

ISO/IEC 11179-4 (Information technology - Metadata registries : MDR) の解説

DOI : <https://doi.org/10.60430/digital.e-learning0009>

独立行政法人情報処理推進機構

デジタル基盤センター

データ環境勉強会

更新日 : 2023-06-01

作成日 : 2021-01-18

目次

1. ISO/IEC 11179について
2. Part 4: Formulation of data definitions について
3. 前提
4. データ定義の要件
5. データ定義の推奨事項

1. ISO/IEC 11179について

- ◆ ISO/IEC11179 (Information technology – Metadata registries : MDR) は、様々な組織で作成、公開、交換、保存されるデータ項目 (data element) について、それに対する説明記述 (descriptions of data) であるメタデータ (Metadata) の構造を規定し、それに基づくインスタンス (個別の具体的なメタデータ) をメタデータ・レジストリ (Metadata Registry) に登録する仕組みを定めた標準

1. ISO/IEC 11179について

- ◆ 標準規格の名称となっているメタデータ・レジストリ（Metadata Registry）は、流通するデータ項目（data element）の特性あるいは属性を定義したデータ定義（メタデータ）を登録する仕組みであり、そこに様々なデータセットの定義情報が登録されることにより、データの理解、再利用、調和、標準化などを行うことが可能になる。
- ◆ ISO/IEC11179は、以下の6つのパートから成るマルチパート標準
 - Part 1: Framework
 - Part 2: Classification
 - Part 3: Registry metamodel and basic attributes
 - Part 4: Formulation of data definitions
 - Part 5: Naming principles
 - Part 6: Registration

2. Part 4: Formulation of data definitions について

- ◆ データ処理と電子データ交換は、データベースに記録された正確で信頼性が高く、制御可能で検証可能なデータに大きく依存する。
- ◆ データを正しく適切に使用すること、および解釈するための前提条件は、データのユーザーと所有者の両方が、データの意味と表現について共通の理解を持っていること。この共通の理解を促進するには、データのいくつかの特性または属性を定義する必要がある。このようなデータの特性は、「メタデータ」、つまり「データを説明するデータ」と呼ばれる。

2. Part 4: Formulation of data definitions について

- ◆ ISO/IEC 11179 Part4 Formulation of data definitions では、メタデータレジストリで指定されるデータ定義の定式化に関する要件と推奨事項を指定している。これらを定義する目的は、データの意味を指定、記述、説明を明確にし、データ要素の標準化または再利用を促進し、データの共有と情報システムの統合を促進すること。
- ◆ メタデータレジストリの構造は、概念的なデータモデルの形式で指定される。メタデータレジストリは、データ要素および関連する概念（「データ要素の概念」、「概念ドメイン」、「値ドメイン」など）に関する情報を保持するために使用される。一般的に、これらはすべて「メタデータ項目」と呼ばれます。このようなメタデータは、データを明確に記述、記録、分析、分類、管理するために必要となる。

※出典：International Organization for Standardization

(<https://www.iso.org/obp/ui/#iso:std:iso-iec:11179:-4:ed-2:v1:en>)

2. Part 4: Formulation of data definitions について

- ◆ ISO/IEC 11179のこの部分で指定されている定義要件と推奨事項は、用語集や言語辞書にある用語の定義に常に適用されるわけではない。言語辞書に適用される要件とメタデータレジストリに適用される要件の間には違いがある。
- ◆ ISO/IEC 11179の要件は、自然言語辞書の要件よりも制限されている。たとえば、言語辞書には、用語または単語の複数の意味をカバーする複数の定義が含まれる場合があるが、データ定義は特定のコンテキスト向けに開発されており、どのコンテキスト内でも複数の意味を持つべきではない。データ定義は、データのコレクション、データ値、データ要素、またはその他のメタデータ項目によって表される概念を説明することを目的としている。

2. Part 4: Formulation of data definitions について

- ◆ 単一の定義を参照定義として確立し、他の言語での定義を同等であると主張することができる。たとえば、ある言語の定義を参照定義として確立し、他の言語の定義を同等と主張することができる。
- ◆ メタデータ項目は、他の非推奨の定義とともに、特定のコンテキスト内で単一の優先定義を持つ場合がある。
- ◆ 多くのデータ定義には、それ自体を定義する必要のある用語が含まれています（たとえば、「料金」、「手当」、「配達」）。これらの用語のいくつかは、異なる産業部門で異なる定義を持っている可能性がある。したがって、ほとんどのメタデータレジストリは、定義で使用される用語の関連する用語集または用語参照を確立する必要がある。

