

【ホワイトペーパー】

「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」



2022年1月15日初版

一般社団法人データ社会推進協議会
技術基準検討委員会 WG2 TG4

目次

Executive Summary	3
第1章 はじめに	4
1.1 背景	4
1.2 問題意識	4
1.3 センシングデータとは	5
1.4 本書の位置づけ	5
1.5 今後に向けて	5
第2章 検討例：センシングデータの品質評価に利用する品質測定量と品質評価	7
2.1 デバイス依存の品質測定量	7
2.2 デバイス依存の品質測定量の評価パラメータの検討	7
2.3 デバイス依存の品質測定量の評価方法とスコアリング基準の検討	9
2.4 デバイス依存の品質測定量によるデータ品質評価（イメージ）	14
第3章 本書について	15
執筆者	15
TG4 参加者一覧	15
団体概要	15
謝辞	16
参考資料	17
【関連規格】	17
【参考文献】	17
【関連ガイドライン等】	18
用語	19
データに関する用語	19
ハードウェアに関する用語	19
品質に関する用語	20
データ品質に関する用語	20
その他(関連用語)	21
略語	22
APPENDIX-1：想定ユースケース	23
APPENDIX-2：想定ユースケースにおける7つの利用シナリオ	24
APPENDIX-3：センシングデータ品質に関するアンケート結果	31
APPENDIX-4：センシングデータの品質モデル（案）	37
ANNEX-1：データ品質に関する国際潮流・国内動向	42
1. 国際潮流	42
2. 国内動向	42
ANNEX-2：データ品質に関する基本的な考え方の整理	43

1.	データ品質とは何か	43
2.	データ品質測定・表示・評価	43
3.	データ品質評価の主体	43
4.	データ品質基準の活用場面	44
5.	先行文献	44
6.	検討対象とするデータ種類	44
ANNEX-3 : SQuaRE の概要		46
1.	SQuaRE とは?	46
2.	SQuaRE の品質モデル	46
3.	品質特性・品質副特性	47
4.	データ品質測定量 (QM)	48
5.	データ品質測定量要素 (QME)	50
6.	データ品質モデルの適合性	50
ANNEX-4 : 故障のしにくさと耐久性の違い		52
この文書について		53

Executive Summary

我が国が目指すべき未来社会の姿である Society5.0 は、フィジカル空間（現実空間）とサイバー空間（仮想空間）が高度に融合したサイバーフィジカルシステム（Cyber Physical System : CPS）によって実現される。

サイバーフィジカルシステムの品質は、収集・利用されるデータの品質に大きく影響される。特に、大量のセンシングデータを収集し、機械学習アルゴリズム（AI）を用いて解析・制御する場合には、リアルタイムに収集されるデータの品質を測定・表示し、複数の関係者間で容易に評価できる仕組みが非常に重要になる。

このため、一般社団法人データ社会推進協議会（略称：DSA）技術基準検討委員会 WG2-TG 4 では、データ品質評価基準の策定を進めている。本書はその一環として、センシングデータの品質評価手法を提案する。

センシングデータの品質評価を行う場合、データの内容のみを評価しても、機器の故障や通信エラーといった、デバイス依存の要因による問題が検知できないことがある。このため、デバイス依存の要因の種類ごとに、評価対象データの品質に与える影響を整理する必要がある。また、品質測定結果をもとに、デバイス依存の要因項目ごとの値の大きさを比較したり、全体のバランス・傾向などを一覧できることが重要である。

そこで本書は、ISO/IEC25021（JIS X25021）が定めるデータ品質モデル（及びデータ品質特性）に基づき、ISO/IEC25024（JIS X 25024）が定めるデータ品質測定量（QM:Quality Measurement）をセンシングデータに適用するために、新たなデータ品質測定量及びデータ品質測定量要素（QME:Quality Measurement Element）の識別を試みた。そのうえで、品質測定量の評価結果を、0.0～5.0 の値域に換算し、評価スコアを算出。評価スコアをもとにランク付けを行うとともに、品質特性ごとに品質レベルが把握できるよう、レーダーチャートなどで提示する方法を提案する。

本書が新たに定めた品質測定量の対象実体は、センシングデータに付与されるメタデータのうち、デバイス（及び、デバイスを制御するソフトウェア）に関する情報である。測定方法は、「設計・情報」「設置・調整」「運用・保守」の3区分で9項目を定義し、それぞれに品質測定量及び、その測定方法となる品質測定量要素を指定する。（2.1章及び2.3章参照）。新たにデバイス依存の品質測定量として定義した項目は、1.デバイスの情報、2.故障にしにくさ、3.耐久性、4.セキュリティの対策、5.通信の安定性、6.設置方法の適切さ、7.システムの安定稼働、8.システムの環境監視、9.アップデートの適切さの9項目である。（2.2章参照）

本書を公開し、センシングデータの品質評価に関心のある関係者からの意見を得ることで、より実用性の高いデータ品質評価基準の策定を目指す。

第1章 はじめに

1.1 背景

第5期科学技術基本計画（内閣府,2016）は、我が国が目指すべき未来社会の姿として、Society5.0を提唱した。Society5.0で実現する社会は、全ての人とモノがつながり、さまざまな知識や情報が共有され、社会の変革（イノベーション）から新たな価値が創造されることで、少子高齢化、地方の過疎化、貧富の格差といった課題を克服できる、人間中心の社会である。

Society5.0は、フィジカル空間（現実空間）とサイバー空間（仮想空間）が高度に融合したサイバーフィジカルシステム（Cyber Physical System : CPS）によって実現される。CPSとは、現実にある人、モノ、環境から多様なデータをリアルタイムに収集し、ネットワーク上でデータの共有・解析を行って、実世界に対する制御・フィードバックを自律的に行う相互作用と協調の仕組みである（今井和雄, 2016）。

CPSには、ネットワークに接続したセンサーデバイス（測定端末）を介して、膨大なセンシングデータを収集し、クラウドサーバーに移転し、リアルタイムで処理する情報システム（IoTシステム）が含まれる。さらに、IoTシステムが収集したデータをAI（人工知能）で解析・制御することで、産業活動や日常生活をさらに自動化・最適化することも期待されている。

このとき、例えば、想定外に低品質のデータが流通すると、誤認・誤操作などにより、予期せぬ事故やトラブルにつながるおそれがある。システムの安全を合理的な労力で保つには、当事者間でデータ品質について共通理解を得られる技術標準が必要である。

1.2 問題意識

本書は、センシングデータのデータ品質モデルと評価方法をまとめたホワイトペーパーである。先行文献を参考に、センシングデータの品質評価に特有の観点を洗い出し、想定ユースケースに基づく実践例を収録した。海外では、UNSD、W3C、IEEEなどの国際団体が、データ品質に関する検討を行う。我が国でも、関連省庁・団体がガイドラインや手引書、用語集、指針を公表する。本書はそれらの先行文献を調査のうえ、IVI会員及びDSA会員を対象に行ったアンケート調査、インタビュー調査の結果も踏まえて、センシングデータ品質の評価に必要な考え方を提案する。

SQuaRE（ISO/IEC 25000シリーズ）は、情報システム及びソフトウェア製品を対象とした品質向上のために体系化された国際標準規格である。品質評価の対象となるシステムの構成要素をハードウェア、コンピュータシステム、ソフトウェア、データ、利用環境などに区分したうえで、データ品質モデルとデータ品質測定量を定める。

しかし、センシングデータの品質評価を行う場合は、データの内容のみを評価しても、機器の故障や通信エラーといった、デバイス依存の要因による問題が検知できないことがある。このため、デバイス依存の要因も評価対象としたうえで、対象データの品質に与える影響を整理する必要がある。

そこでSQuaREのデータ品質モデルとデータ品質測定量を参考に、CPSでのデータ連携・流通に必要なセンシングデータの評価モデルと評価方法を提案する。

1.3 センシングデータとは

センシングデータとは、センサーで検知・測定されたデータの総称である。時間の経過やデバイスの移動、測定対象の変化に伴って、間断なく生成され続けることがある。また、センシングデータはリアルタイムに収集、処理、解析・制御されることもある。従って、データ品質に起因する事故・トラブルの影響が広がり、深刻になるおそれがある。

加えて、データ品質がデバイスの仕様や設置環境、測定方法といった特定の利用状況に左右されやすく、データの内容を評価するだけでは、測定された値の正否を適切に判断できないことがある。例えば、測定値は得られていても、機器の故障や部品寿命、通信エラーなどが生じており、望ましいデータが得られていないことがある。あるいは、測定目的が特異値の検出である場合、得られたデータが検知すべき事象を表現しているのか、何らかの不具合による異常なのかを判断すべきことがある。これらを踏まえた、センシングデータの品質評価を効率的に行う方法の考案が求められている。

1.4 本書の位置づけ

そこで本書では、ISO/IEC25021（JIS X25021）が定めるデータ品質モデル（及びデータ品質特性）に基づき、ISO/IEC25024（JIS X 25024）が定めるデータ品質測定量（QM:Quality Measurement）をセンシングデータに適用するために、新たなデータ品質測定量(デバイス依存の品質測定量)及びデータ品質測定量要素（QME:Quality Measurement Element）の識別を試みた。

対象実体は、センシングデータに付与されるメタデータのうち、デバイス（及び、デバイスを制御するソフトウェア）に関する情報である。測定方法は、「設計・情報」「設置・調整」「運用・保守」の3区分で9項目を定義し、それぞれに品質測定量及び、その測定方法となる品質測定量要素を指定する。

そのうえで、品質測定量の評価結果を、0.0～5.0の値域に換算し、評価スコアを算出。評価スコアをもとにランク付けを行うとともに、品質特性ごとに品質レベルが把握できるよう、レーザーチャートなどで提示する方法を提案する。

なお、本書が提案する新たなデータ品質測定量及びデータ品質測定量要素の定義は、センシングデータ分析における障害事象をもとに、ボトムアップ方式で考案された。従って、SQuaREシリーズが定めるシステム及びソフトウェア品質モデル全体との適合を確認していない。また、既存のデータ品質測定量には修正を加えていない。

1.5 今後に向けて

本書は、ISO/IEC25012及びISO/IEC25024に基づき、デバイス依存の品質測定量の識別を試みるとともに、新たに識別した品質測定量による品質測定の実践例を示した。センシングデータ自体は、CPSのあらゆるシーンでの活用が期待されているが、まずは、工場内に設置したセンシングデバイスから時系列データを取得し、機械学習による分析を行う場面を想定し検討した。（詳細はAPPENDIX-1参照）

センシングデータの品質は、生成、加工、流通、利用の各段階で検討する必要があるが、生成時点での品質が低いと後段でのデータ品質を改善することが困難となる。このため、本書ではまず、検討対象とするセンシングデータを、センシングデバイスから出力された、生成時点のデータに限定した。また、想定利用シナリオは、複数の企業が工場内に共同設置したセンシングデバイスから時系列データを取得し、機械学習による異常検知を行う場面とした。（詳細はAPPENDIX-2参照）

さらに現在、コネクテッドカーから車外に送信されるデータを対象に、デバイス依存の品質測定量及び品質測定量要素が利用可能か検証している。

組織のなかで、データ品質管理に携わる担当者は、誰もが品質評価の主体となりうる。データの品質を決められた尺度・基準で評価し、その評価結果を相互に参照すれば、データの提供者は自身のデータの価値を適切に表現でき、データの利用者は自身の期待・要求を満たすデータを選択できるようになる。データ流通のコスト低減や事故防止、トラブル発生時の原因究明、義務・責任の分担にも有効である。本書を公開することで広く意見を求め、より実用的な評価基準の作成を目指す。

第2章 検討例：センシングデータの品質評価に利用する品質測定量と品質評価

2.1 デバイス依存の品質測定量

今回は、センシングデータの品質がIoTシステムに影響した障害事例を分析し、デバイス依存の品質測定量として、下表に示す9項目を抽出した。

区分	番号	デバイス依存の品質測定量	説明
設計情報	①	デバイスの情報	デバイスに入力された物理量(光、音など)の計測原理、処理方式等の把握レベル
	②	故障のしにくさ	デバイスの稼働レベル
	③	耐久性	寿命部品の低下レベル
	④	セキュリティの対策	セキュリティ対策の実施レベル
	⑤	通信の安定性	通信が途絶、遅延なく動作するレベル
設置・調整	⑥	設置方法の適切さ	条件にあった適切な設置の実施レベル
運用・保守	⑦	システムの安定稼働	安定稼働の計画レベル
	⑧	システムの環境監視	設置状況の把握レベル
	⑨	アップデートの適切さ	適切なソフトウェアバージョンの運用レベル

