

# AI時代の資料データベース

石井 淳平 ( 厚沢部町教育委員会 )

## Database of Museum Inventory Records in the AI Era

Ishii Junpei ( Assabu Board of Education )

### 活用手法

資料台帳 / アクセス可能性 / デジタルアーカイブ / インベントリ

/ Tidy Data / リレーショナルデータベース / AI

Inventory Records / Accessibility / Digital Archive / Inventory /

Tidy Data / Relational Database / AI

博物館の「保管」機能の本質は、資料へのアクセス可能性の確保にある。資料台帳の作成と公開はその一環であり、近年のデジタル技術の発展により、電子化が最も効率的な手段となっている。しかし、2019年時点で資料台帳の整備率は50%未満にとどまり、電子化が進む館と未整備の館との二極化が進行している。本稿では、持続可能な台帳管理のため、表計算ソフトの活用とTidy Dataの原則に基づくデータ整理を提案する。また、Notionのようなクラウド型データベースを用いることで、AIを活用した検索機能や階層的な情報整理を可能にし、資料管理の効率化を図る手法について考察する。博物館法の観点からも、資料台帳の電子公開は博物館の義務として位置付けられるべきである。

## はじめに

博物館の目的である「保管」の意義は、博物館資料へのアクセス可能性を担保することである。その一環として、「目録の頒布」、すなわち資料台帳の作成と公開が博物館の重要な事業となる。近年のウェブ技術の発展やコストの低下により、目録の頒布は電子的に行うことが最も効率的である。この取り組みは、令和4年4月公布の改正博物館法に

より新たに法制化された「デジタルアーカイブ」以前から、博物館が果たすべき役割であった。

しかし、2019年時点で博物館における資料台帳の完備率は50%に達しておらず、約15%の館では台帳が存在すらしていないことが明らかになっている。一方で、電子化された資料台帳を持つ館は着実に増加しており、資料台帳整備の面で博物館の二極化が進行している。

資料台帳作成を進める上で筆者が強調したいのは、インベントリとカタログを区別し、ミニマムな台帳を目指すこと、そして、周到に設計されたデータベースソフトではなく、表計算ソフトを利用することの重要性である。小規模な博物館では、学芸職員の持続的な任用が保証されず、一般事務職員でも資料台帳を引き継いで運用できる体制が求められる。そのため、表計算ソフトを活用し、適切なデータ構造で保存することで、リレーショナルデータベースの機能を部分的に活用しつつ、将来的に本格的なデータベース環境へ移行することも可能となる。当面の次善の策として、表計算ソフトの活用を推奨する。

その際、注意すべき点は機械可読性への配慮である。機械可読なデータの典型であるTidy Dataを作成するためには、いくつかの注意点があるが、最も重要なのは観測項目と観測値を混在させないことである。また、セルの結合などの整飾要素を混在させないことも不可欠である。電子化において最優先すべきは機械の可読性であり、人間の視認性を優先すべきではないことを強調したい。

## 博物館法における資料台帳の位置づけ

### 保管の手段としての目録作成

博物館の目的は、博物館法第2条（註1）に定められているとおり、資料の「収集」、「保管」、「展示」、「調査研究」を行うことである。博物館の各事業や業務は、これらの目的を達成するための手段として位置づけられている。

博物館の事業は、法第3条の各号に定められており、資料台帳に関わる事項は第6号「博物館資料に関する案内書、解説書、目録、図録、年報、調査研究の報告書等を作成し、及び頒布すること」である。このうち、「目録の頒布」は、博物館資料台帳の公開を行うべき法的根拠として位置づけられる。

### 博物館における「保管」の達成とは何か

博物館における「保管」は、博物館の目的の一つである。「保管」の目的を達成するために、博物館では資料劣化の予防措置やクリーニング、補修などが行われている。これらの作業が目指すのは、博物館資料へのアクセス可能性の確保である。

博物館資料の「保管」において最悪の事態は、資料の「紛失」や回復不能な「破損」である。これらの事態が最悪なのは、博物館資料へのアクセス可能性が永遠に失われるためである。したがって、博物館における「保管」とは、アクセス可能性の確保と同義である。UNESCOの勧告が示すように、資料目録の作成は「収蔵品管理の主要な構成要素とされ、そのことが『ミュージアムの保護』につながる」（註2）とされている。

アクセス可能性が確保されない場合、博物館の「保管」機能は失われる。資料が取り出せない、あるいは資料の有無が分からないといった状況は、博物館資料へのアクセスが失われた状態である。さらに、資料そのものが現存していたとしても、採取地、寄贈者、使用年代、利用状況など、資料そのものから導き出せない情報が喪失した場合も、博物

館の「保管」機能が果たされたとは言えない。したがって、博物館における「保管」とは、単に実物資料へのアクセスだけでなく、それに付随するデータの一式を含むものと解される。資料台帳は、博物館の目的である「保管」を達成する手段であり、「目録の頒布」によって広く公開されることで、その「保管」機能をさらに強化するものである。

しかしながら、佐々木秀彦（2002）が指摘するように、「利用者が実物資料の閲覧を含めて全収蔵資料にアクセスできるという考え方は必ずしも確立していない」のが現状である。このことは、博物館の「保管」機能を高めるためには、収蔵資料へのアクセス可能性を拡大することが不可欠であるという目的と手段の理解が十分に浸透していないことにも起因していると考えられる。

## 脱ドメスティックマッチング

博物館資料台帳が公開されていない状況は、資料がドメスティックな（知る人ぞ知る）マッチングにとどまっていることを意味する。つまり、博物館資料と、それを活用できる人が出会えないという問題が生じている。この状況が、我が国の社会教育に負の影響を与えることは明白である。

学習活動において、資料の探索コストはしばしば軽視されがちだが、それが許容できる水準を超えた場合、学習活動からのドロップアウトを引き起こすリスク要因の一つとなる。博物館法第1条が定める博物館の目的や、UNESCOの「ミュージアムとコレクションの保存活用、その多様性と社会における役割に関する勧告」（註3）に示される博物館の社会的な役割を考慮するならば、探索コストの最小化が博物館の重要な使命の一つであることは明白である。

こうした社会的利益を損なう探索コストの低減は、博物館のみならず、産業分野においても公共政策による解決の正当性が指摘されている（註4）。したがって、博物館資料の探索コストの低減は、法制化された「デジタルアーカイブ」以前から、博物館が果たすべき重要な業務であったといえる。

## 日本博物館協会調査による資料台帳の状況

### 資料台帳の整備状況

図1は、日本博物館協会の調査（公益財団法人日本博物館協会 2020: pp. 57-58）（以下「協会調査」）による、博物館の資料台帳の整備状況を示したものである。最新の2019年調査では、資料台帳の整備が「ほとんど全て」と回答した博物館は44.8%であった。この割合は1997年以降、ほとんど変動していない。また、2013年調査から加わった「台帳なし」の回答は、2019年調査において15.3%を占めている（註5）。

