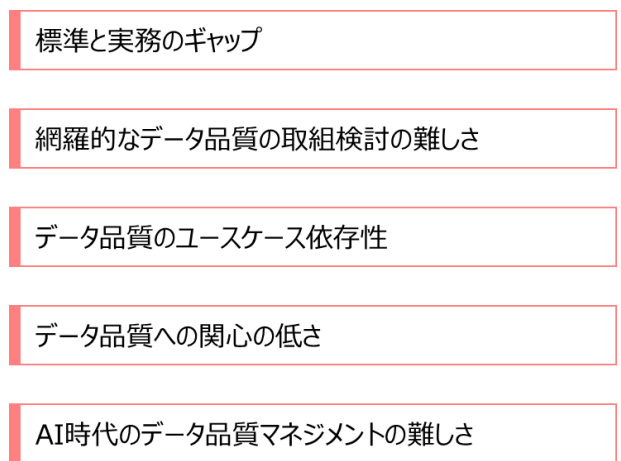


図 3-2 データ品質 SWG で検討された論点

1. 標準と実務のギャップ

データ品質 SWG において、データ品質ガイドのレビュー時に標準と実務のギャップに関する意見が挙がった。現状、データ品質に関しては、ISO8000 シリーズや DMBOK をはじめとする国際標準や体系的なフレームワークが整備されている。しかし、多くの組織では、これらの標準は理論的には理解されていても、実務レベルで十分に適用されていないという課題が存在することが複数のメンバーより指摘された。標準は汎用的な概念として説明されている一方で、個別の事情を持つ現場への概念の落とし込みは簡単に実施することは難しい。

その結果、国際標準や体系的なフレームワークは整備されているものの、現場での実装に至らないケースが多い。例えば、メタデータの管理方法として標準体系を利用すること

が推奨されている中でも、DCAT や Dublin Core などの複数の実装方法があり、利用シーンに応じた適切な実装方法を選択することが難しい。

標準と実務のギャップを埋めるためには、標準をそのまま適用するのではなく、実務者が標準の実装方法を想起できるような取組が必要である。データ品質 SWG では、実際の現場で行われている取組と、その実装上の課題・解決方法を整理して文書化することが、実務者の参考になるという意見があった。

2. 網羅的なデータ品質の取組検討の難しさ

データ品質 SWG において、データ品質取組の網羅的な確認の必要性に関する意見が挙げられた。具体的には、データ品質は単一概念ではなく、複数の品質特性を確保するための取り組みの必要性や技術・組織・ライフサイクル全体の運用など複数のレイヤでの取り組みが必要であり、どの取組を優先的に実施すべきか判断することが難しく、全体像を整理しないまま部分最適的な施策に終始してしまうケースが多い。この背景には、複数の品質特性やレイヤの取組を体系的に整理するフレームワークが組織内で確立されていないことなどの意見がメンバーより挙げられた。

例えば、データ品質の取組を実施することが実務者に求められた場合、自社の AI システム特有のリスクに係る技術的な取組のみを検討してしまい、組織や運用の観点での網羅的な確認ができず、取組に抜け漏れが発生してしまうことが考えられる。

現状、データ品質確保に向けた取組が網羅的に整理されている国内の取組は少ない。AISI のデータ品質ガイドにおいても、網羅的にデータ品質の取組が整理されているが、各組織において取組を簡易的に確認することが難しい。そのため、データ品質 SWG では、データ品質ガイドをベースに、データ品質に関する取組状況が簡易的かつ網羅的に確認できるツールが必要ではないかという意見が挙げられた。

3. データ品質のユースケース依存性

データ品質 SWG において、ユースケースの利用目的や利用データなどの特性によって、重要なデータ品質特性が変わるため、取組を適切に実施することが難しいという意見が挙げられた。AI などの技術によって、データ活用が高度化・多様化しているため、ユースケースによるデータ品質特性の違いが顕著となっている。

例えば、AI モデルの学習においては、一般的にラベルの正確性や偏りの少なさが重要であるが、経営ダッシュボードの利用目的を検討した際には、集計値の整合性が極めて重要である。この違いを整理しないまま一律に基準を適用すると、過剰品質によるコスト増大や、逆に品質不足によるリスク拡大が生じる。

したがって、解決にはユースケース起点で重要となる品質特性を定義するアプローチが必要である。データ品質 SWG では、利用シナリオごとに必要な品質特性を明確化することを目的に、ユースケースの特性（ユースケースの目的、利用データ、関連ステークホルダーの識別、ユーザの懸念・期待など）に応じて、重要となるデータ品質の取組を整理することが必要ではないかという意見が挙げられた。

4. データ品質への関心の低さ

データ品質 SWG において、データ品質は「問題が起きたときに対応するもの」と認識されがちであり、費用対効果がわからず後回しにされるケースがあるという意見が挙げられた。その背景には、品質向上の効果が直接的な売上や利益として可視化しにくいこと、また品質問題が潜在的リスクとして認識されにくいことがある。その結果、品質改善への投資は後回しとなり、障害発生時の対症療法的対応が繰り返される。

例えば、分析結果の誤りが発覚して初めてデータ不備が問題視される、システム統合時にデータ不整合が顕在化して修正コストが膨らむといった事例が挙げられる。これは品質が見えないコストとして蓄積されている状態である。

この課題を解決するには、まず品質問題による損失やリスクを定量的に可視化し、必要となる取組とセットで費用対効果を示すことが重要である。データ品質 SWG では、インシデント事例の整理、修正工数の算定、機会損失の推計などを通じて、品質の経営インパクトを明示することが必要ではないかという意見が挙げられた。さらに、品質指標を KPI に組み込み、経営層を巻き込んだガバナンス体制を構築することで、品質を戦略テーマとして位置付けて、文化醸成（人材育成含む）と制度設計の両面からのアプローチを支援することの重要性も意見として挙げられた。

5. AI 時代のデータ品質マネジメントの難しさ

データ品質 SWG では、進化の早い AI 技術において、データ品質マネジメントを事業者において実施することの難しさに関する意見が挙げられた。AI の社会実装が進む中で、データ品質の重要性は一層高まっている。一方で、AI 技術の進化が早く、RAG や AI エージェントなど次々と出てくる新たな技術に対するリスクに対応する必要がある。また、AI の社会実装が進む中で、AI がフィジカル空間に影響を及ぼすユースケースもあるため、AI の透明性確保の必要性が高まっており、併せてデータの透明性確保も重要となっている。

例えば、特定属性に偏った学習データにより不公平な予測が生じるケースや、誤ったラベル付けがモデル誤学習を引き起こす事例が報告されている。また、生成 AI の活用では入力データの信頼性や出力結果の妥当性評価も新たな課題となる。

解決には、データ収集から前処理、学習、運用、再学習に至る、組織をまたがったデータのライフサイクル全体での品質マネジメントが必要である。データ品質 SWG では、データセットの由来や加工履歴を管理するメタデータ整備、バイアス検証、モデル性能の継続的モニタリングなど、AI 技術の進化に合わせて、重要な取組を支援できるドキュメント・ツールの必要性について意見が挙げられた。

4.

データ品質 SWG における 2025 年度の活動

データ品質 SWG では、3.2 で紹介した各論点の課題解決を目指して活動を行った。2025 年度としては、「AI 時代のデータ品質マネジメントの難しさ」以外の課題解決に資する「ユースケースの整理」、「チェックリストの作成・検証」、「官報データ構造化と AI 出力品質の検証」の 3 つの活動を行った。各活動と論点との対応関係を図 4-1 に示す。また、対応関係は、2025 年度の活動のみで論点の解決ができることを示しておらず、解決に向けた第一歩であり、今後も継続した活動が必要である。

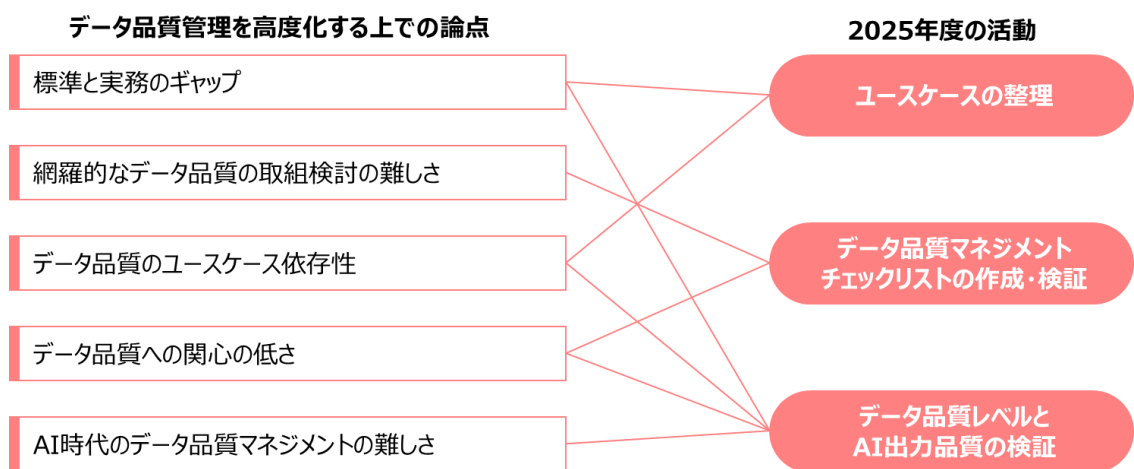


図 4-1 2025 年度の活動と論点の対応関係

各活動の概要を以下に、具体的な内容を本章の各節に示す。

1. ユースケースの整理

実務の中でユースケースに応じてデータ品質をどのように扱うべきかを論じるため、ユースケースの特性に応じた必要なデータ品質の取組整理を試行した。複数のデータ活用のユースケースを対象に、ユースケースにおけるリスク、リスクを踏まえて重要となるデータの品質特性を特定した。複数の分析結果を確認したうえで、ユースケースのどのような特性（事業分野・活用データなど）がデータ品質の考え方に影響を及ぼすかを確認した。

2. データ品質マネジメントチェックリストの作成・検証

データ品質の取組を簡易かつ網羅的に確認するための手段として、データ品質ガイドのプロセスビューをベースにチェックリストを作成した。このようなツールが現場でのデータ品質管理の検討材料となり、関心が高まることを期待する。作成したチェックリストについては、データ品質 SWG、分野別 SWG において適用検証した。

3. データ品質レベルと AI 出力品質の検証

AI 時代のデータ品質を具体的に検証するため、データ品質の違いが生成 AI の応答と安全性に与える影響を確認した。具体的には、高付加価値な公共データとして官報データを取り上げ、データ構造化と AI 出力品質の関係性を検証した。官報データは正確である一方で、人間が見ることを前提とした構造のため、AI のためのデータ整備が必要という点で検証対象として適当であると設定した。

4.1. ユースケースの整理

「標準と実務のギャップ」、「データ品質のユースケース依存性」の論点の課題解決に向けた以下のような意見がデータ品質 SWG より確認された。

- データ品質ガイドだけでは理解が難しく、ユースケースを集めて解説できるとよいのではないか。
- ユースケース別に必要となるデータ品質の取組が変わるのではないか。
- ユースケースを収集することで、ユースケースのパターンに応じて必要なデータ品質の取組が整理できるのではないか。

上記の意見より、ユースケースに応じて必要なデータ品質の取組を整理することが重要であることが確認できた。整理することで、「標準と実務のギャップ」の観点では、データ品質ガイドの内容を具体的に各ユースケースに適用することが可能であり、実務者にとってより参考となる情報を提供できると考えられる。また、「データ品質のユースケース依存性」の観点では、ユースケースをパターン化し、各パターンの主たるデータ品質の取組を整理することで、実務者にとってデータ品質を検討する際の第一歩になると考えられる。

