

埋蔵文化財情報のオープンデータ化—課題と展望

● 茶谷 満

奈良文化財研究所の「全国文化財総覧」の進展により、発掘調査報告書等の情報の入手性は著しく向上した。一方、愛知県埋蔵文化財センターでは報告書 PDF の公開のほか、図・写真等を個別要素化するデータベースを構築しており、高く評価されている。しかし、データ加工及び登録の労力が大きく、更新が停滞しているという課題がある。そこで本稿では、ミニマムなオープンデータとしての公開を提案し、その利活用の展望と課題を整理する。

1. はじめに

近年、私たちを取り巻く環境において生成 AI をはじめとするデジタル技術の進化は目覚ましく、国の政策においてもデジタル庁を中心に DX（デジタルトランスフォーメーション）が推進されている。埋蔵文化財の分野においても、発掘調査現場でのデジタル機器活用から報告書の刊行に至るまで、そのプロセスの大半にデジタル技術が浸透している状況にある。

発掘調査の成果は最終的に印刷物として報告されるが、現在は PDF 形式のデータも同時に制作されるのが一般的である。PDF での公開は、ウェブサイトを通じて遠隔地であっても容易に入手できるメリットとなっている。特に、国立文化財機構奈良文化財研究所（以下、奈文研）が運営する「全国文化財総覧」では、日本各地の発掘調査報告書 PDF を検索・ダウンロードすることが可能となっており（高田 2015、2018）、情報収集における利便性が飛躍的に向上した。

こうした状況下で、愛知県埋蔵文化財センターは、発掘調査報告書内の写真や実測図を PDF から切り離し、個別のデジタル要素として検索・閲覧・再利用可能なシステムを構築しており（堀木 2019、2020）、先進的な取り組みとして高く評価されている（野口 2022a）。ただし、こうした図表単位の要素化は、データを加工してシステムに登録するための多大な労力が必要であり、現状ではデータの追加登録が停滞しているという課題がある。

本稿では、既存のシステムと並行して、こうした埋蔵文化財情報の二次利用を前提とした機械判読に適した形式としてデータそのものを公開・提供していく「オープンデータ化」を提案する。そこで、愛知県埋蔵文化財センターが公開している情報のオープンデータ化における現状と課題を整理し、公共財としての埋蔵文化財情報のさらなる利活用の可能性について展望する。

2. オープンデータと既存のシステム

(1) オープンデータの定義と要件

国の方針として、平成 28 年 12 月に施行された「官民データ活用推進基本法」に基づき、地方公共団体には保有する公共データのオープンデータ化が義務付けられた⁽¹⁾。身近な例では、愛知県、名古屋市など多くの自治体が、行政情報を CSV 形式等のファイルでポータルサイトに公開している。

この「オープンデータ」とは、以下の 3 条件を満たすデータを指す。①営利・非営利を問わず二次利用可能なルールが適用されていること（オープンライセンスの付与）、②機械判読に適した形式であること（特定のソフトウェアに依存しないファイル形式等）、③無償で利用できること（武内・高田 2022）、これは単なる情報の「公開」ととどまらず、利用者がコンピュータを用いて自由に二次利用できることを保証するものである。

①のライセンス明示については、データ利用の際の申請手続きを不要とすることで、デー

タ提供者・利用者双方の手間やコストを低減できる（公益財団法人東京市町村自治調査会 2017）。特に年間数十件に及ぶ申請・承認手続きを考慮すれば、その効率化のメリットは大きい。

③の無償化については、公的資金を用いて作成されたデータ（公共財）を広く国民に還元するという理念に基づいている（野口 2022a）。無償とすることでより多くの利用者が手軽に利用可能となり、利活用の裾野を広げる効果が期待される。