すなわち、博物館に当然備えられているべき資料台帳が整備されている館は半数に満たず、1割以上の館では台帳が存在すらしていないことが明らかとなった。これにより、博物館における台帳整備の進捗は極めて遅く、場合によっては後退している可能性すらある。

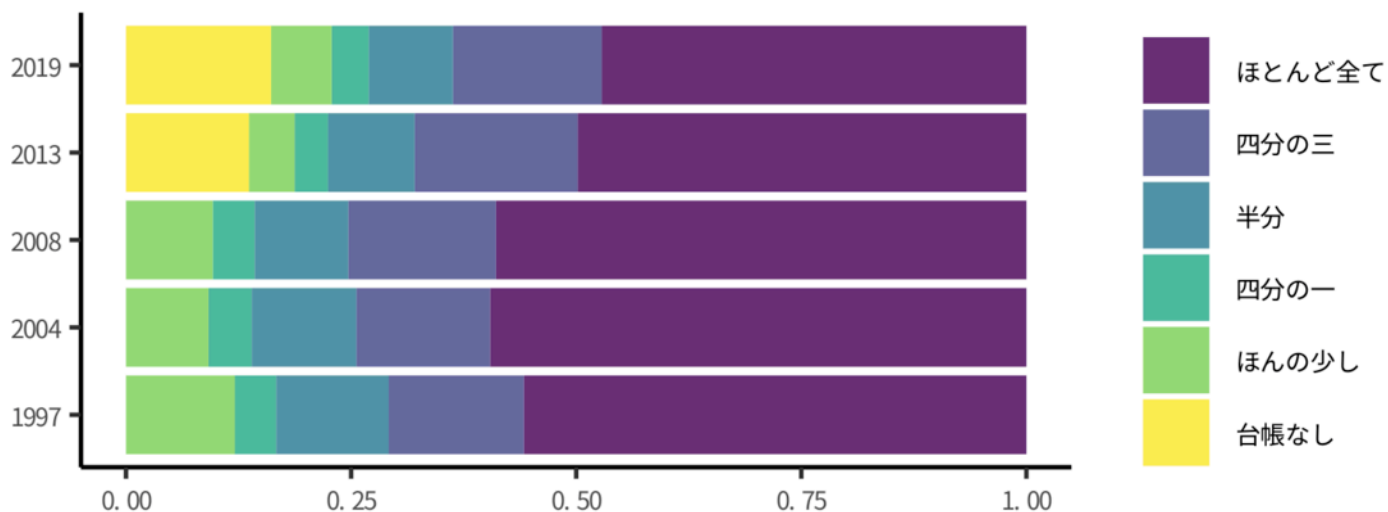


図1 資料台帳記載の収蔵資料の割合

## デジタル化された資料台帳

電子メディアでデータベース化された資料台帳を持つ博物館の割合は、1997年の20.3%から2019年には51.1%へと増加している（図2）。また、データベース化された収蔵資料の割合に関しても、「ほとんど全て」および「四分之三」と回答した館の合計は、1997年の44.5%から2019年には69.7%へと増加している（図3）。一方で、資料台帳の作成自体は停滞しているものの、電子化は着実に進んでいる。

