

データ品質SWG活動報告

データ品質SWG

独立行政法人情報処理推進機構(IPA) デジタル基盤センター

坂元哲平

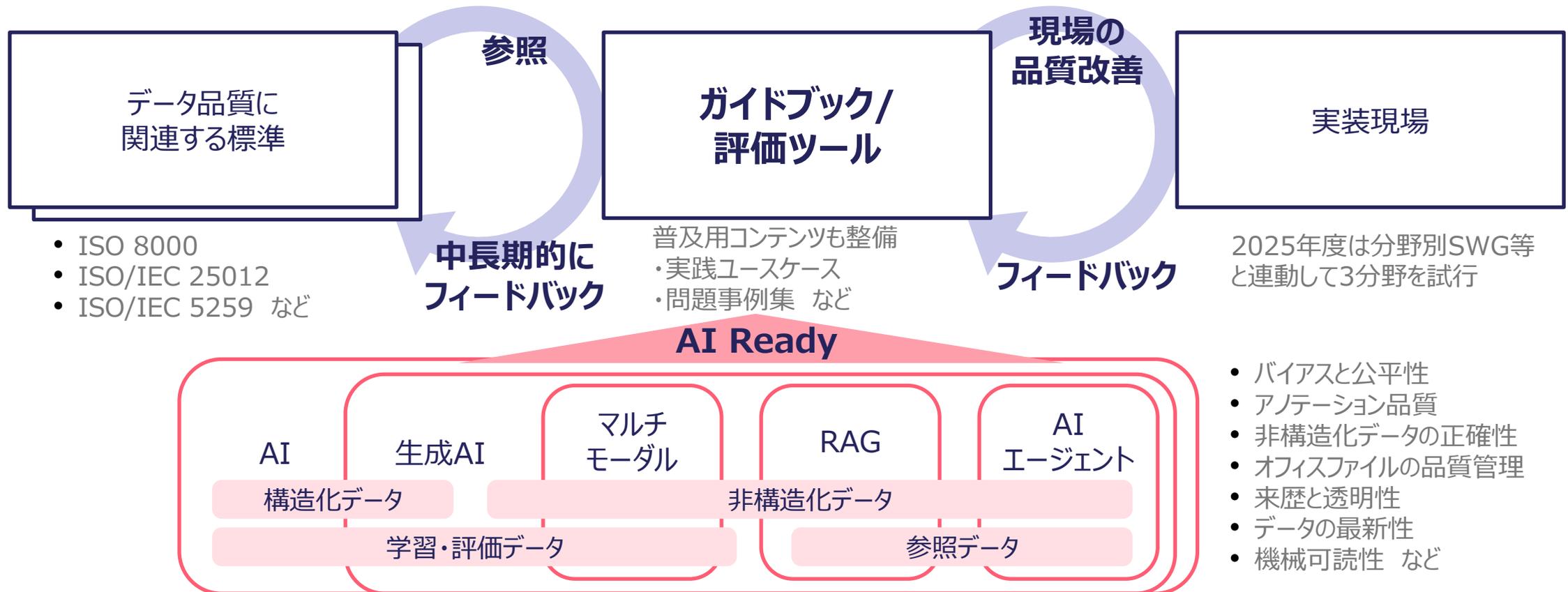
2026年3月10日

AISI事業実証WG 下期報告会

AISI Japan
AI Safety Institute

データ品質SWGの活動の目的

- “Garbage in, Garbage out”と言われるように、データ品質がAIの性能・セーフティの要である
- データが社内外で広く流通する時代において共通的なデータ品質管理の指針が求められる
- データ品質に関する標準類と実装現場を繋ぐ実務的なガイドやツールを整備する



AIセーフティ・インスティテュート (AISI)

2026年3月10日時点、五十音順



独立行政法人情報処理推進機構 (SWGリーダー)