②の機械判読性について、PDF ファイル等はビューワーでの閲覧は可能であるが、他のシステムで利用するにはデータの再入力や加工といった「ひと手間」を要する。

オープンデータが目指すデータモデルには、ティム・バーナーズ＝リーが提唱した「5 スターオープンデータ」という指標がある（表 1）。一般的に、★2 以上が「機械判読可能」、★3 以上が「機械判読に適している」とされ（福島 2020）、自治体や公的機関は、まずデータ活用の基盤となる「★3（CSV 形式等）」を目指すべきとされる。★4 や★5（LOD）に進むことで、統計と地図情報など異なる分野のデータを容易に統合できるようになり、Society5.0 やスマートシティの実現に不可欠な「データの連携」が可能になる（公益財団法人東京市町村自治調査会 2017）。

これに従うと、前述の発掘調査報告書 PDF も、オープンライセンスの付与とインターネット上での公開が適用されていれば「★1」のオープンデータに該当する。中でも抄録や遺構・遺

物一覧表などは、元データが Excel で作成されており、データの構造化がなされているため、CSV 形式への変換・保存は比較的容易である（堀木 2019）。

しかし、★4 以上への対応には課題がある。現時点では、愛知県埋蔵文化財センターでは報告書刊行後に別途データ整備を行う必要があり、多大な手間やコストが生じることになる。

(2)「オープンアクセス」との違い

「オープンデータ」と混同されやすい言葉に「オープンアクセス」がある。第 5 期科学技術基本計画では「オープンアクセス」と「研究データのオープン化（オープンデータ）」の 2 つを含む概念として「オープンサイエンス」と定義されている（内閣府 2016）。

「オープンアクセス」が主に論文などの著作物に対する「情報の到達性（アクセシビリティ）」を重視するのに対し、「オープンデータ」は数値や図面そのものの「再利用性」を核心としている。現在、多くの自治体が「指定文化財一覧」などを国の推奨データセット（CSV 等）に基づき公開している。例えば愛知県では、ウェブマップ「マップあいち」での公開に加え、オープンデータセットとしても同データを公開している⁽³⁾。

発掘調査報告書の公開状況を見ると、愛知県埋蔵文化財センターをはじめとする多くの機関が PDF 形式を採用している。また、自前のリポジトリを持たない自治体等でも、奈文研の「全国文化財総覧」に登録することで全文公開が可能であり、これは「オープンアクセス」の代表例と言える。一方で、各自で報告書 PDF を公開

表 1 5 スターオープンデータの各段階と特徴⁽²⁾

段階	公開の状態（定義）	代表的なデータ形式	特徴・活用度
★★★★★	自社のデータを他分野のデータとリンクさせ、文脈を繋いでいる。	LOD (Linked Open Data)	データ同士が繋がり、高度な分析やマッシュアップが容易な最高レベルの状態。
★★★★	データの識別に URI (Web 上の住所) を用い、外部から直接参照できる。	RDF, SPARQL	Web 上のデータとして、世界中の他のシステムから直接リンクが可能になる。
★★★	特定のソフトに依存しない「非独占的なオープン形式」で公開されている。	CSV, XML, JSON	無料のツールで自由に加工・編集ができ、機械判読に最も適した実用的なレベル。
★★	コンピュータで処理可能な「構造化データ」として公開されている。	XLS (Excel), DOC (Word)	機械判読は可能だが、特定の有料ソフトウェアが必要な場合がある。
★	オープンなライセンス (CC-BY 等) で、形式を問わず Web 上に公開されている。	PDF, JPEG, GIF	人間には読めるが、コンピュータによる自動処理（再利用）には向かない。

しているサイトにおいて「公共データ利用規約」や「クリエイティブ・コモンズ（後述）」等の二次利用規約を明示している事例は管見の限りあまり見受けられない。

（3）愛知県埋蔵文化財センターのデータベースシステム

愛知県埋蔵文化財センターのデータベースシステム（以下、埋文 DB）は、本稿執筆段階で報告書抄録、写真ファイル、遺構ファイル、遺物ファイル、実測図ファイルの5つの主要 DB に加え、年報および紀要の各ファイル DB で構成されている。報告書抄録を基軸として各 DB を相互に関連付けることで、特定の遺物名を検索した際、複数の遺跡から出土した当該遺物の一覧と写真を即座に取得できる環境を実現している（堀木 2019）。