※「故障しにくさ」「耐久性」の違いは、ANNEX-4 参照

2.2 デバイス依存の品質測定量の評価パラメータの検討

デバイス依存の品質測定量要素を定義するために、品質測定量ごとの、品質評価方法を定義するとともに、評価パラメータを決定した。

デバイス依存の品質測定量		評価基準	評価パラメータ (データ品質測定量要素)
1. デバイスの情報	(1)機能に関する情報	センシングデータを生成するセンシングデバイス（センサ、信号処理モジュール等）の機能と性能に関し、詳細な情報が得られるレベルで評価する。	①入力情報が明示
	(2)性能に関する情報		②デバイス情報が明示
			③出力データ情報が明示
			④計測タイミング情報が明示
			⑤計測時間情報が明示
			⑥基本特性情報が明示
			⑦データ種別情報が明示
			⑧システム特性情報が明示
			⑨時刻情報が明示
			⑩その他の情報(特異値、異常値の定義等)が明示
2. 故障のしにくさ		偶発的な故障に関する情報が得られるレベルで評価する。	①評価する単位でのMTBF値が記載
			②規定条件が明示
3. 耐久性		劣化部品(機構部品含む)の特定、特性、寿命に与える要因（環境、使用条件等）に関する情報が得られるレベルで評価する。	①共通項目の記載の有無
			②個別項目の記載の有無
			③適用規格・基準、分類、区分、クラスの記載の有無
			④許容範囲・影響の有無
			⑤推定寿命の有無
			⑥監視・診断の有無
4. セキュリティ対策		脆弱性に関し、オープンで包括的、汎用的に利用可能な基準で評価する。	CVSS V3の評価結果を利用
5. 通信の安定性		通信仕様、準拠する規格、信頼性設計、通信品質に関する情報が得られるレベルで評価する。	①通信仕様及び標準規格の明示
			②通信品質(RSSI/受信強度、データ欠損率等)の明示
6. 設置方法の適切さ	(1)設置に関する情報	設置方法と設置環境に関する情報が得られるレベルで評価する。	①共通項目（温湿度、振動、衝撃等）の有無
	(2)環境に関する情報		②個別項目（計測原理に関連する要因）の有無
(3)その他	③適用規格・基準（IEC,ISO、MIL等）の有無		
			④許容範囲・影響の有無
			⑤監視・診断の有無
			⑥デバイス機能・性能に影響する環境条件の有無
			⑦デバイス機能・性能確認の有無
			⑧通信品質に影響する環境条件の有無
			⑨確認手順の有無
			⑩通信品質確認の有無
			⑪地点情報の有無
			⑫周辺環境調査の有無
			⑬通信環境調査の有無
			⑭教育、トレーニング・認定制度の有無
7. システムの安定稼働		設置調整後の稼働状態の確認方法、メンテナンスに関する情報が得られるレベルで評価する。	①事後保全に関する情報の入手の可否
			②予防保全（定期点検実施）の有無
			③予防保全（リモート監視、自動診断実施）の有無
			④記録・保管・ログ情報の入手の可否
8. システムの環境監視		設置調整後の周辺環境の変化、稼働状態の変化に関する情報が得られるレベルで評価する。	①周辺環境監視の有無
			②通信品質監視の有無
			③通信環境監視の有無
9. アップデートの適切さ		ソフトウェアの更新の状況により評価する。	最適なソフトウェア更新の度合い

2.3 デバイス依存の品質測定量の評価方法とスコアリング基準の検討

品質測定量ごとに品質測定量要素と測定関数（スコアリング）を検討した。

スコアリングは、データ品質が利用者にも与える影響の深刻さ（データの欠損、性能の低下、システムの停止など）を考慮して決定した。

1 デバイスの情報に関する評価方法

デバイスの情報						
評価基準	センシングデータを生成するセンシングデバイス（センサ、信号処理モジュール等）の機能と性能に関し、詳細な情報が得られるレベルで評価する。					
評価方法	下記項目に関する情報が明示がされているかで評価する。 ①入力情報（光、超音波、磁気、ひずみ、振動等） ②デバイス情報（CCD、PD、磁気抵抗、圧電素子等） ③出力情報（アナログ/デジタル、有効桁、単位、統計値、変換値等） ④計測タイミング情報（連続、定時、イベント等） ⑤基本特性情報（感度、ゲイン、リニアリティ、サンプリング周波数・精度等） ⑥システム特性情報（フィルタ、応答特性、温度特性、ノイズ密度等） ⑦時刻情報（精度、方式） ⑧原理に関する項目（太陽光、降雨、霧等）の使用条件、性能低下度合い ⑨その他の情報（特異値・異常値の発生条件、扱い）					
基準	5項目未満	5項目	6項目	7項目	8項目	全て
スコアリング基準	0点	1点	2点	3点	4点	5点

2 故障のしにくさに関する評価方法

故障のしにくさ						
評価基準	センシングシステムの停止、データ欠測、性能低下につながる偶発的な故障に関する情報が得られるレベルで評価する。					
評価方法	データを生成するセンシングデバイスの下記情報が明示されているかどうかで評価する。 ①評価する単位（あるいは部品点数ごと）でのMTBF値（偶発故障）が記載 ※部品点数が多数に渡るため、個々の故障モードは評価しない ②規定条件（信頼水準○○%、周囲温度○○℃、デレージング○○%等）が明示 ※統一条件が明示されていればOKとする					
スコアリング基準	情報なし	/	/	①	/	①&②
	0点	1点	2点	3点	4点	5点

3 耐久性に関する評価方法

耐久性						
評価基準	時間の経過や（劣化部品等）、環境により性能低下が想定される要因や条件、低下度合いに関する情報が得られるレベルで評価する。					
評価方法	下記項目に関する情報が明示されているかどうかで評価する。 ①一般項目（温湿度、振動・衝撃、各種EMC、保護等級、塩水噴霧、耐候性）の適応レベル ②寿命に関する項目の条件、故障モード、性能低下度合い ・劣化部品：温度、電圧、電流等により故障寿命が加速するもの、電解コンデンサ、光学デバイス等 ・バッテリーは通常、消耗品（非故障要因）であるが、使用状況により期間が変動するため含める ③適用規格・基準（MIL、IEC、ISO、JIS、独自等）、分類・区分・クラス ④推定寿命の定義、有無					
基準	情報なし		1項目	2項目	3項目	全て
スコアリング基準	0点	1点	2点	3点	4点	5点

4 セキュリティの対策に関する評価方法

セキュリティの対策						
評価基準	センシングシステムの脆弱性に対し、オープンで包括的、汎用的に利用可能な基準で評価する。					
評価方法	共通脆弱性評価システムCVSS v3の評価結果を用い、その基本評価値の深刻度スコアから評価する。 https://www.ipa.go.jp/security/vuln/CVSSv3.html					
スコアリング基準	リスク評価を実施していない	CVSS深刻度 9.0-10	CVSS深刻度 7.0-8.9	CVSS深刻度 4.0-6.9	CVSS深刻度 0.1-3.9	CVSS深刻度 0.0
	0点	1点	2点	3点	4点	5点

5 通信の安定性に関する評価方法

通信の安定性						
評価基準	センシングデータの欠測につながる通信の安定性に関し、通信仕様、準拠する規格、信頼性設計、通信品質に関する情報が得られるレベルで評価する。					
評価方法	有線の場合には無条件に5点とする。無線の場合について、以下の情報が明示されているかで評価する。 ①通信仕様及び標準規格を明示 ・周波数帯、帯域幅、再送有無、ネットワーク等 ・国際規格：IEEE、国内規格（ARIB、独自規格等） ②通信品質に影響する環境条件（仕様書、取説等）の有無 ③通信品質確認 （手順、確認項目（RSSI/パケットエラー率、PER/データ欠損率、データ欠損率等）、結果）の有無 ④通信環境調査：干渉、ノイズ、フェージング等確認、対策の有無					
基準	情報なし		1項目	2項目	3項目	全て
スコアリング基準	0点	1点	2点	3点	4点	5点

6 設置方法の適切さに関する評価方法

設置方法の適切さ						
評価基準	センシング機器の設置方法と設置環境に関する個別情報が得られるレベルで評価する					
評価方法	下記項目に関する情報が入手できるかどうかで評価する。 (1)設置に関する情報の明示 ①デバイス機能・性能に影響する環境条件（仕様書、取説等）の有無 ②デバイス機能・性能確認（手順、確認項目、基準、結果）の有無 ③地点情報（場所、ID、図面、写真）の有無 (2)環境に関する情報の明示 ④周辺環境調査：バラつき要因（商用電源、振動、障害物等）の確認、対策の有無 (3)その他 ⑤専業者、教育・トレーニング・認定制度の有無					
基準	情報なし	1項目	2項目	3項目	4項目	全て
スコアリング基準	0点	1点	2点	3点	4点	5点

7 システムの安定稼働に関する評価方法

システムの安定稼働						
評価基準	センシング機器の設置調整後の稼働状態の確認方法、メンテナンスに関する情報が得られるレベルで評価する。					
評価方法	下記項目に関する情報が入手できるかどうかで評価する。 ①事後保全（故障してからの対応） ②時間計画保全（定期点検実施） ③状態監視保全（リモート監視、自己診断を実施） ④記録・保管・ログ実施 （チェックリスト/統一フォーマット、初期値、調整値、設置、交換日付等の有無）					
基準	①	②	②&④	③	③&④	②&③&④
スコアリング基準	0点	1点	2点	3点	4点	5点

8 システムの環境監視に関する評価方法

システムの環境監視						
評価基準	センシング機器の設置調整後の周辺環境の変化、稼働状態の変化に関する情報が得られるレベルで評価する。					
評価方法	下記項目に関する情報が入手できるかどうかで評価する。 ①周辺環境監視：バラつき要因（商用電源、振動、障害物等）の確認、対策の有無 ②通信品質監視：RSSI、PER、エラー欠損率等、確認の有無 ③通信環境監視：干渉、ノイズ、フェーディング等の確認の有無 尚、有線の場合には、無条件に4点とし、①のみ加算評価する。					
基準	情報なし			1項目	2項目	全て
スコアリング基準	0点	1点	2点	3点	4点	5点

9 アップデートの適切さに関する評価方法

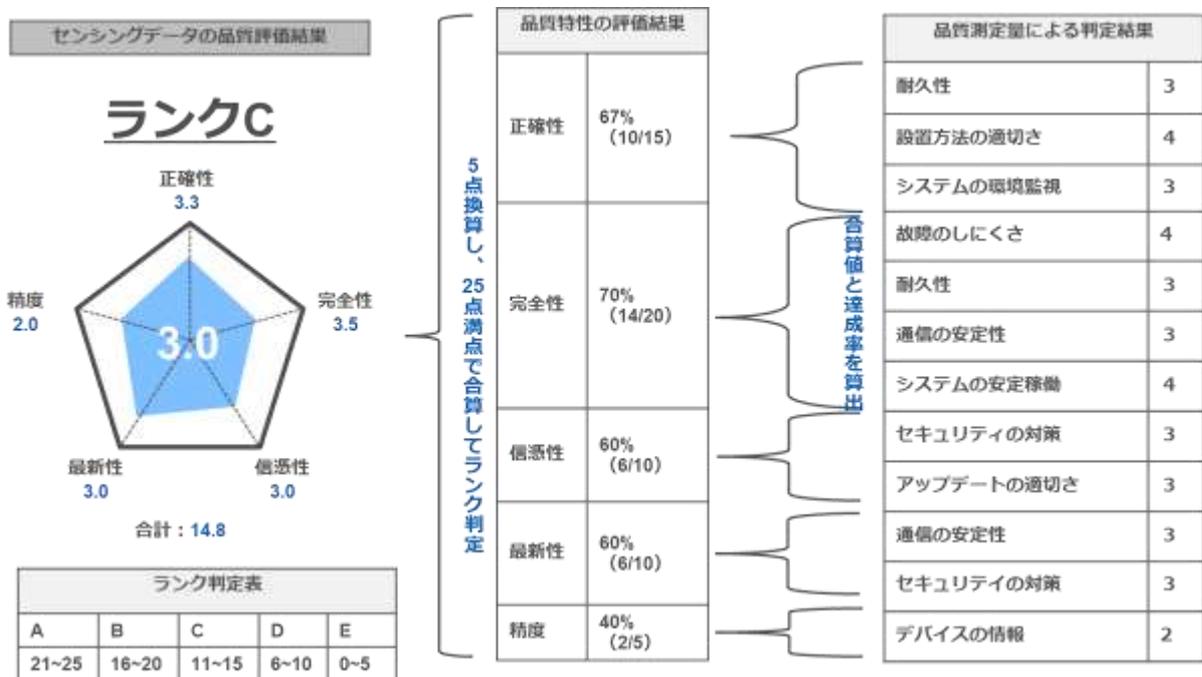
アップデートの適切さ						
評価基準	センシングシステムのソフトウェアの更新の状況により評価する。					
評価方法	問題がないことがわかっている最新版のアップデートがなされているかを確認し評価する					
スコアリング基準	バージョンが確認できない			現状のバージョンは確認できるが、最新のバージョンにアップデートされていなく問題があるかどうかはわからない。		問題がない最新のバージョンが導入されている
	0点	1点	2点	3点	4点	5点

2.4 デバイス依存の品質測定量によるデータ品質評価（イメージ）

実務においては、品質特性ごとの品質測定結果を、項目ごとの値の大小を比較したり、全体のバランス・傾向などを一覧することが重要である。そこで最後に、デバイス依存の品質測定量を用いたセンシングデータのデータ品質の評価イメージを示す。

例えば、品質特性の各項目ごとに、判定に利用する品質測定量要素の判定結果の合算値を、0.0～5.0 点で換算し、品質評価結果として算出することが考えられる。さらに、5 項目の評価時の場合を 25 点満点としてランク判定する。さらに、評価結果はレーダーチャートで表現する。

こうすることで、統合判定結果(ランク付け)だけではなく、各品質特性ごとに品質測定結果を表示しやすくなり、利用者が対象データの品質を容易に判断できるようになると考えた。



第3章 本書について

執筆者

内藤 丈嗣 (オムロン株式会社)

上島 邦彦 (株式会社日本データ取引所)

TG4 参加者一覧

西田 秀志 (オムロンソーシャルソリューション株式会社)

小田 利彦 (オムロン株式会社)

森口 誠 (オムロン株式会社)

内藤 丈嗣 (オムロン株式会社)

清田 幹憲 (凸版印刷株式会社)

田本 孝志 (凸版印刷株式会社)

植田 健治 (大日本印刷株式会社)

上島 邦彦 (株式会社日本データ取引所)

後川 彰久 (日本電気株式会社)

木村 直也 (株式会社日立製作所)

黒政 敦史 (富士通株式会社)

山根 久和 (本田技研工業株式会社)

伊藤 崇 (本田技研工業株式会社)

山崎 克俊 (本田技研工業株式会社)

波多野 健 (株式会社東芝)

畠中 雄一郎 (株式会社東芝)

山田 勇 ((一社)データ社会推進協議会)

清水 響子 ((一社)データ社会推進協議会)

角田 晋也 ((国研)海洋研究開発機構)

松本 高治 ((一社)日本電機計測器工業会)

五味 弘 ((独)情報処理推進機構)