データ品質 SWG の活動として、ユースケースに応じて必要なデータ品質の取組のパターンを整理することを目指す。2025 年度は、第一歩として、ユースケースに応じた重要な品質特性を整理した。また、ユースケースで実際に実施されているデータ品質の取組を確認した。

4.1.1. ユースケースの整理アプローチ

ユースケースに応じた品質特性の整理を行う上では、前段階としてユースケースで懸念されるリスクを特定することが重要ではないかというデータ品質 SWG 内の議論があった。このような議論とデータ品質ガイドの内容を踏まえて、ユースケースに応じた重要な品質特性の整理を図 4-2 に示すアプローチで実施した。具体的に「懸念されるリスクの特定」、「確保すべき品質特性のマッピング」の 2 つのステップの詳細を以下のとおり示す。



図 4-2 ユースケース整理のアプローチ

(1) ステップ1：懸念されるリスクを特定

ステップ1では、ユースケースの内容（データ活用目的、活用されているデータ、ユースケースの提供先など）を踏まえて、以下のとおりにデータ品質ガイドで整理されている「低品質データによって生じるリスク種別」から特に懸念されるリスクを選択する。選択した際には、特に懸念される理由についても記載する。

表 4-1 リスク種別

リスク種別	リスク概要
意思決定の誤り	不正確なデータや不完全なデータは、分析に欠陥をもたらし、その結果、組織のあらゆるレベルにおいて、誤った結論、戦略的な失策、不適切な意思決定を招く可能性がある。
業務効率の悪さ	低品質なデータのクリーニングと修正に時間とリソースを割かなければならず、AI モデルの展開が遅れ、運用効果が低下する。
顧客満足度の低下	不正確または不完全な顧客データは、パーソナライゼーションがうまくいかず、期待に応えられず、サービスや商品に対する信頼を低下させる可能性がある。
コスト増	低品質なデータは、エラー修正、再処理、AI モデルの再トレーニング、誤った予測や推奨による潜在的な金銭的損失によってコストを増加させる。
法的規制リスク	データ保護に関する法律や規制（GDPR、CCPA など）を遵守しない場合、多額の罰金や訴訟につながる可能性がある。
風評被害	低品質なデータによるエラーは、顧客、パートナー、利害関係者との信頼と信用を損ない、長期的なブランド低下につながる可能性がある。
競争上の不利	より質の高いデータを持つ競合他社は、顧客洞察力、業務効率、市場対応力などの主要分野で優れた業績を上げ、組織は遅れをとる。
機会損失	低品質なデータによって洞察やトレンドが見落とされると、新しい市場機会をつかめなかったり、効果的なイノベーションができなかったりする。
モデルのパフォーマンス低下	低品質なデータで訓練された AI モデルは、バイアスや信頼性の低さ、有害な振る舞いを示す可能性があり、倫理的な懸念や実世界での応用における有効性の低下につながる。
セキュリティ・リスク	低品質のデータには、悪意ある行為者が悪用できる脆弱性やエラーが不注意に含まれている可能性があり、潜在的なセキュリティ侵害や機密情報の悪用につながる。
ステークホルダーの不信	データ品質が一貫して信頼性と結果を損なう場合、社内チームや社外の利害関係者は AI システムに対する信頼を失う可能性がある。

(2) ステップ2：確保すべき品質特性のマッピング

ステップ2では、ステップ1で特定したユースケースにおいて特に懸念されるリスクを防ぐた

めに確保すべき品質特性を整理する。品質特性は、以下のとおりデータ品質ガイドにおいて整理されている内容を活用する。また、確保すべき品質特性がどのようにリスクを防ぐかについても記載する。

表 4-2 データ品質特性

品質特性	品質特性の概要
正確性	データが現実世界の値をどの程度正しく反映しているかの度合い
完全性	必要なデータ項目がどの程度漏れなく揃っているかの度合い
一貫性	データの値がデータセット間および時間を通じてどの程度統一されているかの度合い
信憑性	データソースがどの程度信頼できるかの度合い
最新性	データがどの程度最新の状態に保たれているかの度合い
アクセシビリティ	認可されたユーザやシステムがデータにどの程度支障なくアクセスできるかの度合い
標準適合性	データ管理が法令、規制、業界標準にどの程度準拠しているかの度合い
機密性	機密データが不正アクセスや漏えいからどの程度保護されているかの度合い
効率性	品質を損なうことなく、どの程度少ない時間や資源でデータ管理を行えるかの度合い
精度	データがどの程度細かい粒度や詳細さを備えているかの度合い
追跡可能性	データの起源、変換履歴、利用状況をどの程度追跡できるかの度合い
理解性	データが人間および機械にとってどの程度正しく解釈しやすいかの度合い
可用性	必要なときにデータへどの程度アクセスできるかの度合い
移植性	データをプラットフォーム、システム、環境の間でどの程度円滑に移転できるかの度合い
回復性	障害や故障発生時にデータをどの程度迅速かつ正確に利用できるかの度合い
監査可能性	データが監査済みであること、または監査のために関係者が利用可能である度合い
均衡性	データが関連するカテゴリをどの程度偏りなく表現しているかの度合い
多様性	データセットがどの程度幅広い視点、状況、変動を含んでいるかの度合い
有効性	データが特定の AI タスクでの利用要件を満たしているかの度合い
識別可能性	機微情報や個人データがどの程度識別されうる状態にあるかの度合い
関連性	データがユースケースに合致し、必要かつ意味のある情報のみをカバーしているかの度合い
代表性	データが現実世界の母集団や想定される状況をどの程度適切に反映しているかの度合い
類似性	データセット内のサンプル同士がどの程度類似しているかの度合い
適時性	現象が発生してからそのデータが利用可能になるまでの遅れが適切である度合い

ステップ 1,2 の実施イメージを以下に示す。

サービスの概要					品質特性	顧客満足度の低下	風評被害	モデルのパフォーマンス低下	品質特性とリスクの対応理由
ユースケース	概要	活用目的	活用データ	AI					
不正取引防止サービス	不動産掲載データの活用によりECサイト等での後払いにおける不正取引の防止	リスクコンプライアンス	表形式	予測型AI	正確性				不動産掲載データに不足があった場合は、不正取引が防止できない可能性がある。
					完全性	○			
					一貫性	○			…
					信憑性	○			…
					最新性				
					アクセシビリティ				
					標準適合性		○		…
					透明性		○		…
					検閲可能性	○	○		…
					追跡可能性		○		…
					理解性				
					可用性			○	…
					移植性			○	…
					回復性				
					監査可能性		○		…
					均等性				…
					多様性				…
					有効性				
					識別可能性				
					関連性			○	…
					代表性				
					類似性			○	…
					適時性				

図 4-3 ステップ 1,2 の実施イメージ

4.1.2. ユースケースの整理例

(1) AI 活用に向けた合成データ生成

■ ユースケースの概要

本ユースケースは、AI でデータを利用する際に参照データをそのまま用いると信頼性の観点で問題が生じるという課題を解決することを目的に、参照データのオントロジー的な要素をナレッジ化し、合成データを生成するという研究開発事例である。

具体的な活用事例としては、人流データや取引・会計データ等の特徴量をもとに合成データを生成し、それをを用いて AI での予測シミュレーションに利用することなどが挙げられる。また、合成データの生成過程におけるデータ品質上のリスクコントロールをタイムリーにデータアシュアランスする仕組みも検討されている。

本ユースケースのデータ品質の考え方として、データ要件に沿った信頼できる合成データが作成できているかという点が重要視されている。

■ 重要なリスク種別

本ユースケースでは、「顧客満足度の低下」、「法的規制リスク」、「ステークホルダーの不信」の3点が挙げられた。

「顧客満足度の低下」においては、信頼性のある合成データの提供という目的を踏まえて、懸念が高いとされた。具体的には、合成データが参照データの特徴や統計を保持できていない場合に、それをを用いた AI シミュレーションの結果が、参照データに基づくものと大きく異なってしまう、顧客満足度を著しく毀損してしまうリスクがある。

「法的規制リスク」においては、参照データに基づいた合成データ提供というサービス概要を踏まえて、懸念が高いとされた。具体的には、参照するデータが様々であり、各参照データの法的規制に遵守できていない場合、合成データが生成できないリスクがある。

「ステークホルダーの不信」においては、本ユースケースでは合成データの利用者・参照データの提供者などステークホルダーが多いことを踏まえて、懸念が高いとされた。具体的には、生成した合成データの信頼性や品質が期待どおりに確保されておらず、参照データの提供者や合成データ利用者の権利や信頼を毀損してしまうリスクがある。

■ 重要な品質特性

本ユースケースでは、「正確性」、「信憑性」、「標準適合性」、「精度」、「監査可能性」、「有効性」、「識別可能性」の7点が挙げられた。

「顧客満足度の低下」というリスクにおいては、合成データの信頼性・有効性を確保することが重要である。そのため、生成した合成データや参照データの特徴・統計の保持に対して「正確性」・「精度」、有用な合成データの生成に対して「有効性」の確保が必要であることが確認できた。

「法的規制リスク」というリスクにおいては、参照データで懸念されている法的規制リスクがある場合に、それを合成データ化により回避できているかを確認することが重要である。そのため、参照データの法的規制リスクのクリアに向けて「信憑性」「標準適合性」、参照データの機微情報が合成データに含まれていないかの確認に向けて「識別可能性」の確保が必要であることが確認できた。

「ステークホルダーの不信」というリスクにおいては、合成データの利用者・参照データの提供者に対して合理性のある説明可能性を確保することが重要である。そのため、生成した合成データにおいて、ステークホルダーが懸念する機微情報が合成データに含まれていないかの確認のリスクコントロールのために、「監査可能性」の確保が必要であることが確認できた。

コラム：ユースケース(1) データ品質の取組

本ユースケースは信頼される価値創造を促進する「データ・AI スペース」実現に向けた以下2つの重要課題を解決することを目指している。

- ・ 課題①：現実世界のデータの安全安心な流通・利活用
- ・ 課題②：流通・利活用するデータの信頼性の客観的評価と価値づけと制御連携

課題①の解決に向けて、実空間データの活用制約（個人情報・知財・セキュリティ）を回避するために、シミュレーションで活用できる合成データの生成に取り組んでいる。

課題②の解決には、上記の合成データ生成において、実空間データの取得から AI シミュレーション、アウトプットの活用、さらにその成果をビジネス価値へ転換、成果配分まで、一連の流れの全体を通じた品質管理と信頼形成が不可欠である。

そのため、本ユースケースでは、以下の2つのデータ品質の取組を組み込むことで、「安全・安心に使い続けられる信頼性と品質が確保されたデータ・AI スペース」を支える基盤機能の確立を目指している。図 4-4 に、下記2点の取組を組み込んだイメージを示す