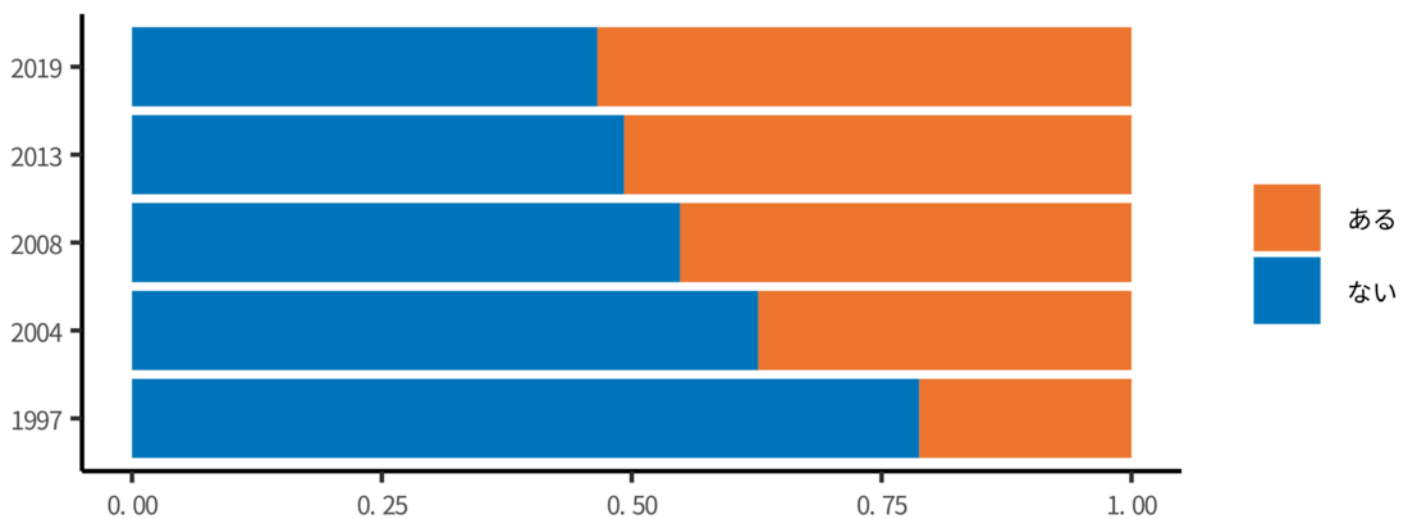


図2 電子メディアにデータベース化された資料台帳

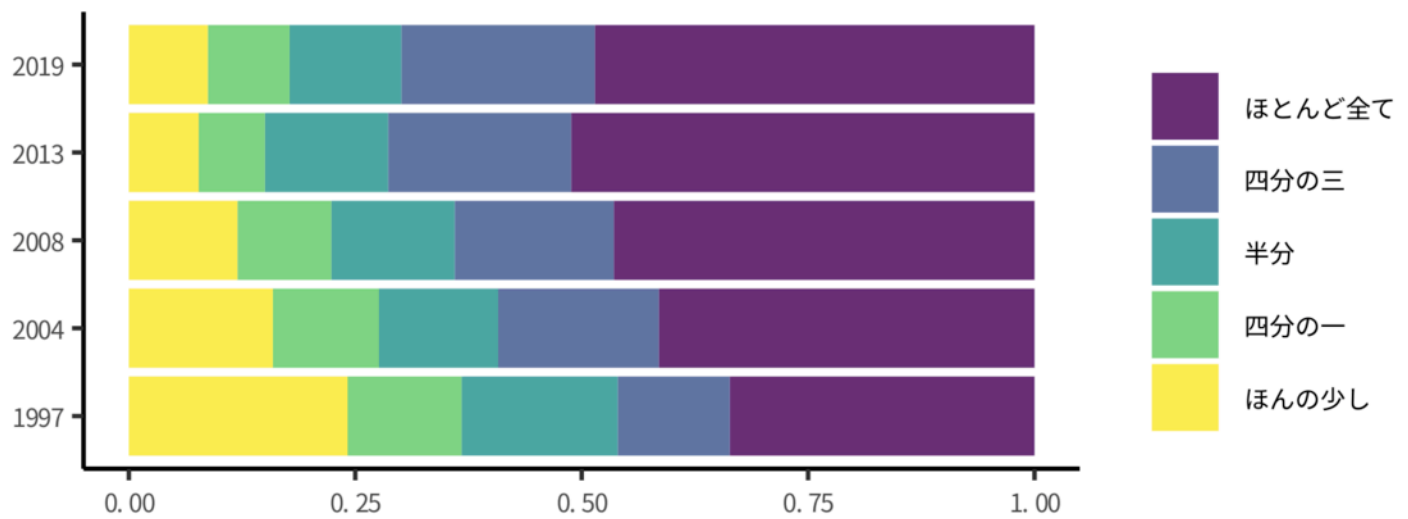


図3 電子メディアにデータベース化された資料の割合

## 二極化する博物館の台帳整備

協会調査から、博物館の基本機能ともいえる資料台帳の整備が極めて低調であり、改善の傾向が見られないことが明らかとなった。一方で、資料台帳のデジタル化率は着実に上昇している。この2つの事実から導き出されるのは、既存の台帳が十分に整備されている博物館では近年、資料台帳のデジタル化が進んでいる一方で、資料台帳が未整備の館では、デジタル化はもちろん、既存台帳の登録業務すら滞っているという現状である。

以上のように、資料台帳のデジタル化を進める博物館が増加する一方、資料台帳そのものの整備が進まない館が過半を占めており、整備状況の二極化が進んでいることがうかがえる。

## 博物館資料とリレーショナルデータベース

### 資料とメタデータをつなぐ資料台帳の役割

資料と、それにまつわる情報の一式をコンピュータを利用して管理・連携する仕組みがデータベース（註6）である。資料台帳は、資料とそのメタデータを結びつける役割を担う。「登録番号」をキーコードとして用いることで、実物資料とメタデータを関連付けることができる。

資料台帳は、必ずしも一つの表（テーブル）で構成される必要はない。「資料名」「受入年月日」「寄贈者」など、資料と一対一で対応する性質のデータは、シンプルな一つの表に格納できる。しかし、博物館資料のメタデータには、借用情報のように、一つの資料に対して複数の履歴を持つものもある。また、写真のように表形式になじまないデータや、専門家による調査記録のように、単なるデータ（註7）とは呼べない形式の情報も、資料に付随するメタデータに含まれる。こうした多様なデータや情報をキーコードを媒介として結びつけ、コンピュータの処理能力を活用して統合的に管理するのが、データベースの役割である。

## リレーショナルデータベースとは何か

リレーショナルデータベースとは、複数のテーブル（表）を組み合わせで構成されるデータベースである。1つのテーブルで表現できる情報には限界があるため、1対多の関係を持つデータ群を効率的に管理するために考案されたデータベースの形式である。

1対多の関係とは、例えば、博物館のデータとその館に所属する学芸員のデータ、または資料のデータと寄贈者のデータの関係などが挙げられる。博物館には複数の学芸員が所属することもあれば、一人の学芸員が複数の館を所管することもあり得る。このようなデータ群を1つのテーブルにまとめることは難しく、完全な表現はできない。そのため、複数のテーブルを用意し、目的に応じて必要な属性を結びつけ、適切な表や帳票を新たに作成することが求められる。これがリレーショナルデータベースの基本的な役割である。

Microsoft AccessやFileMakerのような汎用のデータベースソフトウェアも、リレーショナルデータベースに分類される。一方で、データが比較的単純であり、複雑なテーブル構造を必要としない場合には、リレーショナルデータベースソフトの導入が必ずしも最適とは限らない。そのため、運用するデータの特性に応じて、リレーショナルデータベースが本当に必要かどうかを慎重に検討することが重要である。

## 埋蔵文化財資料の博物館受納とリレーショナルデータベース

リレーショナルデータベースが有効な事例として、博物館資料台帳における埋蔵文化財の登録事務を取り上げる。埋蔵文化財は、発掘調査に伴う整理作業で作成された各種の台帳が存在する状態で受納される。代表的な台帳として、「取り上げ台帳」「整理台帳」「掲載台帳」「収納台帳」などが挙げられる。

「取り上げ台帳」は、遺物の取り上げ単位を基本とし、出土日やグリッドなどの最小限のデータが記録される。この台帳では、1単位が数百点となることもあれば、地点計測を行った遺物では1点が1単位となる場合もある。

「整理台帳」は、「取り上げ台帳」を基に、土器や石器などの種別、器種、分類といった属性情報が追加される。「整理台帳」は「取り上げ台帳」の層位情報などに遺物の属性を加えたものであり、発掘調査における遺物情報を網羅した台帳である。これは文化庁が定める「台帳登録」にあたる（文化庁 2010: pp. 15-16）。「整理台帳」の情報を基に、口径や器高、器面調整といった個体の情報が付加された分析用の台帳や、掲載遺物一覧表などが作成されることも多い。

「収納台帳」は、収納時の荷姿を基礎として作成される。土器の復元個体であれば、梱包されたダンボールなどが1単位となり、土器片や石器の場合は複数の遺物をまとめてコンテナ等に収納するため、コンテナが1単位となる。

これらの台帳のうち、どの台帳を博物館資料の基礎単位とするべきか、また、付随する多くの台帳を資料情報としてどのように活用するかは、担当者の判断を要する課題となる。

実物へのアクセスを考慮すると、「収納台帳」の1単位を博物館資料台帳の1単位とするのが利便性の面で優れていると考えられる。「収納台帳」は荷姿ごとに登録されるため、ダンボールやコンテナ単位で博物館の資料番号を付与すれば、整理が容易になる。一方で、豊富な情報を有する「整理台帳」や「掲載台帳」へのアクセスが、この時点で失われる可能性があることから、「整理台帳」や「掲載台帳」を基に登録する方法も考えられる。しかし、その場合、「整理台帳」単位で資料番号を再付番する手間や、資料番号のほかに「コンテナ番号」の管理が必要となるなど、作業の煩雑さが課題となる。

このような管理の二重化は、「荷姿」と個別の資料が入れ子状の構造を持つことから生じる。こうした1対多の関係にある情報を、目的に応じて適切に呼び出し、表や帳票を生成・管理できる点が、リレーショナルデータベースの優れた機能である（図4）。

博物館資料台帳

登録番号	資料名	……
000001	〇〇年△△遺跡出土遺物	
000002	〇〇年△△遺跡出土遺物	
000003	〇〇年△△遺跡出土遺物	
……		

コンテナ単位で登録

発掘調査整理台帳

整理番号	グリッド	層位	種別	登録番号
001	A-01	3層	土器	000001
002	A-01	3層	石器	000001
003	A-02	2a層	土器	000001
……				