※出典：International Organization for Standardization

(<https://www.iso.org/obp/ui/#iso:std:iso-iec:11179:-4:ed-2:v1:en>)

3. 前提

- ◆ データは特定の目的に使用される。
- ◆ 使用法の違いにより、いくつかの要件と推奨事項の異なる運用上の兆候が必要になる。たとえば、データ定義には通常、さまざまなコンテキストでさまざまなレベルの特異性が必要。「5. データ定義の推奨事項」a) では、さまざまな定義に対してさまざまなレベルの特異性が必要な例を示している。さまざまなコンテキストでの概念間の関係、および概念を区別するために特性がどのように使用されるかについては、ISO 704の5節を参照する。定義は、すべてのユーザーおよび共有データの受信者が理解しやすいように作成する必要がある。

【ISO 704-用語集原則と方法の5節】用語作業の概念の性質 コミュニケーションでは、すべての個々のオブジェクトが区別され、名前が付けられるわけではありません。代わりに、観察と概念化と呼ばれる抽象化のプロセスにより、オブジェクトはクラスに分類されます。さまざまな形式のコミュニケーションで表現される、概念と呼ばれる知識の単位に対応します。（オブジェクト→コンセプト→コミュニケーション）。この国際規格は、表現されているすべての概念を扱っているわけではありません。ただし、専門分野の用語で表される言語のみを使用します。用語の作業については、概念は、特殊な文脈または分野内のオブジェクトの精神的表現とみなされるものとします。概念を抽象的または想像上のオブジェクト（つまり、具体的なオブジェクト、抽象的なオブジェクト、または想像上のオブジェクト）と混同しないでください。与えられた文脈の中で観察され、心理的に概念化され、その後、指定はオブジェクト自体ではなくコンセプト）。オブジェクトとそれに対応する名称の間のリンクまたは定義は、より高いレベルの抽象化である概念を通じて行われます。

4. データ定義の要件

- ◆ 整形式のデータ定義を構築するための要件の理解を容易にするために、説明と例を示します。

a) 単数形で記述される説明

データ定義によって表現される概念は、単数形で表現されるものとする。（概念自体が複数形の場合は例外となる。ただし、日本語では区別することが困難）

例 「記事番号」

- 1) 適切な定義： A reference number that identifies an article. ≡記事を識別する参照番号(ja)
- 2) 不十分な定義： Reference number identifying articles. ≡記事を識別する参照番号(ja)

不十分な定義である理由

不十分な定義(English)では、「article number」が複数の記事を参照している可能性があるため、あいまいな複数形の「articles」を使用している。

4. データ定義の要件

- b) 概念が何であるかを述べる、それが何でないかだけでない説明
定義を構築するとき、概念が何でないかを述べることによつてのみ概念を定義することはできない。

例 「運賃額」

- 1) 適切な定義：ある場所から別の場所に商品を移動する際に荷送人が負担する費用額
- 2) 不十分な定義：梱包、文書化、積み込み、荷降ろし、および保険に関連しない費用

不十分な定義である理由

不十分な定義では、データの意味に何が含まれるかを指定していない。

4. データ定義の要件

- c) 説明的なフレーズまたは文として記述される（ほとんどの言語で）説明
ほとんどの言語においてフレーズは必要。概念の本質的な特徴を含む正確な定義を形成する。
1つ以上の同義語を述べることや、名前の単語を別の順序で単に言い換えるだけでは不十分なので、説明的なフレーズ以上のものが必要な場合は、完全で文法的に正しい文を使用する。