- ・ データ品質・信頼性の客観的でタイムリーなモニタリングと開示
- ・ 評価結果を踏まえたコントロール・改善へのフィードバック

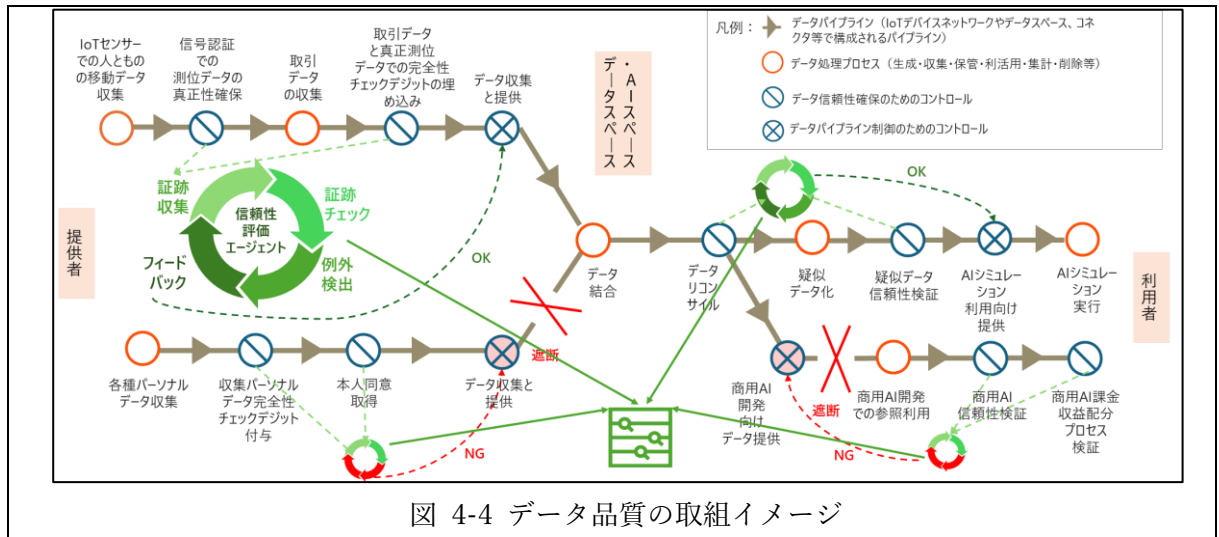


図 4-4 データ品質の取組イメージ

(2) 小売業におけるサプライチェーンマネジメント（SCM）システム

■ ユースケースの概要

本ユースケースは、小売事業者が自社の SCM（出荷管理・財務・商品管理）領域においてデータ活用を推進している事例である。予測型 AI を活用した「運用高度化・安定化」を目的としたサービスであり、現在は本格運用・拡大フェーズに移行している段階である。

具体的な活用事例としては、SCM 関連データを製造委託先・販売委託先へ提供し、取引先・サプライチェーン全体での業務効率化及び在庫の最適化を実現することなどが想定される。

本ユースケースのデータ品質の考え方として、サプライチェーンを通じた様々な企業のデータを利用し、かつ様々な企業のデータを管理しているため、データの揺れやデータの統合が正しくできているかという点が重要視されている。

■ 重要なリスク種別

本ユースケースでは、「意思決定の誤り」、「顧客満足度の低下」、「風評被害」、「ステークホルダーの不信」の 4 点が挙げられた。

「意思決定の誤り」においては、データの間違った処理が自社・他社共に意思決定のミスリードにつながる可能性がある。具体的には、誤ったデータ結合やクレンジング不足によって、精度の低いデータがサプライチェーンの様々な企業に提供され、誤った意思決定がされてしまうリスクがある。

「顧客満足度の低下」においては、本ユースケースは在庫適正化など直接コスト負担につながる可能性がある。具体的には、不完全なデータや予測データが、小売事業者のサービスや商品のスピーディーな提供や品質のばらつきを発生させ、顧客満足度を著しく毀損してしまうリスクがある。

「風評被害」においては、特定分野の事業者を対象とした製品かつ「運用高度化・安定化」フェーズであるため、サービスの信頼性が損なわれることで、長期的に特定分野でのブランドが著しく低下する可能性がある。具体的には、「風評被害」が業界内で広まり、複数の企業における本ユースケースの利用機会を喪失するリスクがある。

「ステークホルダーの不信」においては、サプライチェーンを通して活用されるユースケースであり、複数の事業者からの信頼が損なわれる可能性が高い。具体的には、品質が低いデータによって予測精度が下がった場合には、サプライチェーンへの供給不足などに繋がり、サプライチェーンを通してすべてのステークホルダーの信頼を著しく毀損してしまうリスクがある。

■ 重要な品質特性

本ユースケースでは、「精度」、「信憑性」、「最新性」、「正確性」の4点が挙げられた。

すべてのリスクを通して、SCM サービスの予測型 AI の精度に直接影響を及ぼす品質特性が重要であることが確認できた。「精度」の観点では、データ精度のチェックやサービスへの反映タイミングによって、予測精度に影響を及ぼす可能性がある。「信憑性」の観点では、データソースの特定と検証が重要であり、間違ったデータの入力予測精度に大きな影響を及ぼす。「最新性」の観点では、リアルタイムの予測精度において最新版のデータを活用することが重要である。「正確性」の観点では、データ加工結果の検証・予測データの正答率検証が予測精度に影響を及ぼすため、重要である。

コラム：ユースケース(2) データ品質の取組

本ユースケースにおけるデータ品質の課題として、サプライチェーンを通して組織・部門・システムごとに個別最適化された状態であり、データのサイロ化が広範囲に発生しているケースが多い。また、データ構造が組織・部門・システム横断での活用を想定した設計になっておらず、複数の部門や委託先をまたいだデータ連携・統合が困難な状況にあった。これらの課題は、サプライチェーン全体での予測精度向上や意思決定の高度化を妨げる根本的な障壁となっていた。

上記課題の解決に向け、データ活用領域を「抽出領域」「蓄積領域」「統合領域」「出力領域」の4層に整理し、データ統合基盤内での一元管理体制を構築した。この基盤整備を通じて、以下の取組を実現した。

- 汎用的なデータモデリングと標準化による品質向上
目的ごとに汎用的なデータモデリングを整備し、データ抽出・統合ロジックの標準化を推進。これにより、組織・部門・システム間でのデータ定義の揺れを解消し、データ品質の底上げを実現した。
- 高度な機械学習による需要予測の実現
データ統合とデータ品質の向上を基盤として、高精度な機械学習モデルによる需要予測を実現した。サイロ化解消により、複数部門・委託先のデータを横断した予測が可能となった。
- 信頼できる統合データのデータ HUB としての配信
品質が担保された統合データをデータ HUB として各業務システムへ配信する仕組みを構築。製造委託先・販売委託先へのデータ提供において、一貫性・信頼性の高いデータ流通を実現した。
- メタデータの管理運用
データの定義・来歴・更新状況等のメタデータを体系的に管理・運用することで、デー

データの透明性とトレーサビリティを確保。データ利用者が必要な情報を的確に把握できる環境を整備した。

(3) 自動車の事故発生予測 AI

■ ユースケースの概要

本ユースケースは、実車走行データを基にシミュレーションで事故シナリオ（割り込み・右折等）を多数生成し、画像・映像と位置・速度/加速度などのセンサーデータを組み合わせたマルチモダル構成を用いて、事故発生を予測する事例である。本事例は、予測型 AI を活用した「運用高度化・安定化」を目的としたサービスであり、企画～研究開発フェーズである。

上記では、事故発生を想定しているが、本シミュレーションは、運転技術診断などにも利用されている。本ユースケースのデータ品質の考え方として、自動車に関わるサービスであり、一般市民を含めた多くのエンドユーザに直接影響を及ぼすため、風評被害やサービスへの不信が生まれないかという点が重要視されている。

■ 重要なリスク種別

本ユースケースでは、「意思決定の誤り」、「風評被害」、「顧客満足度の低下」、「機会損失」、「モデルのパフォーマンス低下」、「ステークホルダーの不信」の6点が挙げられた。

「意思決定の誤り」においては、データの間違った処理が運転者の意思決定におけるミスリードにつながる可能性がある。具体的には、誤ったデータによるシミュレーションによる、予測精度の低下によって、事故発生の未来予測が外れてしまい、誤った意思決定がされてしまうリスクがある。

「風評被害」、「顧客満足度の低下」においては、本ユースケースは事故発生の未来予測のため、予測エラーはエンドユーザの不評を招き風評被害をもたらす可能性がある。具体的には、不完全なデータや予測データが、直接的にエンドユーザの意思決定に影響を及ぼし、ユーザの顧客満足度を著しく毀損してしまうリスクがある。

「機会損失」、「モデルのパフォーマンス低下」においては、シミュレーションを通じて自動車走行シーンを生み出すため、現実には即したデータが生成できないと、効果的な開発ができない可能性がある。具体的には、シミュレーションデータが現実と即していない場合、事故予測などのサービスの開発効率が大きく下がるリスクがある。

■ 重要な品質特性

本ユースケースでは、「精度」、「完全性」、「一貫性」、「信憑性」、「正確性」、「追跡可能性」、「監査可能性」、「均衡性」、「多様性」、「有効性」、「代表性」、「類似性」、「適時性」の13点が挙げられた。

すべてのリスクを通して、本ユースケースの最も重要な要素であるシミュレーション作成において必要となる品質特性であることが確認できた。「精度」、「完全性」、「一貫性」、「信憑性」、「正確性」、「均衡性」、「多様性」、「有効性」、「代表性」、「類似性」、「適時性」の観点では、シミュレー

シミュレーションで生成されるデータが物理現象や現実の法規、発生頻度に即していないデータが生成され、正しく学習が行われない可能性があるため、重要である。「監査可能性」、「追跡可能性」の観点では、直接的にエンドユーザに影響を及ぼすサービスであり、サービスの前提となるシミュレーションで生成されるデータがどのような前提や過程を通じて生み出されているかの説明を求められる可能性があるため、重要である。

コラム：ユースケース(3) データ品質の取組

本ユースケースでは、シミュレーションデータを学習に用いている。効果的な学習をするために、オリジナルの車両走行データを踏まえて、様々なバリエーションのシミュレーションデータを生成している。シミュレーションデータの生成にあたって、シミュレーションデータの品質向上に関する取り組みについて紹介する。

まず、現実収集した一連の時系列データを用意し、これをオリジナルのシナリオデータとする。このシナリオデータのバリエーションを作成し、多量の学習用データを生成しようとする意図である。バリエーションの取り方として、以下が検討される。

- オリジナルの車両走行データの動作パラメータを変更

具体例として、車線変更の場面を記録したシナリオデータを正確にトレースしたシミュレーションデータを作成し、このシミュレーションツール上で動く物体の動作パラメータを変化させることでバリエーションを作成する。このようにすることで、ある程度現実に即した品質のデータ作成が可能となる。

- オリジナルの車両走行データに新たな挙動の追加

オリジナルの車両走行データに意図的な動作を加える場合、その変化が急峻になりすぎないように、車線変更動作前後の速度変化を設定し、この間を数学的な曲線補間手法を用いてデータ点列を生成する方式を用いて不自然な動きにならないような工夫も行う。

また、人目によるチェックも重要な品質向上の取組である。全数検査は現実的でない場合もあるが、一部であってもシナリオデータの開始から終了までの一連の動きを人目でチェックし、不自然なところがないかをチェックすることも有効である。

4.1.3. 得られた示唆・今後の改善策

2025年度の活動では、合計3件のユースケースを対象に重要な品質特性の整理を実施した。実施した結果とその結果を踏まえたデータ品質 SWG での議論を通して、以下の点が示唆として得られた。

- 自社のユースケースに応じて重要な品質特性を整理・分析するという観点では、活用しやすいフレームワークであると考えられる。
- ユースケース⇒リスク⇒品質特性を順番に整理することで、自社のユースケースにおいて重要視すべきリスクや品質特性が可視化され、実施すべきデータ品質の取組が明瞭化されることが考えられる。
- 目的ベースでユースケースの重要な品質特性を整理するうえでは、利用しやすい分析手法である。

一方で、改善点につながる以下の点に関する示唆も多く得られた。特に重要な品質特性の特定に向けて検討したフレームワークについて、品質特性にとどまらず、対応策まで示せるフレームワークを準備する必要性に関する意見が多くあった。