取り上げ単位で登録

## 〇〇年△△遺跡出土遺物収納表

登録番号：000001

収納場所：〇〇小学校

整理番号：001

グリッド：A-01

層位：3層

種別：土器

整理番号：002

グリッド：A-01

層位：3層

種別：石器

整理番号：003

グリッド：A-02

層位：2a層

種別：土器

整理番号：004

グリッド：A-02

整理番号：005

グリッド：A-02

整理番号：006

グリッド：A-03

図4 リレーショナル・データベースを利用した発掘調査資料と博物館資料台帳の統合

## ミニマムな博物館資料台帳

## 資料の同定と管理に必要な属性

博物館資料は多様性に富み、図書資料のようにタイトルや著者名などの定型的な属性が一律に付与されることは少ない。「資料名」ですら、外部から与えられるのではなく、受け入れ時に博物館側が付与するものである（村田2016）。

村田良二（2016 前掲）は、博物館資料の記録すべき情報を4つのカテゴリーに分類している。このうち、博物館資料台帳に必須となるのは「識別・同定」のカテゴリーである。これは、「番号や名称、分類といった情報」を指し、数多くの資料の中から当該資料を特定するために必要な属性である。具体的には、「登録番号」と「資料名」が基本となる。さらに、資料受け入れ時にのみ取得可能な一次情報として、「寄贈者」や「寄贈年月日」がある。加えて、資料の

調査や受入時の聞き取りによって明らかになった事柄を簡潔に記した「所見」があれば、資料の同定と管理に必要な情報が揃うことになる。すなわち、「登録番号」「資料名」「寄贈者」「収蔵年月日」「所見」が、ミニマムな博物館資料台帳の必須項目といえる。

なお、「〇〇家文書」のように資料が群をなす場合、台帳の階層化を行うことで、より適切な管理が可能となる。こうした目録の階層化手法として、国際アーカイブ評議会（ICA）のISAD(G) 2nd（国際標準記述規則）がある。東京大学史料編纂所では、このISAD(G) 2nd に準拠した階層化された史料目録を作成している（山田 2019）。

## カタログはあきらめてインベントリをめざせ！！

博物館資料のデータベース作成が進まない理由の一つとして、無意識のうちに帳票形式の精緻なカタログを目標としてしまうことが挙げられる。資料の写真撮影、分類、計測を行い、それらをデータベースに入力する作業には膨大な手間がかかる。前項で述べた「識別・同定」のカテゴリーで構成される台帳は、資料管理に必要な最低限の情報を網羅したものであり、「インベントリ」と呼ばれる（Mairesse, 2023）。インベントリの目的は、収蔵品の同一性の把握と管理上の保存を確実にすることであり、資料の網羅的な情報を格納する「カタログ」とは区別して考える必要がある。インベントリの作成は、博物館にとって最も基礎となる作業であり、確実に行われるべきものである。

この作業には、必ずしもデータベースソフトや資料管理システムは必要ない。極端な話、ペンとノートでも十分であり、表計算ソフトウェアがあればなお適切である。最も重要なのは同一性の保証であり、資料番号は資料に直接マジックで記入しても構わない。また、名称が分からない資料に対して、無理に名称を付与する必要はない。例えば、寄贈された剥製が鳥類だった場合、種名をインベントリに記入する必要はなく、「鳥剥製」のように、誰もが認識できる名称であれば十分である。

## 資料台帳の標準化は必要か？

博物館資料台帳に必要な項目は、館の性質、それに割ける人的リソース、収蔵資料の全体構成によって異なる。資料台帳データベースの流通という観点からは、「他のデータベースとの連携を前提としたデータ作成が必要である」

（犬木 2008）とされている。しかし、資料台帳の項目については、「博物館情報の社会的な共有のためには避けては通れない課題であるが、『標準化』『統一』といった発想では解決が困難であることが明らかになっている」（田島 2007）や、「博物館の資料は図書館のそれとは異なり、一点一点の性質が大きく異なり、また、記録すべきデータも異なるのが特徴である」（後藤 2013）と指摘されている。これらの見解からも、資料台帳の標準化にこだわりすぎることは適切ではないと考えられる。まずは自館の資料の状況に応じて、ミニマムな資料台帳の作成から着手するほうが、挫折のリスクを抑えられるだろう。

なお、博物館資料に関連するメタデータの付与については、以下のような指針やモデルが存在する。

- ・ 国際博物館会議（ICOM）の国際ドキュメンテーション委員会（CIDOC）による「博物館資料の最小限情報分類勧告」（MICMO）（Grant 1994）
- ・ 同じくCIDOCによる「博物館資料情報のための国際指針」（IGMOI）
- ・ 東京国立博物館による「ミュージアム資料情報統合化モデル」（博物館情報処理に関する調査研究プロジェクトチーム2005）
- ・ 山本泰則らが国立民族学博物館や国立歴史民俗博物館の資料情報をもとに提案した「博物館コアメタデータ」（山本ほか2009）

これらの指針やモデルを参考にしつつ、各博物館の実情に即したメタデータの運用を検討することが求められる。



# 表計算ソフトのVLOOKUP関数を利用したデータベース機能

リレーショナルデータベースを利用する最大のメリットは、キーコード（資料番号や遺物番号など）を媒介として、異なるテーブルを結合できる点にある。この機能に近いものとして、表計算ソフトでは「VLOOKUP」関数が提供されている。

表計算ソフトのVLOOKUP関数は、「参照すべき値」「比較すべき値」「引用する列」の3つを引数として使用する。図6の右端にある「登録番号」列に入力されたコードを、収蔵資料台帳から検索し、指定された列の値を取得・結合している。SQLにおける「LEFT JOIN」に似た動作をする関数である。

