

データにも品質があるの？ ISO 8000ってなに？

第2回標準化カフェ



苑田 義明 yoshiaki.sonoda.w8@mhi.com

三菱重工業株式会社

(一財)エンジニアリング協会国際標準部会長

ISO/TC 184/SC 4/WG 13エキスパート

2021/12/15

データの品質が
引き起こす問題

データ品質標準
ISO 8000

このさき、
何が起きるか



データが悪いと、尊い命も失われます

Irmgard Cooper, 85, dies after mistake at London hospital resulted in blood supplies for heart operation being sent back



出典： <https://www.theguardian.com/uk-news/2016/feb/02/irmgard-cooper-bleeds-to-death-surgery-spelling-error-delays-blood-supply>

血液サンプルに手書きの名前
Irmgard Cooper

≠

輸血用血液の患者名に記された名前
Irmgard Cooper

2015年5月7日、ロンドンに住む85歳の女性が、輸血用血液の到着おくれで亡くなりました。

なぜ？

手術中、必要となって緊急手配した輸血用血液が**到着遅延**で間に合わなかった。

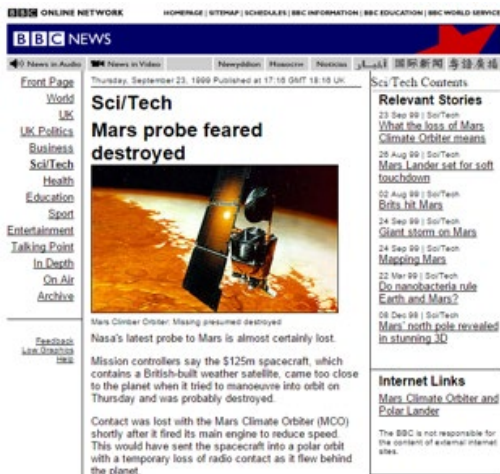
なぜ？

実は、病院には届いていたが、宛先に記された名前が、病院の患者リストになかったので、定められた手続きに従い、**一度送り返されていた**。

なぜ？

血液サンプルに記載された名前が、まちがって綴られていたため、**受け入れ手続きで該当者なし**と判断されていた。

データが悪いと、さまざまな損失が発生します



1999/9/23 Mars Climate Orbiterが墜落！

火星探査機 (NASA)は、探査機と管制システム間の単位系の違いにより、予定軌道より100kmも低い軌道に進入、大気の圧力と抵抗に耐えられず破壊。

3億ドルが一瞬にして失われた！

ヤード・ポンド法



1 重量ポンド・秒
(=4.45ニュートン・秒)

impulse

“32.6”

メートル法



1ニュートン・秒
(本当は4.45ニュートン・秒)

Irmgardさんのケース



火星探査機のケース



データの品質ってなに？ (1/4)

リンゴの品質が優れているとは？

リンゴ

りんごの
品質



良い
優れている

悪い
劣っている

リンゴの品質が優れているとは？

リンゴ

リンゴの
品質特性群

甘さ
sweetness

柔らかさ
mellowness

瑞々しさ
succulent

重さ
weight

大きさ
size

あまければ
甘いほど！

やわらかい
ほうが良い！

瑞々しい
方が良い！

気にしない

気にしない

リンゴの
品質基準

酸っぱいもの！

固い方が良い！

あまり瑞々しく
なくともよい

気にしない

気にしない

気にしない

固い方が良い！

気にしない

重いほど
よい！

大きいほど
良い！

りんごの
最終品質

良い
優れている

いまひとつ

悪い
劣っている

データの品質ってなに？ (2/4)

リンゴの品質が優れているとは？

特性を定める

リンゴ

甘くて、
柔らかくて、
瑞々しくて、
重くて、
小さいリンゴ

リンゴの
品質特性群

甘さ
sweetness

柔らかさ
mellowness

瑞々しさ
succulent

重さ
weight

大きさ
size

あまければ
甘いほど！

やわらかい
ほうが良い！

瑞々しい
方が良い！

気にしない

気にしない

リンゴの
品質基準

評価基準と
判定値○・×

酸っぱいもの！

固い方が良い！

あまり瑞々しく
なくともよい

気にしない

気にしない

気にしない

固い方が良い！

気にしない

重いほど
よい！

大きいほど
良い！

要求品質の
充足度合

りんごの
最終品質

○が60%以上

良い
優れている

○が40~59%

いまひとつ

○が40%未満

悪い
劣っている

データの品質ってなに？ (3/4)

リンゴの品質が優れているとは？

リンゴ

リンゴの
品質特性群

目的
ジュース用

目的
アップルパイ

目的
リンゴ投げ

りんごの
最終品質

甘くて、
柔らかくて、
瑞々しくて、
重くて、
小さいリンゴ

甘さ
sweetness

柔らかさ
mellowness

瑞々しさ
succulent

重さ
weight

大きさ
size

あまければ
甘いほど！

やわらかい
ほうが良い！

瑞々しい
方が良い！

気にしない

気にしない

~~酸っぱいもの！~~

~~固い方が良い！~~

~~あまり瑞々しく
なくともよい~~

気にしない

気にしない

気にしない

~~固い方が良い！~~

気にしない

~~重いほど
よい！~~

~~大きいほど
良い！~~

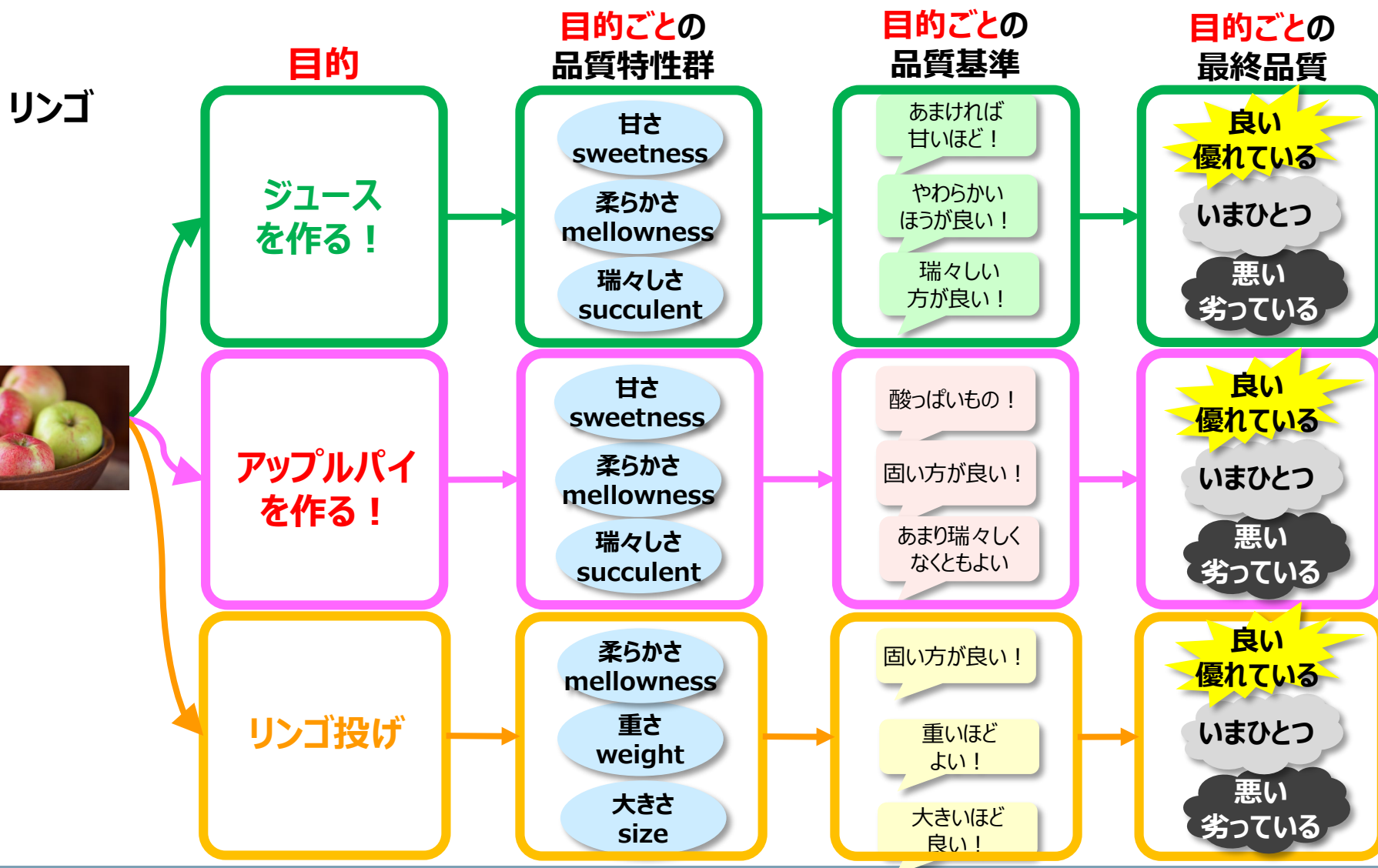
良い
優れている

いまひとつ

悪い
劣っている

データの品質ってなに？ (4/4)