- リスク種別や品質特性について、選択肢の粒度が様々かつ重複もあるように感じられる。
- 今回のフレームワークはユースケースの目的ベースでの整理という点では活用しやすい一方で、ユースケースのデータ種別や条件を踏まえて検討するフレームワークもできるとよい。
- 最終的には、特定した品質特性を踏まえて、対応策を整理できることが望ましい。可能であれば、対応策の概要のみならず、対応策と連携されるツール名も整理できるとよい。
- フレームワークを改善していくうえで複数のユースケースで同様の検討を行い、最終的には、ユースケースと重要な品質特性を一定のパターン化できることが望ましい。
- 継続的な改善が実施できるようなフレームワークにできるとよい。

4.2. データ品質マネジメントチェックリストの作成・検証

「網羅的なデータ品質の取組検討の難しさ」、「データ品質への関心の低さ」の論点の課題解決に向けた以下のような意見がデータ品質 SWG より確認された。

- データ品質ガイドは考え方を網羅的に整理したドキュメントであるが、データ品質の取組を網羅的に整理したドキュメントがない。
- 適切なデータ品質の取組を選択するうえで、データ品質の取組と取組を実施しなかった場合のリスクを結び付けられるとよい。
- 簡易的にデータ品質の取組をチェックできるドキュメントがあるとよい。

上記の意見より、データ品質の取組を網羅的・簡易的に確認できるツールやドキュメントの必要性について確認できた。ドキュメントを作成することで、「網羅的なデータ品質の取組検討の難しさ」の観点では、データ品質の取組が網羅的に整理され、実務者が自社のユースケースに対して必要なデータ品質の取組を検討しやすくなると考えられる。また、「データ品質への関心の低さ」という観点では、整理されたデータ品質の取組とリスクを結び付けることで、データ品質の必要性について一定のアピールになる可能性があると考えられる。

データ品質の取組を網羅的・簡易的に確認できるツールやドキュメントを作成することを目指して、2025 年度では、データ品質ガイドを踏まえてデータ品質の取組を整理した「データ品質マネジメントチェックリスト」を作成した。

4.2.1. データ品質マネジメントチェックリストの概要

データ品質 SWG で作成したデータ品質チェックリスト（以下、本チェックリスト）は、自社のデータ品質の取組をライフサイクルに沿って詳細に評価するチェックリストである。

本チェックリストの主な目的は、現場の担当者が AI システムでデータを取り扱う上で、ライフサイクルごとに推奨される取り組みを実施しているかを確認することである。また、AISI 評価ツ

質の取組を確認することが可能である。

- 「【参考】チェック結果サマリ」シート
「チェックリスト」シートのチェック結果のサマリを可視化している。具体的には、「対象外」と回答した項目以外を母集団として、「実施済み」の回答率をライフサイクルごとに示している。
- 「【参考】品質特性一覧」シート
「チェックリスト」シートの「品質特性」の列について解説している。
- 「【参考】AISI 評価ツール質問項目（データ品質のみ抜粋）」シート
「チェックリスト」シートの「AISI 評価ツール質問カテゴリー」の列について解説している。

「チェックリスト」シートにおいてチェック項目を確認する前に、「チェック対象システム」、「チェック実施日」、「記入者氏名（代表者）」を記入することが望ましい。本項目を記入することで、チェックの振り返りを有効に実施することが可能となる。

「チェックリスト」シートの各項目（列）の内容については、以下のとおりである。チェック実施者は、「チェック項目」とその他の項目の記載内容も踏まえて、「チェック」の欄を記入することが望ましい。「チェック」の欄は4つの選択肢があるため、以下の表の解説を踏まえて、対象システムにおけるチェック項目の適切な実施状況を選択することを想定する。また、「回答補足」において、チェックのエビデンスなどを記載することを推奨する。エビデンスを記載することで、他社への取組実施状況の説明や担当者の変更の場合においても、チェックの振り返りをより有効に実施することが可能である。

ライフサイクル	データ計画、データ取得、データ準備、データ処理、AIシステム、出力評価、結果提供、データ廃棄、横断
チェック項目	データ品質に関して確認することが望ましい項目
優先事項	データ品質SWGの議論を踏まえて抽出した、サプライチェーン企業（取引先/外部データ提供者）に対して優先的に確認したい項目
チェック	チェックの対象システムにおける「チェック項目」の実施状況を以下より選択 「実施済み」：チェック項目を実施している 「検討中」：チェック項目を社内で検討中or実施準備中 「未実施」：チェック項目を実施・検討できていない 「対象外」：チェック項目は対象システムにおいては実施対象外 (可能であれば対象外の原因を備考欄に記載することが望まれる)
回答補足	チェックのエビデンスなどの回答内容を補足
実装例	チェック項目の実装方法の一例
リスク具体例	チェック項目における具体例
品質特性	チェック項目が「実施済み」の場合に満たされる品質特性 ※品質特性は、データ品質ガイドを参照して整理しており、「【参考】品質特性一覧」のシートに解説 ※チェック項目によって複数の品質特性を満たす場合があるが、最も関連深い品質特性を記載
実施者	チェック項目を実施する担当者の役割 ※役割は、データ品質ガイドを参照して整理しており、別表で解説
対象データ	チェック項目が対象としている各データ種別に応じて整理 ※データ種別は、データ品質ガイドを参照して整理しており、別表で解説
EU AI Act Article10 該当箇所	チェック項目とEU AI Act Article 10との対応
AISI評価ツール質問カテゴリー	AISI評価ツールにおける「データ品質」の観点の質問事項との対応

図 4-6 チェックリストの各項目の解説

すべてのチェック項目を「実施済み」にすることがゴールではなく、ユースケースや品質を確保したいデータ種別に応じて必要な項目のみをチェックすることを想定している。チェック項目における必要な項目のみを抽出する際には、「優先事項」列や対象データの列でのフィルタリングなどの活用も想定する。「優先事項」列は、データ品質 SWG の議論を踏まえて、サプライチャー

ン企業（取引先/外部データ提供者）に対して優先的に確認したい項目を抽出している。

本チェックリストの活用方法について、以下のようなシナリオが考えられる。

- AI モデルの学習は行わずに、プロンプトや RAG によって開発されているサービスの場合対象データについて、「全てのデータ」列、「参照データ」列、「AI システム内のデータ」列でフィルタリングして、該当事項について確認する。
- AI 提供は行っておらず、AI 活用を前提としたデータ提供のみのサービスの場合「ライフサイクル」列において、「データ計画」、「データ取得」、「データ準備」、「データ廃棄」でフィルタリングして、該当事項について確認する。
- データ品質の取組を自社で初めて取り組む場合段階的に実施するためにも、「優先事項」列でフィルタリングして、該当事項について確認するところから始める。
- 欧州などでサービスを展開している場合「EU AI Act Article 10 該当箇所」列を踏まえて、該当事項について確認する。

4.2.3. データ品質マネジメントチェックリストの検証

本チェックリストの利便性や網羅性などを検証するために、データ品質 SWG 内、分野別 SWG メンバーへの適用検証を実施した。

(1) 官報データを対象とした検証

AI の社会実装が進む中、公的機関が保有・提供するデータについては、従来の業務利用に加え、AI 利活用を見据えた品質マネジメントの在り方が問われている。

そのため、本検証では、データ品質ガイドで整理した考え方を具体的な実データに適用し、実務上の論点や改善余地を明らかにすることを目的として、官報データを対象とした。具体的なユースケース・検証方法について以下のとおりに示す。

■ ユースケースの概要

官報は、政府情報として正確性・完全性が強く求められる公的情報基盤である。一方で、現行の官報は紙面及び Web 閲覧性を前提として設計されており、構造化、機械可読性、メタデータの整備といった観点では、AI 利活用に向けた課題を内包している。

このように、従来の品質管理は高度に成熟している一方、AI 活用の観点では改善余地が明確であるという特性を有する官報データは、データ品質管理チェックリストの改善点を抽出するためのユースケースとして適切であると判断した。

■ 検証方法

官報データの特性を踏まえ、従来の業務利用を前提とした品質管理と AI-Ready な品質要件との差分を可視化することを目的とし、官報データ管理の「現場」に焦点を当てた実務的な評価を実施した。検証は、官報業務（システム設計・開発・運用管理、編集、製造管理、事務総括）に携

わった経験者 6 名による自己評価（○／△／×＋理由）により実施した。

また、評価に当たっては、官報の実務に適さない項目を事前に除外し、各チェック項目に官報適用例（対象・粒度、適用条件、工程等）を事前に補記することで、官報文脈での前提条件と使いどころをできる限り明確化した。

チェック項目	実装例	官報適用例	評価 (○、△、×、-)	評価の判定理由 (必須)	官報適用例に対する意見等
各データのデータ品質要求レベルは定義されているか。	上記で設定した項目で満たすべきレベルが定義されている。	上記で設定した項目で満たすレベルが定義されているか（訂正記事箇所数、システム稼働率等）。			
データの来歴情報に問題はないか。	各種データの来歴（生成方法、改訂履歴など）を確認している。	来歴（原稿受領→編集→校正→公開のログ、版履歴等）を適切に管理しているか。			

図 4-7 検証で利用したチェックリストのイメージ

■ 検証結果

検証の結果、以下の傾向を確認した。

- 正確性・完全性・来歴管理など、従来の官報業務で重視されてきた品質要件はおおむね満たされている。
- 一方、機械可読インターフェース、構造化、メタデータ、分割・統合・変換の方針といった項目では評価が割れた。
- 特にデータ準備工程において、AI 利活用を前提とした品質要件が明確でないことが課題として顕在化した。
- なお、評価に当たって補記した「官報適用例」については、幅広く活用するためにも有用であるとの意見が多かった。また、As-Is（現状で満たしている品質）／To-Be（AI 活用で目指す品質）に区分し、二層で整理するのが有効であるとの意見があった。
- また、評価者の経験年数・役職により、判断傾向に一定の差が見られた。若手・担当者層は実務リスクを意識して厳しめの評価を行う。また、基盤設計（インターフェース・構造化・メタデータ）については、経験者ほど厳しめの評価を行う傾向が確認された。
- また、継続的改善に向けた運用として、組織が自らのデータ品質の取組状況を確認できる仕組み作り（フィードバックループの構築等）が重要であることを確認した。

■ 考察

官報データを対象としたチェックリスト適用検証により、官報は、正確性・完全性・来歴管理といった従来の業務利用を前提とした品質管理としては成熟している一方で、構造化、機械可読性、メタデータ整備など、AI 利活用を前提とした品質（AI-Ready）の観点では前提設計が十分でないことが明確になった。

本検証の意義は、「AI 活用のために官報をどのように改変すべきか」を直接的に論じるのではなく、「AI 活用時代において、官報（AI 活用に不十分なデータ）の品質をどのように再定義し、どの水準までを品質マネジメントの対象とするのか」という視点を提示した。官報データにおいては、「従来の業務利用を前提とした品質（人による利用を前提とした品質）」から「AI 利活用を見据えた”AI-Ready”な品質」へ、段階的かつ合意形成型でデータ品質マネジメントを高度化していくことが重要であることが確認できた。

また、本検証を通じて、チェックリストの現場適用に向けては、事例併記の標準化、判定基準の段階化、As-Is（現行業務）と To-Be（AI 利活用）を区別した記載方法、工程別整理、フィードバックループの構築などの改善が必要であることが確認できた。

（2）分野別 SWG を対象とした適用検証

分野別 SWG の参加企業に対して、チェックリストの適用検証とチェックリストに対するアンケートを依頼した。本項では、適用検証結果について主に記載する。アンケート結果（5 社）よりチェックリストが以下のような点で評価された。改善すべき点については、4.2.4 の改善策の一部として取りまとめた。