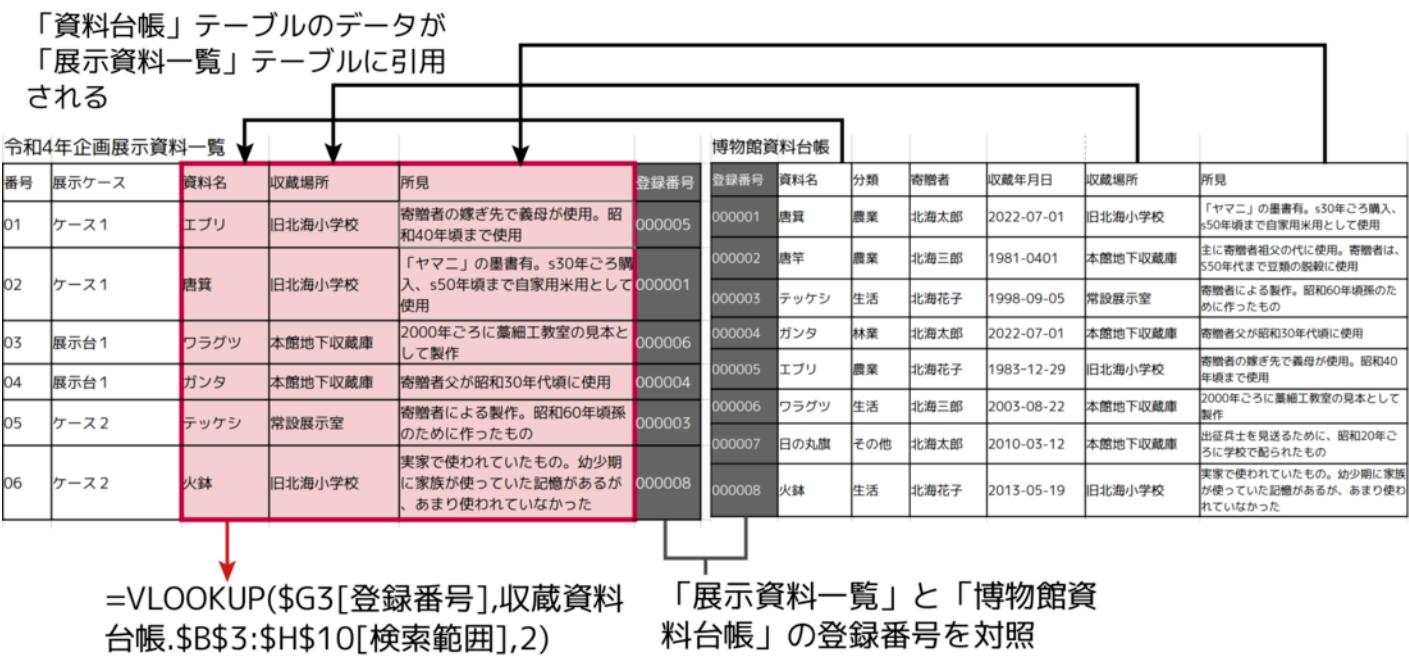


図5 VLOOKUP関数を利用した表計算ソフトによるテーブル結合

しかし、VLOOKUP関数によるJOIN系の機能にはいくつかの制約がある。例えば、キーコードが必ずテーブルの左端に配置されている必要がある点や、filter系の検索関数と組み合わせて使用する際に関数が煩雑になり、可読性が低下する点が挙げられる。また、表計算ソフトはもともとプログラミング言語として設計されていないため、データベースとしての運用には限界がある。

VBAを利用すれば、表計算ソフトをプログラミング環境として活用することも可能ではあるが、構築や維持の難易度が高く、特に理由がない限り、汎用のデータベースソフトウェアを利用する方が、データの持続可能性や運用の効率性の面で優れているだろう。

## Tidy Dataとは何か？

コンピュータでテーブル形式のデータを扱う場合、処理しやすい構造で作成することが重要である。人間が情報を読み取りやすい構造が、必ずしもコンピュータにとって処理しやすい形式とは限らない。むしろ、人間が視覚的に把握しやすい表形式のデータは、多くの場合、コンピュータによる処理には適していない。Hadley Wickhamは"Tiny Data"（註8）の概念を提示し、コンピュータが扱いやすいテーブル型のデータ形式について述べている（Wickham, H. 2014）。

図6の左側の表は、典型的なMessy Data（雑然データ）である。セルの結合は論外としても、列方向に観測値を並べるクロス集計風のテーブルは、人間には認知しやすく、手書きの集計表作成では作業効率上がる場合が多い。しかし、このように観測値が列方向に配置されたテーブルは、コンピュータが処理するには不向きである。一方、図6の右側の表は、Hadleyが提示したTiny Data（整然データ）の要件を満たしており、コンピュータにとって扱いやすい構造になっている。

### Messy Data

列方向に並ぶ観測値

〇〇遺跡発掘調査遺物整理台帳

番号	グリッド	層位	前期	中期	後期	晩期	...
001	A-01	3a層	10	15	4	3	
002			12	11	5	6	
003	A-02	2a層	28	10	2	1	
004		2c層	18	15	4	0	
005			12	17	6	1	
006	A-03	3b層	8	18	17	10	
...							

多用されるセル結合

### Tidy Data

〇〇遺跡発掘調査遺物整理台帳

番号	グリッド	層位	時期	点数	...
001	A-01	3a層	前期	10	
001	A-01	3a層	中期	15	
001	A-01	3a層	後期	4	
001	A-01	3a層	晩期	3	
002	A-02	3a層	前期	12	
002	A-02	3a層	中期	11	
002	A-02	3a層	後期	5	
002	A-02	3a層	晩期	6	
003	A-02	2a層	前期	28	
003	A-02	2a層	中期	10	
...					

Tidy Data 4つの条件のうち最重要2条件

変数 は**1の列**に入力する。

観測値は**1つの行**に入力する。

変数「**時期**」と変数「**点数**」はそれぞれ異なる**列**に格納する。

図6 Tiny Dataとは何か

2つの表の最大の違いは、データの構造にある。左側の表では、「グリッド」や「層位」のような観測項目と、「前期」や「中期」のような観測値が混在している。「前期」や「中期」が観測値であることは、人間にとっては自明であるが、コンピュータは明示的な指示がなければこの構造を理解できない。したがって、Tiny Dataを作成する際に最も注意を払うべき点は、観測項目と観測値を混在させないことである。

「グリッド」や「層位」は発掘時に観測すべき項目であるため、観測項目として列名に設定するのが適切である。一方、「A-01」や「3a層」などは発掘の結果得られたデータであり、観測値とするのが適切である。同様に、「前期」や「中期」は整理作業の結果得られた観測値であるため、列名に設定するのは不適切である。

## Notionを活用した博物館資料台帳

## Notionとは

Notionはクラウド上で利用できるメモアプリである。メモ、タスク管理、データベース、Wiki、プロジェクト管理などを統合できるオールインワンのクラウドツールであり、個人利用から企業まで幅広い用途に対応している。Notionの中核となる機能はデータベースであり、単なるメモであっても、データベースに一元化することができる。Notionに取り込めるコンテンツはテキスト以外に画像や動画があり、また、ウェブコンテンツの埋め込みにも対応している。無料で利用することができ、ウェブブラウザ、PC、スマートフォンなどマルチプラットフォームに対応している。

## Notionの実際

Notionへのデータ移行にはCSVアップロードが利用できる。既存の資料台帳をCSV形式で書き出し、Notionにインポートするだけでデータベースを構築できる。図7は、Notionのテーブルビューで表示した資料台帳である。データベースのフィールドにあたるものは「プロパティ」という名称で管理されている。図7の例では、プロパティは資料台帳としての最低限の情報で構成するようにしており（註11）、インベントリ情報をもっぱらここに格納されている。

 所蔵資料データベース\_230512

Show All

ギャラリービュー

+

○ ≡ ↶ ⚡ 🔍 ...