リンゴの品質が優れているとは？ つまり「目的に合っていること」= **目的適合性**

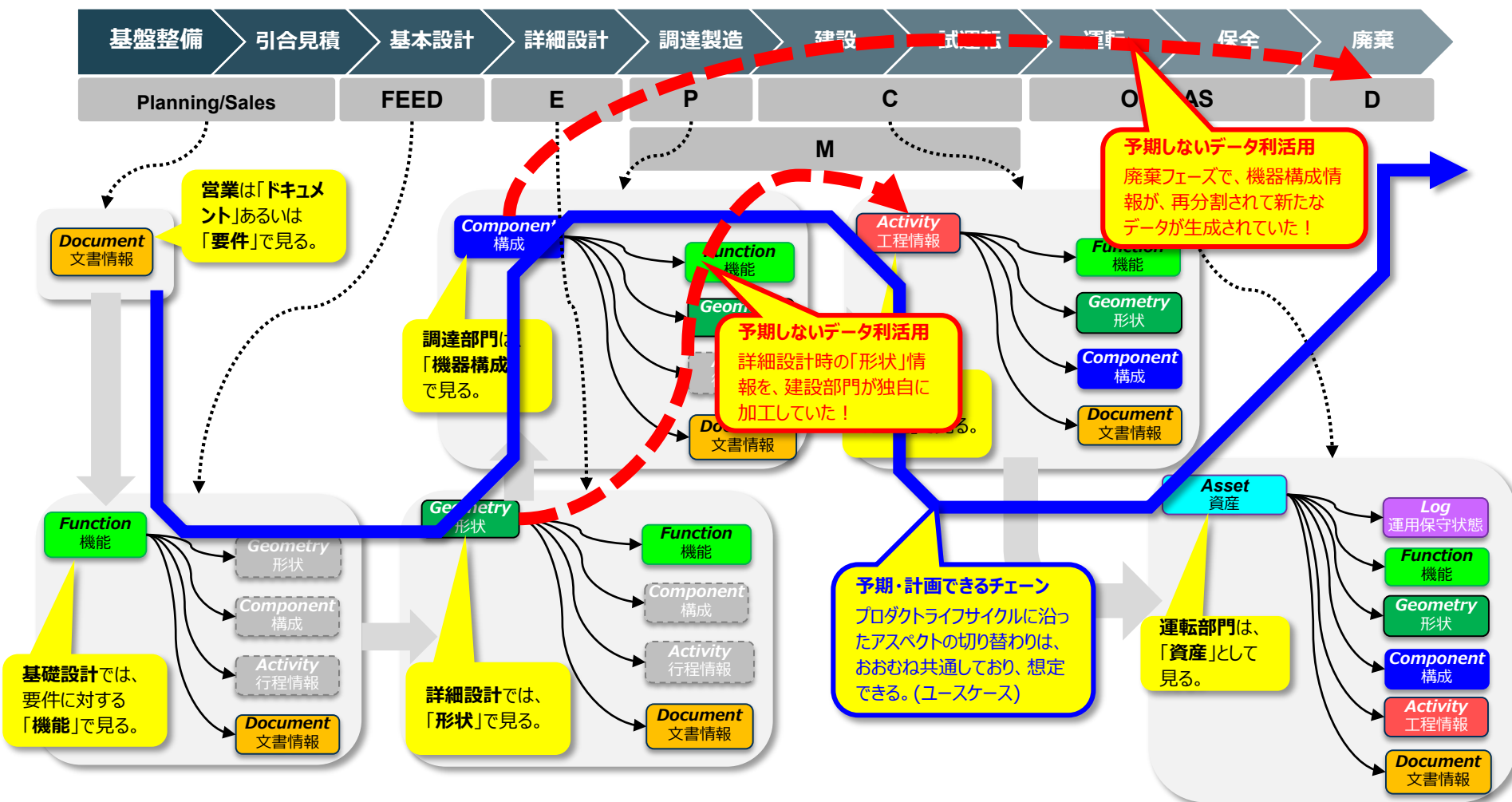


- **目的適合性・合目的性です。**（「普遍的に良いリンゴ」は、無い）
- **同じ対象(リンゴ)でも、目的が違えば、品質の抛り所も異なります。**
 - 品質特性の、種類、数
 - 同じ品質特性でも、定義が異なる
 - 品質特性の、計測方法
 - 判断基準（品質の良し悪しの判断基準）も違います。
- **対象そのものではなく、対象の付属情報も必要です。**
 - 無農薬なのか、誰が作ったのか、どこで作られたのか、etc。

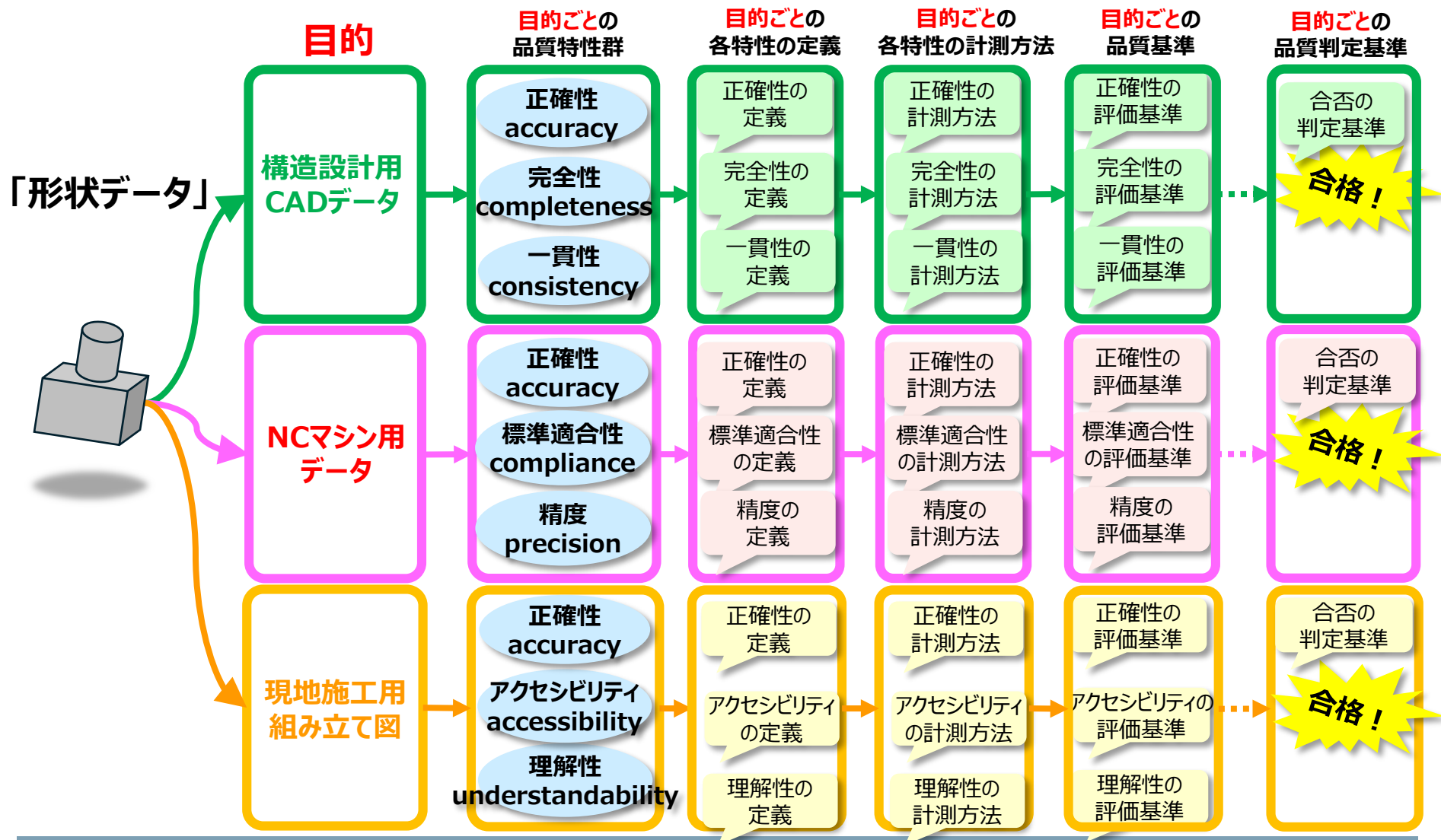
おなじことが、データについても言えます。

データは目的ごとに生成される＝データ品質要求も変わる 三菱重工

- ・ プロダクトライフサイクルのフェーズによって、作業の中心となるデータの着目点（アスペクト）は異なる。
- ・ フェーズが変わるたびに、アスペクトも変わる。
- ・ アスペクトが変われば、データに対する品質要件もまた変わる。

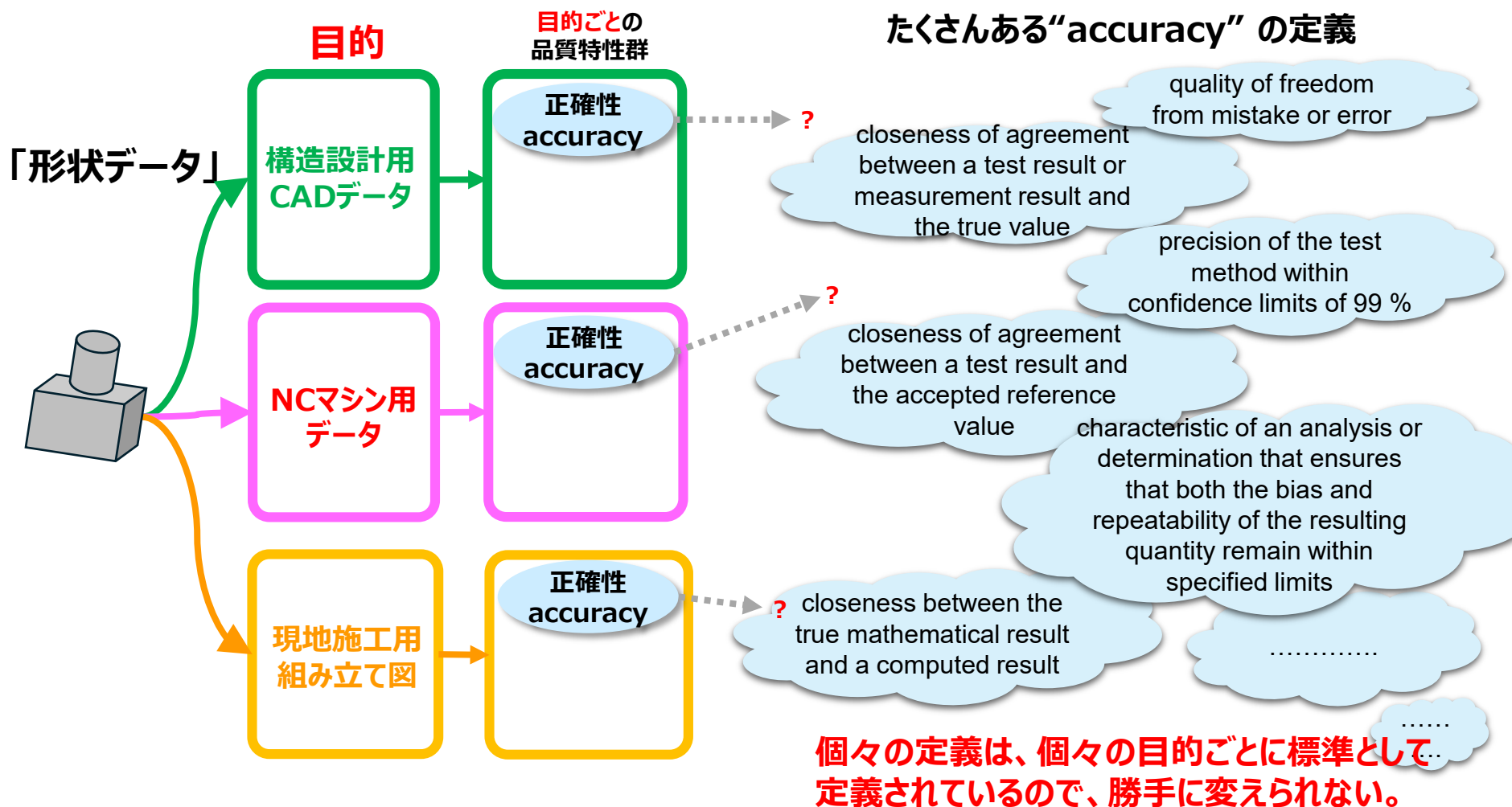


では、「形状データ」のデータ品質が優れているとは？



「正確性」って、目的が違っても同じ？

例えば“**accuracy(正確さ)**”はデータ品質を定める特性の一つだが、ISOだけでも実に多くの定義がある。つまり、「データの品質」は、分野・ユースケースなど、目的の数だけ存在する。（状況依存・文脈依存）