ウイングアーク1st株式会社

株式会社NTTデータ

株式会社NTTデータ バリュー・エンジニア

国立研究開発法人産業技術総合研究所

兼松株式会社

独立行政法人国立印刷局

SAS Institute Japan株式会社

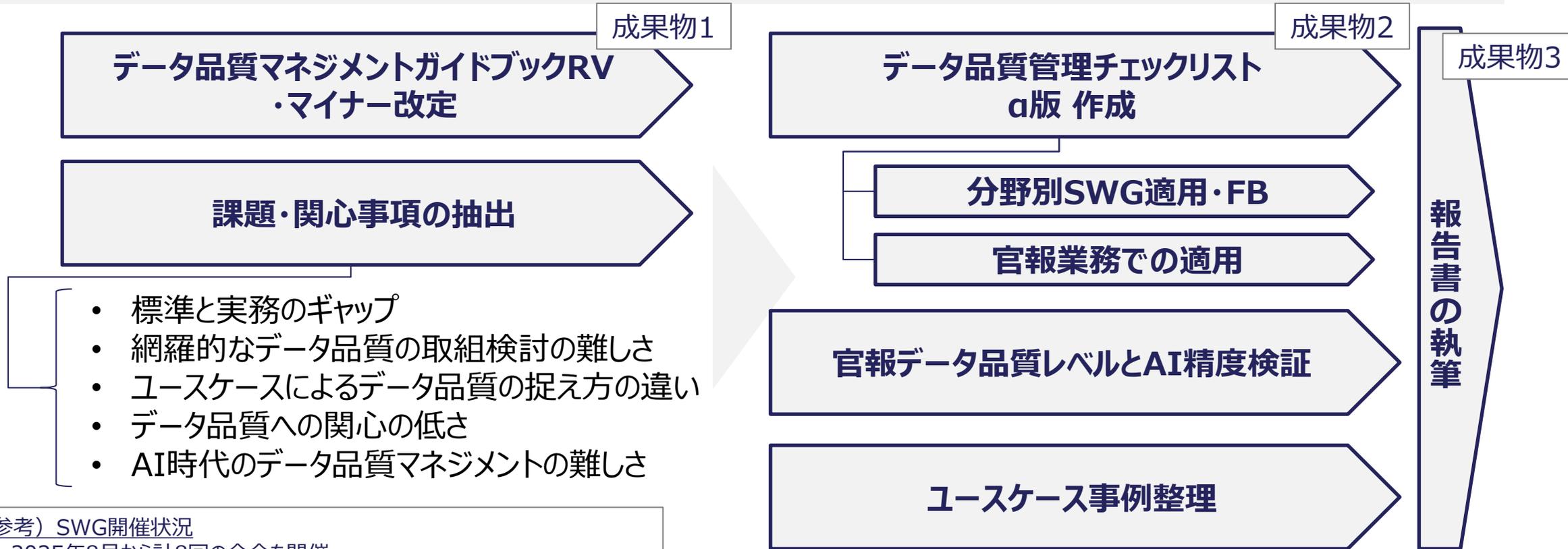
日本電気株式会社

富士通株式会社

PwC Japan 有限責任監査法人

株式会社Preferred Networks

- ガイドブックのマイナー改定や、簡易的なチェックリストの策定、及び現場での適用検証を実施
- 報告書としてそれらの活動状況や、ユースケース事例の整理結果等を記載
- 成果物はAIとデータの設計・運用に携わる実務者を主として、行政担当者、研究者等を広く想定



(参考) SWG開催状況

- 2025年8月から計8回の会合を開催
- 対面／オンラインでのハイブリッド開催
- SWG内外からデータ品質に関する最新動向の共有もあった

- データ品質をライフサイクル、ゲートウェイ、ガバナンスの3つの視点から多角的に評価
- 多様なデータ品質の国際標準を実用的に適用容易なフレームワークに統合
- 今年度は軽微な改訂を実施（表現の見直しや、日本語・機械可読形式の用意等）
- 今後の改善の方向性として、解説書の作成、ユースケース特化、最新のAIの議論の反映等