埋文 DB の構成は、利用者が求める情報へ効率的にアクセスできるよう設計されている。データの基本仕様は発掘調査の電子納品時における整理要項をベースとしており、遺跡 ID や図面・写真のファイル名は半角英数字で規則化され、メタデータの統一が図られている。また、緯度経度の位置情報を用いて国土地理院のウェブマップ上に遺跡のポイントを表示する機能も備えている（堀木 2019）。システム構築におけるインターフェイス開発には多大な労力を要するが、一度運用基盤が確立されれば、その後のデータ追加作業の負担は相対的に軽減される。

このシステムの最大の特徴は、発掘調査報告書という一つの「冊子」に縛られていた情報を「離散的要素（個別の図・表・写真）」として分解し、再利用可能な形で提供している点にある（野口 2020）。刊行物 PDF の全内容から特定の情報を抽出する作業は利用者にとって極めて煩雑である。こうした情報の分解を行うことで、データ提供側である愛知県埋蔵文化財センター自体にとっても、新たな報告書作成時に近隣遺跡の既存データを参照する作業においては、提供者が同時に高度な利用者となる構造を有している。このように、DB は、アナログ情報の単なるデジタル化を超え、情報のあり方や社会との関わり方を変革する DX の先駆的事例と言える。

さて、埋文 DB を前述の表 1 に照らすと、実

用的なデータ活用の基盤となる「★3」に相当する。埋文 DB と ★4 の RDF (Resource Description Framework) を比較すると、情報の「要素化」や「関連付け」という点において強い類似性が認められる一方、データの標準化やウェブ上での相互運用性において明確な相違点が存在する。

両者の最大の共通点は、一連の情報を「バインド（拘束）された形態」から解放し、論理的につなぎ合わせようとする思想にある。埋文 DB が図・表・写真を個別の要素として管理する手法は、RDF がウェブ上のリソースを独立した単位として扱う考え方と共通している。また、埋文 DB が報告書 ID 等をキーとして5つの DB を関連付ける構造は、RDF が「主語・述語・目的語」という3つの要素（トリプル）を用いてリソース間の関係性を記述する枠組みと類似している。さらに、遺跡をまたいだ遺物の横断検索を可能にし、結果をテキストデータとして入手できる点も、異なるデータセット間の横断処理を目指す RDF の目的と共通している。

一方で、実務的な実装方法には相違がある。RDF はすべてのリソースの識別に URI（ウェブ上の住所）を利用し、世界中の他データから直接参照できる仕組みを持っている。対して埋文 DB は、主にシステム内部の「報告書 ID」や「遺跡名」をキーとしており、独自の仕様に基づいている。また、データモデルについても、RDF が分野横断的な連携（Linked Open Data (LOD)）に適した「グラフ構造」を持つのに対し、埋文 DB は「リレーショナル・データベース (RDB)」を用いて表形式のデータを結合する構造となっている。

以上のように、RDF は表 1 の5段階評価のうち URI を用いる「★4」に該当するが、埋文 DB は機械判読可能な形式（CSV 等）での出力を主眼とした、実用的なデータ活用の基盤となる「★3」の段階に位置付けられる。RDF が W3C (World Wide Web Consortium) が規定する世界共通規格であるのに対し、埋文 DB は国土情報としての公共性を意識していることから「地質調査資料整理要項」をモデルとしつつ、独自の PHP/MySQL 環境で行われている（堀木 2019）という点も大きな相違点である。

(4) クリエイティブ・コモンズ (cc)・ライセンスについて

CC-BY ライセンスは、著作権を保持したまま、出典の表示を条件に複製や改変、営利目的での二次利用を広く認めるライセンスである⁽⁴⁾。例えば、愛知県埋蔵文化財センターでは報告書 PDF 等のデータを CC-BY ライセンス (3.0 非移植等) で公開している (堀木 2019)。ただし、データ提供側としては以下の点に注意が必要となる。