例 「エージェント名」

- 1) 適切な定義：他の当事者に代わって行動する権限を与えられた当事者の名前
- 2) 不十分な定義：代表者

不十分な定義である理由

不十分な定義「代表者」はデータ要素名の同義語に近いものであり、定義には適切ではない。

4. データ定義の要件

d) 一般的に理解されている略語のみを含む説明

頭字語や頭字語を含む略語の意味を理解することは、通常、特定の環境に限定される。

同じ略語であっても環境が異なると、誤解や混乱を引き起こす可能性がある。したがって、あいまいさを避けるために、定義では**略語ではなく完全な単語を使用する必要がある**。「ie」や「eg」などの略語が一般的に理解されている場合は、この要件の例外が発生する可能性がある。または、略語が複雑な用語の完全な形式よりも容易に理解され、「レーダーの検出と測距」を表す「レーダー」などの用語として採用されている場合。すべての頭字語は、最初に出現したときに展開する必要がある。

例 「潮の高さ」

- 1) 適切な定義：平均海面（MSL）から特定の潮位までの垂直距離
- 2) 不十分な定義：MSLから特定の潮位までの垂直距離

不十分な定義である理由

頭字語の「MSL」は一般的に理解されておらず、一部のユーザーはそれが何を表しているかを判断するために他のソースを参照する必要があるため、不明確です。

完全な単語がないと、用語集でその用語を見つけるのは難しいか不可能となる。

4. データ定義の要件

e) 他のデータまたは基礎となる概念の定義を埋め込まずに表現する説明

次の例に示すように、2番目のデータ要素または関連する概念の定義は、1次データ要素の適切な定義に表示されるべきではない。

用語の定義は、関連する用語集で提供する必要がある。

2番目の定義が必要な場合は、1番目の定義の最後に注記を付けることができる。

関連する定義には、リレーショナル属性（相互参照など）を介してアクセスできる。

例 「銀行の信用書類番号の発行」

- 1) 適切な定義：発行銀行によって荷為替クレジットに割り当てられる参照番号
- 2) 不十分な定義：発行銀行によって信用書類に割り当てられた参照番号。

信用書類とは、銀行が信用書類を発行した旨を記載した文書であり、これに基づいて受益者は、特定の契約条件の遵守、および規定された文書やそのような手形の提示に対して、支払い、承諾、または交渉を得ることができます。

不十分な定義である理由

不適切な定義には、概念定義が含まれており、用語集に含める必要がある。

5. データ定義の推奨事項

a) 概念の本質的な意味を述べる説明

表現される概念のすべての主要な特性は、コンテキストに関連する特定のレベルで表示される必要がある。本質的でない特性を含めることは避けるべき。

必要な詳細レベルは、システムユーザーと環境のニーズによって異なる。

例 1 「貨物積載シーケンス番号」（意図されたコンテキスト：あらゆる形態の輸送）

- 1) 適切な定義：輸送手段または輸送機器の一部に貨物が積載されるシーケンスを示す番号
- 2) 不十分な定義：貨物がトラックに積み込まれる順序を示す番号

不十分な定義である理由

意図された状況では、貨物はさまざまな輸送モード（トラック、船舶、貨物列車など）で輸送できる。貨物は輸送用のトラックに限定されないため。

例 2 「請求額」

- 1) 適切な定義：請求書に請求される合計額
- 2) 不十分な定義：請求書に記載されているすべての請求可能なアイテムの合計。一方で、手当や割引などの控除、および他方で、保険、輸送、取り扱いなどの料金などの追加を考慮に入れる。

不十分な定義である理由

不十分な定義には、無関係な資料が含まれる

5. データ定義の推奨事項

b) 正確かつ明確な説明

定義された概念の正確な意味と解釈は、定義から明らかになる必要がある。定義は、1つの可能な解釈のみを許可するのに十分明確でなければならない。

例 「出荷受領日」

- 1) 適切な定義：受領者が出荷を受領した日付
- 2) 不十分な定義：特定の貨物が配達される日付

不十分な定義である理由

不十分な定義では、「配信」を決定するものが指定されていません。

「配達」は、意図された目的地で製品を降ろす行為、または意図された顧客が実際に製品を入手する時点のいずれかとして理解することができる。

対象の顧客が自分のサイトで荷降ろしされた製品を受け取らない場合や、顧客が現場で荷降ろしされてから数日後に製品を受け取る可能性も含まれている。

5. データ定義の推奨事項

c) 簡潔な説明

定義は簡潔で包括的でなければならない。「このメタデータレジストリの目的のために」や「説明される用語」などの無関係な修飾句は避けなければならない。

例 「文字セット名」

- 1) 適切な定義：データがエンコードされる表意文字または表意文字の記号のセットに付けられた名前
- 2) 不十分な定義：このメタデータレジストリの目的で、または他の場所で使用されているように、1つ以上のスクリプトでエンコードされたデータを処理するシステムハードウェアおよびソフトウェアの機能のために、データがエンコードされる音声記号または表意文字記号のセットに付けられた名前