Data Quality
Management Guidebook

2025-3-31

AISI Japan
IPA
Digital Transformation Center

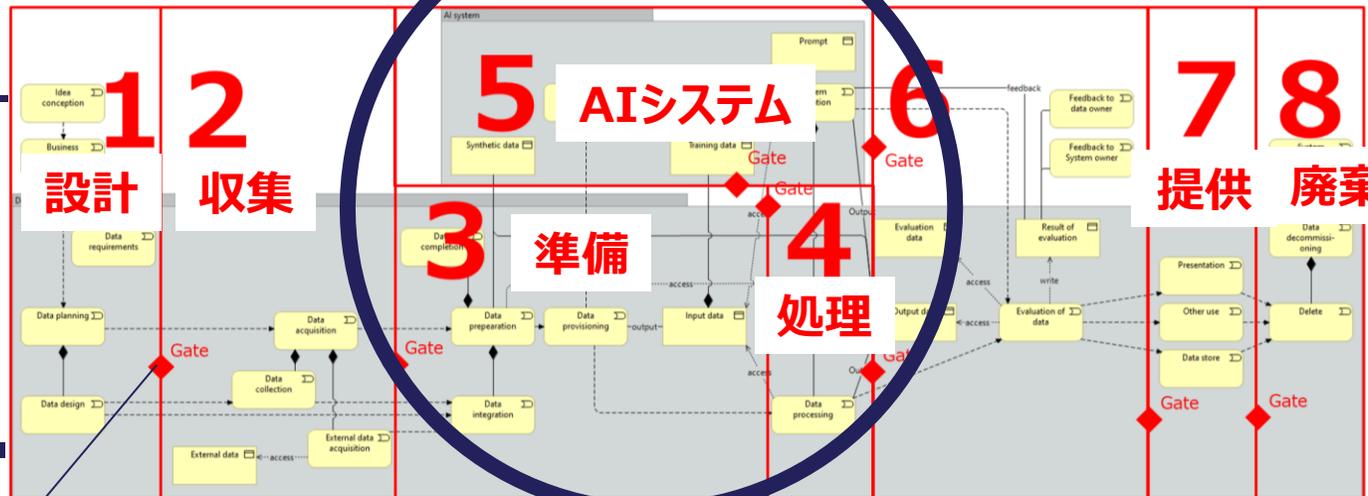
ガバナンス ビュー

組織が持続可能な構造を有していることを確保する
(Ref) ISO 8000-61, ISO/IEC 38505-1, etc.

ライフサイクル ビュー

設計、収集からAI導入、廃棄に至るまでのデータ品質を追跡する

(Ref) ISO/IEC 25024,
ISO/IEC 8183,
ISO/IEC 5259-1, etc.



ゲートウェイ ビュー (品質特性)

主要な意思決定ポイントにおいて品質基準を定義する
(Ref) ISO/IEC 25012, ISO/IEC 5259, etc.

データ品質管理チェックリスト

- データ品質ガイドのデータライフサイクルをベースに、組織が自らのデータ品質の取組状況を網羅的かつ簡易に自己評価できるツールとして、推奨する確認事項一覧（チェックリスト）を策定
- 優先取組項目、実装例、リスク例、主たるデータ品質特性等も付随し、利用しやすい形を目指す

ガイドブック

Data Quality Management Guidebook
- Maximize the value of data and Artificial Intelligence -