①第三者の権利の確認 (最重要事項)：自らが著作権を有していない (第三者の権利が含まれる) コンテンツには、そのままライセンスを適用することはできない。写真や図面の中で、第三者が著作権、肖像権等を有している場合、事前に許諾を得る必要がある (ASP・SaaS・クラウドコンソーシアム 2014)。

②著作権の対象外となるデータの扱い：遺物一覧表の数値データなどは著作権の対象外であり、本来 CC ライセンスの適用対象外であるが、考古学研究においては、根拠となるデータが出典と紐づいていなければ論証が成立しないため、データの信頼性を担保しておくことが重要である。また、元データの誤記や不備がある場合には、それを明記し履歴を管理することも、再利用性を担保する上で不可欠である。

③免責事項の明示：提供した情報に誤りや、それを利用したことで損害が発生した場合に備え、提供側は一切の責任を負わない旨を明示しておく必要がある。情報の正確性、完全性の不保証や、利用者の行動に伴う結果についての免責を「利用規約」に盛り込んでおく必要がある⁽⁵⁾。

④ライセンスの不可逆性：一度 CC ライセンスを付与して公開すると、その後にライセンスを取り消すことはできない⁽⁶⁾。著作権の存続期間中は有効であり続けるため、将来的な利用制限をかける可能性があるデータについては慎重な判断が求められる。

⑤適切な表記と機械判読性の確保：利用者が迷わず正しく利用できるよう、標準的なメタデータやアイコンを使用する。政府のガイドラインなどに準拠した標準的な記述 (例：「CC-BY 4.0 国際」など) を採用し、出典の記載方法を具体

的に例示することが望ましい⁽⁷⁾。

これらの点を踏まえ、まずは「公開できるところから小さく始める」(東京市町村自治調査会 2017) とともに、段階的にライセンス設定を改善し続ける姿勢が重要である。

3. オープンデータの活用

CSV 形式等のオープンデータは、単体では機械判読可能な「離散の要素」の集合に過ぎない。提供側が公開の意義を強調するだけでなく、具体的な活用モデルを提示することが、オープンデータの持つ可能性を広げる鍵となる。

一般的な活用事例としては、避難場所やハザードマップ、ゴミ収集日の通知アプリ等がある。こうした取り組みは、「国民参加・官民協働の推進を通じた諸課題の解決、経済活性化」「行政の高度化・効率化」「透明性・信頼の向上」といった多面的な価値を社会にもたらす⁽⁸⁾。

では、考古学 / 埋蔵文化財の分野ではどのような活用が見込まれるであろうか。遺跡抄録データや遺物一覧表等の構造化されたデータを活用することで、複数の遺跡データを横断的に収集・管理することが可能となる。例えば、地理情報システム (GIS) ソフトを用いた空間分析を行えば、時期による遺物の拡散や土地利用の変遷、広域的な流通網の把握といったマクロな視点での研究を加速させることができる。

また、行政実務においては、開発に伴う事業計画時に周辺の歴史的環境を事前に予測することで、文化財の「予防的保護」や意図しない破壊のリスク低減に寄与できる⁽⁹⁾。さらに、文化財情報を活用した観光ルートの開発や、学校教育でのデジタル教材としての活用など、学習支援への貢献も期待される。つまり、こうしたデータは目的に応じたコンテキスト文脈で抜き出し集約できるため、より柔軟な活用が可能となる。

このように、これまでの考古学が発掘調査報告書を読み込んで様々な情報の抽出を一つの学術的素養としてきたことに対し、個々のデータを独立した要素として解放することは、専門家以外の多くの人が調査成果にアクセスし新たな価値を創造する機会を創出するものと考えられ