- 自社のプロダクトのデータ品質の取組を網羅的にチェックすることが可能であった。
- 各チェック項目の解説要素のうち、特に「実施例」、「実施者」、「対象データ」は、回答の補助となった。
- 「使い方」シートによってチェックリストの作業イメージが明確化された。
- 「データ計画」のチェック項目については、特に自社の気付きにつながった。

適用検証結果としては、分野別 SWG の参加企業のうち、2 社において検証を実施した。

2 社のうち 1 社目では、外部の AI システムを対象に適用検証を行った。外部の AI システムの情報が開示されているケースが少なく、多くのチェック項目が回答対象外となった。この結果を踏まえて、外部の AI システムにおけるデータ品質管理をそのまま評価することは難しい。一方、自社で外部の AI モデルを利用して、RAG を用いた AI システムを開発する際の評価に適していることを、チェックリストの使い方に明記する必要性が確認できた。その結果、2025 年度に作成したチェックリストに対して修正を加えた。

2 社目では、自社で開発している AI システムを対象に適用検証を行った。データ品質管理の取組状況を可視化したレーダーチャートをエラー！参照元が見つかりません。8 に示す。結果が示すとおり、多くのステージにおいて実施率が 80%を超えていることが確認できた。一方で、データ準備・データ処理の項目の達成度が相対的に低いことが確認できた。データ準備・データ処理は、メタデータ管理などのデータ品質における細かな技術要件もあるため、実施できていない項目が多いと考えられる。

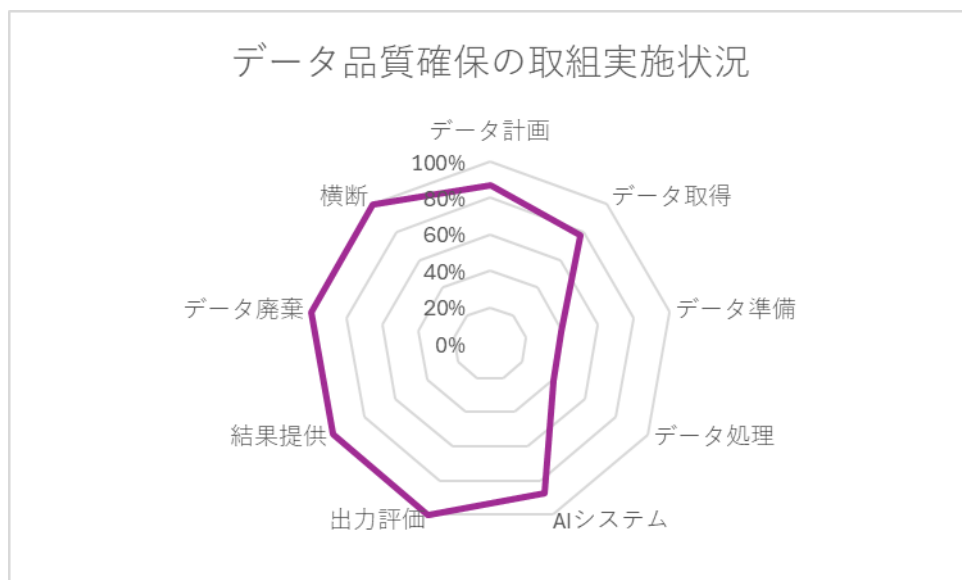


図 4-8 データ品質管理の取組状況の可視化

4.2.4. データ品質マネジメントチェックリストの今後の改善策

本チェックリストについて、データ品質 SWG・分野別 SWG 共にアンケートを実施した結果、様々な改善策が確認できた。アンケートを踏まえて、チェックリストに対する直接的な改善策として以下が挙げられた。

- リスク具体例の精緻化
 リスクの具体例を記載しているが、よりデータ品質の取組の重要性を示すうえでは、リスクの実事例やリスクによる損失をより具体的に示せるとよい。
- 実施者の細分化
 チェック項目の実施者を 2025 年度では 3 者に分類・整理したが、実務的にはより詳細な役割が存在するため、今後はより詳細に実施者を整理することが望まれる。
- 品質特性の再整理
 ISO などの標準を踏まえて、24 個の品質特性を整理したが、実務者にとってわかりづらい分類となっているという意見が多くあった。そのため、標準の記載を踏まえつつ、より実務者にとってわかりやすい品質特性に整理することが望まれる。
- チェック項目の成熟度レベルの導入
 整理したチェック項目の多くは、実施/未実施で測れるものではなく、実施度合いのグラデーションを示すことができる項目が多い。今後は、チェック項目に応じて成熟度レベルを検討して、チェック項目がどの程度のレベルで実施できているかを確認することが望ましい。
- 用語集の作成
 データ品質ガイドをベースに作成したチェックリストであるため、専門用語も多く使用されている。今後、データ品質ガイドの更新で作成される予定である用語集を踏まえて、チェックリストの用語集を作成することが望まれる。

上記の直接的な改善策以外に、今後必要なドキュメントなどに関する改善策などとして、以下

が挙げられた。

- 業界・用途別のカスタマイズガイド
チェックリストの「使い方」シートにおいて、活用方法の例を複数簡易的に記載したが、業界・用途別にチェックリストの使い方を詳細にガイドすることが望まれる。また、ガイドするにあたって、ユースケースとチェックリストを結び付ける概念が必要である。
- チェック項目に関する事例集の作成
チェックリストにおいて、実装例を各チェック項目に対して記載したが、より具体的実装方法を示す事例集を作成することが必要である。具体的には、チェック項目の実装事例を紹介することに加えて、実際のデータ品質確保に向けたツールやサービスまで含めた整理が望まれる。
- ライフサイクル以外の観点でのチェックリストの整理
チェックリストにおいては、データ品質ガイドを踏まえて、データ品質のライフサイクルに沿ったチェック項目を整理しているが、ユースケースに応じた適用を検討した場合、最適な整理方法ではない可能性がある。そのため、チェック項目を整理する適切な軸を検討することが望まれる。

これらの意見や 4.1.3 に示したユースケースに関する改善策なども踏まえて、本チェックリストの改善を適宜実施するとともに、ユースケースとチェックリストを結び付けることを目指して、図 4-9 のようなフレームワークを構築することが考えられる。具体的には、ユースケースの特性を踏まえて特に懸念されるリスクを整理し、整理したリスクをもとに品質特性を特定し、特定した品質特性に応じて、推奨要求事項一覧で必要なデータ品質の取組を抽出し、抽出した取組の事例や実装するためのツールを示すような枠組みを作成することが考えられる。

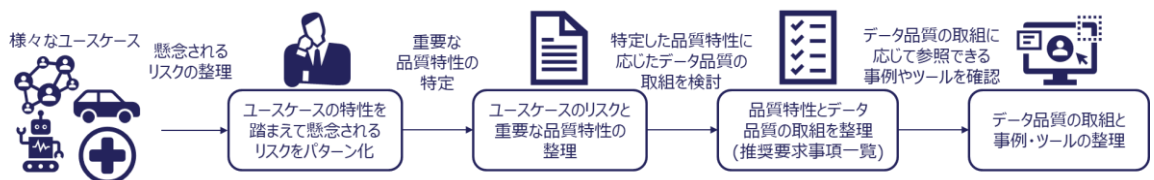


図 4-9 ユースケースとチェックリストの整理フレームワークのイメージ

4.3. データ品質レベルと AI 出力品質の検証

本節では、官報データを生成 AI で利活用する際、同一内容であっても入力仕方（データ形式、ノイズの有無、条・項等の構造情報の付与）によって生成 AI にとっての読み取りやすさが変化し得る点に着目し、入力データの整備水準を段階化して影響を検証する。具体的には、データ品質特性の Understandability（理解性）を「生成 AI が読み取りやすいか（誤読・取り違えが起きにくい）」へ拡張して捉え、PDF/HTML の形式差、クレンジング、タグ付与等の整備度合いを「構造化レベル」として整理する。

本検証は AI モデルやプロンプトの巧拙を競うものではなく、データ側の整備（構造化レベル）

の効果を相対比較することを主眼とする。そのため、条件を可能な範囲で揃えた上で、生成 AI 出力を条文単位で正解データと照合できるよう最小限の正規化を施して評価する。加えて、データ提供者／利用者の役割分担を仮想的に想定し、提供側でどの水準まで整備すれば利活用が進むか、利用側でどの程度の前処理が必要となるかを、結果に基づき議論可能な形で整理する。

4.3.1. 検証の位置付け

本検証の目的は、データ品質ガイドにおいて整理した考え方を、具体的な検証として示し、実務上の論点を明確化することにある。特に、以下の点について、実データを用いた試行を通じて確認・整理することを狙いとする。

- データ品質と生成 AI の出力品質（認識精度）の関係
- 品質特性の実装的な評価の可能性
- 品質向上に伴うコストと効果の関係
- データ提供者と利用者の間での品質水準合意のあり方

官報データを用いた検証は、上記データ品質ガイド検証のための具体的なユースケースの一つとして位置付ける。官報は公的情報として信頼性が高い一方、長文であり、条・項・号・別表等が混在する半構造データでもある。そのため、人間にとっては読めるが、生成 AI が扱う際にはデータ品質（構造化・一貫性・明確性等）が問題になりやすい。さらに、官報は PDF と HTML 等の複数形式で提供され、構造化レベルの違いを段階的に定義しやすいという特徴がある。加えて、法令改正に関する誤認やハルシネーションは、実務上の影響が大きく、社会的にも無視できないリスクとなり得る。以上の理由から、官報はデータ品質ガイド適用事例として適している。

また、本検証は、国立印刷局が官報を「データ提供者」として提供する立場にあることを踏まえ、AI を活用する側の組織を仮想的に想定し、その立場からデータの整備レベルと活用可能性を評価する。これにより、データ提供者側がどの水準まで構造化等を行うと利活用が進むか、また利用者側でどの程度の前処理が必要となるかを、検証結果に基づき議論可能な形で整理する。

先行研究として、デジタル庁の調査研究では、官報等の政府保有データについて、HTML と PNG 等の入力形式を比較し、モダリティによる AI 出力品質の差を検証している²⁵。本検証は、こうした知見を踏まえつつ、官報の利活用場面の一例として「改正箇所の特定」を対象に、入力データの構造化（品質）水準を段階化し、生成 AI 出力への影響を確認する点に特徴がある。

4.3.2. 検証対象ユースケース

官報にはさまざまな情報が掲載されるが、中でもニーズが高い情報の一つが法令の改正情報である。官報には改正法が記載され、被改正法の該当箇所が示される。法令改正は国民社会全体に影響し得ることから、本検証では官報に掲載される法令改正情報を入力として、改正箇所を特定することをユースケースとする。具体的には、生成 AI に官報を入力し、生成 AI が改正対象とな

²⁵ デジタル庁、「政府等が保有するデータの AI 学習データへの変換に係る調査研究を実施しました」
<https://www.digital.go.jp/news/382c3937-f43c-4452-ae27-2ea7bb66ec75>

る条文（第〇条）を同定し、当該箇所に対して行われた改正操作（例：「改める」「削る」「加える」など）や新旧対照表をもとに正しく抽出できるかを確認する。法令が条・項などの階層構造を持つことを踏まえつつ、評価は条文単位で正誤判定を行う。なお、本検証では改正の具体的内容までの特定は対象外とする。

改正箇所の特定制というユースケースは、官報に限らず、社内規程や約款・契約条項など、改定内容が文章として提示される文書においても同型の課題になり得る。すなわち、改定の記述から「どこが変わったか」を取り出し、利用者が参照・確認しやすい形で提示するというニーズは、幅広い文書体系で共通し得る。

