データ品質マネジメントチェックリスト 活用の手引き

本手引きの要点

本手引きは、データ品質マネジメントチェックリスト（以下「チェックリスト」という。）を実務で使い始めるためのものである。チェック項目への回答に先立ち、対象を明確にすると、実際の処理と証跡に基づいて回答しやすい。本書で示す進め方は、利用目的や業務の実態に応じて調整できる。

チェックリストの目的　チェックリストを用いて、データ品質マネジメントの取組状況を確認し、改善が必要な点について、関係者間で共通認識を形成することである。

利用場面　定期点検、委託先・データ提供元との確認、新規導入又は大幅な変更の前後、品質上の問題が起きた後の振り返りなどで利用できる。

関連資料の使い分け　チェックリストは回答の記録に用い、データ品質マネジメントガイドブックは背景や考え方の確認に用いる。本書は、実務上のチェックの進め方、生成 AI 等を用いた補助の方法及び具体例等を示す。

チェックの進め方　以下の 3 ステップでチェックを進めることができる。ステップ 1 で対象を明確にし、ステップ 2 で実際の処理とデータの流れを整理した上で、ステップ 3 で関係するチェック項目に回答する。

表 1 チェックの進め方

ステップ	行うこと
1	チェックの対象範囲を決める
2	プロセスと入出力データを整理する
3	チェックリストに回答する

チェック結果の活用方法　チェック後はデータ品質の維持・向上に向けて、チェック結果を利用することが望ましい。

ステップ 1　チェックの対象範囲を決める

このステップのゴール　「何を、どこからどこまで、いつの状態で確認するか」を短く説明できるようにする。

対象となる業務、データ、工程、期間・バージョンを整理し、回答の根拠を明確にする。

表 2 対象範囲の例

項目	記載内容	記載例
対象システム／業務	チェックの対象とするシステム又は業務等	アンケート回答データを集計し、公表資料を作成するまでの一連のデータ品質マネジメント業務
利用目的・利用者等	利用目的、利用者、利用頻度、利用範囲、誤情報の影響	利用目的は、企業の取組状況や課題の把握、今後の取組の検討及び調査結果の情報提供。利用者は、公表された調査結果を意思決定の参考とする読者（組織内外）。利用頻度は、年 1 回程度であり、成果物は必要に応じて参照する。誤情報の影響として、公表値の誤りが、利用者の状況認識、関係者の検

項目	記載内容	記載例
		討及び提供主体への信頼に影響し得る
主な対象データ	チェックの対象とする中間成果物等のデータ及び成果物となるデータ	回答データ、加工済データ、集計結果、公表資料
確認する業務・工程	チェック対象の開始点、終了点及び対象とする工程（及び対象外とする工程）	調査票の設計から回答データの受領、集計結果の確認・公表まで
確認する期間・バージョン	どの期間、バージョン、更新回又は確認日を確認するか	2026 年度版
特記事項	データ品質について重点的に確認する点	回答の収集や集計等を委託しているため、発注元が、自組織としてデータの品質維持・向上をどのように行っているか又は委託先の取組をどのように把握しているか確認する必要がある 現行業務では AI モデルの学習、推論等を行っていない

迷ったとき　担当者、管理ルール又は証跡が大きく異なる箇所を境界として、チェックの対象範囲を決める。除外した範囲についても記録しておくことで、結果を共有する際の誤解を防ぐことができる。

ステップ 2　プロセスと入出力データを整理する

このステップのゴール　実際の業務の流れをデータライフサイクルプロセスに対応付ける。

まず、実際のファイル、画面、帳票、バッチ処理及びレビュー等を基に、処理の流れ、各処理の内容及び入出力を確認する。次に、関係するデータライフサイクルプロセスを対応付ける。

例えば、以下のように行う。

- チェック対象の計画から成果物の出力、データの廃棄まで、実際の作業を順に並べる
- 各作業に、担当者、使用するシステム・ツール、入力、出力を付ける
- 手順書、ログ、レビュー記録、承認記録等、回答の根拠になりそうな証跡を付ける
- 各作業を、作業内容と入出力に基づいて一つ又は複数のデータライフサイクルプロセスに対応付ける

表 3 プロセス整理表の例

対応プロセス	入力	実際の作業	出力	担当者・ツール	証跡
データ準備	回答データ	不要データの削除、形式統一	加工済データ	集計担当者	削除や形式統一のルール
データ処理	加工済データ	集計	集計表、グラフ	集計担当者・表計算ソフト	作業記録、集計ファイル
出力確認	集計表、グラフ	別担当者による確認、修正	確認済み資料	レビュー担当者	レビュー記録、承認記録

データライフサイクルに対応付ける　仮置きした作業について、チェックリストの「プロセス」やガイドブックのプロセスビュー（運用）を参照し、主に関係するプロセスを対応付ける。対応付けによって、関係するチェック項目を探しやすくする。

作業の目的に応じてプロセスを対応付ける　同じ作業でも、着眼点により位置付けが変わる。例えば、データ形式変換は「データ準備」と捉えられる一方、外部提供形式を整える目的では「結果提供」にも関係する。処理の目的とライフサイクルを対応させるようにする。

ステップ 3　チェックリストに回答する

このステップのゴール　対象システム又は業務の実態と証跡に基づき、回答とその理由を担当者以外の関係者も追える状態にする。

チェック項目を選ぶ　対象システム又は業務の目的、組織のデータ品質要求に合わせてチェック項目を選ぶ。対象外とする項目は、その理由を記録する。また、回答を進める中で対象外と判断することもできる。代表的なユースケース別の確認項目の選び方は、付録 A を参照する。