団体概要

一般社団法人データ流通推進協議会(略称:DTA)技術基準検討委員会は、2017年11月に発足して以来、技術基準検討委員会の傘下に「データ品質」を主題としたワーキンググループ(主査:オムロン株式会社、副主査:株式会社インテック、株式会社日本データ取引所)を設置し、データ品質評価基準の策定を目指して調査・検討活動を行った。2021年4月からは、一般社団法人データ社会推進協議会(略称:DSA)のTG4として同様に活動する(簡単のため、以下「データ品質WG/TG」という)。

データ品質 WG/TG は、2018 年からデータ品質評価に関わる技術規格を分析し、関連企業のニーズを調査するとともに、本委員会が策定すべき評価基準の検討スコープ、想定読者、対象データその他の既往論点を整理してきた。2019 年には、データ品質基準の策定に向けた論点整理を行うとともに、データ取引における基本要件を検討した。2020 年には、センシングデータを主対象とし、データ品質モデルと、その概念モデルに基づいた評価方法の検討を進めた。本書はこれらの活動を踏まえたものである。

謝辞

本書の作成にあたり、2020 年 3 月から 4 月に IVI 会員及び DSA 会員を対象に『センシングデータのメタデータ及び品質指標に関するアンケート調査』を行いました。また、2020 年 12 月から 2021 年 1 月には『センシングデータ品質基準に関するインタビュー』を行いました。ご協力いただいた関係企業の皆様に厚くお礼申し上げます。

参考資料

【関連規格】

- ・ISO/IEC 25012_2008 Data quality model (2008年), (JIS X 25012_2013 ソフトウェア製品の品質要求及び評価-データ品質モデル), ISO/IEC JTC1 SC7
- ・ISO/IEC 25021_2012 Quality measure elements (2012年), (JIS X 25012_2014 システム及びソフトウェア製品の品質要求及び評価-品質測定量要素) , ISO/IEC JTC1 SC7
- ・ISO/IEC 25024_2015 Measurement of data quality (2015年), (JIS X 25024_2018 システム及びソフトウェア製品の品質要求及び評価-データ品質の測定), ISO/IEC JTC1 SC7
- ・ISO 8000-8 Data quality (2015年), - Part8:Information and data quality:Concepts and measuring, ISO SC4 WG13
- ・ISO 19157 Geographic information — Data quality (2013年), (地理情報標準第2版 (JSGI 2.0) 第6章 品質原理), ISO TC211

【参考文献】

- ・Beyond Accuracy - What Data Quality Means to Data Consumers, RICHARD Y. WANG AND DIANE M. STRONG, Journal of Management Information Systems I Spring 1996, Vol. 12, No. 4, pp. 5-34
- ・Data Cleansing Beyond Integrity Analysis (2000年), Jonathan I. Maletic, Andrian Marcus IQ2000
- ・Evolutional Data Quality - A theory Specific View (2002年), Liping Liu, Lauren N. Chi, Proceedings of the Seventh International Conference on Information Quality
- ・A Taxonomy of Dirty Data, WON KIM, BYOUNG-JU CHOI, EUI-KYEONG HONG, Data Mining and Knowledge Discovery, 7, 81-99, 2003
- ・Towards a Vocabulary for Data Quality Management in Semantic Web Architectures, Christian Fürber, Martin Hepp, LWDM 2011
- ・A Framework to Construct Data Quality Dimensions Relationships, Payam Hassany Shariat Panahy, Fatimah Sidi, Lilly Suriani Affendey, Marzanah A. Jabar, Hamidah Ibrahim and Aida Mustapha, Indian Journal of Science and Technology, May 2013
- ・The Six Primary Dimensions for Data Quality Assessment, DAMA UK Working Group ,October 2013
- ・Evaluation the Quality of Social Media Data in Big Data Architecture, ANNE IMMONEN, PEKKA PÄÄKKÖNEN, AND EILA OVASKA, IEEE Access, October 16, 2015,
- ・A Survey on Data Quality Classifying Poor Data, Nuno Laranjeiro , Seyma Nur Soydemir , and Jorge Bernardino, 2015 IEEE 21st Pacific Rim International Symposium on Dependable Computing (PRDC)

- ・A Survey on Data Quality for Dependable Monitoring in Wireless Sensor Networks, Gonçalo Jesus, António Casimiro ,and Anabela Oliveira, Sensors 2017,
- ・A Base Document on Data Quality for Trusted Data – ITU (2018 年), ITU-T TG-DPM WG4 (FG-DPM-I-134)
- ・Data Quality 101:The Ultimate Guide for Data Stewards (2019 年), Elliot King (Melissa Data Corporation)
- ・NII Today 第73号 (2016 年 9 月)
- ・第5期科学技術基本計画 [科学技術基本計画 - 科学技術政策 - 内閣府 \(cao.go.jp\)](http://cao.go.jp)

【関連ガイドライン等】

- ・取引データ品質管理ガイド検討 調査研究報告書 (2009 年), 次世代電子商取引推進協議会
- ・ESS Quality Glossary 2010 Developed by Unit B1 "Quality; Classification (2010 年), (EU 統計局 品質関係用語集、翻訳：独立行政法人 統計センター)
- ・つながる世界のソフトウェア品質ガイド (2015 年), IPA(情報処理推進機構) 技術本部 ソフトウェア高信頼化センター
- ・IEEE PAR 2510 (2017 年), Standard for Establishing Quality of Data Sensor Parameters in the internet of Things Environment, Juan Jose Cazila
- ・機械学習品質マネジメントガイドライン (2020 年), NEDO (新エネルギー・産業技術総合開発機構)
- ・データ品質管理ガイドブック (2021 年), 政府 IT 総合戦略室 デジタル・ガバメント技術検討会議「データ連携タスクフォース」

用語

本書では、各用語を以下のように定義して検討を進めた。

データに関する用語

データ

情報の表現であって、伝達、解釈又は処理に適するように形式化され、再度情報として解釈できるもの。
(日本工業規格 X 0001-1994 情報処理用語－基本用語)

センシングデータ

センサーを用いて取得され、デジタル変換されたデータ。時間の経過やデバイスの移動に伴い、間断なく生成され続けるため、すでに蓄積・保存されたデータ（静的データ）と対比して、動的データと呼ぶこともある。

メタデータ

データに関するデータ。データの取得期間、対象地域といった内容に関する情報や、データの名称、概要、作者名、形式、性質・状態、利用条件、権利帰属などを含む。

センシングメタデータ

センシングデータに特有のメタデータ。センサーの識別情報、性能仕様、設置条件などの補足情報を含む。

情報

発信者から、何らかの媒体を通じて受信者に伝達される一定の意味を持つ実質的な内容

ハードウェアに関する用語

センサー

温度、圧力、磁気、ガス、光、音、運動などの物理量を検知・測定し、電気信号などに変換する素子。

デバイス

特定の機能を持った装置、機器の総称。

ゲートウェイ

デバイスから送信されたデータを処理・制御して、ネットワークに中継する機器・仕組み。

センシングデバイス

センシングデータを出力する装置の総称。センシングデバイスは、内蔵するセンサーで検知・測定した情報をアナログデータからデジタルデータに変換し、ネットワークを通じて他のデバイスと通信することで、他のコンピュータから処理・制御できる。

デバイスにそれらの機能を内蔵する場合と、複数のデバイスやゲートウェイ、アプリケーションなど周辺モジュールの組み合わせで実現する場合がある。

品質に関する用語

品質

品物に本来備わっている特徴や性質。

品質要件

ある品物に対する期待や必要条件を表現すること

品質表示

ある品物の性質・特徴をよく分かるように示すこと。

品質保証

ある品物の性質・特徴が、品質要求や品質表示を満たすことに責任を持つこと。

品質評価

ある品物の性質・特徴が、明示または暗黙の期待や要求を満たすか調べ、判断すること。

品質管理

ある品物の性質・特徴が、品質要求（品質要件）を満たすように取扱い・監督すること。

品質監査

ある品物の性質・特徴が特定の基準や手続を満たすか、検証可能な根拠をもとに評価すること。

データ品質に関する用語

データ品質

指定された状況で使用するとき、明示されたニーズ及び暗黙のニーズをデータの特性が満足する度合い。

データ品質特性

データ品質に影響するデータ品質属性の種類。例えば、正確性、完全性などを指す。

品質測定量

データ品質特性の測定の結果として、値が割り当てられる変数。

品質測定量要素

ソフトウェア測定量又はデータ品質測定量を構成するために使用する測定量。基本測定量および導出測定量がある。

データ品質モデル

データ品質要求事項を仕様化し、データ品質を評価するための枠組みを提供する特性の定義された集合。

プロセス品質

作業手順や手順が、期待や必要条件を満たす度合い。データ自体の品質とは区別する。

利用状況

利用者、仕事、装置（ハードウェア、ソフトウェア及び資材）、並びに製品が使用される物理的及び社会的環境。

周辺環境

利用状況のうち、センシングデバイスが使用される物理的環境を指す。

その他(関連用語)

データカタログ

複数のメタデータを精選したもの。

アプリケーション

特定の目的に沿った動作を行うソフトウェア。ネットワークを通じて取得したデータを抽出・加工・移転したり、分析・可視化するといった機能を持つことがある。

システム

特定の目的を達成するために、相互に影響し合う要素を組み合わせで編成された仕組み。

略語

AI : Artificial Intelligence

DAMA : Data Management Association

IDSA : International Data Space Association

IEC : International Organization for Standardization

IEEE : Institute of Electrical and Electronics Engineers

IoT : Internet of Things

IPA : Information-technology Promotion Agency

ISO : International Organization for Standardization

ITU-T : International Telecommunication Union Telecommunication Standardization Sector

IVI : Industrial Value chain Initiative

NEDO : New Energy and Industrial Technology Development Organization

QM : Quality Measurement

QME : Quality Measurement Element

SQuaRE : Software product Quality Requirements and Evaluation

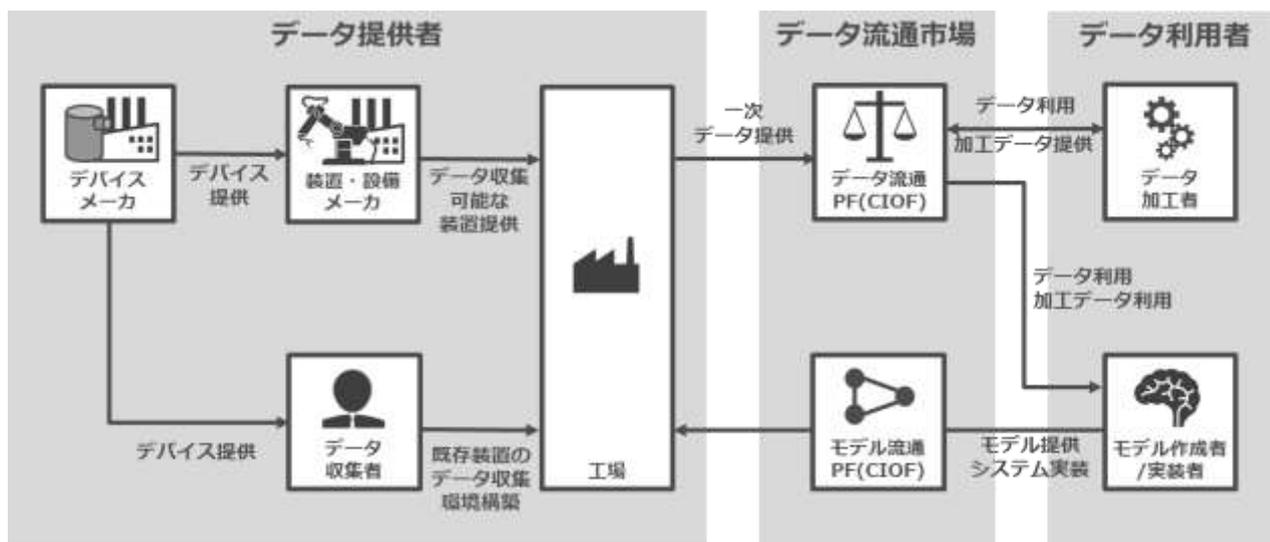
TR : Technical Report

UNSD : United Nations Statistics Division

W3C : World Wide Web Consortium

APPENDIX-1：想定ユースケース

センシングデータの品質測定量の作成にあたり、まず、センシングデータの品質測定が行われるユースケースの検討を行った。今回は、製造分野において、複数の企業が共同で工場内に設置したデバイスから時系列データを収集し、収集したデータを利用する場面を想定した。ステークホルダーは、例えば、機械学習により異常検知や予防保全等を行うために、センシングデータの品質測定を行う。下図に、各ステークホルダーの関係を示す。



データの生成源であるIoTデバイスを作製するプレイヤー
データ品質を決定するIoTデバイスそのものの品質を決定する



観測対象となる装置や設備を作製するプレイヤー
データ品質を決定するIoTデバイスの設置運用に関する品質を決定する
※IoTデバイスメーカと組んでデータ収集者にもなり得る



観測対象となる装置や設備からデータを収集するプレイヤー
データ品質を決定するIoTデバイスの設置運用に関する品質を決定する



観測対象となる装置や設備が使われている工場を運営するプレイヤー
提供するデータのデータ品質をどのレベルにするかを決定する



データを利用できる形に加工するプレイヤー
二次以降のデータ品質を決定する
データ品質の利用者
※データ流通ではデータ利用者でありデータ提供者でもある



データを利用しモデルの作成や実装を行うプレイヤー
データ品質の利用者



データを流通させるプラットフォーム運営者
データ品質を保証する



モデルを流通させるプラットフォーム運営者
モデル品質とモデルのもととなったデータの品質を保証する

APPENDIX-2：想定ユースケースにおける7つの利用シナリオ

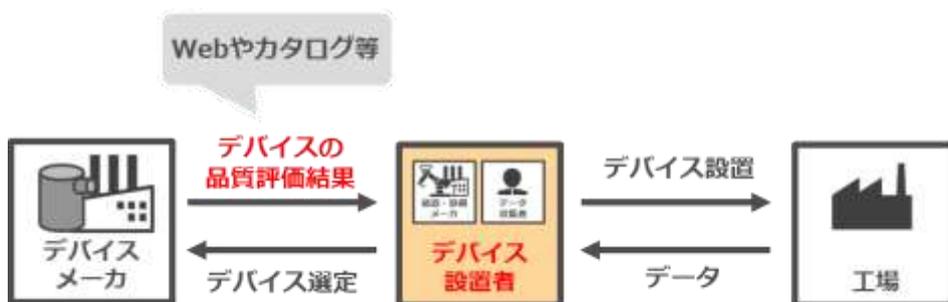
想定ユースケースにおいて、センシングデータのデータ品質基準を利用するシーンとして以下の7つの利用シナリオを設定し検討を進めた。