不十分な定義である理由

不十分な定義では、「...データがエンコードされる」の後に、「文字セット名」の例示が書かれており、定義の説明そのものに無関係な修飾がある。

5. データ定義の推奨事項

d) 独立できる説明

概念の意味は、定義から明らかになるはず。定義の意味を理解するために、追加の説明や参照は必要ない。

例 「学校の場所都市名」

- 1) 適切な定義：学校が置かれている都市の名前
- 2) 不十分な定義：「学校のサイト」を参照してください。

不十分な定義である理由

貧弱な定義は独立した説明となりません。最初の定義の意味を理解する為には、2番目の定義（学校のサイト）の助けが必要となります。

5. データ定義の推奨事項

e) 理論的根拠、機能的用法、ドメイン情報、または手続き型情報を埋め込まずに表現する説明

上記のような説明は必要ですが、定義に無関係な情報を含んでいるため、適切な定義に属しません。有用であると見なされる場合、そのような式は他のメタデータ属性に配置される場合があります（ISO / IEC 11179-3を参照）。ただし、定義の後に例を追加することは許可されています。

- 1) 特定の定義の論理的根拠を定義の一部として含めるべきではない。たとえば、データ要素がキロメートルではなくマイルを使用する場合、その理由を定義に示すべきではない。
- 2) 次のような機能的な用法：「このデータ要素は...に使用しないでください。」のような文章を適切な定義に含めるべきではありません。
- 3) 手続き面についての意見

「このデータ要素はデータ要素 'xxx' と組み合わせて使用されます」のような記述は、定義に表示されるべきではない。代わりに、ISO / IEC 11179-3で指定されている「関連データ参照」と「関係のタイプ」を使用して記述する。

5. データ定義の推奨事項

例 「データフィールドラベル」

- 1) 適切な定義：インデックス、シソーラス、クエリ、データベースなどのフィールドの識別
- 2) 不十分な定義：インデックス、シソーラス、クエリ、データベースなどのフィールドの識別、これは、要約、テーブル内の列などの情報の単位に提供されます。

不十分な定義である理由

不十分な定義には、機能の使用法に関するコメントが含まれている。「これは...に提供されます」で始まるこの情報は、必要な情報である場合、定義から除外し、別の属性に配置する必要があります。

5. データ定義の推奨事項

f) 循環論法の回避説明

2つの定義を相互に定義してはならない。定義は、その定義として別の概念の定義を使用するべきではない。相互に定義することにより、概念が別の概念の助けを借りて定義され、次に、与えられた概念の助けを借りて定義される状況が発生する。

例 定義が不十分な2つのデータ要素

- 1) 「従業員ID番号」
従業員に割り当てられた番号
- 2) 「従業員」
従業員ID番号に対応する人

十分な定義である理由

各定義は、その意味について他の定義を参照している。
どちらの定義にも意味はない。

5. データ定義の推奨事項

- g) 関連する定義に同じ用語と一貫した論理構造を使用する説明
類似または関連する定義には、共通の用語と構文を使用する必要がある。

例 どちらの定義も同じ論理構造と同様の用語を使用している例

1) 「商品発送日」

特定の当事者によって商品が発送された日付

2) 「在庫日」

特定の当事者が商品を受け取った日付

十分な定義である理由

同じ用語と構文を使用すると、理解が容易になる。

そうでなければ、ユーザーは、同義語と変数構文の使用によって何らかの違いが暗示されるかどうか疑問に思う。

5. データ定義の推奨事項

h) 定義されているメタデータアイテムのタイプに適している説明

メタデータレジストリ内のさまざまなタイプのメタデータアイテム（データ要素の概念、データ要素、概念ドメイン、値ドメインなど）はそれぞれ異なる役割を果たす定義。