新規

≡ 分類番号

Aa 資料名

+

フィルターを追加

Aa 資料名	≡ 分類番号	≡ 校番	≡ 寄贈年月日	≡ 寄贈者住所	≡ 寄贈者氏名
郷土資料館展示原稿	1110	1			
 館城出土和鏡	4000	1	2024/08/05		
 アツシ	5130	1	2013/04/08		
袖無	6130	1	1982/10/08	当路	
前だれ	6130	2	1982/10/08	当路	
半袖	6130	3	1982/10/08	美和	
前だれ	6130	4	1982/10/08	美和	
軍服上着	6130	5	1987/04/24	松園町	
軍服ズボン	6130	6	1987/04/24	松園町	
軍服ベルト	6130	7	1987/04/24	松園町	
軍服帽子	6130	8	1987/04/24	松園町	

図7 Notionのデータベース画面

## 高い拡張性

Notionの優れた点は、データベース内で情報を無限に拡張できることである。図8は、データベースに紐付いた個別ページを示しており、この部分には無限に情報を記入できる。

所蔵資料データベース	
Show All	...
分類番号	分類番号 Aa 資料名
郷土資料館展示原稿	1110
館城出土和鏡	4000
館城出土和鏡	4000
アツシ	5130
袖無	6130
前だれ	6130
半袖	6130
前だれ	6130
軍服上着	6130
軍服スボン	6130
軍服ベルト	6130
軍服帽子	6130
二重まわし	6130
二重まわし	6130
はんでん	6130
テッケシ	6130
上着ホームスパン	6130
はんでん	6130
袖無	6130
軍靴	6140
カンジキ	6140
カンジキ	6140
金カンジキ	6140
金カンジキ	6140
カンジキ	6140
ござ打ち下駄	6140
ござ打ち下駄	6140

## 覚書

水辺に牡丹花が描かれる図案で、上部にはひし形の窓を設けて中に鳶の図象が3つ描かれる。家紋でいう「三つ盛り鳶」に近い。

ともに寄贈された4000-2は直径22.5cm、本資料は直径17.5cmでかなり大きさに違いがある。江戸時代には島田髷のような大きなつくりの髪型が生まれ、そのために鏡も大型化したといわれる。大小の鏡をセットで用い、大きい鏡を顔の前に、小さい鏡を後頭部にもっていき、合わせ鏡のようにして髪型を整えた。

銘は「藤原義勝」とある。同時に寄贈された4000-2には「天下一 藤原吉長」の銘がある。江戸時代前期には「天下一」の銘が乱造されたため、幕府がこれを禁止するほどで、江戸時代後期には規制がゆるみ、「天下一」が復活する。よって、この鏡も江戸時代後期の作の可能性が高い。

## 広報掲載

広報厚沢部2024年10月号記事で紹介しました。

[www.town.assabu.lg.jp](http://www.town.assabu.lg.jp)

<https://www.town.assabu.lg.jp/uploaded/attachment/5187.pdf>

図9 個別ページに記入されたメモと広報へのリンク

Notionの個別ページでは、入れ子状にページを作成することもできる。つまり、ページ内にコンテンツを列記するだけでなく、別のページを埋め込むことも可能である（図10）。この機能を活用することで、階層的な情報整理が容易になり、大量の資料や複雑なデータを直感的に管理できる。また、各資料ごとに個別ページを作成するだけでなく、事業のレポートを資料に紐付けたり、金属製品のメンテナンスなどの保存業務日報を個別資料に関連付けたりすることも可能である。



## 所蔵資料データベース

Show All

...

+

↑ 分類番号

Aa 資料名: 粉鉢

+ フィルタ

Aa 資料名

≡ 分類番号

木製片口粉鉢

6220

📄 木製片口粉鉢

6220

+ 新規ページ

コメント

🗨️ コメントを追加...



2025年1月15日撮影

(字鶴 氏父寄贈)

### 認定こども園はぜるでの粉引きの様子

📄 認定こども園 きな粉づくり体験

図10 ページ内に埋め込まれた別ページ

資料を利用した体験学習や資料調査の様子は、コンテンツが長大になりがちのため、別ページを作成することが適切であろう。図11は、町内の認定こども園で行われたきな粉づくりの様子を記録したもので、当日の様子を写真や動画で記録している。厚沢部町郷土資料館から貸し出された粉鉢と石臼を使用しているため、図11の子ページは石臼の個別ページにも埋め込まれている。このような高い拡張性と柔軟な運用が、Notionの魅力である。

## 所蔵資料データベース

Show All

+

↑ 分類番号

Aa 資料名: 粉鉢

+ フィルタ

Aa 資料名

≡ 分類番号

木製片口粉鉢

6220

木製片口粉鉢

6220

+ 新規ページ

## 認定こども園 きな粉づくり体験

- 2025年1月16日
- さんに目立てしてもらった臼とともに使用



図11 別ページに記入された資料活用の様子

## AIの利用と検索










博物館の資料データベースにおいて、検索は頭の痛い問題である。メジャーな資料であればあるほど、ローカルな名称が数多く存在し、資料台帳の一貫性を損なう原因となっている。シソーラスの混乱は、博物館資料の登録作業における大きな負担となっている。

一方、近年大幅に能力を向上させた大規模言語モデルは、語の厳密な一致を必ずしも必要としない。独自の意味空間を備え、たとえば「テッケシ」「テカエシ」「テッカエシ」のように同一の対象を指す表現のゆれは、接近した意味空間にマッピングされる。一方で、「テッケシ」と文字列的に似た「テッポウ」は、相対的に離れた意味空間に位置する。大規模言語モデルのこうした特性は、シソーラスの統一が困難な博物館資料の検索に非常に適している。

Notionの有料サービスでは、AIが提供されており、Notion内のコンテンツを対象とした対話的な検索やチャットが可能である。その一例として、厚沢部町郷土資料館に保管されている戊辰戦争・箱館戦争に関連する資料を探索させた結果を図12に示す。

戊辰戦争や箱館戦争に関連する資料を探してください。

収蔵庫に以下の箱館戦争・戊辰戦争関連資料が保管されています：

- 箱館戦争の資料 
- 箱館戦争と大野藩に関する資料 
- 古文書だより函館の箱館戦争新資料展の資料 
- 南蝦夷戦争記・箱館日誌 
- 戊辰関係資料を含む北海道志 
- 二股戦争之圖 
- 戦記箱館戦争に関する林家文書解読講座の資料 
- 戊辰己巳戦死者墓地関係資料 
- 館砦志士戦没之図写真 

これらの資料はすべて収蔵庫に保管されており、分類番号SKまたは8810で整理されています。

図12 Notion AIが探した戊辰戦争・箱館戦争関連資料

このように、Notion AIをはじめとした大規模言語モデルは、正規表現のような厳密性を必要とせず、既存のコンテンツから語彙の関係性を把握し提示するため、冗長であっても多くのテキストを含むコンテンツがデータベースに内包されていることが重要となる。AIを活用した資料整理のためにも、多量のコンテンツをインベントリ情報とともに保管できるNotionのようなサービスは、博物館や資料館のデータベース管理に適したツールとなる。従来の資料整理では、統一されたシソーラスの整備や厳密なキーワード設定が求められ、それが運用の負担となっていた。しかし、Notion AIを活用することで、必ずしも事前に厳格な分類を行わなくても、後から検索や分類の精度を向上させることが可能となり、多様な情報を日常的に入力し続けることが、有用なデータベースの構築につながる。この方法の延長線上において、博物館資料データベースは作業日誌に近づいていくだろう。