シナリオ①：データを収集する機器（IoTデバイス）の選定条件に利用する。

(シナリオ内容)

Who	IoTデバイス設置者 （装置・設備メーカーまたはデータ収集者） （IT SIerなどのデータ収集のために工場の設備や装置にIoTデバイスを設置するプレイヤー）
When	設置するIoTデバイスの選定時（デバイスのカタログを閲覧し設置するIoTデバイスを選定するとき）
Where	Web、カタログ等
What	IoTデバイスの品質評価結果 （「センシングデータの品質評価基準」をもとに数値化されたデバイスの品質評価結果）
How	上記を参考に、工場の設備や装置に設置するIoTデバイスとしてどれが適しているかの優劣を判断する

(データ流通フロー)

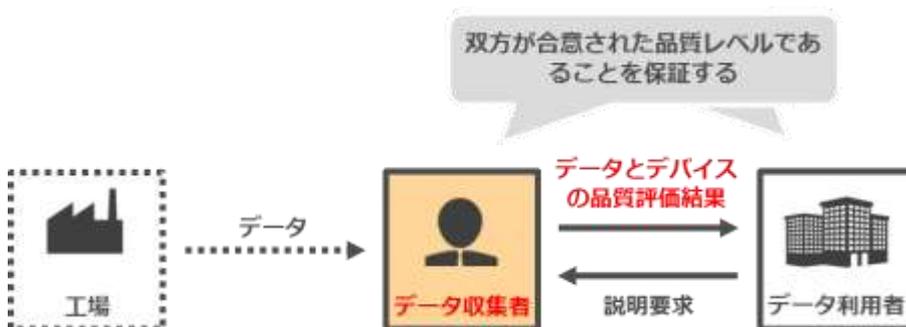


シナリオ②：データの品質保証に利用する

(シナリオ内容)

Who	データ収集者 (IT SIerなどの工場の設備や装置からデータを収集するプレイヤー)
When	収集したデータがどれほどの品質か保証する際 (データ収集とデータ利用が別組織のとき)
Where	自社または企業間、データを取引する市場
What	データとデバイスの品質評価結果 (「センシングデータの品質評価基準書」をもとに作られたデータ利用者からも評価可能なデータの品質評価結果とデータの品質を決定するデバイスの品質評価結果)
How	上記を参考に、工場の設備や装置から収集されたデータの品質がデータ利用者とデータ収集者の間で合意された品質レベルであることを保証する
Who	データ収集者 (IT SIerなどの工場の設備や装置からデータを収集するプレイヤー)
When	収集したデータがどれほどの品質か保証する際 (データ収集とデータ利用が別組織のとき)
Where	自社または企業間、データを取引する市場
What	データとデバイスの品質評価結果 (「センシングデータの品質評価基準書」をもとに作られたデータ利用者からも評価可能なデータの品質評価結果とデータの品質を決定するデバイスの品質評価結果)
How	上記を参考に、工場の設備や装置から収集されたデータの品質がデータ利用者とデータ収集者の間で合意された品質レベルであることを保証する

(データ流通フロー)

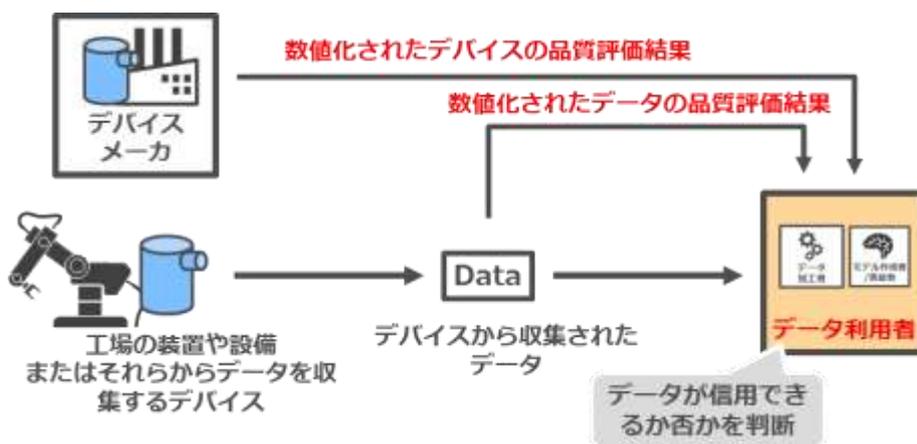


シナリオ③：データ利用前の評価に利用する

(シナリオ内容)

Who	データ利用者 (AI venderなどの工場の装置や設備から得られたデータを利用するプレイヤー)
When	データを利用する際 データを利用している際
Where	自社等
What	数値化されたデータの品質評価結果とデバイスの品質評価結果 (「センシングデータの品質評価基準」をもとに第三者からも理解可能な数値化されたデータの品質評価結果とデータの品質を決定するデバイスの品質評価結果)
How	上記を参考に、データ利用時にそのデータが信用できるか否かを評価し判断する

(データ流通フロー)

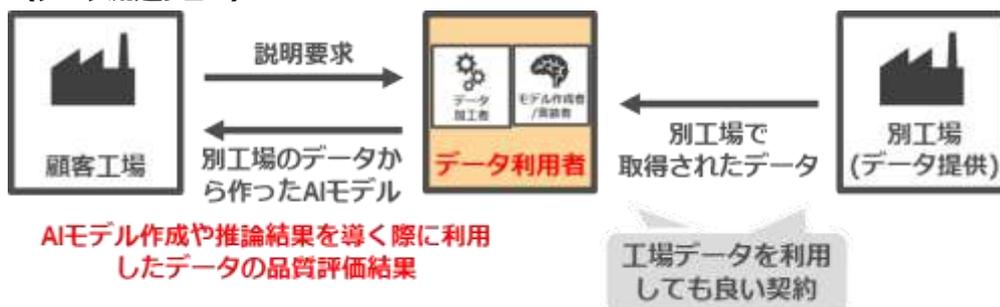


シナリオ④：データ利用前の評価に利用する。

(シナリオ内容)

Who	データ利用者 (AI venderなどの工場の装置や設備から得られたデータを利用するプレイヤー)
When	データを利用しシステムやAIモデル作成した際 (作成したAIモデルや実装したシステムの推論結果がどのようなデータから作られたかを顧客に示す際)
Where	自社や顧客工場 (システムとAIモデルの納入先)
What	AIモデル作成や推論結果を導く際に利用したデータの品質評価結果 (「センシングデータの品質評価」をもとに数値化され第三者からも評価可能なデータの品質評価結果とデータの品質を決定するデバイスの品質評価結果)
How	上記を参考に、どのようなデータからAIモデルを開発したか、または推論結果が導き出されたかを示す。 (顧客等のエンドユーザーに対して出自となったデータに問題がないことを示す)

(データ流通フロー)



シナリオ⑤：データ売買におけるデータの選定基準に利用する。

(シナリオ内容)

Who	データ購入者 (データをAIモデル開発などで利用するためにデータを購入するプレイヤー)
When	データカタログ閲覧時 (データやその価格を示したもの)
Where	データ取引市場 (データを売買することが可能なPF)
What	データとデバイスの品質評価結果 (「センシングデータの品質評価基準書」をもとに作られたデータ購入者からも評価可能なデータの品質評価結果とデータの品質を決定するデバイスの品質評価結果)
How	上記を参考に、購入するデータが自分たちの要求する品質を満足しているか判断する

(流通処理フロー)



シナリオ⑥：データ売買におけるデータの付加価値に利用する。

(シナリオ内容)

Who	データ提供者 (装置や設備を持つ工場を経営し、そこで取得されたデータを販売するプレイヤー)
When	データカタログ掲載時 (データやその価格を示したもの)
Where	データ取引市場 (データを売買することが可能なPF)
What	データとデバイスの品質評価基準 (データ購入者とデータ提供者の共通価値化された「センシングデータの品質評価基準」そのもの)
How	上記を参考に、品質の高いデータを提供できることを示すことで高付加価値化する

(流通処理フロー)



シナリオ⑦：データ売買におけるデータ利用で作られたモデルやサービスの品質保証に利用する。

(シナリオ内容)

Who	モデル提供者 (モデル)
When	モデルカタログ掲載時 (モデルやその価格を示したもの)
Where	モデル取引市場 (モデルを売買することが可能なPF)
What	データとデバイスの品質評価基準 (モデル購入者とモデル提供者の共通価値化された「センシングデータの品質評価基準」 そのもの)
How	上記を参考に、どのようなデータからAIモデルを開発したかを示すことで、モデルの品質 が高いことを示し高付加価値化する。

(流通処理フロー)



APPENDIX-3 : センシングデータ品質に関するアンケート結果

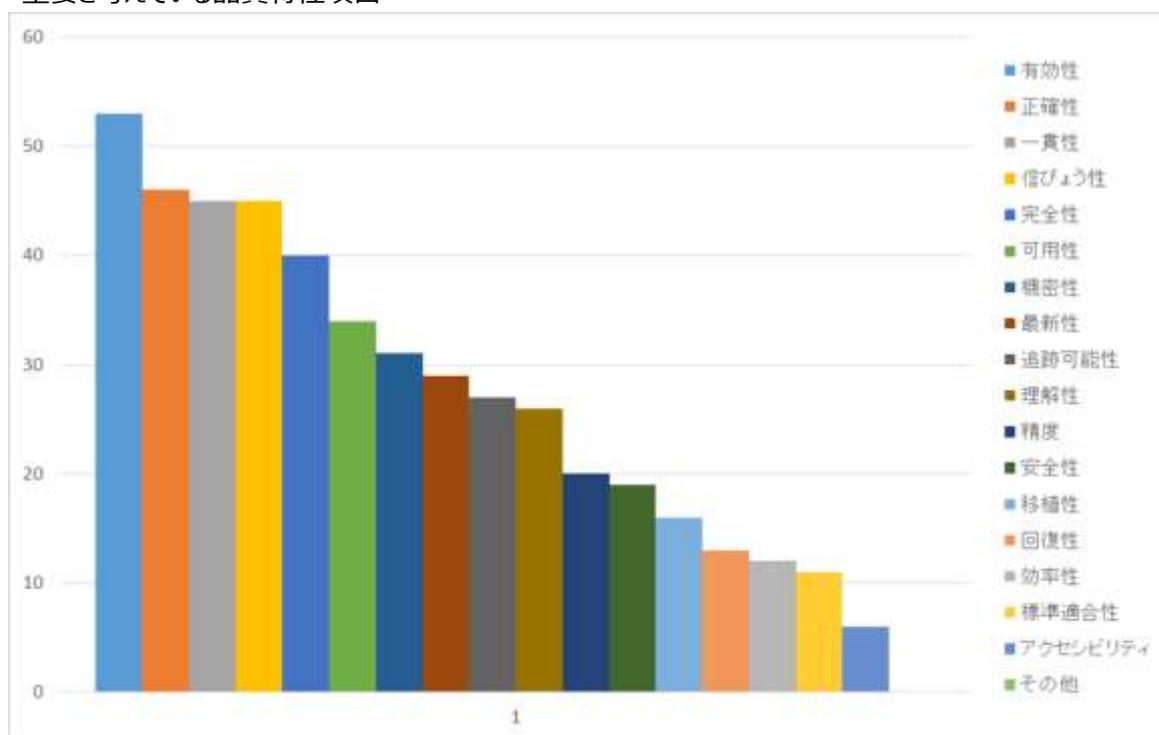
センシングデータのデータ品質を評価する上で特に重要となる品質特性項目を把握するため、2020年4月に、一般社団法人データ流通推進協議会（DTA）と一般社団法人インダストリアル・バリューチェーン・イニシアティブ（IVI）の会員を対象としたアンケート調査（回答数 98 名）を行った。本調査のうち、以下 2 設問を本書の参考とした。

- ① センシングデータを利用する立場として、特に重要と考えている品質特性とその理由
- ② センシングデータのデータ品質に影響を与えるデバイスに関する品質測定項目に関し、特に重要と考えている項目とその理由

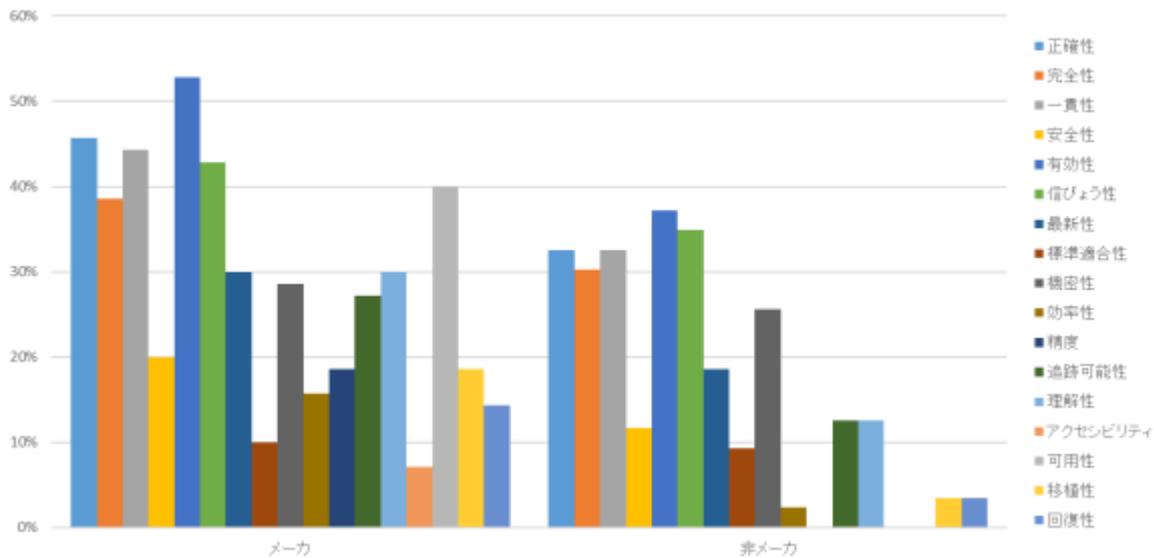
調査の結果、回答者は、有効性、正確性、一貫性、信憑性、完全性の 5 項目を特に重要視していることが確認できた。また、センシングデータを評価する際には、「デバイスの性能や機能」「設置方法」「時間管理」が重視されていることを確認できた。以下にアンケート結果の詳細を掲載する。