なお、本ユースケースは、官報本文に明示された改正記述に基づき正解データ（人手アノテーション）を比較的整備しやすいこと、また高度な法令解釈を前提とせず主として「記載された情報の読み取り」として評価できることから、データの整備状況（形式や前処理等）の違いが生成 AI の出力に与える影響を観察する題材として適している。

4.3.3. 検証の設計方針

本検証では、官報データを生成 AI で利用する場面を想定し、改正箇所の特定制を題材として、生成 AI がどの程度正しく情報を抽出できるかを確認する。さらに、入力データの品質水準（データ形式やクレンジング等）の違いが、生成 AI の出力品質（正誤、誤認・取り違え等）に与える影響を比較できる形で整理する。

特に本検証では、データ品質ガイドに記載されているデータ品質特性のうち Understandability（理解性）に注目する。従来、Understandability は「人が理解しやすいか」という観点で捉えられることが多かったが、生成 AI の利用が一般化する中では、これを「生成 AI が読み取りやすいか（誤読・取り違えが起きにくい）」という観点へ拡張して捉える必要がある。本検証では、この生成 AI にとっての理解性を、データ形式の選択（例：PDF/HTML）、クレンジング（ノイズ除去）、タグ付与等のデータ整備によって段階化し、これを本検証における「構造化レベル」と呼ぶ。以降の項では、この構造化レベルを複数水準として定義し、同一条件の下で AI の出力品質を比較する。

また、本検証は AI モデルやプロンプトの巧拙を競うものではなく、あくまでデータ側の整備の効果を相対比較することを主眼とする。このため、プロンプト設計は必要最小限の指示（出力形式や対象範囲の明確化等）に留め、強い役割指定や過度な誘導によって出力品質を押し上げる工夫は評価対象外とする。

最後に、評価指標・正誤判定の方法については、条文単位で正解か否かを判定できるよう、必要最小限の前処理（例：漢数字の表記ゆれ、空白・改行、余計な出力の除去等）を施した上で照合する。

4.3.4. 検証対象データ

本検証では、官報に掲載される法令改正情報を対象データとする。官報は PDF 形式に加え、HTML 形式でも提供されており、同一の内容であってもデータ形式や構造の違いが生じる。本検

証では、この形式差を含め、入力データの整備状況が生成 AI の読み取りに与える影響を比較するため、官報データを複数の形態で用いる。

対象とするのは、法令改正に関する官報の記事（計 16 件）である。本検証で対象とする官報データは、国立印刷局から提供を受けた。対象とした法令は、法律、政令、府省令、規則の 4 種であり、それぞれ数件ずつ用意した。

なお、用いた官報データは、その日の記事全文ではなく、当該の改正法部分が抽出されたものである。これは本検証のために特別に抽出したのではなく、官報発行サイトでも提供されている形式と同一である。

HTML 形式の官報データについて、HTML タグによる構造化や意味付けは行われておらず、画面表示のためのタグが付与されている。

官報には 2 種類の改正方式、「改め文方式」と「新旧対照表方式」がある。改め文方式は「第一条中『A』を『B』に改める」のように文章で修正箇所を指示する伝統的な形式で、正確性が高い一方、理解しにくい。一方、新旧対照表方式は修正前後の条文を左右に並べ、変更箇所を網掛け等で強調する形式で、視覚的に分かりやすい。記載方式の違いによる AI 解釈が異なることに着目して出力品質の評価を行った。

例えば、下記が「改め文方式」の例である。

<p>法律 第八十一号 租税特別措置法及び東日本大震災の被災者等に係る国税関係法律の臨時特例に関する法律の一部を改正する法律 （租税特別措置法の一部改正）</p> <p>第一条 租税特別措置法（昭和三十二年法律第二十六号）の一部を次のように改正する。 第八十八条の八を削る。 第八十九条を次のように改める。</p>

図 4-10 改め文方式の例

改正後	改正前
目次	目次
第一章～第一章の四 [略]	第一章～第一章の四 [同上]
第二章 福祉の保障（第七条- <u>第三十六条の三十の六の十四</u> ）	第二章 福祉の保障（第七条- <u>第三十六条の三十の六の三</u> ）
第三章～第四章 [略]	第三章～第四章 [同上]
附則	附則
<p>第七条の四十五 法第十九条の二十二第四項に規定する厚生労働省令で定める方法は、個人番号カード（行政手続における特定の個人を識別するための番号の利用等に関する法律第二条第七項に規定する個人番号カードをいう。<u>第十七条の三第二項第一号及び第三十六条の三十の六の六第二項第一号</u>において同じ。）を提示する方法とする。ただし、当該方法によることができない状況があるときは、書面により提示する方法とする。</p>	<p>第七条の四十五 法第十九条の二十二第四項に規定する厚生労働省令で定める方法は、個人番号カード（行政手続における特定の個人を識別するための番号の利用等に関する法律第二条第七項に規定する個人番号カードをいう。<u>第十七条の三第二項第一号</u>において同じ。）を提示する方法とする。ただし、当該方法によることができない状況があるときは、書面により提示する方法とする。</p>

図 4-11 新旧対照表方式の例

なお、どの官報がどの改正方式に当たるかは下記に示す。

表 4-3 検証に利用した官報一覧

No.	種別	官報公布	法令番号	法令名	改正方式
1	法律	12月5日	法律第八十一号	租税特別措置法及び東日本大震災の被災者等に係る国税関係法律の臨時特例に関する法律の一部を改正する法律	改め文方式
2		12月10日	法律第八十二号	更生保護制度の充実を図るための保護司法等の一部を改正する法律	改め文方式
3		12月22日	法律第八十八号	地方交付税法及び特別会計に関する法律の一部を改正する法律	改め文方式
4	政令	11月12日	政令第三百六十九号	行政書士法の一部を改正する法律の施行に伴う関係政令の整理に関する政令	改め文方式
5		11月27日	政令第三百八十八号	老朽化マンション等の管理及び再生の円滑化等を図るための建物の区分所有等に関する法律等の一部を改正する法律の施行に伴う関係政令の整備に関する政令	改め文方式
6		11月28日	政令第三百九十五号	独立行政法人男女共同参画機構法の施行に伴う関係政令の整備及び経過措置に関する政令	改め文方式
7		12月12日	政令第四百十二号	脱炭素成長型経済構造への円滑な移行の推進に関する法律及び資源の有効な利用の促進に関する法律の一部を改正する法律の施行に伴う関係政令の整備等に関する政令	改め文方式
8		12月17日	政令第四百二十三号	出入国管理及び難民認定法等の一部を改正する法律の施行に伴う関係政令の整備等及び経過措置に関する政令	改め文方式
9		12月26日	政令第四百四十六号	漁業災害補償法施行令の一部を改正する政令	改め文方式
10	府省令	11月12日	環境省令第二十二号	資源循環の促進のための再資源化事業等の高度化に関する法律施行規則	改め文方式
11		11月27日	国土交通省令第百十四号	航空法施行規則及び民間の能力を活用した国管理空港等の運営等に関する法律施行規則の一部を改正する省令	改め文方式
12		11月28日	内閣府令第百号	児童福祉法施行規則の一部を改正する内閣府令	新旧対照表方式
13		12月10日	総務省令第百八号	住民基本台帳法施行規則等の一部を改正する省令	新旧対照表方式
14		12月23日	総務省令第百十三号	普通交付税に関する省令の一部を改正する省令	新旧対照表方式
15	規則	10月17日	国家公安委員会規則第十八号	風俗営業等の規制及び業務の適正化等に関する法律施行規則等の一部を改正する規則	新旧対照表方式
16		12月26日	公正取引委員会規則第十二号	公正取引委員会事務局組織規程の一部を改正する規則	新旧対照表方式

4.3.5. データ準備

(1) 正解データ作成

本検証では、生成 AI の抽出結果を評価するための正解データを作成した。正解データは、官報データの原文を人手で読み解き、官報本文中に明示されている改正操作（例：「改める」「削る」「加える」）や新旧対照表を手掛かりとして、改正対象となる条文（第〇条）を単位に整理したものである。

例えば、官報 No.1 の一部は以下のように正解データとして整理する。

法律 第八十一号
租税特別措置法及び東日本大震災の被災者等に係る国税関係法律の臨時特例に関する法律の一部を改正する法律
(租税特別措置法の一部改正)
第一条 租税特別措置法(昭和三十二年法律第二十六号)の一部を次のように改正する。
第八十八条の八を削る。
第八十九条を次のように改める。
第八十九条 削除



公布された法令の 法令番号	公布された法令名	改正対象法令の法令 番号	改正対象法令名	改正対象の条 文番号
法律 第八十一号	租税特別措置法及び東日本大震災の被災者等に係る国税関係法律の臨時特例に関する法律の一部を改正する法律	昭和三十二年法律第二十六号	租税特別措置法	第八十八条の八
法律 第八十一号	租税特別措置法及び東日本大震災の被災者等に係る国税関係法律の臨時特例に関する法律の一部を改正する法律	昭和三十二年法律第二十六号	租税特別措置法	第八十九条

図 4-12 正解データの整理イメージ

正解データ作成にあたっては、官報や法律に必ずしも精通していないメンバー2名で作業を進める前提とし、専門家の判断を極力要しない形で作業手順を設計した。具体的には、まず改正操作に関する一般的な考え方(典型的な改正パターン、表現の種類、整理の観点)を文献調査や有識者へのヒアリング等により把握した上で、並行して、実際の官報データを確認し、ボトムアップにパターンを洗い出す作業を行った。これにより、例えば「条文レベルの改正」と「文言レベルの改正」のように粒度が異なるケースがあること、条文が移動する場合にどのように条文単位の正解として整理するか、また別表等を含む場合に表をどの範囲まで対象とするか、といった論点を早期に把握できた。以上の整理を踏まえて、(1)改正パターンの洗い出し、(2)正解データ作成仕様の文書化、(3)仕様に基づく作業の実施、(4)想定外パターンの記録と仕様への反映(初期は短サイクルで相互チェック)を反復しながら、正解データを整備した。

また、正解データ作成の効率と再現性を高めるため、改正箇所は可能な範囲で粒度を細かく切り出して整理し、必要に応じて後から集約できる形を採った。例えば、条レベルの改正と文言レベルの改正が混在する場合でも、まずは細かい単位で改正操作や対象箇所を記録しておくことで、後段で条文単位へ集約して評価に用いることができる。細かい粒度で作成することは一見負荷が高いように見えるが、作業そのものが単純化し、判断のぶれを抑えやすいこと、また後からの集約は可能である一方で、粗い粒度から細分化することは困難であるという利点がある。さらに、「改める」のような基本パターンについては、文書作成ソフトでの色付け等、機械的にサポートできる工夫を取り入れ、手作業の負担軽減を図った。

(2) 品質レベル別データの作成

本検証では、官報データを生成 AI に入力するにあたり、同じ内容であっても提示のされ方(デ

ータ形式・ノイズの有無・構造情報の付与)によって生成 AI の読み取りやすさが変化し得る点に着目し、入力データを複数の水準に整理した。以降、これらの水準を本検証における構造化レベルと呼ぶ。

構造化レベルの考え方は、オープンデータにおける「公開・利活用のしやすさを段階化する」発想（例：5 スターオープンデータモデル）等を参考に、生成 AI にとっての理解性（Understandability）という観点で、より機械処理しやすい形へ段階的に整備していくという整理に基づく。本検証の品質レベル定義は以下のとおりである。