回答する　実施状況を確認し、チェックを付ける。回答補足には、確認した事実、証跡、必要に応じて確認日及び確認者等を記載する。確認できた事実と証跡を記載し、未確認の事項は「要確認」として明確にする。

表 4 回答補足の書き方例

避けたい記載	追跡できる記載
「対応済み」とだけ記載	集計手順書 v2.1 に基づき実施。2026 年 8 月 25 日のレビュー記録で承認を確認
「問題なし」とだけ記載	入力件数と出力件数を照合し差分 0 件。照合結果は作業記録に保存
「対象外」とだけ記載	外部へのデータ提供を行わないため対象外

対象外について　対象業務に関係する取組が未実施の場合は、「未検討」又は「検討／準備中」とする。「対象外」は、チェック項目が想定するデータ、処理又は活動が対象範囲に存在しない場合に用いる。

チェック区分のカスタマイズ　チェック区分を段階化する等、自組織向けにカスタマイズすることも可能である。カスタマイズする場合は、原区分との対応関係が分かるようにし、集計結果にもカスタマイズの内容を示す。

判定条件を整理する　チェック項目ごとに、各チェック区分に該当すると判断する条件を整理するとよい。対象システム又は業務では、具体的にどのような事実や根拠となる証跡があれば、「実施済み」、「未検討」等と判断されるのかを整理する。チェックリストの原文を基準とし、ガイドブックで背景・目的を確認した上で、対象業務で確認する事実や証跡を具体化する。

チェック後　改善事項を管理する

改善事項を決める　データ品質マネジメントのチェックは継続的に実施し、データ品質の維持・向上につなげる。チェックを通じて今後さらに改善できる点を把握することで、データの品質を継続的に高めていくことができる。

優先度の検討　未実施項目を一律に扱わず、対象システムやデータの利用目的、リスク、影響を踏まえて改善の優先度を決める。検討に当たっては、「データ品質マネジメントガイドブック」の「2.9 原則」に示される以下の考え方を参考とする。

- 利用者とそのニーズを理解する。
- 目的に適合していることを確保する。
- 品質とコストのトレードオフを考慮する。
- エラーは発生するものと受け止め、完全性を追い求めすぎない。

結果の適切な解釈　実施率やレーダーチャートの形状だけを用いて、システム間、業務間又は組織間の優劣、成熟度又は適合性を単純に比較しない。結果は、対象範囲、回答補足、対象外の理由及び改善事項と合わせて解釈する。「未検討」は現状を示す区分であり、改善の優先度は、利用目的と未対応による影響を踏まえて判断する。

状況に応じた進め方

他部門等がチェックする場合　業務担当者へのヒアリングと証跡の確認を組み合わせ、チェック担当者が回答をまとめることもできる。まず対象システム又は業務を理解し、必要に応じて複数の担当者へ実態を確認する。

表 5 第三者確認における主な確認相手と質問例

確認相手	確認しやすい質問
処理の担当者	実際に使ったファイルは何か。手作業はどこか。誤りに気付くのはどの時点か。
データ担当者	項目定義、コード表、欠損・重複・外れ値の扱いはどこに記録されているか。
システム担当者	入力制御、権限、ログ、バックアップ、変更履歴をどのシステムで確認できるか。
レビュー担当者	何を根拠に承認するか。差戻しや訂正の記録は残るか。
委託先管理者	契約上の責任分界、納品時の検証、エラー時の対応、証跡の保管はどうなっているか。

AI システムを含まないデータ業務への活用　チェックリストは AI システムのためのデータ品質マネジメントを体系化したものであるが、その項目の多くは、AI システムを含まないデータ業務のデータライフサイクルにも活用できる。AI システムを含まない場合は、まずチェック項目の原文が求める活動が対象システム又は業務に存在するかを確認する。例えば、統計処理や分析処理に読み替えると、新たな気付きを得られる場合がある。

表 6 読み替え例

AI 前提の表現	AI システムを含まない場合における読み替え例
推論時入力	定期更新や成果物作成時に投入するデータ
モデル出力	集計結果、グラフ、ダッシュボード、文書案等の中間・最終成果物
バイアス	地域、業種、期間、文書出典等のカテゴリの偏り
RAG	検索・参照に用いる文書、データベース、マスタ、ナレッジ

チェックリスト利用上の注意

チェックリストと本書は、データ品質マネジメントの取組状況を確認し、改善の対話を始めるためのものである。利用に当たっては、次の点に留意すること。

- 個別データの完全性、AI システムの安全性、法令・規格への適合性又は認証取得を保証するものではない。
- 法令、規制、契約、規格等については、利用時点の最新情報を確認し、必要に応じて専門担当者へ相談すること。
- 評価結果は、実施率、回答補足及び改善事項を合わせて判断する。
- 関連資料について、チェック項目の背景や考え方はデータ品質マネジメントガイドブックを、回答区分や各項目の内容はチェックリスト本体を確認すること。版及び公表先は巻末の「参考資料」に記載している。

付録 A　利用状況・確認目的に応じた項目の選び方

本編の「利用場面」はチェックを実施するタイミングを示している。本付録では、対象となる AI・データの利用形態等に応じて、確認するチェック項目を選ぶ際の考え方を示す。