る。

4. 課題と展望

(1) データの追加、更新作業

オープンデータ化は一過性の作業ではなく、継続的なデータの追加・更新を通じて、システムの可用性や情報の完全性を機関として保証し続ける体制（予算・人材）の構築が不可欠である（公益財団法人東京市町村自治調査会 2017）。既存の報告書抄録等を CSV 形式にまとめ、ウェブサイトへ追加・更新する作業は比較的少ない労力で実施可能である。しかし、従来のデータベースシステムのように図や写真を 1 点ずつ要素化して登録する作業は多大な作業工数を要するため、更新頻度に格差が生じる懸念がある。

この課題を解決するには、発掘調査データの電子納品の在り方を再定義し、調査段階から機械判読を前提としたデータ作成を求めることで、公開プロセスの簡略化を図ることができるであろう。既存の埋文 DB は利用者が目的に応じて自由にデータを加工・分析できる環境を加工済みの情報を提供することで利用者の利便性を高めているが、オープンデータは利用者が目的に応じて自由にデータを加工・分析できる環境を提供するものである。この「加工済みデータ」と「素材としてのオープンデータ」を利用者が用途・スキルに合わせて選択できる環境が現時点での最適解といえよう。

(2) 属人的データの峻別

オープンデータ化においては、寸法や重量といった「機械的な属性データ」と、遺構のコンテキストや時代決定等の「(考古学的) 解釈」を峻別する必要がある。土層の記述や遺物観察には、調査者の知識や経験が反映されており、これこそが報告書の学術的成果の中核をなすものである。こうした情報を「離散的要素」として切り離すことは、コンテキスト（文脈）が削ぎ落とされるリスクを孕む。同様に、考古学の（伝統的な）手実測は、単なる図化・記録作業ではなく、実物を詳細に観察して主観的な思い込みを修正しながら解釈を深める重要なプロセスであり、後進への知識継承の場としての役割

も担っている（渡部 2019）ためである。

一方で、第三者が異なる視点からデータを再構成し、新たな解釈や評価を発見する機会を創出するメリットも併せ持っている。デジタルデータとしての客観性をどう担保し、専門家の判断という不確実な要素をどう記述するか（メタデータの整備）が、データの真正性確保の鍵となる。

(3) 懸念される事項

データのオープン化には多面的なリスクが懸念されている。具体的には、遺跡の所在地特定による盗掘（違法な発掘）や不法侵入といった文化財の物理的破壊のリスク、不動産取引や開発計画における「障害」とみなされることによる私権への影響が挙げられる（野口 2022b）。また、研究面では、出典を明記しない不適切な引用や、意図しない文脈での利用、さらには揚げ足取りのような「悪意のある引用」による研究者のキャリアリスクへの懸念も根強い（池内 2018）。

データ提供側としては、システムやサーバーの維持管理に加え、PhotoCD 等の旧形式データや低解像度画像のマイグレーション（データ移行）にかかる予算・人員の確保が継続的な課題となる。CSV 等の汎用形式を基軸としつつ、組織内でミニマムなマネジメント体制を維持し続けることが重要である。

(4) 展望

以上のような現状や課題を踏まえ、今後、愛知県埋蔵文化財センターとしては既存の DB システムを維持しつつ、並行して CSV 形式などのオープンデータを積極的に公開し、ユーザーが用途に応じて自由にデータを選択・利用できる「ハイブリッドな提供環境」を整備することが現時点で最善と考える。そのためには、システムとオープンデータカタログサイト双方の利用規約（公共データ利用規約等）を整理し、利用者に明示することが先決となる。

また、発掘調査の設計段階から公開を逆算して設計する「オープンデータ・バイ・デザイン」⁽¹⁰⁾ の考え方を導入することも重要である。これにより、報告書刊行後の再加工コストを削減するとともに、将来的には埋蔵文化財保護行政の新たなモデルとなる可能性がある。将来的に

これらインターネット上の膨大なデジタルデータが自動的に紐づけられるようになって、人工知能（AI）等を用いた定量分析や自動分類が一般化すると、専門家はこうした膨大なデータと社会あるいは一般ユーザーとの間を橋渡しする「解釈の通訳者」としての役割がより強く求められるようになる（渡部 2019）。そのためには考古学を学ぶ者として遺跡・遺物を観察し、解釈する知識等は変わらずに必要となろう。