1. アンケート①の結果

・重要と考えている品質特性項目



・品質特性項目に対するメーカー/非メーカーの違い



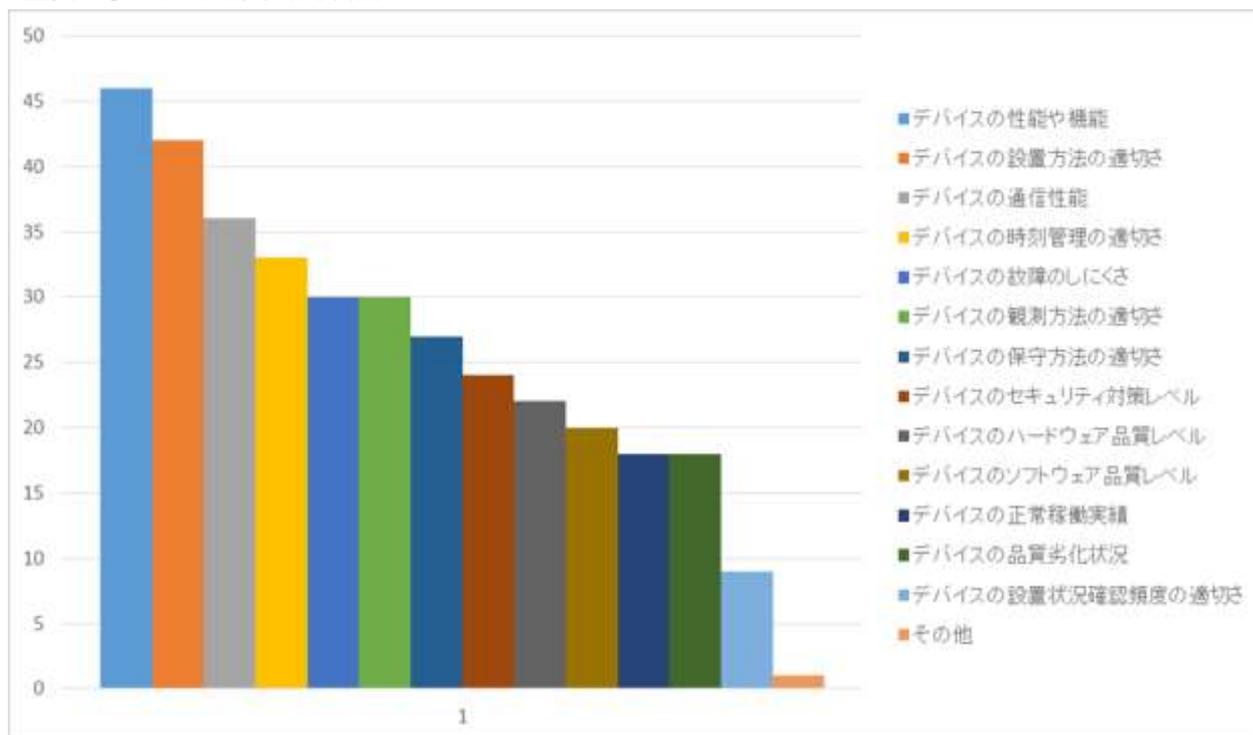
・重要と考えた理由

- 正確性（タイムスタンプおよびデータ自体）。分析する際にデータの品質が担保できていない改ざん可能性のあるデータをいくら分析したところでその分析結果は信用できない。
- データの品質が担保されてこそ機械学習や AI であるため。
- 品質レベル（数値ではなく定性的でよい）
- その他の選択肢は、要件によりケースバイケースで重要度が変わるか、または、データ取得後でないと判断できないと思われるが、選択したセキュリティに関する項目は、常時必要と思われる。
- 情報分析に必要な時間に対して、データのクレンジング処理に必要な時間が多大であるため。
- データを活用したい人が、容易にデータを扱えることが最も大事なことだと考えている。容易に扱えなければ、結局、紙でデータを扱っていたときのように苦勞をかけてしまうことになり、意味が無くなってしまう。同時にデータが増えることでその正しさをチェックすることも難しくなるため、この対策も重要。
- システムの継続性と問題が発生した時の追跡性（原因追及ができること）
- データのクレンジングなどが特に重要と考えるので。（必要なデータだけを取り出さないと間違った結果に誘導されるので。）
- データの集めやすさと利用やすさが重要と考える
- システム化を考慮すると、複数のセンサー等から収集したデータ間の相関関係の信頼性評価が重要となる。特にその中で時間や順序に関する情報の信頼性に関して、品質項目として評価可能になっていると、データを利用する際に効果的である。また、データ収集システムに存在する、収集サイクルや伝送サイクルとタイムスタンプ等の関係が、データの品質に影響する。それらに関する情報も、データ解析の際に基本情報として利用できることが望まれる。

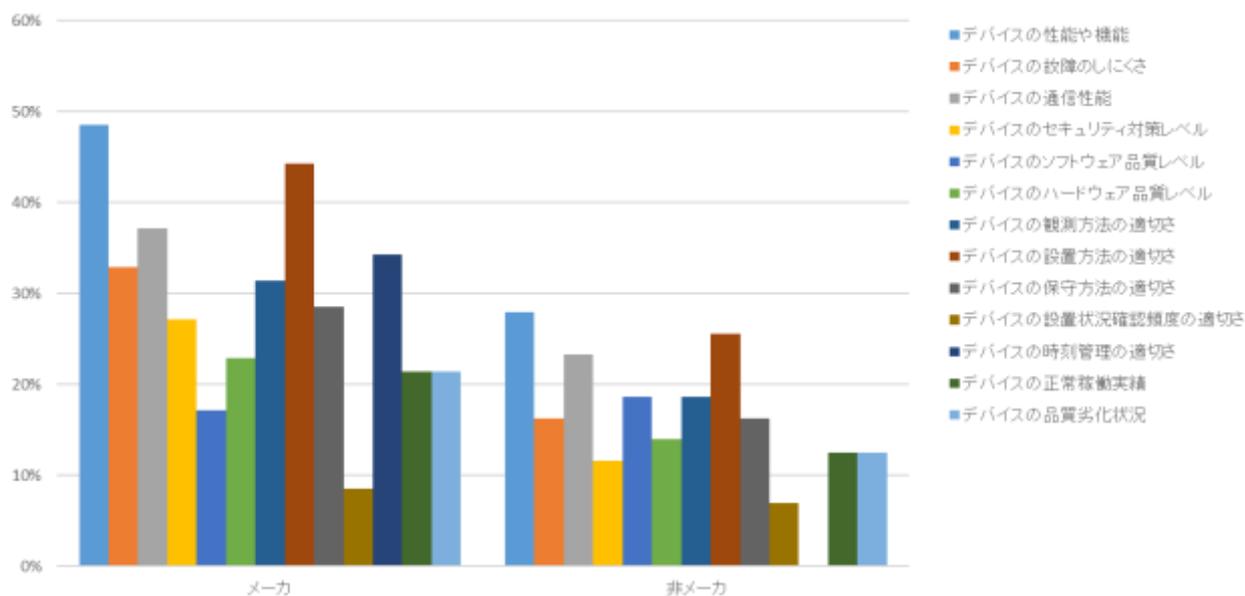
- 正しいデータを取得できないと正しい分析ができない
- 製品・部品の性能、耐久などに関する相関関係
- 多数のデバイスから収集したデータの同期性（タイムスタンプが統一されていること）
- データ収集蓄積によって、他社に対する競争力を得ることが出来るか（独自性）
- 顧客のやりたいことが曖昧のことが多く、PoC からはじまるケースが多いのですが、まずどのようなデータがそもそもあるかを把握できないと始められない
- 有効性はほかの特性を包含するとも考えられる重要な品質の基準
- 計測値であれば単位が揃っている。製造オーダとの紐付き情報がある。製造物に識別 ID が付与されており、かつデータに紐づいている
- 非常に多くのデータを収集しているが、ごみデータが含まれており、ごみも含めて分析するので分析コストが上がるといった
- 悪循環が生じている。また、IoT 化などで データ収集や蓄積が容易になった分 セキュリティ対策が一層重要になってきていると感じる。
- データを活用する目的の多くは現場の品質や生産性を現在よりもさらに高めることを目指しているはずである。
- そこで活用するデータの質が不十分だと誤った判断が導かれるリスクがある。
- 妙なデータを信じたために品質や生産性を低下させてしまうという本末転倒が起きないようにするという観点で上記項目を選択した。
- 【アクセシビリティ】は、「障害のある方が利用しやすいか」など利用者の属性を限定するのではなく、「データを利用したい方が、アクセスしやすいか」が重要だと考えています。また、アクセシビリティを考慮する際には、品質特性自体を測定するのではなく、品質特性を満足するために必要十分なメタデータが入手可能であることが必要だと考えています。
- デバイスに関する各種のメタデータが、どの品質特性を満足する情報なのか整理されていると助かります。

2. アンケート②の結果

・重要と考えている品質測定項目



・品質測定項目に対するメーカー/非メーカーの違い



・重要と考えた主な理由

- 保守方法。 センサーは故障をするものであるから、いかに故障した際にそれを人が理解し、交換して正常稼働に戻せるかが重要。
故障に気づかずに利用し続けているのは、一期一会のセンサデータの全体の品質悪化につながる。
- デバイスの管理によって取得できるデータのばらつきを抑えるため
- その他の選択肢は、要件によりケースバイケースで重要度が変わると思われるが、選択したセキュリティに関する項目は、必須要件と思われる。
- 分析内容の正確性に影響するため
- 集めているデータが間違っているのは、判断を間違えてしまう。そのため、容易に正しくデータを収集でき、故障しにくいことが重要だと考えている。
- 正しくデータが取れているか。再現性はあるのかなど。
- 時間が重要
- センサーと計測対象との関係が最も重要。そのため、観測方法・設置方法・保守・点検・劣化判断を重要と考えた。またシステム化を考慮すると、データ収集時刻（時刻・通信周期等関係）に関する項目も重要である。当然デバイス自身の性能類も計測対象に合わせて設定すべきである。
- データを同時刻に合わせないと分析できない
- 予兆保全などに活用する際の信頼性が重要と考えたから
- 時刻と場所、これによって位置と時間でユニークに特定出来るため
- 性能や機能はそれらが満たされない限り採用することはないので当然の要求であり、それらが持続的に担保されることが重要
- いつどこで得たデータなのか、データの質がどのレベルにあるのかということを把握することが重要である。また、デバイスの購入や設置にかかるコスト対効果の妥当性が必ず問われる。
- 観測・設置の方法が適切であること、また、時刻管理が適切であることがとくに重要だと思います。デバイスに対する品質要求と所要コストはトレードオフになると思います。デバイスの普及を促すには、品質レベルの低い安価な製品も必要で、それらのデバイスが「悪さをする」のを防ぐのに最小限求められる指標が重要だと考えました。