チェックリスト

プロセス	項目 (参考日本語訳)	チェック	主なリスク種類	リスク具体例	品質特性	実装者	実装例
データ対策	データ活用するシステムのサービス品質は明確に定義されているか。(スピード、コストなど)		業務効率の悪化	低品質なデータのクレンジング/補正/検証/リソースを割けなければならない。AIモデルの信頼性が落ち、運用効果が低下する。	信頼性	AIシステム開発者	システム全体のチーム内で共有
データ対策	システムに必要なデータはストロップされているか。		業務効率の悪化	低品質なデータのクレンジング/補正/検証/リソースを割けなければならない。AIモデルの信頼性が落ち、運用効果が低下する。	完全性	AIシステム開発者	システムの目的や用途、チーム内
データ対策	各データについて品質管理項目が定義されているか。		業務効率の悪化	低品質なデータのクレンジング/補正/検証/リソースを割けなければならない。AIモデルの信頼性が落ち、運用効果が低下する。	信頼性	AIシステム開発者	システムに活用するデータの完全性、正確性、可用性
データ対策	各データのデータ品質要求レベルは定義されているか。		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す。その結果、組織の意思決定に悪影響を及ぼす。また、法的リスク、倫理的なリスク、不適切な意思決定を招く可能性がある。	信頼性	AIシステム開発者	上記で設定したレベル
データ対策	データアーキテクチャと設計方針が文書化されているか。		業務効率の悪化	低品質なデータのクレンジング/補正/検証/リソースを割けなければならない。AIモデルの信頼性が落ち、運用効果が低下する。	信頼性	データ管理者	データモデルやデータフロー
データ対策	必要なデータは定義され、利用可能か。		競争上の不利	より高いデータ品質を持つ競合他社は、顧客獲得力、業務効率、市場対応力などの主要分野で優位性を確保し、組織は競争力を失う。	可用性	AIシステム開発者	システムに必要な状態にあるデータ
データ対策	データの活用に関する法的規制をアウツしたか。		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。	コンプライアンス	データ管理者	リストアップした用途の制約や条件
データ対策	データプライバシーや機密性が高いものは含まれていないか。		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。	機密性	データ管理者	社内データの場合、社外データの場合は匿名化/仮名化
データ対策	データに関する法的要件は満たされているか。		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。	機密性	データ管理者	適用するデータの用途

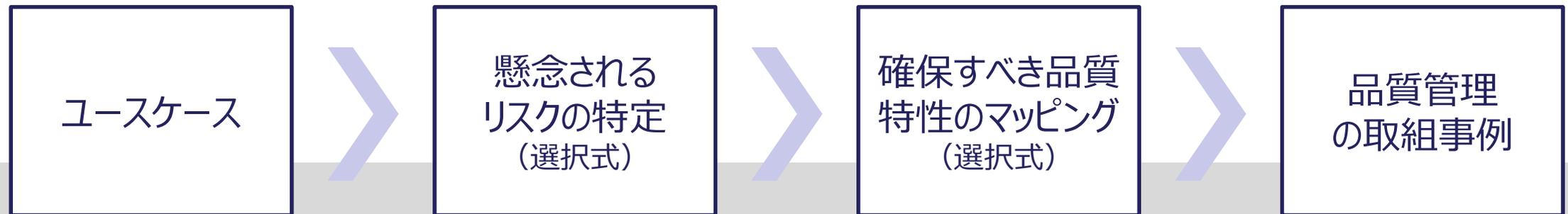


- スプレッドシートベース
- データ品質管理の勘所を抑えた100問超の推奨するチェック項目
- 取組状況はライフサイクル別にレーダーチャートで参考として可視化
- 完全な順守を目的とするものではなく、認識を高め、実務者の支援が目的

- No.
- プロセス
- 推奨確認事項
- 優先事項
- チェック
- エビデンス
- 実施例
- リスク例
- 主たる品質特性
- 実施者
- 対象データ
- EUAI法該当箇所
- AISI評価ツール該当箇所
- 備考

- データ品質は利用目的によって規定されることから、ユースケースに応じた品質管理の整理が求められる
- リスク特定と品質特性の紐付けという、2ステップの整理フレームワークを適用し、そのうえで実際の品質管理の取組を紹介する形で整理した

ユースケース整理のアプローチ



このアプローチに従って4つの事例を報告書に掲載

- AIシミュレーション用の合成データ生成
- 小売業におけるサプライチェーン管理（需要予測AI）
- 自動車事故発生予測AI
- 不動産掲載データを用いた不正取引防止AI