フィルタリングは確認項目を探すための開始点である。ステップ 1 で定めた対象範囲、ステップ 2 で整理したプロセス、利用者への影響及びリスクを踏まえ、必要な項目を追加又は対象外とする。

表 7 利用状況・確認目的に応じた項目の選び方

想定する状況	主なフィルタリング条件・確認方法
AI モデル自体の学習や追加学習は行わず、プロンプトの工夫や RAG による外部データの参照によってサービスを構築する場合。	「学習データ」列を除き、対象データや対象プロセス等で絞り込む。RAG を利用する場合は、検索・参照するデータの取得、準備、管理及び提供に関係する項目も確認する。
AI サービスの開発・提供は行わず、AI 等での利用を想定したデータの作成、整備又は提供を行う場合。	「プロセス」列を用いて、自組織が担当するデータライフサイクル上の工程に関係する項目を絞り込む。例えば、データ計画、データ取得、データ準備、データ廃棄等を確認する。
データ品質マネジメントに初めて取り組む等、どの項目から確認すべきか判断することが難しく、まず主要な項目から確認を始めたい場合。	「優先事項」列で「〇」とされた項目を開始点とする。ただし、利用目的、リスク及びデータ種別に応じて、優先事項以外の項目も確認する。
EU AI Act 等の海外法令・制度への対応が必要となる可能性があり、関連項目を確認したい場合。	「EU AI Act 該当箇所」列等を用いて関連項目を確認する。この対応付けは実務上の整理のための参考情報であり、法的解釈又は公式見解ではない。実際の対応では、最新の法令原文、関連ガイダンス及び専門家の助言等を確認する。

複数のユースケースに該当する場合　該当する条件を組み合わせて確認する。フィルタリングで除外された項目についても、利用目的や影響が大きい場合は確認対象に含めること。

付録 B　生成 AI によるチェックリスト回答支援

生成 AI は、システム又は業務の整理や回答案の作成を補助するために利用できる。ここでは対話型の生成 AI を想定している。

本付録で示すプロンプト例は、チャットに入力した上で、山括弧（＜ ＞）内を対象に合わせて置き換えて使用する。また、【対象の概要】は、ステップ 1 の対象範囲メモの整理結果を記載することを想定している。

利用上の注意　利用前に、組織の情報セキュリティ方針及び利用するサービスの条件を確認し、入力が認められた情報だけを使用すること。本プロンプト例は、生成 AI が正しい回答を保証するものではない。生成 AI の出力は検討の出発点とし、項目の原文、対象範囲、実態及び証跡と照合して最終判断する。

処理単位の調整　生成 AI の性能、入力可能なファイル形式、入力可能な情報量及びファイルの読取り精度は、利用するモデルやサービスによって異なる。必要に応じて、一度に処理するチェック項目数を減らすなど、処理単位を小さくする。サービスによっては Excel ファイルに直接記入したり、転記しやすい形式で出力したりすることができる。

AI へ渡す資料　最新版のデータ品質マネジメントガイドブック（以下「ガイドブック」という。）、最新版のチェックリストを用意する。必要に応じて、手順書、ログ、レビュー記録、出力データ等も追加する。ファイルを渡せない場合は、関係する箇所だけをプロンプト本文へ貼り付ける。同じチャットで続けて実行する場合も、使用する資料名や参照範囲を各プロンプトに記載すると、適切な回答を得やすい。

プロンプト例 1　データライフサイクルプロセス案を作る

```
＜対象システム又は業務名＞で実際に行われている活動を、データライフサイクルプロセスに整理すること。

【添付ファイル】
- データ品質マネジメントガイドブック
- データ品質マネジメントチェックリスト
- 処理フロー、手順書又はヒアリングメモ等、＜対象システム又は業務名＞の作業が分かるもの

【対象の概要】
- 対象システム／業務:
- 利用目的・利用者等:
- 主な対象データ:
- 確認する業務・工程:
- 確認する期間・バージョン:
- 特記事項:

【出力内容】
- 「該当するデータライフサイクルプロセス」「実際の活動」「主な入力データ」「主な出力データ」を列とする整理表

【ルール】
- 出力は活動の整理に限定する。
- 情報がない事項は「要確認」とし、確認質問を示す。
- 境界が曖昧な作業は、主な位置付けと関連するプロセスを併記する。
- データ品質マネジメントガイドブック、データ品質マネジメントチェックリストの内容に整合させること。
```

プロンプト例 2　確認するチェック項目の候補を出す

```
＜対象システム又は業務名＞について、データ品質マネジメントチェックリストで確認すべきチェック項目の候補を抽出すること。

【添付ファイル】
- データ品質マネジメントガイドブック
- データ品質マネジメントチェックリスト
```

- 処理フロー、手順書又はヒアリングメモ等、＜対象システム又は業務名＞の作業が分かるもの
- データライフサイクルプロセスの整理結果

【対象の概要】
- 対象システム／業務:
- 利用目的・利用者等:
- 主な対象データ:
- 確認する業務・工程:
- 確認する期間・バージョン:
- 特記事項:

【出力内容】
1. 確認候補となるチェック項目とその理由
2. 対象外となる可能性がある項目とその理由

【ルール】
- 出力は確認候補となるチェック項目の選定に限定する。
- 情報がない事項は「要確認」とし、確認質問を示す。
- チェック項目は、添付されたチェックリストに実在する項目だけを使用すること。項目 ID 及びチェック項目の原文は、添付したチェックリストの記載をそのまま使用すること。
- データ品質マネジメントガイドブック、データ品質マネジメントチェックリストに整合させること。