既存のシステムの更新とオープンデータの公開という両輪の取り組みにより、愛知県埋蔵文化財センターが所有する埋蔵文化財情報は「真の公共財」としてさらに広く利活用されていく

ことが期待される。野口淳氏が指摘するように、「ミニマムな構成であっても公開されていれば外部利用者はアクセスし、検索し、利用できる」（野口 2020b）という事実が重要である。公開につなげる姿勢を最優先に、今できるスモールステップに着手していきたい。

謝辞

本稿をまとめるにあたり、堀木真美子氏、渡部展也氏に多大なる助言を賜った。ここに記して感謝の意を表する。

註

- (1) https://www.digital.go.jp/resources/open_data（参照 2026-05-08）
- (2) <https://5stardata.info/ja/>（参照 2026-05-08）
- (3) <https://maps.pref.aichi.jp/opendata.html>（参照 2026-05-08）
- (4) <https://creativecommons.jp/licenses/>（参照 2026-05-08）
- (5) https://www.aist.go.jp/aist_j/press_release/pr2022/pr20221018/pr20221018.html（参照 2026-05-08）
- (6) <https://creativecommons.jp/faq/#16>（参照 2026-05-08）
- (7) 前掲 (3) 公共データ利用規約参照
- (8) 前掲 (3) オープンデータ基本指針
- (9) https://www.aist.go.jp/aist_j/press_release/pr2022/pr20221018/pr20221018.html（参照 2026-05-08）
- (10) 前掲 (3) オープンデータ基本指針

参考文献

- 池内有為 2018 「日本における研究データ公開の状況と推進要因、阻害要因の分析」『Library and Information Science』No.79
- 公益財団法人東京市町村自治調査会 2017 『基礎自治体によるオープンデータ化と利活用の可能性に関する調査研究報告書』
- 高田祐一 2015 「「全国遺跡報告総覧」の機能と期待される効果」『カレントアウェアネス-E』No.287
- 高田祐一 2018 「全国遺跡報告総覧における学術情報流通と活用への取り組み」『カレントアウェアネス』No.337
- 武内樹治・高田祐一 2022 「日本全国における文化財オープンデータの現状と課題」『デジタルアーカイブ学会誌』Vol.6, No.2
- 特定非営利活動法人 ASP・SaaS・クラウドコンソーシアム 2014 『平成 25 年度 防災・災害情報のオープンデータ化・二次利用促進のためのガイド』
- 内閣府 2016 『科学技術基本計画』第 5 期 (<https://www8.cao.go.jp/cstp/kihonkeikaku/index5.html>)
- 野口 淳 2020 「発掘調査報告書とデータの公開利用—「記録保存」と情報のフロー、再現性・再利用性—」『デジタル技術による文化財情報の記録と利活用 2』奈良文化財研究所
- 野口 淳 2022a 「考古学・埋蔵文化財情報のオープン化」『デジタル技術による文化財情報の記録と利活用 4』奈良文化財研究所
- 野口 淳 2022b 「考古学・埋蔵文化財 GIS データの標準化、ファイルフォーマット、オープン化」『デジタル技術による文化財情報の記録と利活用 4』奈良文化財研究所
- 野口 淳 2023 「文化財データ・データベース・デジタルアーカイブスの公開について考える」『デジタル技術による文化財情報の記録と利活用 5』奈良文化財研究所
- 福島幸宏 2020 「文化財を真の公共財とするために」『デジタル技術による文化財情報の記録と利活用 2』奈良文化財研究所
- 堀木真美子 2019 「調査データの活用—整理と公開—」『デジタル技術による文化財情報の記録と利活用 1』奈良文化財研究所
- 堀木真美子 2020 「報告書掲載データと Web コンテンツ」『デジタル技術による文化財情報の記録と利活用 2』奈良文化財研究所
- 渡部展也 2019 「考古学・文化財におけるデジタル計測と情報活用の現状」『経済史研究』第 23 号 大阪経済大学