アンケート時に利用した TG で定義したセンシングデータの品質特性は以下の通りである。

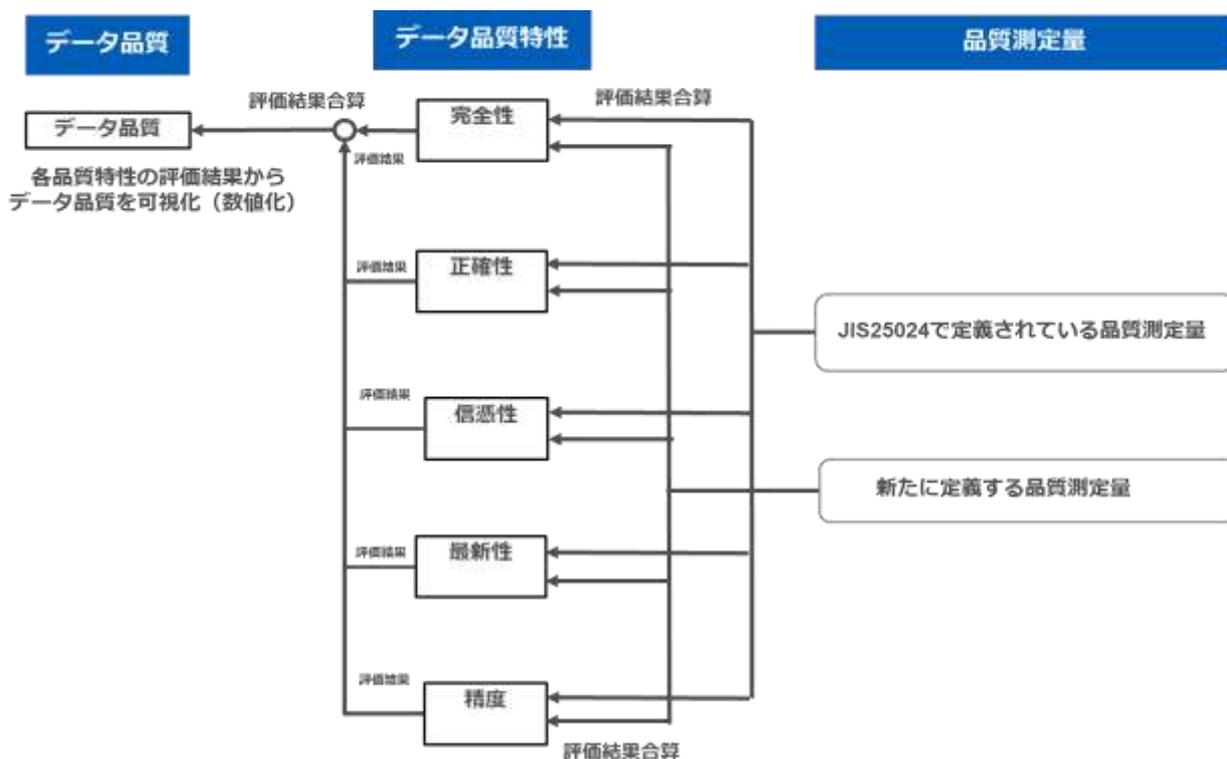
品質特性	センシングデータの品質特性の定義	補足説明(具体的な例)
正確性	センシングデータが、どの程度真値に近く測定できているかを示す度合い	・データに外れ値がないか？
完全性	センシングデータが、欠損や抜けがなく、完全にそろっているかを示す度合い	・データに欠損がないか？
一貫性	センシングデータ間において、矛盾なく整合が取れているかどうかを示す度合い	・同時刻、同じ計測パラメータでの計測データか？
信憑性	センシングデータの内容がどの程度信頼できるかを示す度合い	・データが改ざんされていないか？
最新性	センシングデータが遅延なく計測対象の変化に追従できているかを示す度合い	・データのリアルタイム性は？ ・データ入手までの時間は？
アクセシビリティ	センシングデータへのアクセスのしやすさを示す度合い（障害発生時）	
標準適合性	センシングデータに関する法令や各種基準に適合しているかを示す度合い	・業界で定めた規格への適合性は？
機密性	センシングデータの利用が、正当な利用者のみに限られているかを示す度合い	・利用許可を得たデータか？
効率性	センシングデータの取り扱い時に、メモリ等のリソースを最適に利用しているかを示す度合い	・センシングデータの表現の冗長性は？
精度	センシングデータの詳細度（有効桁数）を示す度合い	・サンプリング、分解能は？
追跡可能性	センシングデータへのアクセスやセンシングデータの変更を容易に追跡できるかを示す度合い	・データ変更の確認が可能なレベルは？
理解性	センシングデータの内容を容易に判読できるかを示す度合い	
可用性	センシングデータ利用時に、容易に利用できるかを示す度合い	
移植性	センシングデータの他システムへの利用時に、容易に利用できるかを示す度合い	
回復性	センシングデータが、トラブル発生時にどの程度正常に回復できるかを示す度合い	

APPENDIX-4 : センシングデータの品質モデル (案)

アンケート調査でとくに重要視された品質特性（完全性、正確性、一貫性、信憑性、最新性）に精度を加えた5項目の品質特性について、データ品質の評価フレームワークを検討した。

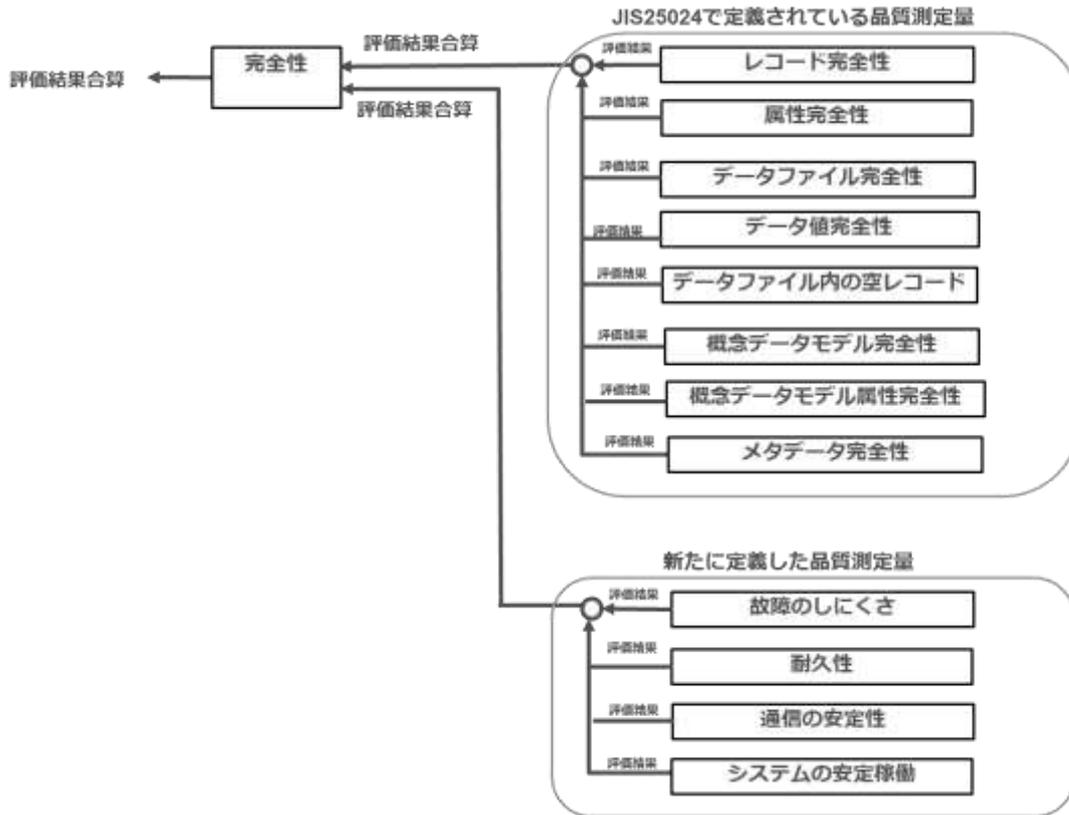
各図に示すように、センシングデータのデータ品質は、既存の品質測定量（QM）に加え、新たに定義したデバイス依存の測定量の合算で評価できるのではないかと考える。

【重要な5つのデータ品質特性の評価フレームワーク】

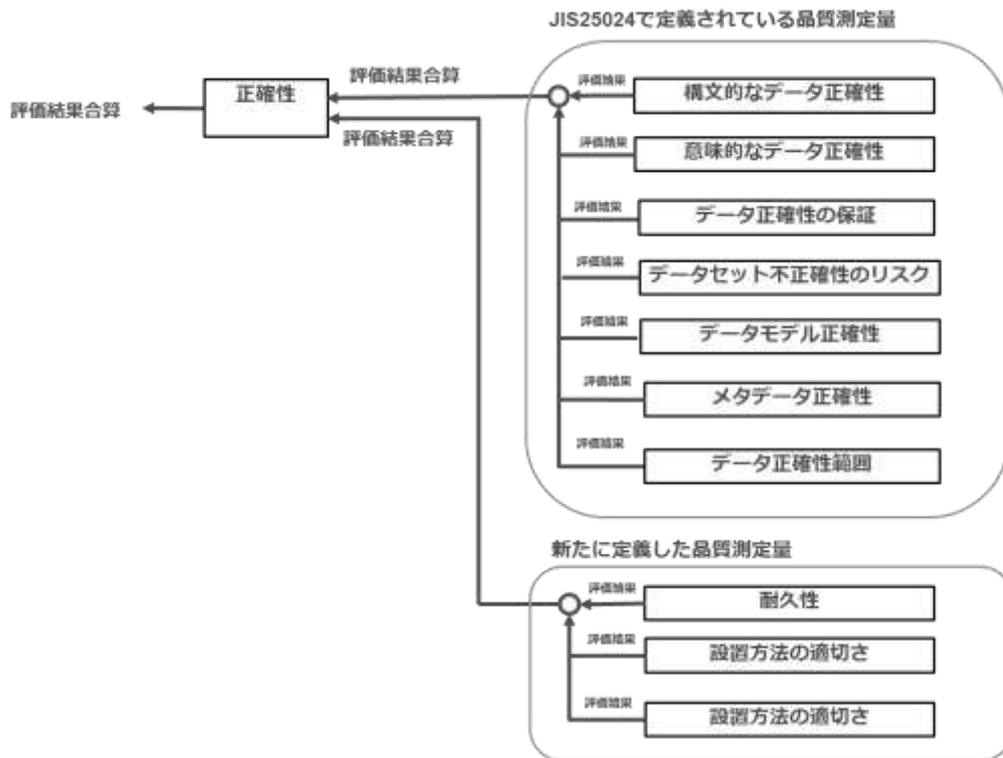


【各品質特性ごとの品質特性の評価フレームワーク】

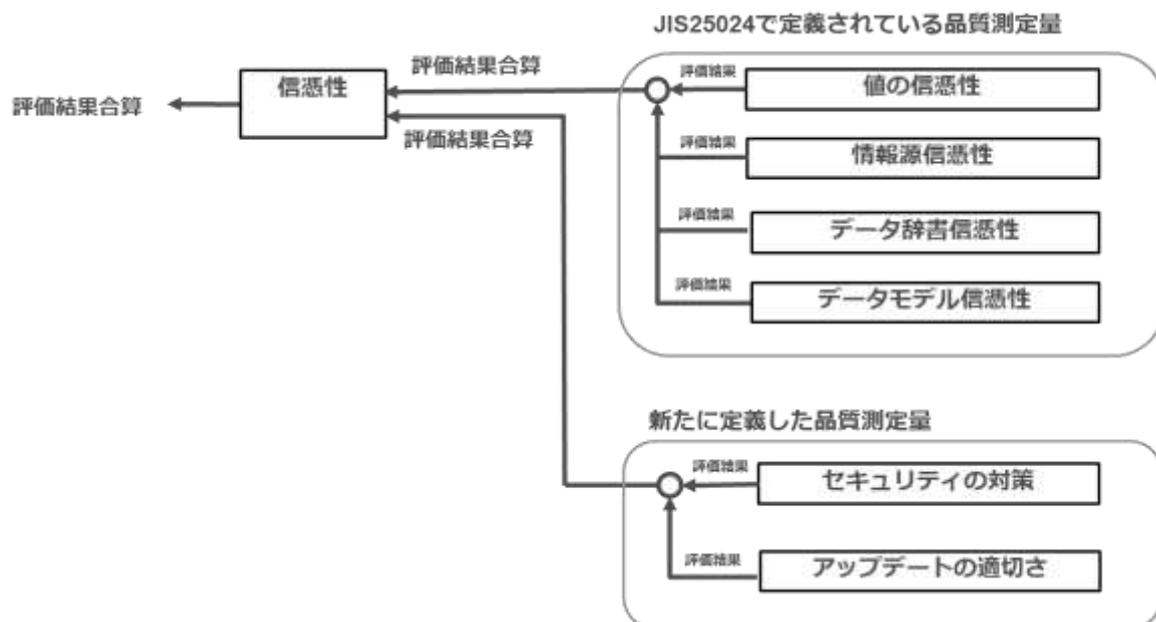
『完全性』



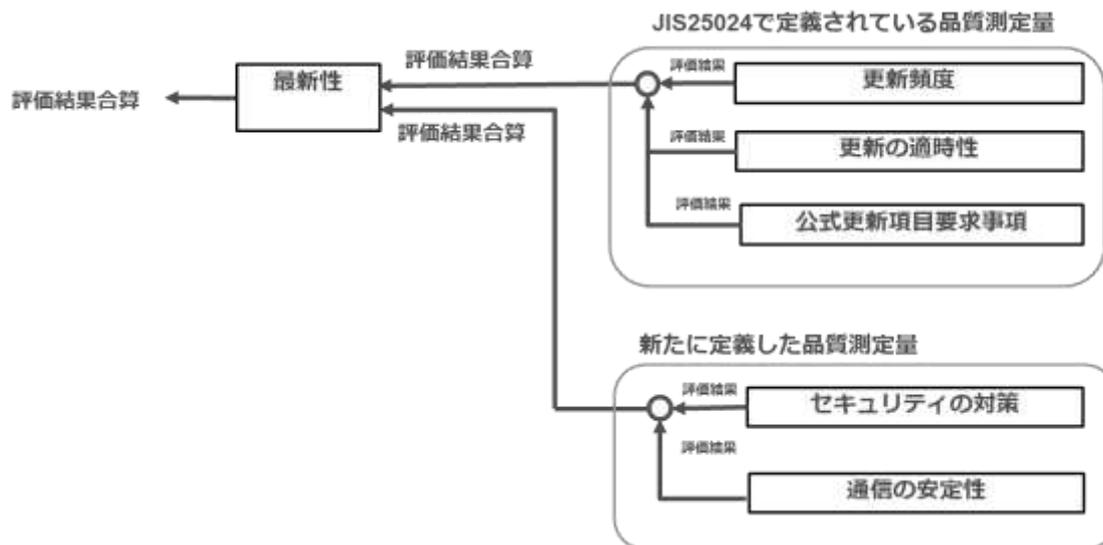
『正確性』



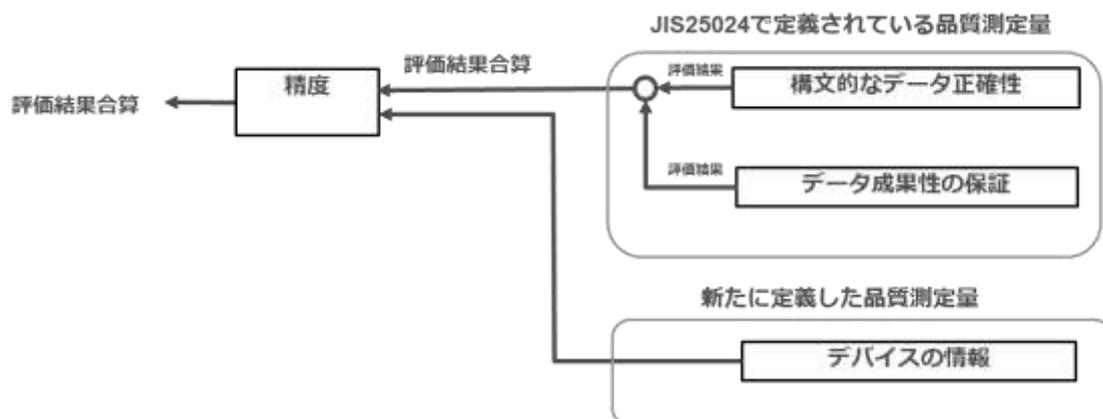
『信憑性』



『最新性』



『精度』



TG で定義したセンシングデータの品質特性の定義 (APPENDIX-3 で記載) に基づき、5 項目 (正確性、完全性、信頼性、最新性、精度) のセンシングデータの品質特性を評価するために利用すべきと考えたデバイス依存の品質測定量は以下通りである。