プロンプト例 3　チェック基準案を作る

＜対象システム又は業務名＞について、データ品質マネジメントチェックリストに回答するためのチェック基準案を作成すること。

【添付ファイル】
- データ品質マネジメントガイドブック
- データ品質マネジメントチェックリスト
- 処理フロー、手順書又はヒアリングメモ等、＜対象システム又は業務名＞の作業が分かるもの
- データライフサイクルプロセスの整理結果
- 確認候補となるチェック項目

【対象の概要】
- 対象システム／業務:
- 利用目的・利用者等:
- 主な対象データ:
- 確認する業務・工程:
- 確認する期間・バージョン:
- 特記事項:

【出力内容】
1. データ品質マネジメントチェックリストに掲載された各項目について、当該チェック区分と判断するための判定条件
2. 判定条件を設定した理由

【ルール】
- 出力は判定条件の作成に限定する。
- チェックリストの原文に示された要求をそのまま判定条件に反映する。
- 情報がない事項は「要確認」とし、確認質問を示す。
- データ品質マネジメントガイドブック、データ品質マネジメントチェックリストに整合させること。

プロンプト例 4　チェックリストの回答案を作る

＜対象システム又は業務名＞について、データ品質マネジメントチェックリストの回答案を作成すること。

【添付ファイル】
- データ品質マネジメントガイドブック
- データ品質マネジメントチェックリスト
- 処理フロー、手順書又はヒアリングメモ等、＜対象システム又は業務名＞の作業が分かるもの
- データライフサイクルプロセスの整理結果
- 確認候補となるチェック項目
- 確認候補となるチェック項目の判定基準

【対象の概要】
- 対象システム／業務:
- 利用目的・利用者等:
- 主な対象データ:
- 確認する業務・工程:
- 確認する期間・バージョン:
- 特記事項:

【出力内容】
- 項目 ID（チェックリストの「No.」列に該当）
- 回答案（チェックリストの「チェック」列に該当）
- 回答補足（確認できた事実と証跡）（チェックリストの「回答補足（エビデンスなど）」列に該当）
- 未確認事項及び追加質問

【ルール】
- 判断に必要な情報又は証跡が不足している場合は、「チェック」列を空欄とし、「回答補足（エビデンスなど）」列に「要確認」と記載すること。「要確認」をチェック区分として使用しないこと。
- チェック項目は、添付されたチェックリストに実在する項目だけを使用すること。項目 ID 及びチェック項目の原文を変更又は新規作成しないこと。
- データ品質マネジメントガイドブック、データ品質マネジメントチェックリストに整合させること。

プロンプト例 5　ネクストアクション案を作る

＜対象システム又は業務名＞について、データ品質マネジメントチェックリストのチェック結果から、ネクストアクション（データ品質を向上・維持するための改善策）を整理すること。

【添付ファイル】
- データ品質マネジメントガイドブック
- **回答済みの**データ品質マネジメントチェックリスト
- 処理フロー、手順書又はヒアリングメモ等、＜対象システム又は業務名＞の作業が分かるもの
- データライフサイクルプロセスの整理結果

【対象の概要】
- 対象システム／業務:
- 利用目的・利用者等:
- 主な対象データ:
- 確認する業務・工程:
- 確認する期間・バージョン:
- 特記事項:

【出力内容】
1. 改善を検討するチェック項目と確認された＜対象システム又は業務名＞の課題点

2. 課題に対して、データ品質を向上・維持するための改善案
3. 改善案の優先度案とその理由
4. 担当候補、期限案及び完了を確認する証跡

【ルール】
- 実施状況の判断は実態と実在する証跡に基づくこと。
- 証跡不足と取組の未実施を区別すること。
- 担当者又は期限を資料から確認できない場合は、具体的な人名や日付を推測せず、「担当要設定」「期限要設定」とすること。担当候補を示す場合は、個人名ではなく役割名で示すこと。
- データ品質マネジメントガイドブック、データ品質マネジメントチェックリストに整合させること。

付録 C　チェックリストの適用例

付録 C では、実際にチェックリストの適用検証を行った二つの事例について、本編で示した 3 ステップをどのように実施するかを示す。

二つの事例は、ともに AI システムを含まないデータライフサイクルである。また、データ管理組織がチェックリストに回答するのではなく、他組織がデータ管理組織へのヒアリングを通じてチェックリストを適用した事例である。

C.1　適用例の見方

各事例では、本編で示した 3 ステップの実施例と併せて、データ品質マネジメント上のポイント及びチェックリスト適用時のポイントについて、他の事例にも応用できる考え方を中心に示す。

なお、AI システムを含まない場合は、チェック項目の背景・目的を理解し、一般的なデータ品質マネジメントの観点に読み替えて確認している。

また、チェックリストの判定結果、事例固有の改善案等の詳細については、適用検証レポートを参照する。

表 8 適用例に示す内容

ステップ	本付録で示す内容
ステップ 1　チェックの対象範囲を決める	チェックの対象範囲の設定例と、その範囲とした考え方
ステップ 2　プロセスと入出力データを整理する	実際の活動とデータライフサイクルプロセスの対応例の抜粋及び対応付けの考え方
ステップ 3　チェックリストに回答する	チェック項目の選定例、確認した事実・証跡、判定結果及び判断根拠