「品質特性群」は、目的が違っても同じ？

- 基本的には、目的が異なれば、特性群の種類も数も異なる。
- 同じ特性(名)でも、目的が異なれば、定義は異なる可能性が高い。

共通の品質特性群は定められるか？
ソフトウェア製品のデータ品質を定める
ISO/IEC 25012には**15**の特性群

ISO 8000-8の開発時、
様々なデータ品質を調査(2010)

目的

ソフトウェア
製品
ISO/IEC
25012

正確性
accuracy

完全性
completeness

一貫性
consistency

信ぴょう性
credibility

最新性
currentness

アクセシビリティ
accessibility

標準適合性
compliance

機密性
confidentiality

効率性
efficiency

精度
precision

追跡可能性
traceability

理解性
understandability

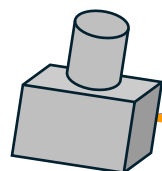
可用性
availability

移植性
portability

回復性
recoverability

10
のデータ品質標準

67
のデータ品質特性



「形状データ」

...

...



- 品質管理のスコープ(範囲)は様々。さまざまな範囲で存在。五月雨式に増えていく。
→ 「全体」のプロセス品質はどうなってる? 「全体」のビジネス品質は?

現実世界

全体の品質が担保されているか判断するには
スコープによらない共通のビジネスプロセス品質の考え方が必要。

ISO 9000ファミリ

製品・サービス

情報世界

データ品質に関しても、
スコープによらない共通のデータ品質の考え方が必要。

ISO 8000シリーズ

データ

目的

データ品質
クライテリア

データ品質
クライテリア

各分野毎の品質定義

購入購買部門

設計部門

エネルギー

交通インフラ

などなど

データ品質は、部分と部分、部分と全体が関係しており、相互に影響する。

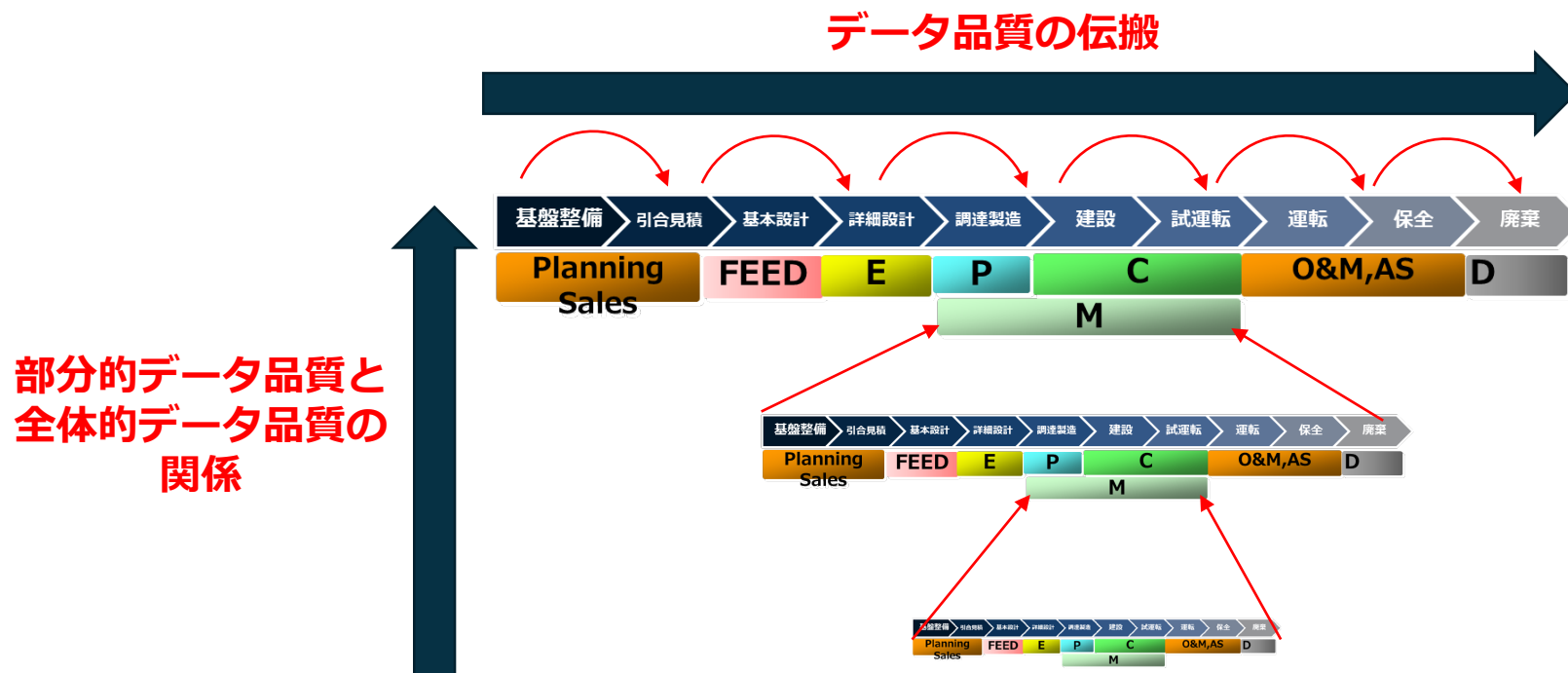
- 「部分と部分の関係」⇒ 時間軸での伝搬など。

前段のフェーズのデータ品質の劣化は、後段のフェーズに伝搬・蓄積していく。

- 「部分と全体の関係」⇒ 入れ子関係での相互影響

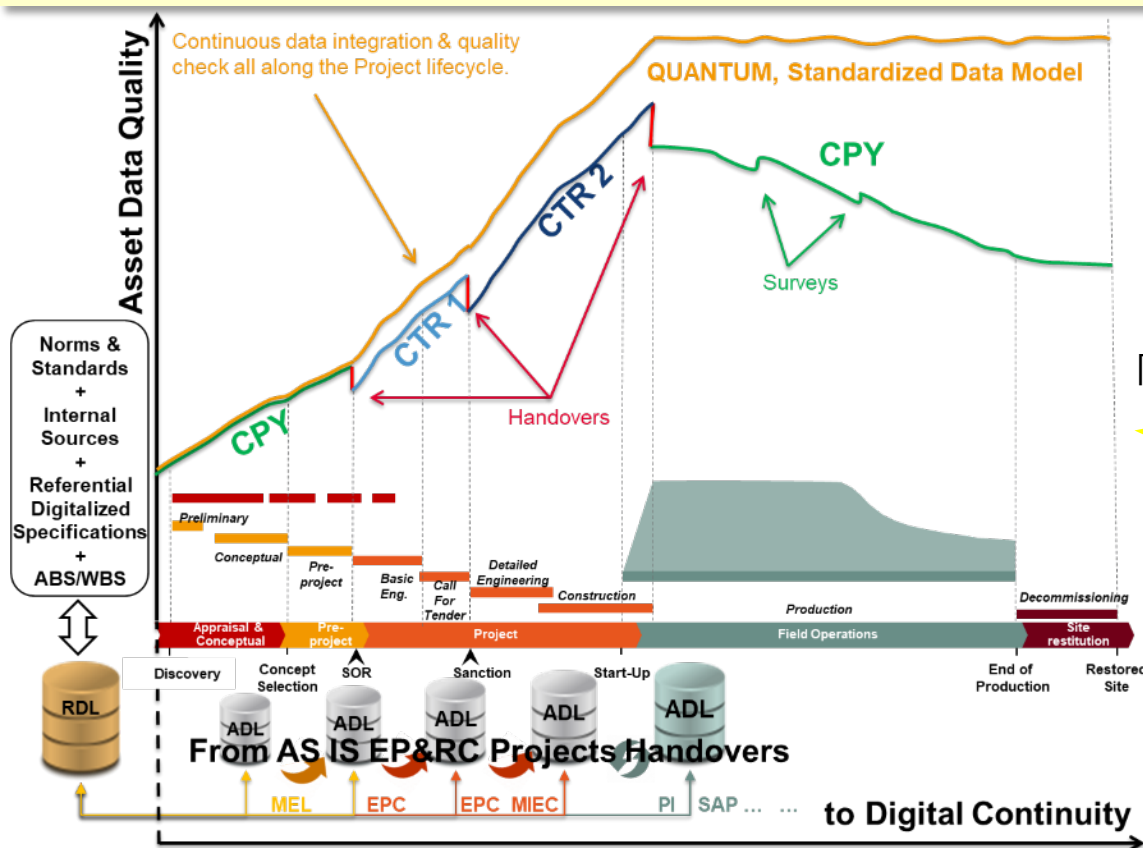
部分のデータ品質の劣化は、プロジェクト全体のパフォーマンスに影響を及ぼす。

・つまり、部分・全体、分野・フェーズ毎の違いにかかわらず、一貫したデータ品質マネジメント・ガバナンスの考え方が必要



データ品質の相互影響の例：データハンドオーバー

- ・ 構造化されていない単なるデジタルデータは、ステークホルダー間でハンドオーバーされるたびに、情報の欠落がおき、データ品質が低下する。データ品質の低下は、工事・製品の品質低下へと直結する。
- ・ 結果、数週間後・数か月後に問題が発覚、設計見直し・工事やりなおし・性能未達・工事遅延・コストオーバ、ひいては事故の発生、死亡事故、環境汚染、企業のブランド失墜、そして事業撤退へつながる。
- ・ 2012年、Shellは社内標準EISを公開、オランダUSPIを中心にハンドオーバー業界標準CFIHOSの開発を開始。
- ・ 2019年、CFIHOSはIOGP(石油ガスオーナーオペレータ協会) が引継ぎ、現在v1.5を発行。
- ・ 石油ガススーパーメジャー 5 社は、入札時の要件としてCFIHOSを求めることを宣言している。