品質特性	デバイス依存の 品質測定量	品質評価時に利用すべきと判断した理由
正確性	③ 耐久性	経年変化をしていなければデータの外れ値等はお出にくい
	⑥ 設置方法の適切さ	設置方法が適切であれば外れ値等はお出にくい
	⑧ システムの環境監視	設置状況が初期状態から変化していなければ外れ値等はお出にくい
完全性	② 故障のしにくさ	故障しなければデータの欠損は発生しない
	③ 耐久性	経年変化をしていなければデータの欠損は発生しない
	⑤ 通信の安定性	通信が途絶しなければデータの欠損は発生しない
	⑦ システムの安定稼働	安定的に動作していればデータの欠損は少ない
信憑性	④ セキュリティの対策	セキュリティ対策レベルが高ければ攻撃によるデータの改竄はされにくい
	⑨ アップデートの適切さ	ソフトウェアの適切なアップデートはデータの改竄防止に効果がある
最新性	④ セキュリティの対策	セキュリティ対策レベルが高ければデータの改竄による遅延が生じにくい
	⑤ 通信の安定性	通信が途絶しなければデータの欠損による遅延が生じにくい
精度	① デバイスの情報	デバイスの情報を把握することがデータの精度を評価する上で重要な

ANNEX-1 : データ品質に関する国際潮流・国内動向

1. 国際潮流

ISO/IEC JTC1（第1合同技術委員会）が公表する「ISO/IEC 25000」シリーズは、情報システム及びソフトウェア製品を対象とした品質評価のための規格体系で、SQuaREと略称される。

2007年の技術報告（ISO/IEC TR 25021:2007）を受けて、2008年に「ISO/IEC 25012: ソフトウェア製品の品質要求及び評価（SQuaRE）－データ品質モデル」第1版が発行され、近年では2019年にレビュー・確認された。

他の国際団体も検討を行っている。例えば、国連統計局が2010年に国際データ品質フレームワーク（National Data Quality Framework : NQAF）を定める。W3Cは、2016年に「Data Quality Vocabulary」と題するデータセットの品質を説明するための用語定義と概念フレームワークを定めた。IEEEは、2017年に「P2510 - Standard for Establishing Quality of Data Sensor Parameters in the Internet of Things Environment」を立ち上げた。ITU-Tは、2019年に技術報告書「Framework to support data quality management in IoT」を公表した。

データの種類を限定した標準規格もある。例えば、「ISO 19157 Geographic information — Data quality(地理情報標準第2版 (JSGI 2.0) 第6章 品質原理)」は、地理情報に関する用語や基準を定める。EU統計局が公表した統計用語集「ESS Quality Glossary 2010」のように、拘束力は持たないものの、多くの読者に参照されることを期待した文書もある。

技術規格化せず、調査報告書や提案・提言、ガイドブックとして刊行される文書もある。例えば、非営利団体 DAMA International が刊行するデータマネジメントの知識体系「DMBOK」(Data Management Body of Knowledge) には、データ品質に関する言及がある。また、Q. Ethan McCallum「Bad Data Handbook: Cleaning Up The Data So You Can Get Back To Work」(オライリー社, 2013) や Iliot King「Data Quality 101: The Ultimate Guide for Data Stewards」のように、専門家による著作も刊行されている。

2. 国内動向

総務省統計局が「公的統計の品質保証に関するガイドライン」(初版: 2010年、改定: 2019年) を定めるほか、総務省が「統計表における機械判読可能なデータ作成に関する表記方法」(2020年) を公表している。

国土交通省国土地理院は、「空間データ品質評価に関するガイドライン」(初版: 2004年) を定め、品質評価手順書を提供している。AI分野では、NEDO(新エネルギー・産業技術総合開発機構) による「機械学習品質マネジメントガイドライン」(2020年) がある。

行政ガイドラインの他にも、様々な手引書、用語集、指針が公表されている。IPA/SCE(情報処理推進機構ソフトウェア高信頼化センター) が「つながる世界のソフトウェア品質ガイド」(2015) がある。また、政府CIO補佐官で構成されるデジタル・ガバメント技術検討会議「データ連携タスクフォース」が、「データ品質ガイドブック」を公表している。

他にも、独立行政法人統計センターでは、先述した「EU統計局 品質関係用語集」を翻訳するなど、関連情報の提供を行っている。また、気象庁では「気象観測統計指針」を定め、統計に関する一般事項や用語の定義、品質管理の方法を公表している。

ANNEX-2：データ品質に関する基本的な考え方の整理

信頼性のある自由なデータ流通には、持続的で総合的なデータ品質の確保が欠かせない。もっとも、ひとりで「データ品質評価」といっても、プロダクト品質、プロセス品質、マネージメント品質、リソース品質など、データ品質の確保には様々な観点がある。ひとりで「データ品質評価」といっても、プロダクト品質、プロセス品質、マネージメント品質、リソース品質など、多様な評価観点が介在する。データ品質評価を行う前に、当事者間で何を論点とし、何に着目すべきか見極めることが望ましい。そこで本章では、データ品質に関する基本的な考え方を要約するとともに、次章以降で検討する論点を整理する。

1. データ品質とは何か

本書は、データ品質とは、「特定の使用条件で、明示された要求（または暗黙の期待）をデータの内容が満たす度合い」であると定義する。品質とは、品物に本来備わっている特徴や性質を指す。品質の良し悪しは、品物の性質・特徴のみでは決まらず、使用者の期待や目的によって判断される。例えば、軽いものは加速しやすく、重いものは安定しやすい。速さを求める使用者は軽量化を高品質だと考え、安定を求める使用者は低品質だと考える。同様に、望ましいデータ品質を確保するためには、データの内容のみではなく、データ利用者による要求や期待の考慮が欠かせない。

2. データ品質測定・表示・評価

データ品質の測定、表示、評価その他の作業は、類似作業と混同されやすい。そこで本書は次のように区別する。

データの提供者は、データ品質を一定の基準に沿って測定し（品質測定）、第三者に分かりやすく示し（品質表示）、品質表示に対して責任を負う（品質保証）。

データの利用者は、データに対する期待や必要条件を表現し（データ品質要求）、得られたデータの性質・特徴が、明示された要求（または暗黙の期待）を満たすか調べ、判断する（データ品質評価）。また、データ品質が特定の基準や手続を満たすか、検証可能な根拠をもとに審査する（データ監査）。

また、データ提供者と利用者が協力して、品質保証の範囲・条件を限定したり、品質要求を満たすようにデータを取扱い・監督する（データ品質管理）。

これら一連の業務工程において、共通の尺度・基準でデータ品質を測定・表示・評価できれば、データ提供者は自身のデータの価値を適切に表現でき、データ利用者は自身の期待・要求を満たすデータを選択できるようになる。データ流通のコスト低減や事故防止、トラブル発生時の原因究明、義務・責任の分担にも有効である。

3. データ品質評価の主体

Industrial Data Space「White Paper」によれば、データ品質は当事者の共同で確保され、その方法には、1.データ提供者による保証、2.仲介者による評価、3.公衆による評価、4.未評価の4通りがある。また、データ品質管理に責任を持つ役割は、1.データ提供者、データ利用者、仲介者、不在の4通りがあるとす。利害の異なる複数の当事者が、データ品質に関する相互理解を得るためには、データの品質に影響する要素を明確する必要がある。

4. データ品質基準の活用場面

産業・分野を横断したデータ品質基準があれば、組織・法人の担当者が、データ品質について合意形成するための条件などを概観できる。また、データ種類ごとに品質測定の方法を検討したり、グレード・等級等の簡易なデータ品質表示を行う参考情報となる。例えば、次のような場面で活用することが期待される。

1. データ提供者の活用場面：

- a. データ品質測定：自身で管理するデータの品質を一定の規則に沿って測定する。
- b. データ品質保証：データ取引の契約・交渉の過程で、相互に合意できる品質基準を定める。
- c. データ品質表示：データカタログに記載する品質情報の項目定義や、評価方法の記述に用いる。

2. データ利用者の活用場面：

- a. データ品質要求：事前にデータ品質要件をデータ提供者に通知し、品質表示を求める。
- b. データ品質評価：データ提供者から受け取ったデータの品質を評価し、自身の利用目的を満たすか判断する。
- c. データ監査：データ品質の要求、評価、保証、監査、改善などを行う実務担当者が、具体的なチェック項目や品質表示を作成したり、使用する。

3. 双方の活用場面

- a. データ品質管理：自身で管理するデータの品質を保証する範囲・条件を限定したり、品質要求を満たすようにデータを取扱い・監督する。

5. 先行文献

海外では、国連統計局、W3C、IEEE、ITU-Tなどの国際団体が、データ品質に関する検討を行う。我が国でも、関連省庁・団体がガイドラインや手引書、用語集、指針を公表する。本書はそれらの先行文献を調査のうえ、主としてISO/IEC 25000シリーズ（略称：SQuaRE）を参考に、センシングデータ品質の評価に必要な品質測定量（QM）の定義を提案する。

SQuaRE（ISO/IEC 25000シリーズ）は、情報システム及びソフトウェア製品を対象とした品質向上のために体系化された国際標準規格である。品質評価の対象となるシステムの構成要素をハードウェア、コンピュータシステム、ソフトウェア、データ、利用環境などに区分したうえで、ISO/IEC 25012でデータ品質モデルを定め、ISO/IEC 25024でデータ品質測定量を定める。

6. 検討対象とするデータ種類

前提として本書は、企業が取り扱う秘匿性の高いセンシングデータを対象とし、行政機関や学術団体が公表するデータや、他の種類のデータは直接の検討対象としない。過去に蓄積・公開されたデータとちがひ、公衆による品質評価が行いづらく、データ提供者とデータ利用者の合意形成が課題となりやすいからだ。一方で、データの加工度合いによる限定

は行わない。そのデータが観測値であれ、補正值であれ、推計値であれ、メタデータであれ、センシングデータの品質評価に基本的な考え方は共通であると想定する。

また、本書では、センシングデバイスから出力された、生成時点のデータを検討対象とする。センシングデータの品質評価は、1.センサーで測定したデータをデバイスが出力するまで（生成時点）、2.ゲートウェイを介して複数のアプリケーションで抽出・加工されるまで（加工時点）、3.組織や法人をまたいでデータが移転されるまで（流通時点）、4.分析・可視化され、その結果を利用者が参照するまで（利用時点）の4段階に分けて考えることができる。当然ながら、組織や法人をまたいだデータの流通が起きることもある。

	1.生成時点	2.加工時点	3.流通時点	4.利用時点
対象データ	デバイスが出力	アプリケーションで抽出・加工	・組織・法人をまたいだ移転 ・ネットワークを介した移転	分析・可視化される
品質表示	デバイス事業者	アプリケーション・プラットフォーム事業者	・ネットワーク事業者 ・アプリケーション・プラットフォーム事業者	アプリケーション・プラットフォーム事業者
品質評価	・ネットワーク事業者 ・アプリケーション・プラットフォーム事業者 ・最終利用者	・ネットワーク事業者 ・他のアプリケーション・プラットフォーム事業者 ・最終利用者	・他のアプリケーション・プラットフォーム事業者 ・最終利用者	最終利用者

複数の事業者がデータ流通に関与することで、評価対象となるデータの品質が、段階ごとに変化する可能性がある。とはいえ、データ流通段階ごとの留意事項をすべて評価対象とすることは効率的でない。データ品質に影響する要素は、データ自体の品質、デバイスの品質、ネットワーク品質、ソフトウェア品質、事業者・サービスの品質、取扱い品質など多岐に渡るからだ。

例えば、生成時点のデータが加工時点で扱われるとき、ゲートウェイの挙動やネットワーク品質によって、次の段階に受け渡すデータが変質することはありうる。さらに、加工時点から流通時点では、複数のアプリケーションが異なる処理を行ったり、そのデータの権利帰属が転々とすることがある。

このように、データ流通の段階ごとに留意すべきデータ品質要件は異なり、データ利用者が手元で取り扱うデータの品質は、前段階におけるデータ品質管理の積み重ねとして現れる。利用時点に至って、前段階でどのような変化が生じたのか不明確になると、最終利用者がデータ品質を十分に信頼できないこともある。そもそも、生成時点のデータ品質が低いと、後段階でデータ品質を改善することは困難になりやすい。そこで本書では、センシングデバイスから出力された、生成時点のデータを検討対象とした。

ANNEX-3 : SQaRE の概要

1. SQaRE とは？

SQaRE (ISO/IEC 25000 シリーズ) は、情報システム及びソフトウェア製品を対象とした品質向上のために体系化された国際標準規格である。製品品質一般部門、品質モデル部門、品質測定部門、品質要求部門、品質評価部門、拡張部門の 5 部門で構成される。このうち、ISO/IEC 25012 でデータ品質モデルを定め、ISO/IEC 25024 でデータ品質測定量を定める。

図：ISO/IEC2500 シリーズの規格構成



2. SQaRE の品質モデル

論点を明確にするために、SQaRE は、ISO/IEC 2502n: 品質測定部門において、標準規格の適用対象とする広義の人間 - コンピュータシステムを、情報システム、コンピュータシステム、ソフトウェア製品、データ、利用者・利用環境の 5 要素に区別する。そのうえで、品質測定の対象として、「利用時品質」「製品品質」「データ品質」の 3 つを考える。

まず、情報システムの構成要素を、品質モデルの対象となる対象コンピュータシステム (Target Computer System)、非対象とするコンピュータシステム、通信システム (Communication System) を区別する。