C.2 事例 1 アンケート調査データの集計・分析及び調査資料作成

この事例から確認できること

本事例では、企業等を対象としたアンケート調査を実施し、回答データを集計・分析して、レポートやデータ集を作成する業務を取り上げる。回答の収集、クレンジング、集計等の一部を外部委託しており、Web フォーム、表計算ファイルを印刷した書面等の複数の方法で回答データを取得している。

アンケート調査では、適切に調査工程が管理されていない場合、結果の解釈を誤るおそれや、回答データの集計結果が実態を適切に反映しないおそれがある。適切に管理されていない例としては、経年比較する際に回答者構成や調査条件が異なることや、設問構造や回答内容に応じた処理が適切に行われていないことが挙げられる。また、データの収集、加工、集計に複数の主体が関与する場合には、主体ごとに処理方法や判断基準が異なることで、作業間に不整合が生じることもある。

以上を踏まえ、次の点に着目してデータ品質マネジメントを確認する。

特に着目する点	本事例で確認すること
調査結果を適切に解釈・比較できる状態にする	業種、企業規模等の回答者構成や補正の影響を把握するとともに、設問、調査対象、回答方法、集計条件等の実施回間の変更内容とその影響を確認する
回答データを分析に利用できる状態にする	設問分岐への対応、無効回答の判定等は一定の判断を伴うため、処理ルールや判断基準が整理された上で実施されているか確認する
委託先と発注側を通じてデータ品質を確保できる状態にする	委託先と発注側の役割、データの受け渡し、クレンジング・集計、確認・差戻し・再確認を整理し、回答データから公表値までの処理と確認を追跡できる状態かを確認する

C.2.1 ステップ 1 チェックの対象範囲を決める

チェック対象とする業務の成果物がどのような目的で利用されるのか、利用者は誰になるのか、チェックの対象となるデータや工程はどこになるのか等を以下のように整理した。

表 9 事例 1 のチェックの対象範囲

確認事項	記載内容
対象システム／業務	企業等の取組状況や課題をアンケートにより把握し、回答を集計・分析して、内部での検討及び外部への情報提供を行う業務。
利用目的・利用者等	利用目的は、企業の取組状況や課題の把握、今後の取組の検討及び調査結果の情報提供。利用者は、公表された調査結果を意思決定の参考とする組織内外の読者。アンケート調査は年 1 回程度実施し、調査結果は必要に応じて参照する。誤情報の影響として、公表値の誤りが、利用者の状況認識、関係者の検討及び提供主体への信頼に影響し得る
主な対象データ	企業等から取得した回答データ、クレンジング済みデータ、集計済みデータ、成果物のレポート等。
確認する業務・工程	調査の計画から、回答データの取得、クレンジング等のデータ準備、集計・分析、結果の確認、レポート・データ集の作成・提供、データの管理までを対象とする。
確認する期間・バージョン	2026 年度調査

確認事項	記載内容
特記事項	回答の収集や集計等を委託しているため、発注元が、自組織としてデータの品質維持・向上をどのように行っているか又は委託先の取組をどのように把握しているか確認する必要がある。

データ品質マネジメントとチェック時のポイント　調査結果の解釈・比較に影響する条件を把握する

継続的なアンケート調査では、調査対象、設問・選択肢、回答方法、集計条件、回答者構成等が、調査結果の解釈や実施回間の比較に影響する。このため、調査結果を適切に解釈・比較できるよう、これらの条件や変更内容を把握することが重要である。

データ品質マネジメントのチェックを行う際には、欠損や鮮度といったデータ自体の状態だけでなく、調査結果の解釈や実施回間の比較に影響する調査条件も確認対象に含める。これらの条件は調査の計画段階で定められることが多いため、利用目的や利用者を踏まえ、計画段階を含めて確認範囲を設定する。

C.2.2　ステップ 2　プロセス及び入出力データを整理する

対象業務で実際に行われている活動を確認し、データライフサイクルプロセスへ対応付けた。以下に対応付けの抜粋と対応付けの考え方を示す。

表 10 事例 1 のデータライフサイクルプロセスへの対応抜粋

プロセス	実際の活動	対応付けの考え方
データ計画	設問及び選択肢等を定め、関係者によるレビューを行う。 必要な有効回答数及び属性ごとの回答状況の確認方法を定め、必要に応じて回答者構成の補正について方針を定める。	設問構造、調査で取得するデータの条件及び利用する際の留意点等を事前に定めているため「データ計画」。
データ取得	発注側から受領した設問等の仕様書に基づき、委託先は調査対象となる企業等へ回答方法を案内し、Web 等から回答データを収集する。	Web 回答データを収集しているため、委託先の作業であるが「データ取得」。
データ準備	委託先はクレンジングを行うとともにローデータ及び設問・選択肢との対応を示すレイアウト表等を発注側へ納品する。 発注側はクレンジングが正しく行われたかどうかを確認する。	本業務で目的とする処理が行えるよう、データの形式を変換し、その結果を確認しているため「データ準備」。