## AIの時代の資料台帳

すでに述べたように、Notionのデータベース機能を活用すれば、各資料に関連する多様なメディア（画像、動画、スキャンデータ、音声記録など）を一元管理できる。これにより、単なるテキスト情報だけでなく、視覚的・聴覚的なデータも含めた統合的なアーカイブが可能となる。加えて、AIの対話的検索機能を利用することで、従来のキーワード検索に依存しない、柔軟な情報探索が実現する。「明治時代の民具に関する資料を探したい」といった自然言語による問い合わせに対して関連する情報を提示することができる。この機能は、特に専門的な知識を持たない利用者にとって有用であり、博物館や資料館の情報公開の在り方にも変化をもたらす可能性がある。



## まとめ

### 台帳は小さく確実に

博物館資料台帳は、博物館法が定める目録であると同時に、資料管理の基礎となり、全資料へのアクセスを確保するための重要なツールである。あらゆる事態に備えるために、膨大な項目を設けたり、周到に設計されたデータベースを利用したりすることは、必ずしもデータの持続可能性を保証するわけではない。項目数の多さや多様なテーブル構成は、入力時の負担を増大させるだけでなく、作業に熟練した担当者が異動や退職によって不在となった際に、引き継ぎが円滑に行われず失敗するケースも多い。そのため、小規模な博物館では、可能な限りシンプルな台帳を基本とするのが望ましい。特に、民俗資料を中心とする郷土資料館などでは、「博物館コアメタデータ」（山本ほか 2009 前掲）を参照し、必要な項目を慎重に検討するのが適切であろう。

### 表計算ソフトを活用する

一枚の表形式に収まるようなデータであれば、必ずしもデータベースソフトを使用する必要はない。データベースソフトの操作は、Office系のソフトウェアと比較すると決して簡単とは言えず、高い習得コストが求められる。また、担当者の異動やソフトウェアの更新停滞などにより、維持が困難になる可能性もある。一方で、Office系の表計算ソフトは特別な習熟を必要とせず、一般の事務職員でも操作が可能である。さらに、LibreOffice(註9)のようなフリー・オープンソースソフトウェアを利用できる点も、表計算ソフトを積極的に活用する理由の一つとなる。

また、表計算ソフトからCSVなどのプレーンなファイル形式に書き出すことで、将来的に本格的なデータベースへ移行したり、GISや他のデータベース環境にインポートしたりすることも容易になる。データベースの持続可能性に不安がある施設では、非データベースソフトの活用も視野に入れるべきだろう。

### Tidyなデータを作成する

人間が視覚的に認知しやすい構造が、コンピュータで扱いやすい構造とは限らない。Tidy Dataの概念は慣れなければ難しく感じるが、「観測値は何か」を念頭において、観測項目を列方向に並べていくことで達成される。当然ながら、表計算ソフトのセル結合機能などは「百害あって一利なし」であり、使ってはならない(註10)（奥村2013）。表計算ソフトを利用する最大の弊害は、データ構造と整飾要素が容易に混在しうることである。データベースソフトではこのようなことは原則として発生しない。

Tidy Dataは機械がデータを判読するためのデータ構造であり、人間にとって読みやすいものではない。人間の視認性よりも機械の判読性を優先する理由は、データの処理を機械を通して行うことがもっともデータを有効に活用できる手段だからである(註12)。

また、Notionのような多様なコンテンツを内包するメディアの利用においても、核となるのはTidyなデータである。Tidyなデータ構造を維持することは、AIの活用だけでなく、データの一貫性を保ち、長期的な運用を容易にする上でも重要である。特に、博物館や資料館のように、多様な形式のデータ（テキスト、画像、動画、音声、関連資料のリンクなど）を管理する場合、データの整理が不十分だと、検索性が低下し、適切な情報の抽出が困難になる。

そのため、データベースの設計段階でTidy Dataの原則を考慮する必要がある。たとえば、各資料のメタデータ（タイトル、作成者、作成年代、関連する歴史的背景など）は一貫したフォーマットで整理することや、日付や数値データの形式を統一すること、自由記述の欄と構造化データの欄を明確に分けることで、後のデータ処理がスムーズに行えるようになる。将来的には人間がTidyなデータを全く意識せず、AIがデータ整形や構造化を行うことも考えられるが、現時点では、人間が適切にデータを整理し、管理するための基本的なスキルを身につけることが不可欠である。

## 資料台帳作成と電子データの公開は博物館の義務である

目録の頒布は博物館法に定められた博物館の事業である。博物館法第3条は「例示」であり、必ずしもこの全てを行うものではないとされる（栗原2022）。しかし、博物館の目的である「保管」を達成するためには目録（資料台帳）の作成が不可欠であり、資料活用のための方策として目録頒布は旧法においても博物館の基礎に位置づけられるべき業務である。さらに、公立博物館では地方自治法第2条第14号(註13)の規定に照らしても、目録頒布を可能な限り効率的な方法で行うよう努めなければならないだろう。資料台帳作成はもとより、その電子データのウェブ公開も、もはや「義務」といえる時代を迎えている。

## 引用文献

François Mairesse, 2023. Dictionary of museology: Routledge

Grant, Alice ; Nieuwenhuis, Joséphine ; Petersen, Tony.1994. Minimum Information Categories for Museum Objects (MICMO) : Proposed guidelines for an international standard.

1995. International Guidelines for Museum Object Information: The CIDOC Information Categories([url{https://cidoc.mini.icom.museum/wp-content/uploads/sites/6/2020/03/guidelines1995.pdf}](https://cidoc.mini.icom.museum/wp-content/uploads/sites/6/2020/03/guidelines1995.pdf). 2022年6月6日取得)

Wickham, H, 2014. Tidy data. Journal of Statistical Software, 59 (10). doi:10.18637/jss.v059.i10

犬木努 2008 「5. データベースの現状と課題」『新しい博物館学』全国大学博物館学講座協議会西日本部会, pp. 162-164

奥村晴彦 2013 「「ネ申Excel」問題」『情報教育シンポジウム2013論文集』2, pp. 93-98

景山隆 2022 「経営継承におけるマッチングデータベース その必要性と理論的基礎（他分野や海外の例も参照しつつ）」『農政調査時報』第587号,一般社団法人全国農業会議所, pp. 29-37

栗原祐司 2022 『基礎から学ぶ博物館法規』 同成社, pp. 34-46

後藤真 2013 「博物館資料情報のLinkedOpenData化へ向けたモデル試作\CID{00661}花園大学歴史博物館資料を題材に\CID{00661}」『情報処理学会研究報告』 Vol.2013-CH-97 No.5, pp. 1-6

佐々木秀彦 2002 「公共財としての博物館資料～アクセスを保証する資料整備・公開体制の構築：人文系博物館を中心に～」『博物館学雑誌』 第27巻第1号（通巻35号）, pp. 13-24

田良島哲 2007 「5-2.東京国立博物館の文化財情報管理とデジタルアーカイブ」『映像情報メディア学会誌』 Vol61, No.11, pp. 1586-1588