- レベル 1（PDF：ベースライン）：縦書きの PDF を入力とする条件。なお、画像化された PDF ではなく、文字情報（テキスト層）を保持した PDF を用いる。
- レベル 2（HTML：形式差の比較）：同一内容について、官報発行サイト等で提供される HTML 形式のデータを入力とする条件。
- レベル 3（HTML＋クレンジング）：HTML データを対象に、生成 AI の読み取りに寄与しない要素を除去するなど、一定のクレンジング（ノイズ除去）を施した条件。現時点では、HTML のタグを一括で除去する簡易的なクレンジングを実施している。クレンジングには、BeautifulSoup ライブラリを利用する。
- レベル 4（HTML＋クレンジング＋構造付与）：レベル 3 に加え、条・項等の構造を表現するタグ付与など、入力データに追加の構造情報を付与した条件。本検証では、官報のドキュメントを Markdown 形式に変換し、条文のタイトル部分をセクションのタイトルとして強調し、また、見出し（条名直前の括弧書き）と条名（「第〇条」等の記載）の順番を入れ替えることにより、その条文が何を行っているかの関連付けが分かりやすくなることを期待する。

レベル 2 とレベル 4 の違いについて補足する。官報データに付与されている HTML タグは、意味的な構造を表現することを主目的としたものではなく、主として表示上の体裁を整えるために付与されている。一方で、一部の HTML タグには、結果として意味的な構造を間接的に示すものも含まれている。すなわち、HTML タグには意味的に有用な情報とそうでない情報が混在している。これに対し、レベル 4 では、検証期間および検証リソースの制約を踏まえつつ、簡易的な構造化を実施している。言い換えれば、HTML に混在する情報のうち、意味的に有用な要素の一部を抽出・再構成し、生成 AI が文書構造を把握しやすい形に整理している。

4.3.6. 検証方法

本項では、4.3.5 で作成した構造化レベル別データを用いて、生成 AI により改正箇所の特定を行い、その出力を正解データと照合して出力品質評価（精度評価）する手順を述べる。評価の基本単位は条文とし、同一タスク・同一条件の下で、構造化レベルの違いが出力に与える影響を比較する。

(1) アーキテクチャ

検証の処理フローは概ね以下のとおりである。

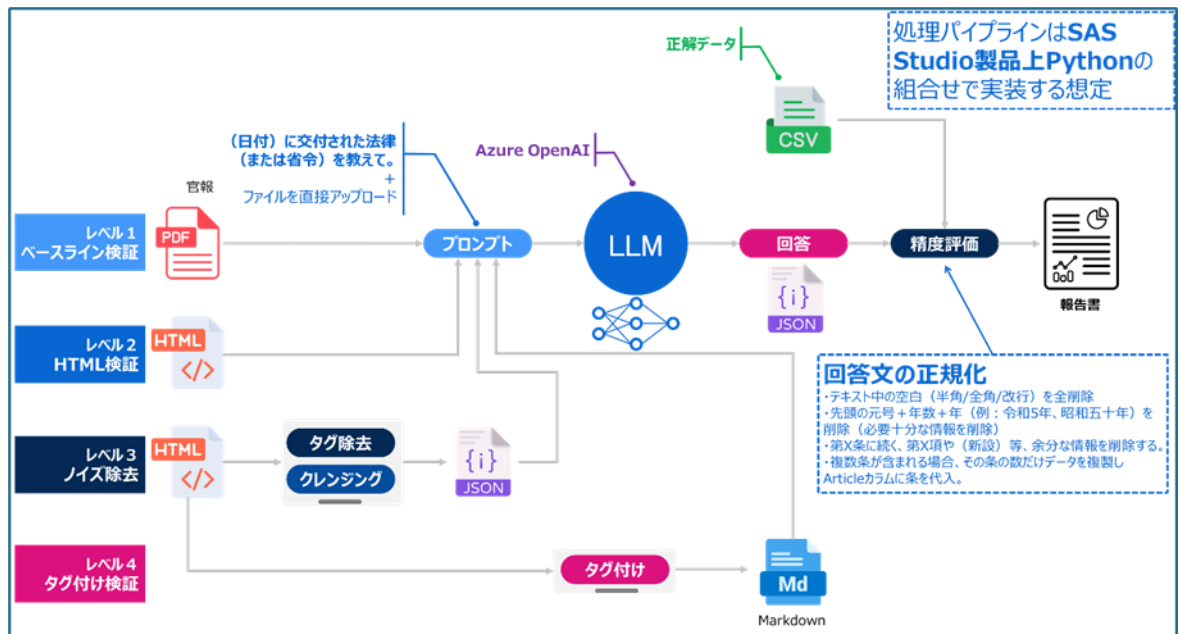


図 4-13 アーキテクチャ図

1. 官報データ（改正法部分）を取得
2. 構造化レベル（レベル1～4）に応じた前処理を実施
3. 生成AIへ入力し、改正箇所（条文）を抽出
4. 生成AIの出力を整形（必要最小限の正規化）
5. 正解データと照合し、条文単位で正誤判定
6. 結果を集計し、構造化レベル間で比較

処理・集計は、SAS 製品と Python を組み合わせて実施した（SAS® Studio Version: Stable 2026.02）。生成AIの呼び出しは Azure OpenAI を用いる。

(2) AI モデル

利用した生成AIモデルは Azure OpenAI 上の GPT-5.2-chat (Model version: 2025-12-11) である。モデルの主な設定値は以下のとおりである。

- Temperature : 1.0
- Max tokens : 128,000
- Top_p : 1.0
- Reasoning_effort: none

※本検証では、モデルやパラメータの最適化を目的とせず、構造化レベル差の影響を比較する観点から、設定は原則として固定した。

(3) プロンプト

プロンプトは、改正箇所の特定制というユースケースに必要な最小限の指示に留め、役割付与や強い誘導により認識精度を押し上げる工夫は行わない方針とした。入力には構造化レベルごとに整

備した官報データとし、出力は条文単位で列挙する形式とした。代表的なプロンプト例を以下に示す。

- 指示の要旨：官報本文中の改正記述を手掛かりに、改正対象の条文を抽出すること
- 出力形式：JSON
- 禁止事項：なし

```
"type": "input_text",
"text": f"{html_content}¥n 官報本文中に明示されている改正条を、正しく抽出し、スキーマに従って構造化してください。スキーマの解説は次の通りです。new_law_id→公布された法令番号、new_law_name→公布された法令名、amendment_law_id→改正された法令番号、amendment_law_name→改正された法令名、article→改正された法令における改正箇所の構造化された条。改正された条ごとに構造化（レコード化）してください。数字はすべて漢数字で抽出してください。",
```

図 4-14 プロンプトのイメージ

(4) 評価指標／正誤判定

評価は条文単位で実施する。生成 AI が抽出した条文が正解データに含まれる場合を正、含まれない場合を誤として判定し、構造化レベルごとの結果を比較する。

- 指標：Precision（適合率）、Recall（再現率）、F 値
 - Precision（適合率）：予測したデータのうち、実際にどれだけ正解していたかを示す指標
 - Recall（再現率）：実際に正解であるデータのうち、どれだけ正しく予測できたかを示す指標
 - F 値：適合率と再現率の調和平均であり、全体のバランスを示す指標

いずれの指標も 0～1 の値をとり、1 に近づくほど望ましい結果を示す。

正誤判定にあたっては、照合の安定性を確保するため、比較の前に必要最小限の前処理（正規化）を施した。具体例として、漢数字の表記ゆれの吸収、不要な空白・改行の除去、条番号表現の揺れの補正、生成 AI が付加した余計な装飾の除去等を行う。なお、実施した正規化は以下のとおりである。

- テキスト中の空白（半角/全角/改行）を全削除
- 先頭の元号＋年数＋年（例：令和 5 年、昭和五十年）を削除（必要十分な情報を削除）
- 第 X 条に続く、第 X 項や（新設）等、余分な情報を削除する。
- 複数条が含まれる場合、その条の数だけデータを複製し Article カラムに条を代入。

4.3.7. 結果と考察

本項では、4.3.6 で述べた手順に基づき、構造化レベル別に得られた生成 AI の出力を条文単位

で評価し、結果の概要と考察の観点を整理する。

(1) 結果概要

評価データ全体での構造化レベル別の評価結果（適合率、再現率、F 値）を下記の表に示す。

表 4-4 レベルごとの評価結果

構造化 レベル	改正 条文 数	真陽 性 TP	偽陽 性 FP	偽陰 性 FN	適合率 Precision	再現率 Recall	F 値
1: PDF	333	151	38	182	0.799	0.453	0.579
2: HTML	333	221	47	112	0.825	0.664	0.735
3: HTML+ク レンジング	333	187	37	146	0.835	0.562	0.671
4: HTML+ク レンジング+ 構造付与	333	214	43	119	0.833	0.643	0.725

また、改正方式「改め文方式」と「新旧対照表方式」に分けて集計した構造化レベル別の評価結果を下記の表に示す。

表 4-5 改正方式ごと・レベルごとの評価結果

構造化 レベル	改正 方式	改正 条文 数	真陽 性 TP	偽陽 性 FP	偽陰 性 FN	適合率 Precision	再現率 Recall	F 値
1: PDF	改め 文方 式	277	105	37	172	0.739	0.379	0.501
2: HTML	改め 文方 式	277	175	43	102	0.803	0.632	0.707
3: HTML+ク レンジング	改め 文方 式	277	141	34	136	0.806	0.509	0.624
4: HTML+ク レンジング+ 構造付与	改め 文方 式	277	168	38	109	0.816	0.606	0.696
1: PDF	新旧 対照 表方 式	56	46	1	10	0.979	0.821	0.893

2: HTML	新旧 対照 表方 式	56	46	4	10	0.920	0.821	0.868
3: HTML+ク レンジング	新旧 対照 表方 式	56	46	3	10	0.939	0.821	0.876
4: HTML+ク レンジング+ 構造付与	新旧 対照 表方 式	56	46	5	10	0.902	0.821	0.860

(2) 考察

本検証の考察は、次の観点に沿って整理する。

全体の傾向として、構造化レベル別に評価結果に違いがみられることが確認された。評価指標別には、適合率の構造化レベル間で差が最大 0.036 に対し、再現率は構造化レベル間で差が最大 0.210 と大きな差が生まれた。このことから、今回の検証シナリオにおいて構造化レベルは特に再現率に影響を与えることがわかる。再現率に注目すると、レベル 2 が最も結果が良く、相対的にレベル 4 も近い数値である。

改正方式「新旧対照表方式」では、評価結果の違いがあまり見られず、どの構造化レベルでも高い水準となった。「新旧対照表方式」では結果の違いが見られないのは、改正された法律が既に分かりやすい構造になっている可能性がある。特に PDF 形式が最も良い結果になっているのは、新旧対照表方式は AI（本検証ではマルチモーダルモデルを利用）が視覚的に理解しやすかったことも一つの要因と考えられるが、要因の切り分けは今後の分析課題である。

一方、「改め文方式」では構造化レベルが上がることで、AI 出力品質が改善する傾向がみられた。再現率では、レベル 2 と 4 の結果が高く、構造化することで検知漏れを防ぐ効果が期待できる。

上記より、次の観点で整理する。

- 構造化レベル差による影響の有無と傾向（改め文方式）
 - レベル 1（縦書き PDF）→レベル 2（HTML）で改善が見られる。これは、縦書き PDF の読み順・境界解釈の不安定さに比べ、HTML のほうが条文同定の手掛かりを保持しやすい可能性を示唆する。
 - レベル 2（HTML）→レベル 3（クレンジング）で改善が見られない。クレンジングにより不要情報が除去される一方で、条文同定の手掛かりとなっていた構造情報（タグ等）まで失われた可能性がある。すなわち、Understandability（理解性）を高める施策は「ノイズ除去」だけでなく、「構造の保持・付与」という観点と一体で検討する必要がある。