データ品質マネジメントとチェック時のポイント　回答データを分析に利用できる状態にする

アンケート調査のデータは、取得した状態のままでは分析に利用できない場合があるため、設問構造や回答内容を踏まえ、分析に利用可能な形に整理する必要がある。特に、設問分岐がある場合のデータ構造の整理や、異常値・不整合が認められる回答を無効回答として取り扱うか否かの判断などについては、場当たり的な対応を避け、一貫した処理を行う。このため、あらかじめ処理ルールや判断基準を整理した上で、クレンジングを実施する。

データ品質マネジメントのチェックに当たっては、こうした処理ルールや判断基準がデータ計画段階又は当該工程において整理されているか、また、実際の作業がそれらに基づいて一貫して実施されているかを確認する。

C.2.3 ステップ 3 チェックリストに回答する

C.2.3.1 チェック項目を選ぶ

対象業務に関係する項目を抽出するとともに、対象業務に該当する活動が存在しない項目については「対象外」とする。

表 11 事例 1 でチェック対象外とした項目の例

No.	チェック項目	扱い	本事例での判断
2.DAC-14	合成データを利用する場合にはその旨が明示されており、明示されていない合成データが含まれていないことを確認している。	対象外	合成データを利用していないため。
3.DPE-04	AI モデル構築のためのデータの分割(学習・検証・評価等)が、目的に応じて論理的かつ適切に行われていることを確認している。	対象外	本業務には AI モデル構築は含まれず、学習・検証・評価データの分割は必要ないため。
4.DPO-04	処理の並列化・分散化が、適切に設計されている。	対象外	集計対象のデータ量が小さいため、並列・分散処理は該当しない。

チェック時のポイント AI を前提とした項目をデータ品質マネジメントに活用する

次のチェック項目は AI システムを前提とした表現を含むが、アンケート調査では、調査対象となる集団と実際の回答データの分布の関係を確認する項目として読み替えることができる。

表 12 事例 1 における AI システムを前提とした項目の読み替え方

No.	チェック項目	本事例における読み替え例
2.DAC-15	取得するデータが、AI システムの対象となる集団の属性や分布を適切に反映していることを確認している。	「AI システムの対象となる集団」を「調査・分析の対象となる企業等の集団」と捉え、業種や企業規模等の属性別に回答の集まり方を確認する項目として読む。

アンケート調査では、回答者の構成と調査対象となる集団の構成に乖離がある場合、その差を調整する手法の一つとして、ウェイトバック集計(調査対象となる集団の構成比等に基づき回答に重みを付けて集計する方法)等がある。本事例では、回答データと調査対象となる集団の分布を比較した上で、ウェイトバック集計を行った場合の集計結果への影響を検証し、その影響が軽微であることを確認するなど、回答データの特性を踏まえて補正等の対応の要否を判断していた。

C.2.3.2 回答する

抽出したチェック項目について、実際の業務内容や確認した資料等を踏まえ、いずれかのチェック区分を選択し、その判定根拠を回答補足に記載する。

本事例における具体的な回答例を以下に示す。

表 13 事例 1 のチェックリスト回答例

No.	チェック項目	チェック区分	回答補足の記載例
2.DAC-10	取得したデータが、定義されたデータ品質要求レベルを満たしていることを確認している。	実施済み	発注側でも受領したローデータを確認し、設問の分岐条件に合わない回答やロジック上の不整合等がある場合は委託先へ確認・差戻しを行っている。必要に応じて修正したデータの再提出を受けている。
3.DPE-16	データセットやサンプリング結果における各カテゴリの分布が、利用目的に照らして適切な水準となっていることを確認している。	実施済み	業種・企業規模等の属性別に回答状況を確認し、回答の過不足を把握している。また、主要な設問について回答者構成を補正した場合と補正しない場合の結果を比較し、補正の要否を判断している。
3.DPE-04	AI モデル構築のためのデータの分割(学習・検証・評価等)が、目的に応じて論理的かつ適切に行われていることを確認している。	対象外	対象業務ではアンケート回答データの集計・分析を行っており、AI モデルの構築や、学習・検証・評価用にデータを分割する活動は行っていないため、本項目は対象外とした。

データ品質マネジメントとチェック時のポイント　委託先と発注側の確認をつなぎ、判断の根拠を残す

回答データの収集、クレンジング、集計等を委託する場合でも、発注側が調査仕様や処理条件を定め、納品されたデータや集計結果を確認し、必要に応じて差戻しや再確認を行うことで、委託先と発注側の双方を通じてデータ品質を確保する。

データ品質マネジメントのチェックを行う際には、委託先が処理を実施していることだけでなく、発注側が何を確認し、どのような場合に差戻しや再確認を行っているかを確認する。回答補足には、確認した事実や証跡を記載し、回答データから集計結果、公表値までの処理と確認を追える状態にする。

C.2.4　チェック結果を基に改善案を検討する

本事例では、既存の確認作業を共通の仕様及び証跡に結び付け、再現可能な状態にすることが改善の中心であると整理された。また、機械判読可能な形式での提供を既存の公表手順に組み込み、利用ニーズと運用負担を踏まえて段階的に進めることも案として挙げられる。

C.3 事例 2 複数データを統合したダッシュボードの作成・利用

この事例から確認できること

本事例では、複数の公開データや組織内データを取得・加工・統合し、地域ごとの状況を可視化するダッシュボードを作成する業務を取り上げる。データソースごとに更新時期、粒度、項目定義及びファイル形式が異なり、取得後に形式変換、統合、補完等の処理を行っている。

