

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-

野口 淳（奈良文化財研究所）

Excavation Reports and the Disclosure of Archaeological Data: The Flow, Reproducibility and Reusability of Information Under the Policy of "Preservation in Records"

Noguchi Atsushi (Nara National Research Institute for Cultural Properties)

- ・情報のフロー／Information flow・再現性／Reproducibility
- ・再利用性／Reusability・整然データ／Tidy data

はじめに

考古学・埋蔵文化財保護のための遺跡発掘調査では、多様かつ多量の情報が記録される。本来、それらは「記録保存」の対象であるべきなのだが、実際には報告書として整理・編集・刊行されたものだけが残される傾向が強いのが現状である。これは学術研究のために、または文化財の保護・活用のために望ましい状況とは言えない。現状における問題を確認し、データの流通・利用の観点から展望をまとめた。

1. 『過去をつなぎ合わせる』

考古学・埋蔵文化財保護のための遺跡発掘調査（以下、発掘調査）では、多様かつ多量の情報が記録される。これらは情報の質、単位、記録方法・手段・媒体も様々である。また遺跡や遺物などの状態を直接計測・記録する一次情報から、調査者の観察、専門知識・経験による解釈を含む二次情報までを含む。一方で、対象ごと、あるいは記録の単位・方法ごとの個別的な情報もある。

発掘調査の目的は、第一に遺跡に残された様々な物質的資料・痕跡とその状況・関係を記録することである。しかしそれだけにとどまらず、個別的な情報を整理し、つなぎ合わせることによって、過去の場景やそこに継続した人類の活動、社会、文化を復元することが、高次の目標となる。遺跡に残された

資料・痕跡と状況・関係は、過去に存在していたものの全体ではなく、意図的に残された、あるいは偶然残った部分であり、その種類や量・状態は個別の条件に拠る。そしてほとんどの場合、それらはきわめて断片的であり、本来の脈絡や背景情報は失われている。

そこでそれらの断片を適切に取り出し、つなぎ合わせて過去の復元を目指すことが考古学の基本的な操作ということになる。『考古学研究法』の邦題で翻訳出版されている、V.G.Childe の古典的名著の原題が、“Piecing together the Past”（Childe 1956/近藤訳1965）であることは象徴的である。ばらばらの破片より、整理分類され接合復元された土器は情報量が多い。さらに多くの情報、証拠がつなぎ合わされることで、過去に関する知識が蓄積される。

2. 考古学におけるデータ・情報の組織化

伝統的または正統的（orthodox）な考古学のゴールは、そうした「過去の断片のつなぎ合わせを総合化した単著（monograph）をまとめることである。そしてその前段階として、論文（research paper）や報告（report）を出版することが必要となる。このため、個別的な断片的なデータより、体系的に総合化された意味内容の豊富な情報にプライオリティが認められる。

この枠組みは、情報科学における「可視化の階層」とも合致する。離散的要素であるデータ、それらを

つなぎ合わせた情報、情報を組織化した知識、知識の活用としての叡智は、発掘調査・整理作業における記録、それらを編集し記述・説明する報告書、特定の論点について分析・考察する論文、総合的な歴史叙述を行う单著に照応する（図1）¹⁾。より組織化された情報は、意味内容も増加する。上位の階層ほど、単位量あたりの情報の密度が増大・凝集するので、価値も高くなる。専門家・研究者の取り組みは、発掘調査・整理作業の記録を組織化し、上位階層へと昇華することを目指している。

一方で、この図式には、もう一つ重要な含意がある。各階層は単系の序列として位置づけられるのではなく、裾の広がったピラミッドを構成することに着目すると、上位階層は、十分な量の下位階層要素の上に成り立つことが理解される。单一の、またはきわめて限定的な記録だけでは歴史叙述に至るのは難しい。それ故に考古学は多量の記録・情報もある

が、これら膨大な記録の蓄積が日本考古学のピラミッドの裾野を支えている。

3. 「考古学ビッグデータ」の可能性

データ・情報の蓄積、すなわちストックの観点からみて日本考古学は大きなアドバンテージを有している。しかしデータ・情報の利用、それを可能にする流通、すなわちフローの観点（野口 2019b）から見た時にはどうであろうか。

現状では、より多くの関係者がアクセス・利用可能な形で流通している日本考古学のデータ・情報は、そのほとんどが発掘調査報告書（以下、報告書）の形態をとる。報告書は、定型的な構成をもち²⁾、印刷された冊子を媒体とする（文化庁文化財部記念物課監修2010）。近年では『全国遺跡報告総覧』の公開運用をはじめデジタル化・電子化も進むが、公式的には印刷物が基本であり、デジタル化されたPDF形