ネクストステップ

- リスク種別や品質特性の選択に実務視点では悩む点があるため、改良の余地がある
- 今後、事例が蓄積されれば、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

- 全般的なデータ品質管理を問うチェックリストの検証を分野別SWGと国立印刷局にて実施
- 具体的なデータ品質について、改正法令抽出AIに対してデータ品質レベルと精度の検証を実施

目的	チェックリストの検証		データ品質とAI精度の検証
対象	分野別SWG	官報データ	
内容	分野別のAIシステム開発にチェックリストを適用した際のギャップ分析、現場適用の容易性等を検証	官報データの品質管理において、チェックリストとのギャップ分析、現場適用の容易性等を検証	官報データに記載された法令改正について、データ品質レベルとAIの読み取り精度の段階的検証
検証協力	分野別SWG協力企業	国立印刷局	国立印刷局、SAS、IPA

AIの進化に対応し、国際協調を視野に、実務・実証を基盤とした データ品質マネジメントの高度化と普及を推進する

短期的な取組み (令和7年度)

- データ品質マネジメントガイドブックの改訂（内容拡充、ガイド性向上）
- 簡易的な品質評価ツール（チェックリスト等）の試作と評価
- データ品質マネジメントガイドブック、簡易ツールに基づく適用検証の実施

中期的な取組み (令和8年度～9年度)

- 複数分野での横断的導入実証による活用モデルの確立
- マルチモーダルデータやマルチエージェント等の新たな利用形態への対応
- 評価ツール機能の本格化

長期的な取組み (将来的なビジョン)

- 組織内での継続的なデータ品質評価の運用モデル確立
- データ品質の議論と社会的な枠組みの連動を推進
- 日本発のデータ品質評価フレームワークや指標定義の国際的な議論へ反映

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

- ◆ 適用検証の促進
 - 2025年度の成果を起点にして、SWG内外で適用検証を推進する
- ◆ 新たな利用形態への対応
 - マルチモーダルAI、Agentic AI、Physical AIなど利用形態拡大に応じたデータ品質管理を扱う
- ◆ ユースケースの明確化
 - データ流通時の品質表示、合成データの品質管理など特定のユースケースを深化する
- ◆ 実務サポート・ツール開発
 - 静的ドキュメントではなく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
 - SWG活動に連動して、ガイドブック等の更新や、フィードバックの受け皿を作る

AISI

Japan AI Safety Institute

官報データを対象とした適用検証

データ品質SWG

独立行政法人国立印刷局

阿部 隆弘

官報データは、推奨する確認事項一覧（チェックリスト）の 改善点の抽出に適切

- ◆ 官報は政府情報として、正確性・完全性が強く求められ、多方面で活用されている重要な情報資源
- ◆ 現在、ガバメントAIのためのデータセットとして官報データの活用が期待されているが、紙面およびウェブ閲覧性を重視した現行設計では、構造化・機械可読性・メタデータが不足しており、AI 活用には大きなギャップが存在
- ◆ 官報データにおけるギャップを明確化することで、チェックリストの実効性および改善余地を評価

“現場”に焦点を当て、自己評価を実施

- ◆ 官報業務の実務に適さない項目（DPL-15～17等）は事前に除外し、評価範囲を適切に限定
- ◆ 各チェック項目について、官報における適用例（対象・粒度、適用条件、工程等）を補記した上で、○／△／×とその理由を記述する形式で評価を実施
- ◆ 評価は、官報のシステム設計・開発・運用管理、編集、製造管理、事務総括などに携わった経験者が担当

現場でそのまま使える手触りとなるよう官報適用例を補記

官報適用例についても意見収集

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。			

検証用のチェックリスト（イメージ）

現場適用や幅広い活用には、「ケースに応じた事例の記載」が有用

主な意見等

- ◆ すべての状況に対応可能な万能型チェックリストを作成することは現実的ではない。
- ◆ 実施者によって判断や解釈にバラつきが生じるのは避けられず、評価のブレを減らすためには事例提示が有効
- ◆ 現場（利用者）からのフィードバックを継続的に収集する仕組み（ループ）を構築してはどうか。
- ◆ チェックリストを工程別に分割して提供すれば、利用者の負担軽減が図れるのではないか。
- ◆ 事例に加えて、NG例の記載も有用ではないか。
- ◆ 官報業務全体のBPRの検討にも活かしたい。