TOTAL CIO の宣言
「我々は、もう“TOTAL” 社内標準は必要ない」
「我々が必要なのは、業界ステークホルダーの
貢献により開発された業界標準である」

「品質」とは、つまり「**目的適合性**」である。

ISO 8000-8, 5.1～5.3

ISO 9000

「品質」とは、
ある**対象**が本来備える特性の集合が、
要求事項を満たす度合い

「要求事項」とは
明示されている、通常暗黙のうちに了解されている、
又は義務として要求されているニーズ若しくは期待

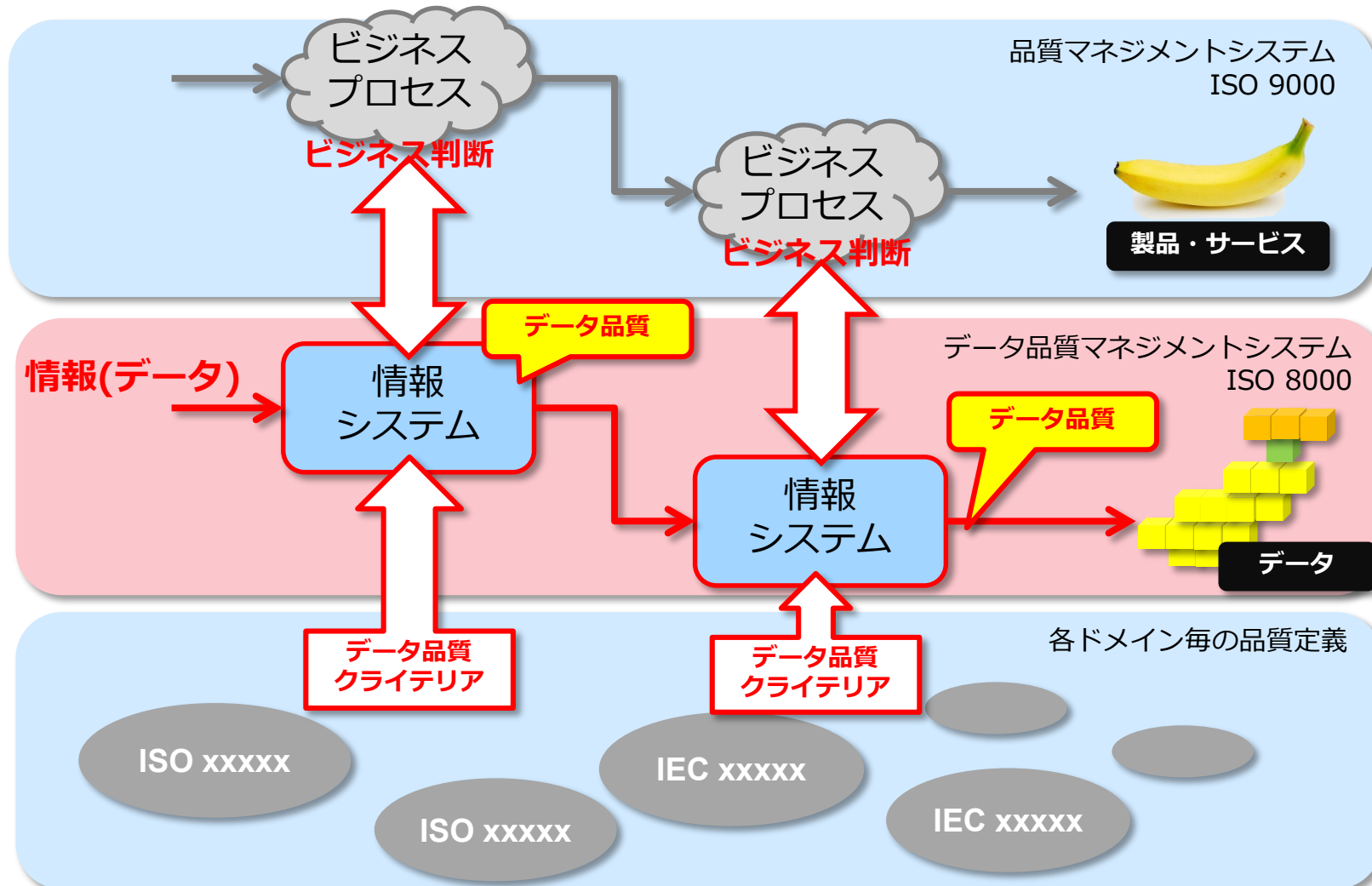
ISO 8000

「データ品質」とは
ある**データ**が本来備える特性の集合が、
データ仕様を満たす度合い

「データ仕様」とは
表現形式と特定の事象・事物を正しく反映するための要件
(業務要件に対応し、データ項目や出力・伝達のタイミングなど)を取り決めたもの。

ISO 9000, ISO 8000, 分野別データ品質標準

ISO 9000はビジネスプロセスの品質管理規格だったが、ISO 8000は計算機によりデータ品質を計測評価することを念頭に置いている。



要件そのものではなく
どのような要件を定
めるべきかを、3種類
で規定

構文的品質

(syntactic quality)

メタデータ、データ
件などの規定構文に対して
データが規格適合して

意味的品質

(semantic quality)

データが、特定の目的のため
に適切かつ意味のあるものと
みなされる度合い

実用的品質

(pragmatic quality)

データが、特定の目的のため
に適切かつ意味のあるものと
みなされる度合い

スコープによらない共通のデータ品質の考え方
ISO 8000シリーズ

データ品質
要件

データ品質
要件

データ品質
要件

データ品質
要件

データ品質
要件

データ品質
要件

個別分野毎のデータ品質

目的1
ソフトウェア製品
のデータ品質
ISO/IEC 25012

目的2
Vプロセス
のデータ品質
ISO/IEC 15288

目的3
ITSシステム
のデータ品質
ISO/TR 21707

目的4
航空宇宙・自動車
のデータ品質
ISO 8000-311

目的n
地理情報の
データ品質
ISO 19157

「みんなちがって、みんないい」
スコープごとのデータ品質の考え方
さまざまな国際・業界標準

ISO 8000の「データ品質」は

- ・ **systematic**(系統的) / **systemic**(体系的)に、
- ・ **syntactic**(構文的品質) / **semantic**(意味的品質) / **pragmatic**(合目的品質)
な要件や手続き、評価方法を定める。
- ・ **個別の分野の、個別なデータ品質要件を定めるものではない。**

構文的品質

(syntactic quality)

メタデータによって宣言される要件などの、規定された構文(シンタックス)に対して、データが規格適合している程度を示す。

(ISO 8000-8 5.1, Annex B)

- ① 構文的品質を定める**完全なルール群**
- ② 情報を表現するための**構文的規約**、
などなど。

意味的品質

(semantic quality)

データが、表現しようとしているものと一致している程度を示す。

(ISO 8000-8 5.2)

- ① 対象領域の**概念モデル**、
- ② **データの表現**、
- ③ **データの検証を行う方法の説明**、
などなど

合目的品質

(pragmatic quality)

データが、特定の目的のために適切かつ意味のあるものとみなされる程度を示す。

(ISO 8000-8 5.3, Annex C)