- ▶ レベル 4 は、HTML と同程度の AI 出力品質が得られる。説明上の便宜として比喩的に表現すれば、レベル 2 は意味的に有用な情報とそうでない情報が混在した、いわば「玉石混合」の状態にあり、レベル 4 はその「玉」の一部を抽出・再構成したものである。本検証の範囲において、レベル 2 とレベル 4 の結果に大きな差は認められなかった要因として、レベル 4 における構造化の過程で一部の情報が欠落した可能性が考えられる。このことから、本検証シナリオにおいては、ノイズ（「石」）が含まれていたとしても重要な情報（「玉」）が保持されていることと、構造化によって一部の情報が欠落することの影響が同程度となる場合があることが示唆される。

ただし、本検証では同一の法令改正について改正方式のみを変えて比較しているわけではなく、改正方式による差異と断定することはできない。今後、同一内容について改正方式を変えた場合を比較する等、追加的な検証を行うことが望ましい。

また、本検証においては、生成 AI 特有の出力結果のばらつき（非決定性）については十分に評価できていない。また、使用する AI モデルの違いによる影響についても、本検証の対象外としている。これらを踏まえると、本検証で得られた結果の再現性および一般化可能性については一定の制約があると考えられる。したがって、出力の安定性の評価やモデル間比較を含めた再現性の向上については、今後の継続的な検討課題とする。

構造化レベルにおいては、本検証では Markdown 化を試みた。さらなる方向性として、法令標準 XML スキーマ²⁶のような形式で高度に意味付けした文書として整備を進めることが考えられるが、整備コストが大きいと、少量のサンプルで効果を確認しながら段階的に検討することが望ましい。

本検証では、PDF 形式より HTML 形式のほうが AI の出力品質が良いと想定し、実際に一定のデータにおいては出力品質改善を確認した。このことは、官報のような半構造データに対し、機械可読性の高い形式（構造情報を保持した HTML）の提供が、生成 AI 利活用の基盤として有効である可能性を示唆する。提供者側で整備（例：形式の提供、最低限の構造情報の保持）を担保するほど、利用者側が同様の前処理を個別に繰り返す必要が減り、社会全体としての重複作業の低減につながり得る。一方で、改正方式によって必要な整備状況が変わることが検証により確認できたため、どの整備水準までを共通提供として、どの整備内容用途に応じて利用者側で補完するかは、効果・コスト・対象範囲を踏まえて検討する必要がある。

加えて、必ずしもすべての整備が AI 出力品質の向上に繋がるわけではなく、データ特性やユースケースに応じた品質設計が必要であることが示唆される。

²⁶ e-gov, 法令標準 XML スキーマ

<https://laws.e-gov.go.jp/docs/law-data-basic/419a603-xml-schema-for-japanese-law/>

本章では、第4章までに示した2025年度の活動結果及び得られた示唆を踏まえ、データ品質SWGとして継続的に検討・改善していくべき論点と、その方向性を整理する。ここで述べる内容は、次年度の体制・予算・関係者合意等の前提に依存し得るため、確定的な計画ではなく、現時点での課題認識と検討の射程を示すものである。

5.1. 2025年度の到達点

(1) ユースケース整理方法の検討

ユースケースの整理方法をリスクと品質特性の観点から検討し、複数ケースを報告書にまとめた。今後、事例が蓄積されれば、事例集として活用できる可能性があるほか、一定のパターン化により実務への適用をガイドしやすくなることが期待される。

(2) データ品質チェックリストの作成・検証

SWG内でのレビュー及び分野別SWGからの意見収集を踏まえ、データ品質チェックリストを作成した。ライフサイクルに沿って観点を整理し、データ品質マネジメントを俯瞰できる形にした点は一定の成果である。一方で、項目量や運用負荷、具体性の粒度など改善余地も残っており、今後の継続的な改訂が必要である。なお、本チェックリストは、現時点ではあくまで参考として公表するものであり、統一的な遵守を求めるものではない。各組織が自組織の状況に応じて試行的に活用し、得られた示唆を今後の改訂に反映することを想定している。

データ品質ガイド及びデータ品質チェックリストを現場運用に活用するための検証を行い、実務へ落とし込む際には「ガイドを実務に翻訳する作業」が必要であること、またその際にドメインとデータ品質管理双方の知識が求められることが明確になった。

(3) 具体的な適用検証（データ品質レベル検証）による示唆の獲得

包括的な品質管理の検討と並行して、官報データを対象として具体的な適用検証（データ品質レベル検証）を実施した。その結果、入力データの整備水準の違いによって、生成AIの出力品質が変化し得ることを確認した。特に、整備水準を高めることで出力品質が改善する場合がある一方、単純な前処理が必ずしも改善に直結しない場合も見られ、データ品質の検討は対象データやユースケースに応じて設計する必要があることが示唆された。これは、抽象的なデータ品質の重要性の提示に留まらず、どの整備が効果に寄与し得るかを比較可能な形で示すものであり、提供者・利用者の役割分担や整備水準の合意形成、さらには投資対効果・運用負荷の議論に資する判断材料となる。

(4) ガイドブックのマイナー改定（利活用性の向上）

データ品質ガイドについて、細部の表現や構成を見直したことにより、初版と比べて内容面以外での読解上の支障が軽減され、可読性が向上したものと考えられる。また、従来の PDF 形式から Markdown/Word 形式に変更することで、利活用性の向上を図った。これにより、データ品質ガイド自体を参照・再利用しやすくなるだけでなく、AI による読み取りと活用にあたっての障壁も低減され、今後の活用拡大が期待される。

5.2. 見えてきた課題の整理

(1) 包括性と具体性のギャップ

整備したデータ品質ガイド及びデータ品質チェックリストは包括的であり、汎用的に参照可能な枠組みを提供している。一方で、実務で必要となる具体的な管理方法（手順、判断基準、記録様式、運用設計）に踏み込むには限界があり、現場適用にはドメインに応じた追加の翻訳作業が必要となる。具体性を高めるには、対象ユースケースの限定や、適用状況に基づく反復的改善が不可欠である。

(2) ユースケースの優先順位付け

データ種別、業務領域、利用目的により、品質課題の出方や有効な管理手段が大きく異なる。包括的整理を維持しつつ実務に踏み込むためには、社会的影響度や実効性等の観点でユースケースの優先順位を付け、重点領域を定めた上で深掘りするアプローチが必要になる。

加えて、ユースケースをまたがる横断的な論点として、データの流通時点の品質をどのように可視化するかが重要な論点となっている。例えば、来歴、更新頻度、欠損・誤りの状況、利用条件等を、受け渡しに耐える形で表現・共有する必要がある。このため、個別実装を拘束するのではなく、品質情報の提示項目と記述粒度をそろえるためのテンプレートの整備が求められる。

(3) 実効性の向上

データ品質ガイドやデータ品質チェックリストの実効性の向上には、運用負荷を下げ、継続的改善を回すことが重要である。そのためには、評価・記録・収集の効率化が必要である。特に、チェックリストの項目量が多いこと、複数データソースを横断して情報収集が必要となることから、一定の自動化ニーズが見込まれる。ただし、データ品質の向上に向けた具体的な実装方法（ツール選定、業務プロセスの設計等）は、各組織が独自に工夫・差別化を図る領域でもある。そのため、SWG としては個別最適な実装を約束するのではなく、共通部分（評価・記録の最小セット、参照可能な手順例、用語・観点の整合）を整理し、実装の指針として提供することが現実的である。

(4) 評価レイヤの整理と拡張

品質評価には複数のレイヤが存在する。例えば、(a)マネジメントシステム（運用・体制・プロセス）の評価、(b)メタデータの評価、(c)データそのものの評価である。現行チェックリストは主として(a)に近い。一方で、(b)及び(c)については、SWG として十分に踏み込めていない領域が残

っている。(c)は完全性等の基本指標に留まらず、バイアス評価やプライバシー評価など、単純集計では困難な評価が存在する。SWGとして特定の指標について深堀し得るか、どのような観点・方法論があり得るかを整理し、ガイドできる状態に近づける必要がある。

5.3. スコープ拡大（新たな利用形態への対応）

AIの能力及び利用シーンは拡大しており、データ品質の論点も広がりつつある。本SWGで整備したデータ品質ガイドは、データ品質マネジメントに関する一般的な事項を中心に整理している。そのため、AIの利用形態が拡大している状況においても、下記に例示する新たな利用形態に対して一定程度は適用可能と考えられる。一方で、データの取り扱い方や品質リスクの現れ方が変化し得るため、既存観点の適用可能性を検証しつつ、必要な観点の追加・再整理を進めることが重要である

(1) マルチモーダル AI への対応

テキスト・画像・音声等、データ種別ごとに品質管理の勘所が異なる。また、複数モデルを統合して扱う際には、モデル間整合やアノテーション方針、来歴管理等、新たな観点が重要となる。既存ガイドの適用可能性を確認しつつ、観点の再整理が必要である。

(2) Agentic AI への対応

Agentic AIでは、各種ツール連携が前提となり、参照データの品質に加えて、ツール間を繋ぐメタデータ、ログ、入出力の追跡可能性等の管理が一層重要になる。コンテキスト管理をデータ品質管理の射程にどこまで含めるかも含め、整理が必要である。

(3) Physical AI への対応

実世界と相互作用するロボット等では、センサーデータの品質、地図・環境データ、リアルタイム処理、物理的整合性など、品質管理項目が拡大する。従来のコンピュータ内完結の前提では捉えにくい論点が増えるため、物理的整合性を含む品質要素の位置付けを含めて検討することが求められる。

5.4. 継続的取組の方向性

今後の検討を継続的に前進させるため、以下の運営上の方向性を確認することが求められる。

- 適用検証の促進：2025年度のデータ品質チェックリストの検証、データ品質レベルの検証を起点にして、データ品質SWG内外で適用検証を推進する。
- ユースケースの明確化：チェックリスト等の成果物について、想定する適用場面や最低限の満たし方（何が満たされればよいか、何を記録すべきか）を整理し、試行結果・フィードバックを踏まえて段階的に改訂する。

- 実務サポート・ツールの開発：データ品質ガイドやデータ品質チェックリストの静的なドキュメントに加えて、実際に動くツールを検討する。
- リビングドキュメントとしての最新化：実務適用や検証を通じて得られたフィードバックを受け、抽象論に留まらない実務的知見を蓄積し、データ品質ガイド・データ品質チェックリストを適宜改訂する。
- フィードバックの受け皿の明確化：関係者からの指摘・改善提案を受け取れる仕組みを整備し、反映プロセスを透明化する。
- 社会実装と国際議論への接続：得られた成果について、想定する適用場面と満たし方を明確化し、必要に応じて、政策・法令・標準・調達条件等の枠組みとの連動も視野に入れた整理を進める。そのうえで、データガバナンスや標準化に関する国内外の議論に積極的にインプットしていく。

以上により、2025年度の成果を出発点として、実務適用性の向上と新たな利用形態への対応を両立しつつ、継続的な改善サイクルを回していくための論点と方向性を整理した。

今年度の検証を通じて、AI出力品質や運用負荷がデータ品質に左右されること、またデータの流通・受け渡しにおいて品質を可視化する重要性が改めて確認された。データ品質は、AIの安全性を「後から担保する」のではなく、設計・運用の初期段階から組み込むべき要件である。データ品質SWGとしては、共通の観点と実務で使える形の整理を継続し、関係者と連携しながら改善を進める。