公益財団法人日本博物館協会 2020 『令和元年度日本の博物館総合調査報告書』, pp. 57-58

博物館情報処理に関する調査研究プロジェクトチーム 2005 「ミュージアム資料情報構造化モデル」 東京国立博物館 ([url{https://webarchives.tnm.jp/docs/informatics/smmoi/}](https://webarchives.tnm.jp/docs/informatics/smmoi/). 2022年6月6日取得)

文化庁文化財部記念物課 2010 『発掘調査の手引～整理・報告編～』

村田良二 2016 「博物館におけるコレクション情報の組織化 情報標準と東京国立博物館の事例」『情報管理』 vol59, no.9, pp. 577-586

山田太造 2019 「歴史データをつなぐこと\CID{00661}目録データ\CID{00661}」『歴史情報学の教科書 歴史のデータが世界をひらく』 文学通信, pp. 23-35

山本泰則・安達文夫 2009 「博物館資料情報統合検索のためのコアメタデータ」『じんもんこん2009論文集』 情報処理学会, pp. 287-294

## 註

註1 博物館法 第二条 この法律において「博物館」とは、歴史、芸術、民俗、産業、自然科学等に関する資料を収集し、保管（育成を含む。以下同じ。）し、展示して教育的配慮の下に一般公衆の利用に供し、その教養、調査研究、レクリエーション等に資するために必要な事業を行い、あわせてこれらの資料に関する調査研究をすることを目的とする機関（社会教育法による公民館及び図書館法（昭和二十五年法律第百十八号）による図書館を除く。）のうち、地方

公共団体、一般社団法人若しくは一般財団法人、宗教法人又は政令で定めるその他の法人（独立行政法人（独立行政法人通則法（平成十一年法律第百三号）第二条第一項に規定する独立行政法人をいう。第二十九条において同じ。）を除く。）が設置するもので次章の規定による登録を受けたものをいう。

註2 「ミュージアムとコレクションの保存活用、その多様性と社会における役割に関する勧告」（2015年11月20日,UNESCO）

註3 UNESCOの「ミュージアムとコレクションの保存活用、その多様性と社会における役割に関する勧告」17節では「ミュージアムは社会全体に語りかけるゆえに社会的な繋りと団結を築き、市民意識の形成また集団的アイデンティティを考える上で、重要な役割を持つ重要な公共空間である。ミュージアムは、恵まれない立場のグループを含め、すべてに開かれた、あらゆる人々の身体的・文化的アクセスを保証する場であるべきである。」と述べ、博物館の社会的役割を強調する。

註4 景山隆は円滑な農地継承について「摩擦的成本である探索コストが、負担し得る水準以上の場合には、就農プロセスや就農後の安定経営からドロップアウトする危険がある」とのべ、探索コストの増加が農業の新規参入や安定経営の妨げとなる可能性に言及し、「探索コストが十分に大きい場合、単に当事者、つまりミクロ経済主体である継承希望者の利益だけでなく、マクロ的、社会的利益をも損なうこととなる」とし、このことが「取引コストを構成する、探索コストの最小化を図る公共的政策の正当性が存在する」とする（影山2022）。

註5 2013年から「台帳なし」が選択肢に加わったことで、これまで無回答だった館が「台帳なし」を選択するようになった結果、見かけ上「ほとんど全て」の比率が低下した可能性もある。

註6 日本工業規格,X 0001-1994「情報処理用語-基本用語」の定義によれば、「複数の適用業務分野を支援するデータの集まりであって、データの特性とそれに対応する実体との間の関係を記述した概念的な構造に従って編成されたもの」とされる。

註7 日本工業規格,X 0001-1994「情報処理用語-基本用語」の定義によれば、データとは「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。」とされる。同様に情報とは「事実、事象、事物、過程、着想などの対象物に関して知り得たことであって、概念を含み、一定の文脈中で特定の意味をもつもの」であるから、調査記録などは表現形式によってはデータではなく情報に区分すべきケースも多い。

註8 Hadleyが提示したTidy Dataの概念は次の4つの要素で構成される。1. 個々の変数 (variable) が1つの列 (column) をなす。2. 個々の観測 (observation) が1つの行 (row) をなす。3. 個々の観測の構成単位の類型 (type of observational unit) が1つの表 (table) をなす。4. 個々の値 (value) が1つのセル (cell) をなす。（『整然データとは何か』<https://id.fnshr.info/2017/01/09/tidy-data-intro/>, 2022年6月22日取得）

註9 LibreOfficeは、The Document Foundationが支援するオープンソースのオフィスソフトでMozilla Public License v2.0で提供されている。ワープロ、表計算、プレゼンテーション、ベクターグラフィック、データベース数式作成で構成される。MicroSoftのOfficeソフトウェア群のファイル形式を互換する。

註10 表計算ソフトのセル結合の極限として「エクセル方眼紙」がある。セルをグリッド状に配置し、セル結合を駆使して思い通りの帳票を実現する手段である。エクセル方眼はもはや電子データと呼ぶに値せず、表計算ソフト上でも値の引用ができなくなるなど不具合が多い。筆者は日常的に国や都道府県とデータのやり取りを行っているが、エクセル方眼に類似した様式の使用を強要されることが多い。このことは我が国の電子化の水準や意識をよく示していると思われ、暗澹たる気持ちになる。

註11 筆者の勤務する厚沢部町郷土資料館の資料データベースのフィールドは分類番号、枝番、寄贈年月日、寄贈者住所、寄贈者氏名、年代、旧番号、一括資料記号、備考である。資料管理の目的にはこれで十分である。

註12 人間にとって100万件のデータから必要なデータを見つけ出すことは不可能に近いが、機械には極めて容易である。

註13 地方公務員法第2条第14条 地方公共団体は、その事務を処理するに当つては、住民の福祉の増進に努めるとともに、最少の経費で最大の効果を挙げるようにしなければならない。

---

2025-02-27 「図1 資料台帳記載の収蔵資料の割合」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所 <https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図2 電子メディアにデータベース化された資料台帳」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所  
<https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図3 電子メディアにデータベース化された資料の割合」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所  
<https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図4 リレーショナル・データベースを利用した発掘調査資料と博物館資料台帳の統合」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所  
<https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図5 VLOOKUP関数を利用した表計算ソフトによるテーブル結合」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所

<https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図6 Tidy Dataとは何か」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所 <https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図7 Notionのデータベース画面」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所 <https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図8 Notionのデータベース内の個別ページ」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所

<https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図9 個別ページに記入されたメモと広報へのリンク」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所

<https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図10 ページ内に埋め込まれた別ページ」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所 <https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図11 別ページに記入された資料活用の様子」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所

<https://sitereports.nabunken.go.jp/cultural-data-repository/96>

2025-02-27 「図12 Notion AIが探した戊辰戦争・箱館戦争関連資料」 『文化財データリポジトリ 石井淳平「AI時代の資料データベース」『デジタル技術による文化財情報の記録と利活用7』』 奈良県 奈良文化財研究所

<https://sitereports.nabunken.go.jp/cultural-data-repository/96>