官報を対象としたデータ品質とAIの精度検証

データ品質SWG

SAS Institute Japan株式会社

プルン ワスキト

高付加価値な公共データの例として官報を取り上げ、
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する。



改正された法令を教えてください

AI

(AIの回答)

1

PDF

縦書き・段組等
を含む
文字情報は保
持されるが、解釈
が不安定になり
得る

2

HTML

テキスト化されて
いるが、構造情
報や周辺情報の
混入により誤読
が起こり得る

3

HTML
+ノイズ除去

HTMLタグを除
去しクリーンな状
態とする

4

HTML
+構造付与

条・項等の境界
を明示
(例：タグ付
与)

品質レベル (データの整備水準)

本検証で明らかにしたいこと

- 構造化レベルの違いによるAI回答の正確性とリスクの差 (有効性・安全性)
- 精度向上と準備コストのトレードオフ (費用対効果)

本検証はガイド全体の有効性を評価するものではなく、ガイドが扱う主要論点の一つである「機械可読性／構造化 (Understandability)」を、官報を例に具体化し、WG内の議論材料を提供することを目的とする。

官報16件の中で、**改正方式が2種類**「改め文方式」と「新旧対照表方式」存在して、**その違いによるAI解釈が異なること**に着目して精度評価を行った。

No	名前	官報公布	種別	改正方式
1	法律第八十一号	2025/12/05	法律	改め文方式
2	法律第八十二号	2025/12/10	法律	改め文方式
3	法律第八十八号	2025/12/22	法律	改め文方式
4	政令第三百六十九号	2025/12/24	法律	改め文方式
5	政令第三百八十八号	2025/11/12	政令	改め文方式
6	政令第三百九十五号	2025/11/27	政令	改め文方式
7	政令第四百十二号	2025/11/28	政令	改め文方式
8	政令第四百二十三号	2025/12/12	政令	改め文方式
9	政令第四百四十六号	2025/12/17	政令	改め文方式
10	環境省令第二十二号	2025/12/26	政令	改め文方式
11	国土交通省令第百十四号	2025/11/12	府省令	改め文方式
12	内閣府令第百号	2025/11/27	府省令	新旧対照表方式
13	総務省令第百八号	2025/11/28	府省令	新旧対照表方式
14	総務省令第百十三号	2025/12/10	府省令	新旧対照表方式
15	国家公安委員会規則第十八号	2025/12/23	府省令	新旧対照表方式
16	公正取引委員会規則第十二号	2025/12/26	規則	新旧対照表方式

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。**

改め文方式

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（昭和三十三年法律第二十六号）の一部を次のように改正する。