- ① **品質評価軸**に関する定義
- ② 側面における既知の相互依存性、
- ③ 品質側面の計測に利用されるデータの収集方法の定義、などなど

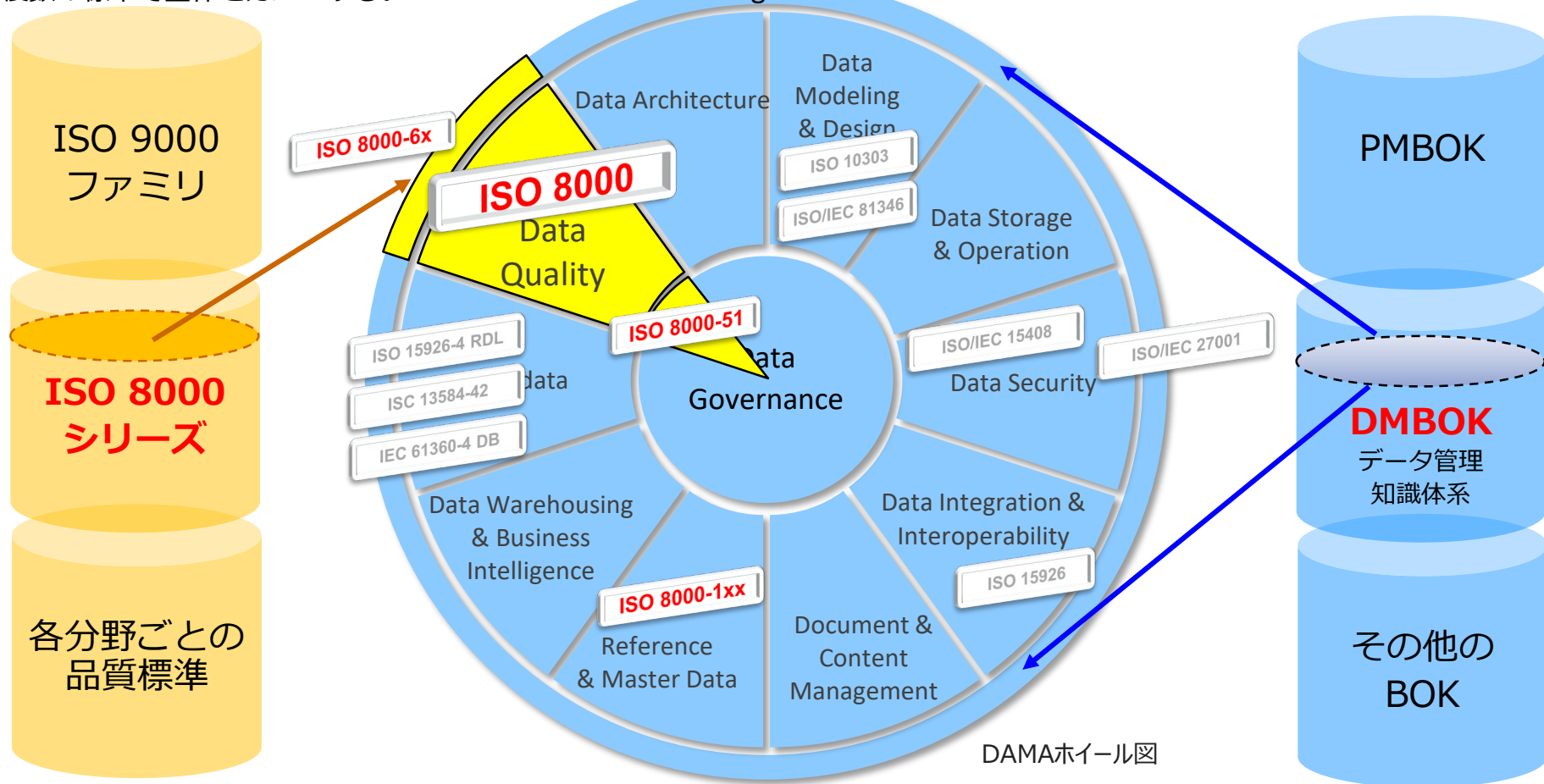
検証 (Verification)
により計測

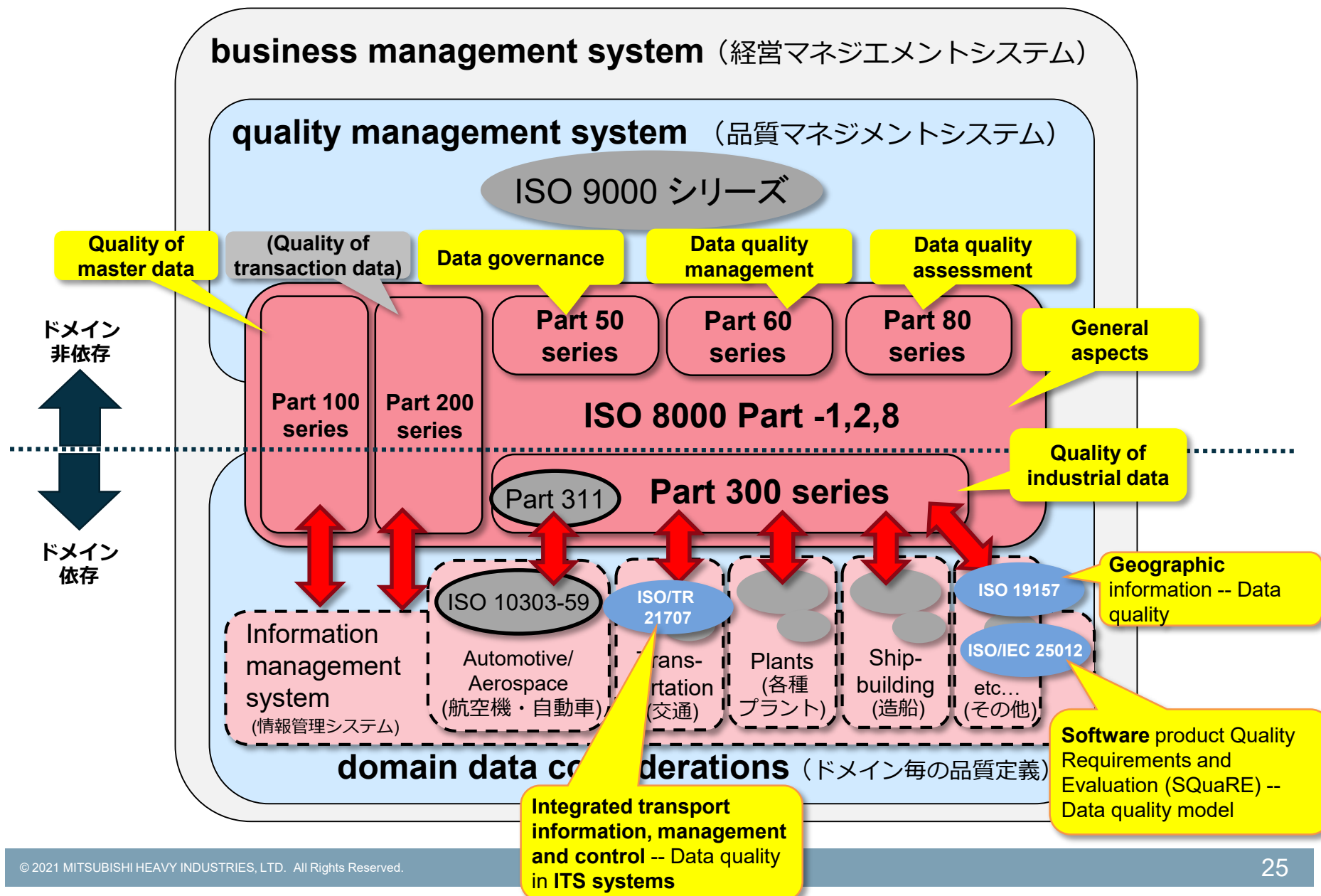
妥当性評価 (Validation)
により計測

- ・データ品質：NATOにおける機器調達の手書(ISO 22745)運用適正化のためにISO 8000シリーズ開発開始
- ・データマネジメント：DAMA(Data Management Association)が包括的・網羅的ガイドを策定・発行（ed 2）

あらゆるデータの品質の要件と計測方法、
データ品質マネジメントを国際標準化
複数の標準で全体をカバーする。

データ管理に関して、原理原則からはじめて、
包括的・網羅的にまとめたもの





【議題3.2.1】 TMB TF18 基盤規格(Foundational Standards)

<背景>

- 2018年9月、TMBは基盤規格に関してTF18を設立し、定義の開発、市場ニーズの把握、SWOT分析等の検討を開始。

<審議内容・結果> 決議番号：49/2019

- ①「水平規格類」の共通定義を承認
 - 多数の委員会又はセクターに関連がある、若しくは標準化規格類全体で一貫性を確保することが極めて重要なテーマを取り扱う規格類。
 - 注1) 水平規格類は、基本原則、概念、図記号、用語又は一般的特性を提示することができる。
 - 注2) 一部の水平規格類は、新しい革新技術の実効性があり、公平で、責任ある監督の基盤となる。
- ②ISO/CSはIEC/COと協力し、規格作成者及びユーザ向けのガイダンスを策定。
- ③今後の進め方
 - ISO/CS及びIEC/COに、2020年末までに水平規格類の使用及び価値に関する事例研究を提供する情報及び委員会を集めるよう要請。
 - 2021年中旬に状況の見直しを行うとともに、次のステップを決定。

Horizontal deliverable

deliverable dealing with a subject relevant to a number of committees or sectors or of crucial importance to ensure coherence across standardization deliverables.

Note 1 horizontal deliverables may provide fundamental principles, concepts, graphical symbols, terminology or general characteristics.

Note 2 some horizontal deliverables provide the foundations of workable, fair and responsible oversight of new innovative technologies.

ISO 8000シリーズは、水平規格です。

●ISO Horizontal deliverablesの数：84標準（開発・改訂中含む、2021/11/3現在）

うち、Data qualityに関するHorizontal deliverableは以下の3つ

ISO/DIS 8000-1 Data quality — Part 1: Overview

ISO 8000-2:2020 Data quality — Part 2: Vocabulary

ISO 8000-8:2015 Data quality — Part 8: Information and data quality: Concepts and measuring