このようなケースでは、例えば、ダッシュボードの一部の指標に想定と異なる値が表示されることがある。その原因を確認すると、データごとに対象年度や地域コードの粒度が異なることや、データの統合に当たってコードの補完等の処理が行われていることが影響している場合がある。また、こうした加工・統合の方法が、担当者個人の判断や経験に依存している場合もある。

よって本事例では、次の点に注目してデータ品質マネジメントを確認する。

特に着目する点	本事例で確認すること
異なる条件のデータを比較できる状態にする	利用するデータの対象期間・バージョン、更新時期、粒度、項目定義等を把握し、ダッシュボード上で比較できるよう整理する
元データから表示値までのつながりを追える状態にする	形式変換、コード付与、補完、結合、集計等の処理を整理し、元データから表示値までの対応関係を確認する
更新時にも同じ処理を再現できる状態にする	データの取得、統合、補完等のルールや証跡を整理し、継続的に更新・確認できる状態かを確認する

C.3.1 ステップ 1 チェックの対象範囲を決める

本事例では、ダッシュボードそのものだけでなく、利用するデータの選定、取得、形式変換、コード付与、粒度の統一、BI ツールへの取込み、集計・可視化、表示値の確認及び共有・版管理までを対象とした。

表 14 事例 2 のチェックの対象範囲

確認事項	記載内容
対象システム／業務	地域に関する複数の統計・業務データを集約・可視化し、地域ごとの状況把握や、必要に応じて外部関係者へ説明する際のダッシュボード作成・運用業務。
利用目的・利用者等	利用目的は、複数データの可視化・比較による状況把握及び分析支援。利用者は主として作成者とし、必要に応じて外部関係者にも提示する。利用は、定期的又は必要に応じて行う。誤情報は、分析、状況把握及び外部説明に影響し得る。
主な対象データ	人口、経済・労働、企業数、各種認定・評価情報、試験関連統計等について、公的統計サイト、公開サイト、組織内の担当部署等から取得する CSV、Excel、PDF 等のデータ。
確認する業務・工程	利用する指標・データの計画から、データ取得、形式変換・統合、BI ツールでの処理・可視化、表示結果の確認、共有・版管理までを対象とする。
確認する期間・バージョン	2026 年 6 月時点
特記事項	複数のデータソースを利用しており、更新頻度、粒度、項目定義、ファイル形式が異なる。

データ品質マネジメントとチェック時のポイント　比較するデータの条件を把握する

複数のデータを統合する場合は、利用するデータの取得元、対象期間・バージョン、更新時期、粒度、項目定義等を整理する。例えば、ダッシュボードの場合、対象期間が異なるデータを利用していることをあらかじめ把握及び整理しておけば、処理の不整合の低減や、利用者の誤解の回避につながる。

データ品質マネジメントのチェックを行う際にも、チェック対象となるデータの範囲を適切に整理する。完成したダッシュボードだけでなく、利用するデータの範囲やその特徴を把握することで、以降のチェックで重点的に確認すべき点を明確にできる。

C.3.2　ステップ 2　プロセス及び入出力データを整理する

対象業務で実際に行われている活動を確認し、データライフサイクルプロセスへ対応付けた。以下に対応付けの抜粋と対応付けの考え方を示す。

表 15 事例 2 のデータライフサイクルプロセスへの対応抜粋

プロセス	実際の活動	対応付けの考え方
データ準備	Excel、CSV 等のデータを、後続処理で利用可能な表形式に変換する。	本業務で目的とする処理が行えるよう、データの形式を変換しているため「データ準備」。
	PDF 等から取得したデータについて、OCR や生成 AI 等を補助的に利用して、Excel の表形式への変換や行・列の変換を行う。	本業務で目的とする処理が行えるよう、データの形式を変換しているため「データ準備」。OCR や生成 AI はデータ形式変換の補助に利用しているため、「データ準備」に位置付ける。
データ処理	加工済みデータを BI ツールへ取り込み、結合・集計・割合計算等を行う。	本業務で目的とする集計処理を行っているため「データ処理」。
出力評価	BI ツールの表示値、合計値、割合、グラフ等に違和感がある場合、元テーブルやレコードまで遡って原因を確認する。	公開前の最終確認として行っているため「出力評価」。

データ品質マネジメントとチェック時のポイント　加工・統合の流れを追い、元データと表示値をつなぐ

形式変換、コード付与、補完、結合、集計等の処理について、入力と出力を整理し、元データから表示値までのつながりを確認するとよい。どの処理で値・意味が変わり得るかが整理され、最終的な数値に不整合が生じたときに原因を特定しやすくなる。

データ品質マネジメントのチェックを行う際には、実際の活動や入出力をデータライフサイクルの各プロセスに対応付けて整理する。例えば、データの「変換」に該当する処理が複数の工程に含まれるなど、どのプロセスに対応付けるべきか判断に迷う場合もある。このような場合には、作業名のみで判断するのではなく、当該処理の目的や前後の入出力を踏まえて整理することが有効である。実務上の活動を共通のフレームワークに基づいて整理することで、関係者間で認識を共有しやすくなるとともに、改善策を検討する際にも、各プロセスに対応する管理策や改善事例を参照しやすくなる。