第八十八条の八を削る。

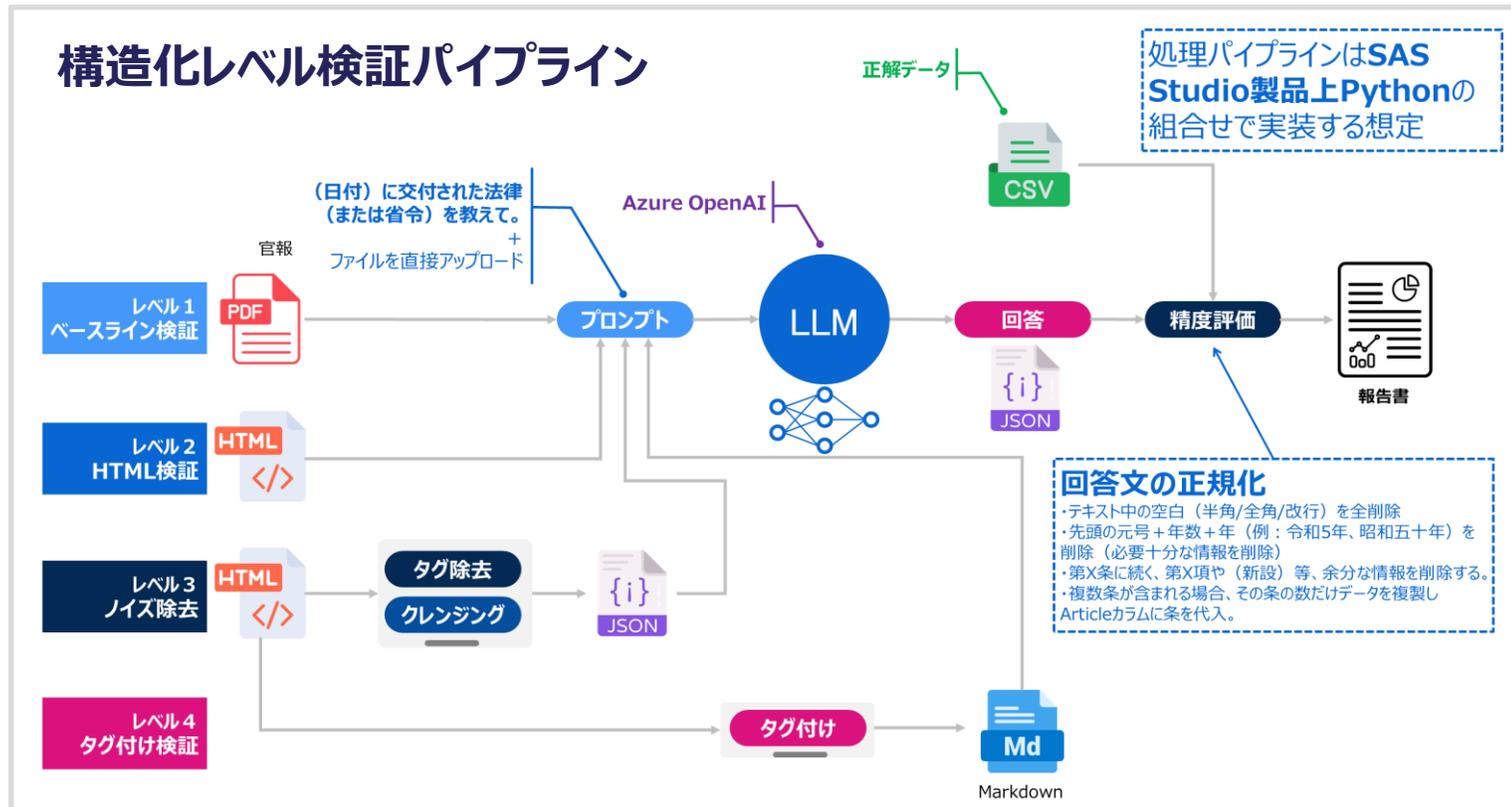
第八十九条を次のように改める。

改め文方式は「第1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。**

新旧対照表方式

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-第三十六条の三十の六の十四）	第二章 福祉の保障（第七条-第三十六条の三十の六の三）
第三章～第四章 [略]	第三章～第四章 [同上]
附則	附則
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。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。	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。第十七条の三第二項第一号において同じ。）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。

官報データを複数形式の入力データとしての**検証パイプライン**を構築し、
構造化の有効性（事前準備した正解データとの差異）を検証



テストデータ

- PDF/HTML/HTML(タグ除去)/Markdown
- Azure OpenAIのAPIでJSON出力機能を使い正解データと同じ構造で出力
- プロンプト：**
データ+ 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…

マッチング

- それぞれの回答データと正解データを比較しTP/FP/FNを算出
- 空白や必要十分な情報は機械的に正規化
- AIからの出力結果に幅があるので、3回テストデータ生成・マッチングを行い平均を算出