●“Data quality”を含むISO標準数：53標準（開発・改訂中含む、2021/11/3現在）

うち、ISO 8000シリーズは27標準（約50%）

Reference	Document title	Developing committee
ISO/IEC AWI 5259-1	Data quality for analytics and ML — Part 1: Overview, terminology, and examples	ISO/IEC JTC 1/SC 42/WG 2
ISO/IEC AWI 5259-2	Data quality for analytics and ML — Part 2: Data quality measures	ISO/IEC JTC 1/SC 42/WG 2
ISO/IEC AWI 5259-3	Data quality for analytics and ML — Part 3: Data quality management requirements and guidelines	ISO/IEC JTC 1/SC 42/WG 2
ISO/IEC AWI 5259-4	Data quality for analytics and ML — Part 4: Data quality process framework	ISO/IEC JTC 1/SC 42/WG 2
ISO/IEC 25012:2008	Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model	ISO/IEC JTC 1/SC 7
ISO/IEC 25024:2015	Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Measurement of data quality	ISO/IEC JTC 1/SC 7
ISO/TS 10303-1520:2018	Industrial automation systems and integration — Product data representation and exchange — Part 1520: Application module: Product data quality definition	ISO/TC 184/SC 4/WG 12
ISO/TS 10303-1521:2018	Industrial automation systems and integration — Product data representation and exchange — Part 1521: Application module: Product data quality criteria	ISO/TC 184/SC 4/WG 12
ISO/TS 10303-1522:2018	Industrial automation systems and integration — Product data representation and exchange — Part 1522: Application module: Product data quality inspection result	ISO/TC 184/SC 4/WG 12
ISO/TS 10303-1523:2019	Industrial automation systems and integration — Product data representation and exchange — Part 1523: Application module: Shape data quality criteria	ISO/TC 184/SC 4/WG 12
ISO/TS 10303-1524:2018	Industrial automation systems and integration — Product data representation and exchange — Part 1524: Application module: Shape data quality inspection result	ISO/TC 184/SC 4/WG 12
ISO/TS 10303-1821:2014	Industrial automation systems and integration — Product data representation and exchange — Part 1821: Application module: Externally conditioned data quality criteria	ISO/TC 184/SC 4/WG 12
ISO/TS 10303-1822:2014	Industrial automation systems and integration — Product data representation and exchange — Part 1822: Application module: Externally conditioned data quality inspection result	ISO/TC 184/SC 4/WG 12
ISO/AWI TS 10303-1848	Industrial automation systems and integration - Product data representation and exchange — Part 1848: Application module: Annotated 3d model data quality criteria	ISO/TC 184/SC 4/WG 12
ISO/AWI TS 10303-1849	Industrial automation systems and integration - Product data representation and exchange — Part 1849: Application module: Annotated 3d model data quality inspection result	ISO/TC 184/SC 4/WG 12
ISO/AWI TS 10303-1850	Industrial automation systems and integration - Product data representation and exchange — Part 1850: Application module: Triangulated shape data quality	ISO/TC 184/SC 4/WG 12
ISO/AWI TS 10303-1851	Industrial automation systems and integration - Product data representation and exchange — Part 1851: Application module: PMI data quality	ISO/TC 184/SC 4/WG 12
ISO/TS 8000-1:2011	Data quality — Part 1: Overview	ISO/TC 184/SC 4/WG 13
ISO/DIS 8000-1	Data quality — Part 1: Overview	ISO/TC 184/SC 4/WG 13
ISO 8000-100:2016	Data quality — Part 100: Master data: Exchange of characteristic data: Overview	ISO/TC 184/SC 4/WG 13
ISO 8000-110	Data quality — Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification	ISO/TC 184/SC 4/WG 13
ISO 8000-110:2009	Data quality — Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification	ISO/TC 184/SC 4/WG 13
ISO/PWI 8000-114	Data quality — Part 114: Master data: Application of ISO/IEC 21778 and ISO 8000-115 to portable data	ISO/TC 184/SC 4/WG 13
ISO 8000-115:2018	Data quality — Part 115: Master data: Exchange of quality identifiers: Syntactic, semantic and resolution requirements	ISO/TC 184/SC 4/WG 13
ISO 8000-116:2019	Data quality — Part 116: Master data: Exchange of quality identifiers: Application of ISO 8000-115 to authoritative legal entity identifiers	ISO/TC 184/SC 4/WG 13
ISO/CD 8000-117	Data quality — Part 117: Application of ISO 8000-115 to Quality Blockchains	ISO/TC 184/SC 4/WG 13
ISO 8000-120:2016	Data quality — Part 120: Master data: Exchange of characteristic data: Provenance	ISO/TC 184/SC 4/WG 13
ISO 8000-130:2016	Data quality — Part 130: Master data: Exchange of characteristic data: Accuracy	ISO/TC 184/SC 4/WG 13
ISO 8000-140:2016	Data quality — Part 140: Master data: Exchange of characteristic data: Completeness	ISO/TC 184/SC 4/WG 13
ISO/DIS 8000-150	Data quality — Part 150: Data quality management: Roles and responsibilities	ISO/TC 184/SC 4/WG 13
ISO/TS 8000-150:2011	Data quality — Part 150: Master data: Quality management framework	ISO/TC 184/SC 4/WG 13
ISO 8000-2:2020	Data quality — Part 2: Vocabulary	ISO/TC 184/SC 4/WG 13
ISO 8000-2:2020/Amd 1:2021	Data quality — Part 2: Vocabulary — Amendment 1	ISO/TC 184/SC 4/WG 13
ISO/TS 8000-311:2012	Data quality — Part 311: Guidance for the application of product data quality for shape (PDQ-S)	ISO/TC 184/SC 4/WG 13
ISO/CD 8000-51	Data quality — Part 51: Data governance: Exchange of data policy statements	ISO/TC 184/SC 4/WG 13
ISO/TS 8000-60:2017	Data quality — Part 60: Data quality management: Overview	ISO/TC 184/SC 4/WG 13
ISO 8000-61:2016	Data quality — Part 61: Data quality management: Process reference model	ISO/TC 184/SC 4/WG 13
ISO 8000-62:2018	Data quality — Part 62: Data quality management: Organizational process maturity assessment: Application of standards relating to process assessment	ISO/TC 184/SC 4/WG 13
ISO 8000-63:2019	Data quality — Part 63: Data quality management: Process measurement	ISO/TC 184/SC 4/WG 13
ISO/DIS 8000-64	Data quality — Part 64: Data quality management: Organizational process maturity assessment: Application of the Test Process Improvement method	ISO/TC 184/SC 4/WG 13
ISO/TS 8000-65:2020	Data quality — Part 65: Data quality management: Process measurement questionnaire	ISO/TC 184/SC 4/WG 13
ISO 8000-66:2021	Data quality — Part 66: Data quality management: Assessment indicators for data processing in manufacturing operations	ISO/TC 184/SC 4/WG 13
ISO 8000-8:2015	Data quality — Part 8: Information and data quality: Concepts and measuring	ISO/TC 184/SC 4/WG 13
ISO/TS 8000-81:2021	Data quality — Part 81: Data quality assessment: Profiling	ISO/TC 184/SC 4/WG 13
ISO/DTS 8000-82	Data quality — Part 82: Data quality assessment: Creating data rules	ISO/TC 184/SC 4/WG 13
ISO/TR 21707:2008	Intelligent transport systems — Integrated transport information, management and control — Data quality in ITS systems	ISO/TC 204/WG 9
ISO 19160-3:2020	Addressing — Part 3: Address data quality	ISO/TC 211/WG 7
ISO 19157:2013	Geographic information — Data quality	ISO/TC 211/WG 9
ISO 19157:2013/Amd 1:2018	Geographic information — Data quality — Amendment 1: Describing data quality using coverages	ISO/TC 211/WG 9
ISO/CD 19157-1	Geographic information — Data quality — Part 1: General requirements	ISO/TC 211/WG 9
ISO/TS 19157-2:2016	Geographic information — Data quality — Part 2: XML schema implementation	ISO/TC 211/WG 9
ISO/AWI 19157-3	Geographic information — Data quality — Part 3: Data quality measures register	ISO/TC 211/WG 9
ISO/WG 30435	Human Resource Management — Workforce Data Quality Standard	ISO/TC 260/WG 2

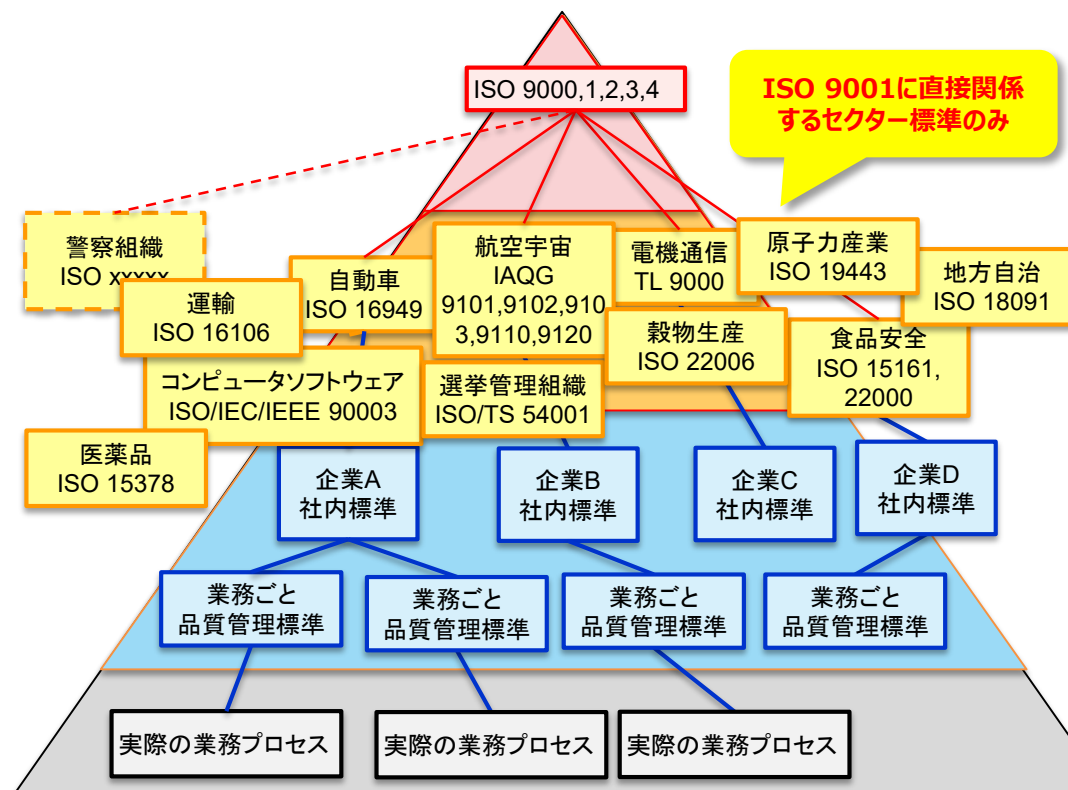
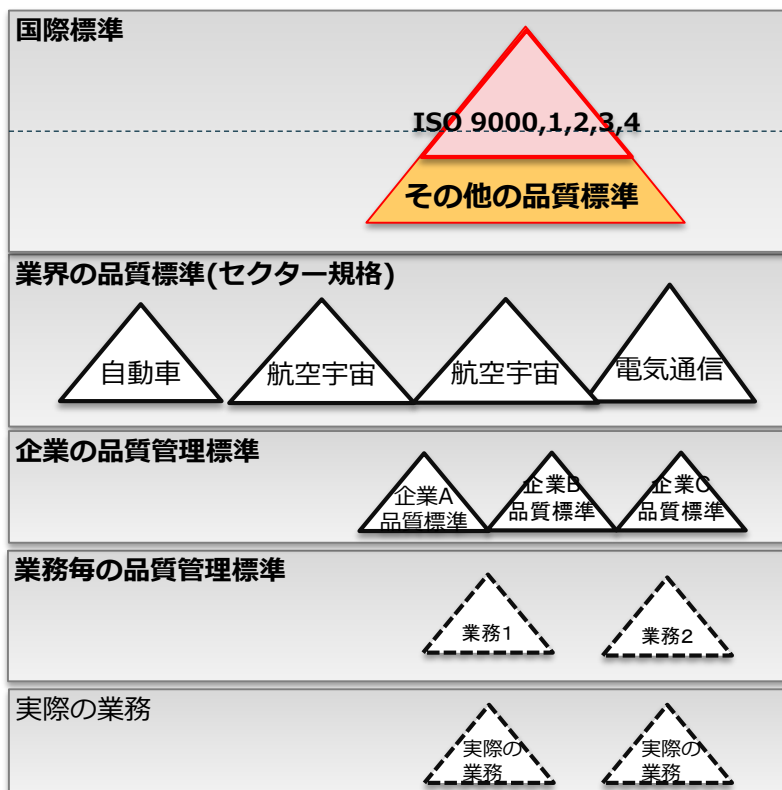
普遍的要件を定めた9000シリーズと、個別分野毎のセクター規格の階層構造