また、対象コンピュータシステムの構成要素を、品質モデルの対象となる対象ソフトウェア、非対象ソフトウェア、コンピュータハードウェア、データに区別し、対象ソフトウェアのみを「製品品質モデル」の論点とする。そして、「データ品質モデル」を考える場合は、対象データ (Target Data) と非対象データを区別する。さらに、「利用時の品質モデル」を考える場合は、対象とするデータ、ソフトウェア、コンピュータシステムより上位の情報システムやその利用環境、利用者 (一時利用者、二次利用者または間接利用者) が利用時品質に影響を与えるとする。

3. 品質特性・品質副特性

ISO/IEC 2502n: 品質測定部門では、どの場合でも、品質モデルが複数の「品質特性」から構成されると考える。品質特性とは、正確性、完全性、一貫性といった用語で定義される性質を指し、下表に示すように、下位概念に複数の「品質副特性」を持つと考える。

例えば、システム／ソフトウェア製品品質モデルには、「機能適合性」という品質特性があり、これは「機能完全性」「機能正確性」「機能適切性」といった副特性から構成される。利用時の品質モデルも同様に、例えば「満足性」という品質特性は、「実用性」「信用性」「快感性」「快適性」といった副特性から構成される。

データ品質モデルは、「正確性」「完全性」「一貫性」といった品質特性がある。ただし、既定の品質副特性は定義されていない。

品質モデル	品質特性	品質副特性
システム／ソフトウェア 製品品質	機能適合性	機能完全性、機能正確性、機能適切性
	性能効率性	時間効率性、資源効率性、容量満足性
	互換性	共存性、相互運用性
	使用性	適切度認識性、習得性、運用操作性、ユーザー防止性、ユーザーインターフェイス快感性、アクセシビリティ
	信頼性	成熟性、可用性、障害許容性（耐故障性）、回復性
	セキュリティ	機密性、インテグリティ、否認防止性、責任追跡性、真正性
	保守性	モジュール性、再利用性、解析性、修正性、試験性
	移植性	適応性、設置性、互換性
利用時の品質	有効性	有効性
	効率性	効率性
	満足性	実用性、信用性、快感性、快適性
	リスク回避性	経済リスク緩和性、健康・安全リスク緩和性、環境リスク緩和性
	利用状況網羅性	利用状況完全性、柔軟性
データ品質	正確性	-
	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	-
	機密性	-
	効率性	-
	精度	-
	追跡可能性	-
	理解性	-
	可用性	-
	移植性	-
	回復性	-

4. データ品質測定量 (QM)

ISO/IEC 25024 (「システム及びソフトウェア製品の品質要求及び評価 (SQuaRE) データ品質の測定」) は、ISO/IEC 25021 で定義されたデータ品質モデルに基づいて、データ品質測定の考え方を定義する。

まず、品質特性 (および品質副特性) を何らかの方法で定量化したものを「品質測定量 (Quality Measurement: QM)」と呼ぶ。例えば、品質特性「正確性」の品質測定量には、構文的なデータ正確性、意味論的なデータ正確性、データ正確性の保証、データセット不正確性のリスク、データモデル正確性、メタデータ正確性、データ正確性範囲などがある。これらの品質測定量の数値を求めることで、品質特性が定量的に表現できるようになる。ただし、品質測定量には「システム依存の要素」と「データ固有の要素」があり、「システム依存の要素」は、品質測定の対象データの性質のみでは、妥当な品質評価を行えないことがある。

図：データ品質モデルを構成する品質特性および品質測定量

品質モデル	品質特性	品質測定量
-------	------	-------

データ品質	正確性	構文的なデータ正確性、意味論的なデータ正確性、データ正確性の保証、データセット不正確性のリスク、データモデル正確性、メタデータ正確性、データ正確性範囲
	完全性	レコード完全性、属性完全性、データファイル完全性、データ値完全性、データファイル内の空レコード、概念データモデル完全性、概念データモデル属性完全性、メタデータ完全性
	一貫性	関連するインテグリティ、データ様式一貫性、データ不一致のリスク、アーキテクチャ一貫性、データ値一貫性網羅率、意味論的な一貫性
	信憑性	値の信憑性、情報源信憑性、データ辞書信憑性、データモデル信憑性
	最新性	更新頻度、更新の適時性、公式更新項目要求事項
	アクセシビリティ	ユーザアクセシビリティ、装置アクセシビリティ (★)、データ様式アクセシビリティ (★)
	標準適合性	値及び／又は様式の規制標準適合性、技術に起因する規制標準適合性 (★)
	機密性	暗号化使用率、非脆弱性 (★)
	効率性	効率的なデータ項目様式、使用可能到達性、データ様式効率性 (★)、データ処理効率性 (★)、浪費スペースのリスク (★)、レコード重複によって占有されたスペース (★)、データ更新の時間遅延 (★)
	精度	データ値の精度、データ様式の精度 (★)
	追跡可能性	データ値の追跡可能性、データアクセスの追跡可能性、データ値の項目追跡可能性
	理解性	シンボル理解性、意味論的な理解性、マスタデータ理解性、データ値理解性、データモデル理解性 (★)、データ表現理解性 (★)、連結されたマスタデータの理解性 (★)
可用性	データ可用性比率 (★)、データ利用可能の確率 (★)、アーキテクチャ構成要素の可用性 (★)	

	移植性	データ移植性率 (★)、将来のデータ移植性 (★)、アーキテクチャ構成要素の移植性 (★)
	回復性	データ回復性率 (★)、定期的なバックアップ (★)、アーキテクチャ回復性 (★)

※ (★) はシステム依存の品質測定量

5. データ品質測定量要素 (QME)

品質測定量を表現する計算方法は、「測定関数」によって定義される。測定関数は、単一または複数の測定方法からなる「品質測定量要素 (Quality Measurement Element: QME)」から構成され、一般に、0.0~1.0 (又は 1.0 を超える) 範囲内の値に正規化される。例えば、「構文的なデータ正確性」は、品質測定量要素「A = 構文的に正確な、関係する値を持つデータ項目の数」と「B = 構文的な正確性を要求されるデータ項目の数」で構成され、測定関数「 $X=A/B$ 」で定義される。

具体的には、100 のデータ項目 (フィールド) を持つデータファイルのなかで、95 のデータ項目に「システム」という値があり、5 のデータ項目に「システム」という値があるとすると、このとき、このデータファイルの「構文的な正確性」は 0.95 (= 95/100) と計算できる。

6. データ品質モデルの適合性

なお、ISO/IEC 25024 が品質評価の対象データとするのは、主として「幾つかの種類のアプリケーションに利用されるコンピュータシステム」に保存されたデータであって、「体系化された様式で保持される幾つかの種類のデータ」であることには注意を要する。

言い換えれば、ISO/IEC 25024 の規定のみでは、データ品質を十分に表現できないことがある。例えば、構造化されていないデータや、複数の主体および／またはコンピュータシステムの間を転々流通し、そのたびに性質変化する可能性のあるデータを取り扱う場合には、データ品質測定量や品質測定量要素の追加・変更が求められることもある。

この問題に対応するために、ISO/IEC25024 (JIS X25024) は、「2 適合性」において、次のように述べている。

- a) JIS X 25012 で規定されたように明示又は評価されたデータ品質特性を選択する。
- b) データ品質特性を測定しなければならない対象実体を選択する。
- c) 対象実体に関して各々のデータ品質特性に対して、箇条 8 で規定した適切なデータ品質測定量を選択する。
- d) データ品質測定量を修正する場合、いずれの変更に対しても根拠を提供する。
- e) この規格に含まれないが、使用した品質測定量又は QME があれば、一覧表示する。
 - a)及び b)の順序は、逆に適用することができる。

変更したデータ品質測定量又は新しいデータ品質測定量を使用する場合には、利用者は対象実体、測定方法及び JIS X 25012:2013 又は他に使用された品質モデルに関連したデータ品質特性を明示しなければ

ならない。この規格は、データライフサイクルの間に定義されたデータに関連した、品質測定量の完全な一覧は提供していない。利用者は、適用された技術に応じて、他の何らかの品質測定量を識別してもよい。この規格に含まれる幾つかの品質測定量が、経験的に検証されていなかったとしても、また、たとえそれらの一部が、業界において、まだ最高のものであると確認されていないとしても、この規格は、データ品質測定量を改善するためのよい基盤であり、よい機会である。

出所：日本規格協会「システム及びソフトウェア製品の品質要求及び評価（SQuaRE）—データ品質の測定 JIS X 25024：2018（ISO/IEC 25024：2015）（IPJA/JSA）」

本書は、ISO/IEC25024（JIS X25024）が定めるデータ品質モデルにおいて、既存のデータ品質測定量をセンシングデータに適用するために、新たなデータ品質測定量の識別を試みた。対象実体は、センシングデータに付与されるメタデータのうち、デバイス依存の障害要因に関する情報であると見なした。測定方法は、「設計・情報」「設置・調整」「運用・保守」の3区分で9項目を定義し、それぞれに品質測定量要素を指定した。

なお、本書が提案する新たなデータ品質測定量及びデータ品質測定量要素の定義は、センシングデータ分析における障害事象の積み上げ（ボトムアップ）で考案されたことから、SQuaRE シリーズが定めるシステム及びソフトウェア品質モデル全体との適合を確認していない。

例えば、「デバイスの情報」に定める「（1）機能に関する情報」「（2）性能に関する情報」は、品質測定の対象実体をハードウェアとした場合、ISO/IEC 25023：2016（JIS X 25023：2017）が定める「機能完全性」の品質測定量要素として定義できる。また、対象実体をメタデータとした場合は、ISO/IEC25024（JIS X25024）における「メタデータ正確性」に相当すると考えられる。

ANNEX-4 : 故障のしにくさと耐久性の違い

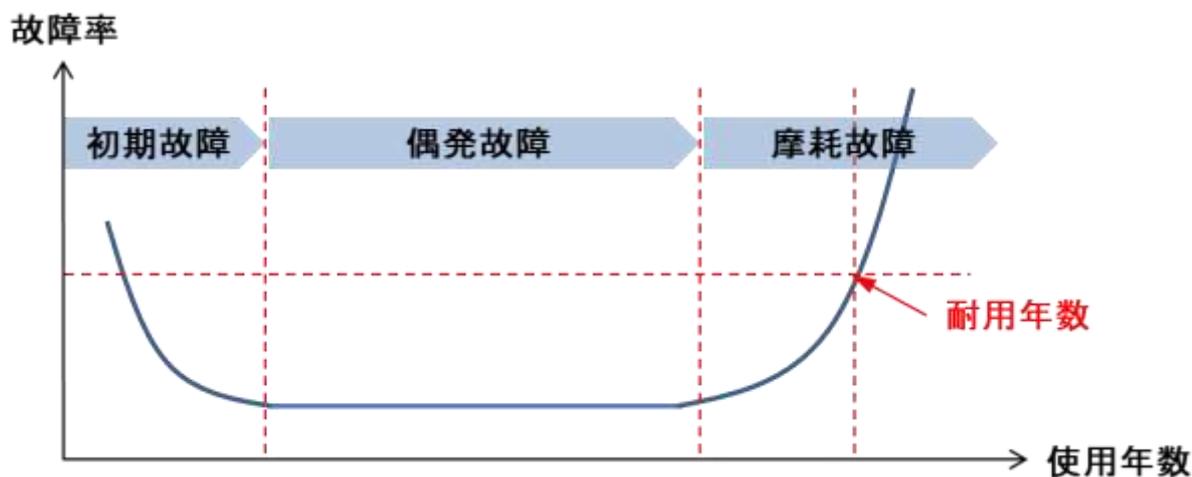
一般的に電気機器の故障率と使用年数の関係は下図に示すような曲線を描き、故障の要因として初期故障、偶発故障、摩耗故障の3段階に分けられる。

初期故障は製造、検査段階で除去が可能であるが、偶発故障は耐用年数期間内において予期できない突発的な故障であり技術的な対策が難しく、統計的な取扱いに基づく施策しかとることができない。

また、摩耗故障は劣化や摩耗により耐用年数の終末付近で発生するもので、故障率が時間の経過と共に急激に増加する。

このため、摩耗故障の始まる前に部品や機器の交換による予防保全を行うことでシステムの稼働率を維持できる。

以上の理由により、統計的な取扱いに基づく施策しかとることができない故障に関するデバイス依存の品質測定量を故障のしにくさと定義する。また、予防保全を行うことでシステムの稼働率を維持できる故障に関するデバイス依存の品質測定量を耐久性と定義する。



使用年数と故障率の関係

この文書について

- 名称 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
- ファイル名 220118-D69-sensing-data-quality-evaluation-standards-wp-tecst.pdf
- 掲載 URL <https://data-society-alliance.org/survey-research/data-quality-evaluation-standards/>
- 概要
一般社団法人データ社会推進協議会(DSA)技術基準検討委員会 WG2-TG 4 は、データ品質評価基準の策定を進めています。本書はその一環として、センシングデータの品質評価手法を提案するホワイトペーパーです。センシングデータの品質評価に関心のある関係者から意見を得ることにより、一層実用性の高いデータ品質評価基準策定を目指すことを目的としています。
- 基本情報
 - DSA 基準文書区分 ホワイトペーパー
 - 作成者 一般社団法人データ社会推進協議会 4011005007414
 - 公開者 一般社団法人データ社会推進協議会 4011005007414
 - 著作権者 一般社団法人データ社会推進協議会 4011005007414
 - 発行日 2022年1月15日
 - 公開日 2022年1月18日
 - 作成アプリケーション Microsoft Word
 - 公開形式 PDF
 - 公開ファイル容量 2,906KB
 - ページ数 54 ページ
- 利用条件
 - 本書を利用したこと、利用しなかったことにより直接または間接に生じた損害に対して、DSA は一切の責任を負いません。
 - 本書を組織や団体として活用される際は、DSA へご一報いただければ幸いです。

本書に関するお問い合わせ

一般社団法人データ社会推進協議会(DSA) 4011005007414

E-mail info@data-society-alliance.org

ホームページ <https://data-society-alliance.org/contact/>