「政令の条」と「直前の括弧」が関連していることに着目して、
この**順番を入れ替える**ことによりその**付属関係を明確にする**。

元の官報

令和七年十一月十二日

内閣総理大臣 高市 早苗

****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

内閣は、行政書士法の一部を改正する法律（令和七年法律第六十五号）の施行に伴い、この政令を制定する。

(特定商取引に関する法律施行令の一部改正)

****第一条**** 特定商取引に関する法律施行令（昭和五十一年政令第二百九十五号）の一部を次のように改正する。

別表第二第十五号中「第一条の二第一項又は第一条の三」を「第一条の三第一項又は第一条の四第一項」に改める。

(中小企業等経営強化法施行令の一部改正)

****第二条**** 中小企業等経営強化法施行令（平成十一年政令第二百一号）の一部を次のように改正する。

第十二条第二項第一号中「第一条の二第一項及び第一条の三第一項」を「第一条の三第一項及び第一条の四第一項」に改める。

(有限責任事業組合契約に関する法律施行令の一部改正)

Markdown_improved

令和七年十一月十二日

内閣総理大臣 高市 早苗

****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

内閣は、行政書士法の一部を改正する法律（令和七年法律第六十五号）の施行に伴い、この政令を制定する。

****第一条**** **(号 外)** **(特定商取引に関する法律施行令の一部改正)**
特定商取引に関する法律施行令（昭和五十一年政令第二百九十五号）の一部を次のように改正する。

別表第二第十五号中「第一条の二第一項又は第一条の三」を「第一条の三第一項又は第一条の四第一項」に改める。

****第二条**** **(中小企業等経営強化法施行令の一部改正)** 中小企業等経営強化法施行令（平成十一年政令第二百一号）の
第十二条第二項第一号中「第一条の二第一項及び第一条の三第一項」を「第一条の三第一項及び第一条の四第一項」に改

****第三条**** **(有限責任事業組合契約に関する法律施行令の一部改正)**
有限責任事業組合契約に関する法律施行令（平成十七年政令第二百六十九号）の一部を次のように改正する。

() の括弧書きは直後の **()** の第○条に対応している。

「**新旧対照表方式**」で精度を見ると、構造化レベルによる**精度の違いが極めて小さい**が、**「改め文方式」**では、構造化レベルを上げると、**精度が改善する傾向**にある。

値が0~1の範囲で、1に近いほど良い

レベル	改正方式	改正条数	TP	FP	FN	適合率 Precision	再現率 Recall	F-score
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_tagremoved	改め文方式	277	141	34	136	0.806	0.509	0.624
4: Markdown_improved	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_tagremoved	新旧対照表方式	56	46	3	10	0.939	0.821	0.876
4: Markdown_improved	新旧対照表方式	56	46	5	10	0.902	0.821	0.860

特に**F1値（適合率[Precision]と再現率[Recall]のバランスを評価する指標）**では、レベル2と4の精度が最も高く、**構造化することでハルシネーションを防ぐ効果が期待**できる。

構造化レベルによって、精度が変化したことを確認できたが、
全体でHTML構造(レベル2)がどの評価指標でも最も精度が良かった。

値が0~1の範囲で、1に近いほど良い

構造化レベル	改正条数	TP	FP	FN	適合率 Precision	再現率 Recall	F値
1: PDF	333	140	45	193	0.758	0.419	0.540
2: HTML	333	208	53	125	0.797	0.626	0.701
3: HTML_tagremoved	333	165	60	168	0.732	0.494	0.590
4: Markdown_improved	333	197	72	136	0.732	0.593	0.655

Precision HTML > PDF > HTML(tag除去) = Markdown
Recall HTML > Markdown > HTML(tag除去) > PDF
F値 HTML > Markdown > HTML(tag除去) > PDF

HTMLが精度の高い結果となった理由として、**HTMLタグによって構造化されており、それがAIにもコンテキスト処理しやすくなっているのではないかと考えられる。**

- 適合率 : 正解と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
- 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、全体の精度を示す指標