- ISO 9000を頂点とするISO 9000シリーズは、すでに必然として階層化されている。ISO 9000の普遍的要件に沿って、各分野毎により具体化されたセクター標準が多く開発・発行されている。例えば自動車産業と食品産業という違いがあっても、ISO 9000という共通の標準にそって具現化されているため、同じ程度にそのビジネスプロセスは信頼できると言える。
- 最近の状況では、2021/7にカナダ（SCC）が新規PC設置提案

NP Quality management systems — Guidelines for the application of ISO 9001 in policing organization

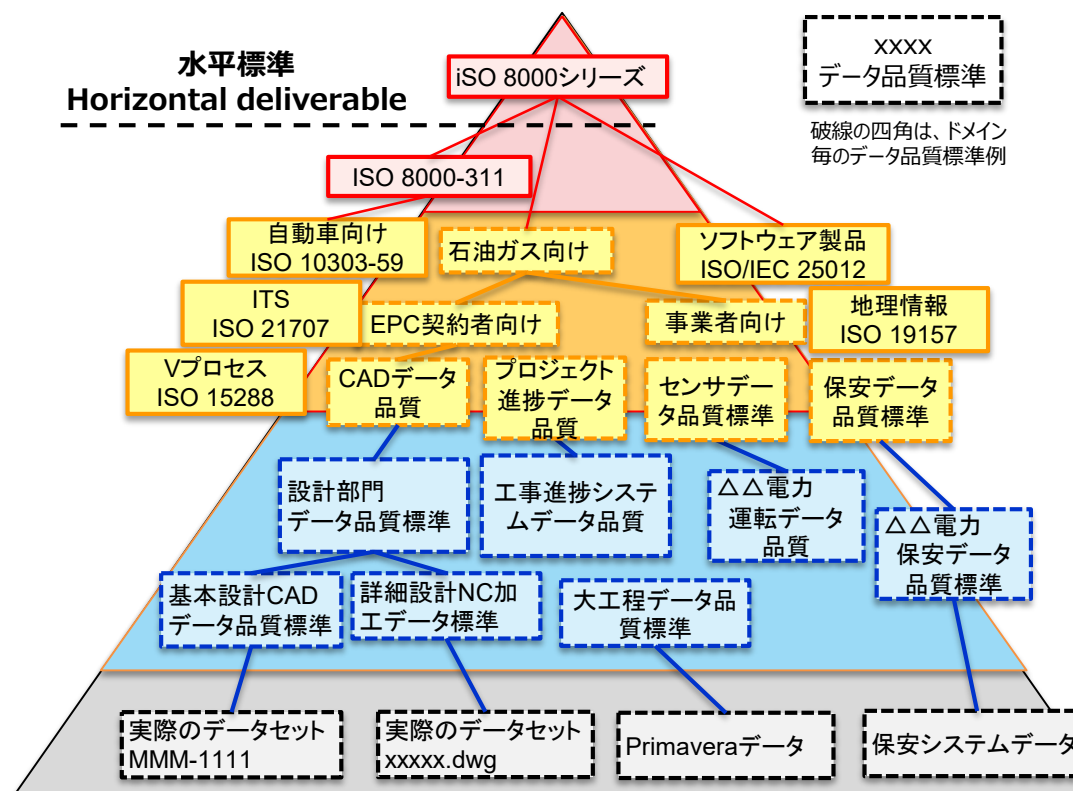
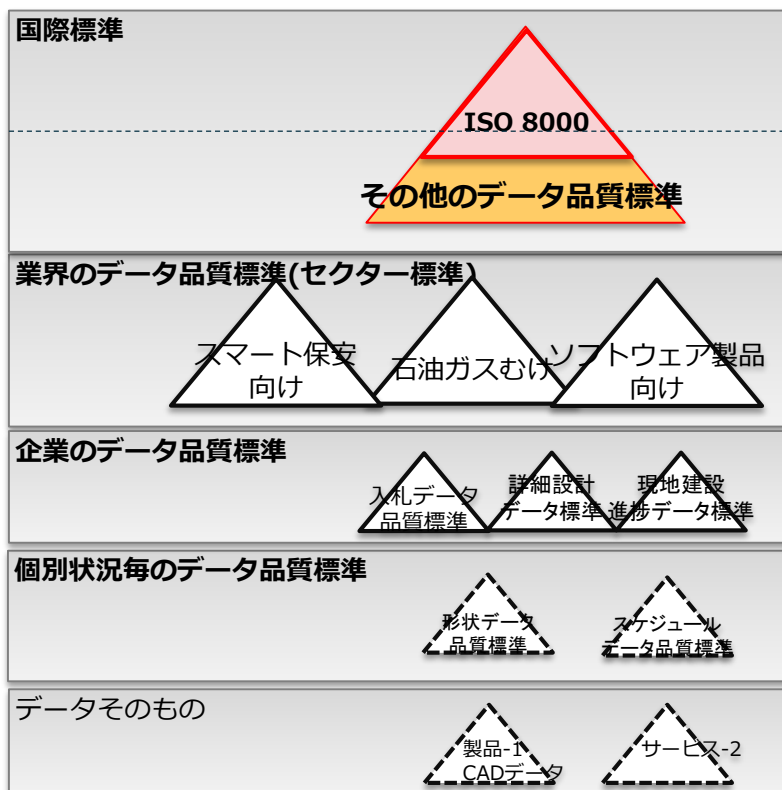
（品質管理システム—警察組織におけるISO 9001の適用のためのガイドライン）

がなされ、いままさに投票中。



データ品質標準は国際標準(水平・個別) から、業界標準、企業標準と階層化される。

- デジタル経済圏でのデータ連携では、相互（のデータが）が信頼できることが必須要件。一方、個別の事業分野毎にデータ品質の考え方は異なる。多様性を維持しつつも、全体をどのように必要に応じて相互連携できるようにするかが大きな課題。
- データ品質に関しても、個々のプロダクトライフサイクルフェーズ毎・ステークホルダ毎・システム毎に異なるデータ品質の標準や考え方があり、それらが構成するプロダクトライフサイクル全体のデータ品質もまた、これら個別のデータ品質の充足状況が、より上位のデータ品質の充足へと連携する必要がある。（そうでなければ、個別の状況でのデータ品質評価ができて、全体としてのデータ品質のガバナンスが効いているのかどうかの判断がつかない。）
- いずれか一つの状況依存のデータ品質標準をもって全体に変えることはできないし、また異なる複数のデータ品質標準の和を取ることもまたできない。個々のデータ品質標準は、それそのものとして扱い、より普遍的要件を定める水平標準が必要。



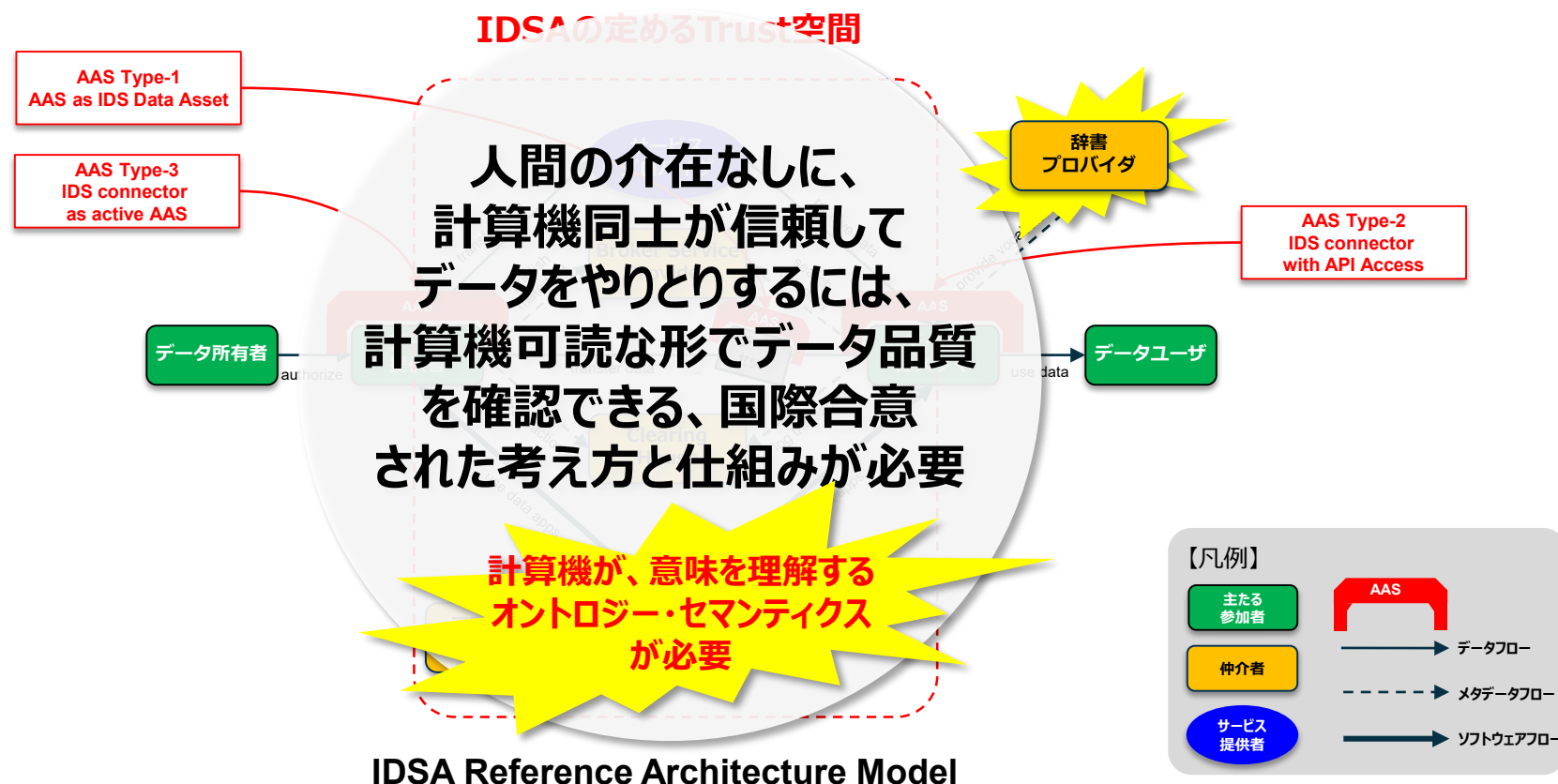
- ・「体重が70kgある」というデータは、それだけでは、「重い」のか「軽い」のか、「適切」なのか「不適」なのかのよしあしが判断できない。
- ・つまり、データの品質を判断するよりどころ（ルール・要件）は文脈依存である。
- ・「健康状態を判断するための体重」と「力士候補を選定するための体重」は、同じ70kgでも、全く意味が異なる。
- ・また、計測するための方法や単位、また良し悪しを判断するクライテリアがあって、はじめて意味をもち、品質について、良い・悪いが言える。