C.3.3 ステップ 3 チェックリストに回答する

C.3.3.1 チェック項目を選ぶ

対象に関係する項目を抽出するとともに、対象に該当する活動が存在しない項目については「対象外」とする。

表 16 事例 2 でチェック対象外等とした項目の例

No.	チェック項目	扱い	本事例での判断
3.DPE-04	AI モデル構築のためのデータの分割(学習・検証・評価等)が、目的に応じて論理的かつ適切に行われていることを確認している。	対象外	AI モデル構築は行っていないため。
3.DPE-11	データに、不適切な内容や公序良俗に反する内容が含まれていないことを確認している。	対象外	対象のダッシュボードは数値や図表の表示のみであるため。
4.DPO-04	処理の並列化・分散化が、適切に設計されている。	対象外	利用用途及びデータ量を踏まえると、並列化・分散化による処理性能の向上を必要としていないため。

チェック時のポイント AI を前提とした項目をデータ品質マネジメントに活用する

AI を前提とした項目は、対象業務に対応するデータ・処理・活動を確認し、共通するデータ品質マネジメントの観点を活用する。

例えば、次の項目のように、対象業務の表現に読み替えて確認できる。

表 17 事例 2 における AI システムを前提とした項目の読み替え方

No.	チェック項目	本事例における読み替え例
3.DPE-12	AI システムの構築に必要なデータが、機械可読な形式又はインターフェースで提供されている。	「AI システムの構築に必要なデータ」を「ダッシュボードの分析・連携・処理に必要なデータ」と捉え、PDF 等を Excel 等の後続処理可能な形式へ変換できているかを確認する。
6.EOO-02	信頼できるデータや根拠と整合しない出力が行われていないことを確認している。	「出力」を BI ツールの表示値、合計値、割合、グラフ等と捉え、元データや算出条件と整合しているかを確認する。

読み替えた上で原文が想定する活動が存在しない場合は、その項目を「対象外」とする。背景・目的に共通する一般的なデータ管理活動がある場合は、読み替えを行うことで幅広くデータ品質マネジメントの取組をチェックすることができる。

C.3.3.2 回答する

抽出したチェック項目について、実際の業務内容や確認した資料等を踏まえ、いずれかのチェック区分を選択し、その判定根拠を回答補足に記載する。

本事例における具体的な回答例を以下に示す。

表 18 事例 2 のチェックリスト回答例

No.	チェック項目	チェック区分	回答補足の記載例
1.DPL-02	AI システムに必要なデータがリストアップ・整理されている。	実施済み	「AI システムに必要なデータ」を、ダッシュボードの作成に必要なデータとして読み替えた。人口、企業数、各種認定・評価情報等、利用するデータは整理され、取得元や取得条件についても資料等で把握している。
3.DPE-01	データ統合に関する方針及びルールが策定されている。	検討／準備中	都道府県名、2 桁コード、5 桁コード等を対応付けたマスタを作成し、BI ツール上で複数データを結合している。コードがないデータには県名からコードを付与するなど、実務上のルールは存在するが、統合方針・結合ルールとして文書化されていない。
4.DPO-03	データパイプラインの冪等性が担保されている。	未検討	同じ入力・条件で処理を再実行した場合に同じ結果となることを確認する仕組みや手順は定められていない。

データ品質マネジメントとチェック時のポイント　統合・補完のルールを残し、更新時にも再現できる状態にする

本事例では、都道府県コードを用いたデータの結合を行っている一方、コードの優先順位や補完ルールは文書化されていない。更新時にも一貫した方法で処理できるようにするとともに、担当者以外でも同様の処理を行えるよう、統合・補完のルールを明確化し、必要に応じて文書化することが重要である。

データ品質マネジメントをチェックする際には、実際に処理が行われていることだけでなく、その方法やルールが整理され、更新時にも同じ考え方で処理を再現できる状態にあるかを確認する。

C.3.4　チェック結果を基に改善案を検討する

本事例では、業務規模に応じた最小限の管理を整備し、担当者の知見を再利用可能な形にすることが改善の中心であると整理された。具体的には、次の取組が挙げられる。

- データソース、取得日、対象期間、利用条件及び更新頻度を一覧化し、ダッシュボード上の表示と対応付ける。
- 統合、補完、単位・コード変換、除外等の処理手順を記録し、主要な集計値について再実行可能な確認手順を設ける。
- 更新時に、件数、欠損、重複、値域、前回値との差異等を確認する簡易な検証を実施する。

参考資料

- AISI「Data Quality Management Guidebook」：Version 1.02、2026 年 5 月 14 日公表
  https://aisi.go.jp/assets/pdf/260514_Data_quality_management_guidebook_en.pdf
- AISI「データ品質マネジメントチェックリスト」：Version 1.00、2026 年 5 月 14 日公表
  https://aisi.go.jp/assets/file/Data_quality_management_checklist.xlsx
- AISI「データ品質マネジメントチェックリスト適用検証レポート」：Version 1.00、2026 年 11 月 18 日公表
  https://aisi.go.jp/assets/pdf/261118_Data_quality_management_checklist_practical_application_evaluation_report_ja.pdf

発行主体

本書は、AISI データ品質サブワーキンググループの活動の一環として発行されたものである。

改訂履歴

- 1.00 版（2026-11-18）
  - 初版公開。

免責事項

本書に記載した情報の正確性及び信頼性の確保には最大限努めているが、その内容について明示又は黙示を問わず、いかなる保証も行うものではない。いずれの組織も、本書の利用に起因して生じたいかなる損失、損害、請求についても責任を負わない。すべての情報は現状有姿で提供されており、その内容に基づいて行われる一切の行為については、読者が全面的な責任を負うものとする。