		要件①	要件②	要件③	要件④
データ品質 水平標準	ISO 8000-8	文脈（オントロジ）を定めよ	適合性(良し悪し)を判断する属性を定めよ	適合性の計測方法定めよ	適合性の判断方法を定めよ
データ品質 個別標準	健康状態の判断データ 品質個別標準	健康管理分野オントロジ辞書と、データモデル。	「体重」プロパティで判断する。	「体重」プロパティは数字（小数点以下1桁）IS単位でkgではかる。JISxx等級の体重計で計測	70kg以下 → 健康 71kg以上 → 肥満
	相撲取りの候補選定 データ品質個別標準	相撲オントロジ辞書とデータモデル。	「体重」プロパティで判断する。	「体重」プロパティは数字（小数点以下2桁）単位は貫目で測る。相撲取り専用体重計で計測	30貫目以下 → 不適 31貫目以上 → 適
	ISO/IEC 25012	...	15の評価軸	...	
	ISO 21707	...	...	...	

ざっくりとした水平標準と個別標準の関係を顕そうとした例なので、厳密さに欠けることがあります。

データ品質は、デジタル経済圏への参加条件

- ・ **GAIA-X**は2019年活動開始、欧州デジタル経済圏を支える、**データ主権**を取り入れたデータ連携基盤。
- ・ **IDSA**は2014年活動開始、GAIA-X の元でデータ主権を技術として実装する
- ・ **Industry4.0**の中核技術のひとつである**AAS(Asset Administration Shell)**は、その基盤としてオントロジー辞書(IEC CDD)の利用を前提としている。IDSAはAASを拡張適用していることから、おのずとオントロジー辞書が必要となる。



改竄できないように処置されたデータセット

Immutable data set (**off-ledger data set**)

例えばレコード全体のCRC32 Hash
を利用し、データの正しさを示す

7B384A82

immutableな
data setであること

On-ledger identifiers

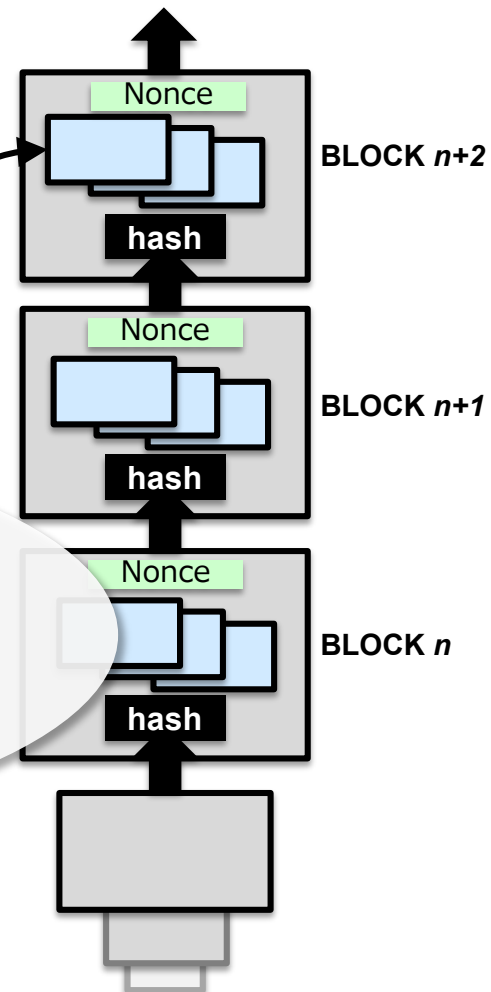
7B384A82Corning.LS:352075

Verifier of
Off-ledger data set
(ISO 8000-117)

Quality identifier
(ISO 8000-115)

識別子	バリュー
0188-1#05-000845#1	Corning.LS:352075
0188-1#05-003612#2	Tube
0188-1#05-009981#1	Centrifuge
0188-1#05-002791#5	Conical Bottom
0188-1#05-005578#1	8 / Pk
0188-1#05-003751#1	48 / Cs
0188-1#05-011121#7	Falcon®
0188-1#05-008412#1	225 ml
0188-1#05-002133#1	Gamma irradiation
0188-1#05-009122#1	7500 xg
0188-1#05-004221#3	Clear
0188-1#05-073156#1	Blue
0188-1#05-000118#1	Plug Seal Screw
0188-1#05-043751#75	Polypropylene (PP)
0188-1#05-006131#2	61 mm
0188-1#05-097895#1	0.010 ounces
0188-1#05-000845#1	137 mm

データ品質ブロックチェーン



ISO/CD 8000-117 Quality Blockchains

多様なデータを、ブロックチェーンを利用して
データ品質を担保しながらやりとりする
ための標準が審議されています。

※ご注意

- ・ 2021/10/12, CD承認。
- ・ 本図は、審議途中のアイデアを発表者が独自に図示化したもので、今後の議論で変わる可能性があります。
- ・ "0188-1#05-000845#1"などの辞書項目の識別子は、説明のために識別子の例として作成したもので、標準として定められている正規の識別子とは関係ありません。

データ品質とは

- 目的適合性（データがデータ仕様を満たす度合い）
- 同じデータでも、目的が違えば、データ品質の要件もまた異なる。
- データそのものではなく、データの付属情報(メタデータ) も必要です。
- 部分から部分(時間軸)、部分から全体に伝搬します。

ISO 8000シリーズとは

- ISO 9000シリーズのデータ版
- systematic(系統的)/systemic(体系的)な分野非依存の、水平標準です。
- 計算機によるデータ品質計測を念頭においています。
- syntactic(構文的) / semantic(意味的) / pragmatic(実用的)の3つの側面で、汎用的なデータ品質要件や手続き、評価方法を定めています。(ISO 8000-8)

今後は

- 計算機同士が、直接データ取引するようになります。（デジタル経済圏）
 - 計算機処理できるデータ品質要件と仕組みへの対応が必要になります。
 - そのカギは、オントロジーとセマンティクスです。
- 対応できないと、デジタル世界での「地図に存在しない島」になります。**

ご清聴ありがとうございました。

ご質問はありますか？

MOVE THE WORLD FORWARD

**MITSUBISHI
HEAVY
INDUSTRIES
GROUP**

Part	Cat.	Status	Publish	Development Stage	Title
General aspects of data quality					
1	TS		2011/12		Overview
1ed2	IS	DIS		DIS ballot ended on 2021/7/31	
2ed4	IS	PUB	2020/6/8	CR10 approved on 2021/6/4	Vocabulary
3	canceled				(taxonomy) Removed by Resolution F on 76th Chicago meeting (2018/11)
8	IS		2015/11		Information and data quality: Concepts and measuring
Data governance					
51	IS	DIS		CD approved on 2021/10/12	Exchange of data policy statements
Data quality management					
60	TS		2017/10/13		Overview
61	IS		2016/11/17		Process reference model
62	IS		2018/9/12		Organizational process maturity assessment: Application of standards relating to process assessment
63	IS		2019/12/18		Process measurement
64	IS	DIS		DIS ballot ended on 2021/10/14	Organizational process maturity assessment: Application of the Test Process Improvement method
65	TS		2020/6/25		Process measurement questionnaire
66	IS		2021/10/5		Assessment indicators for data processing in manufacturing operations
150	TS		2011/12/7		Quality management framework
150ed2	IS	DIS		DIS ballot initiated on 2021/9/23	Roles and responsibilities
Data quality assessment					
81	TS		2021/5/20	Registered for formal approval on	Data profiling
82	TS	PUB		DTS approved on 2021/3/8	Creating data rules

Part	Cat.	Status	Publish	Development Stage	Title
Quality of master data					
100	IS		2016/10		Exchange of characteristic data: Overview
100ed2	IS			NP ballot closed on 2020/5/27	Exchange of characteristic data: Overview
110	IS		2009/11		Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification
110ed2	IS	FDIS		FDIS ballot ended on 2021/9/24	
114	IS	NP		NP ballot ended on 2021/7/27	Portable data - Application of ISO/IEC 10646, ISO/IEC 21778 and ISO 8000-115 to the exchange of master data
115	IS		2018/04		Exchange of quality identifiers: Syntactic, semantic and resolution requirements
116	IS		2019/9/2		Application of ISO 8000-115 to the formatting of Authoritative Legal Entity Identifiers (ALEI) for individuals and organizations
117	IS	DIS		CD Ballot ended on 2021/9/1	Identifiers on blockchains
120	IS		2016/10		Exchange of characteristic data: Provenance
130	IS		2016/10		Exchange of characteristic data: Accuracy
140	IS		2016/10		Exchange of characteristic data: Completeness
Quality of transaction data					
210	IS	(NP)		NP ballot will start on 2021/12	Sensor data quality
Quality of industrial data					
311	TS		2012/4/12		Guidance for the application of product data quality for shape (PDQ-S)