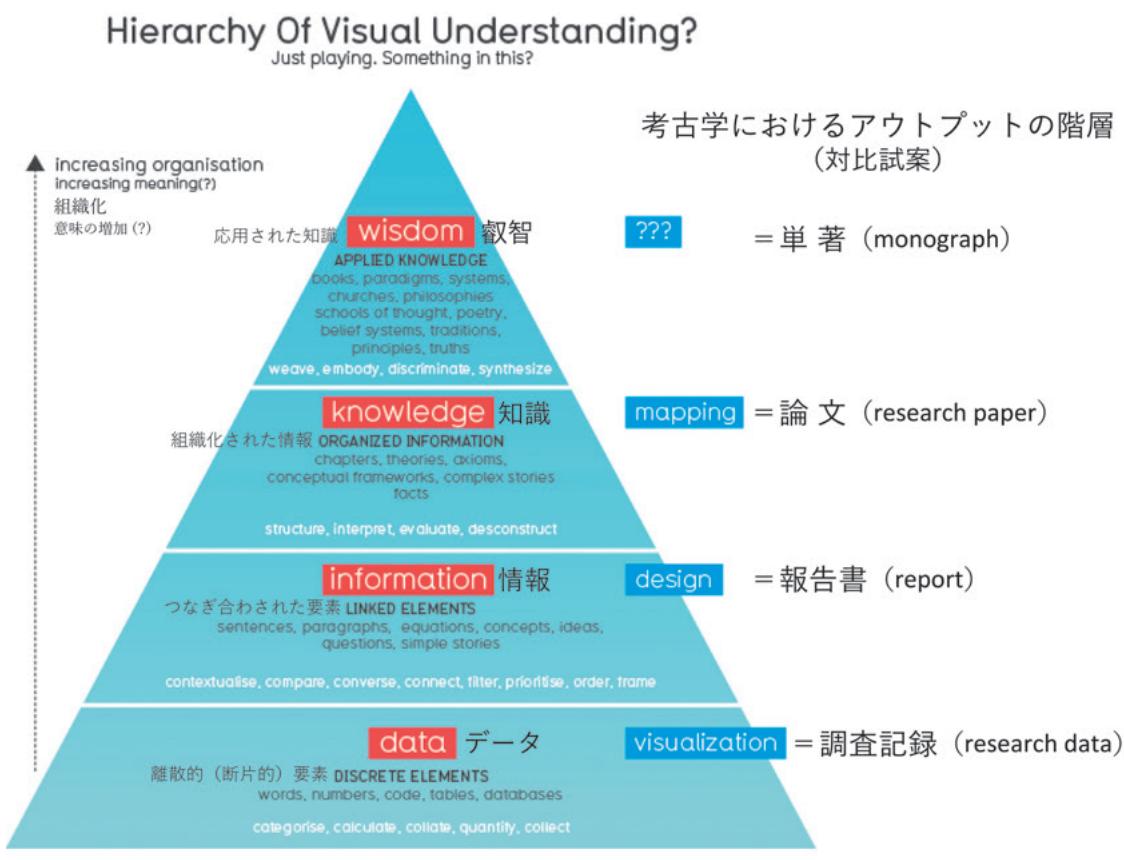


図1 情報科学における「可視化の階層」と考古学におけるアウトプット

式のものは、バックアップ（高精度）、普及公開（低精度）と位置づけられている（文化庁文化財第二課埋蔵文化財部門2019）³⁾。

こうした「報告書」の正確な刊行数は定かではないが10～15万点と推計されている（高田2019c）。過去の刊行物は遡及してデジタル化する必要があるが、近年はデジタル製版が基本なので、コンピューターを利用した情報処理可能な状態での公開・流通が進むだろう。そのデータ量は約97億文字、図・写真各500万件、表・グラフ50万件と推計されている（高田前掲。）。これらが利用可能な状態になる時、日本考古学は、人文科学のデジタル化（デジタル・ヒューマニティーズ）、データ・サイエンス利用のトップランナーとなる潜在的 possibility を有している（高田2019a, b）。

4. 報告書というボトルネック1：データ利用・参照の狭窄

ところで、報告書は原則的に発掘調査の事業単位（遺跡・地点・調査区）でまとめられる。学術研

究、あるいは普及公開などテーマ別・目的別に、複数の遺跡の情報を編集したり、集合的にデータを取り扱い分析するものではない。これを図1の構造に当てはめると、第3階層（情報＝報告書）と第4階層（データ＝調査記録）の関係が固定的であることを意味する。一方、上位の階層では、たとえば論文（＝知識）はテーマ・目的によって多数の報告書を選択的に参照することが可能であり、有機的に情報を組み合わせて多くの知識を作り出すことができる。単著と論文の関係も同様であり、さらに単著は下位の階層（報告書）の情報を参照利用することもできる（図2）。

しかし第3階層と第4階層では、報告書と遺跡・地点・調査区単位のデータセットとが一対一の固定的な関係となる。上下の階層をつなぐ結合（ノード）が多いほど、情報の組織化が進み意味内容が増加すると考えると、上位の階層群に対して、調査データと報告書の間での組織化・意味内容の増加が限定されていることになる。これは、複数の研究者がそれぞれ情報を参照・組織化してまとめる論文や単著に

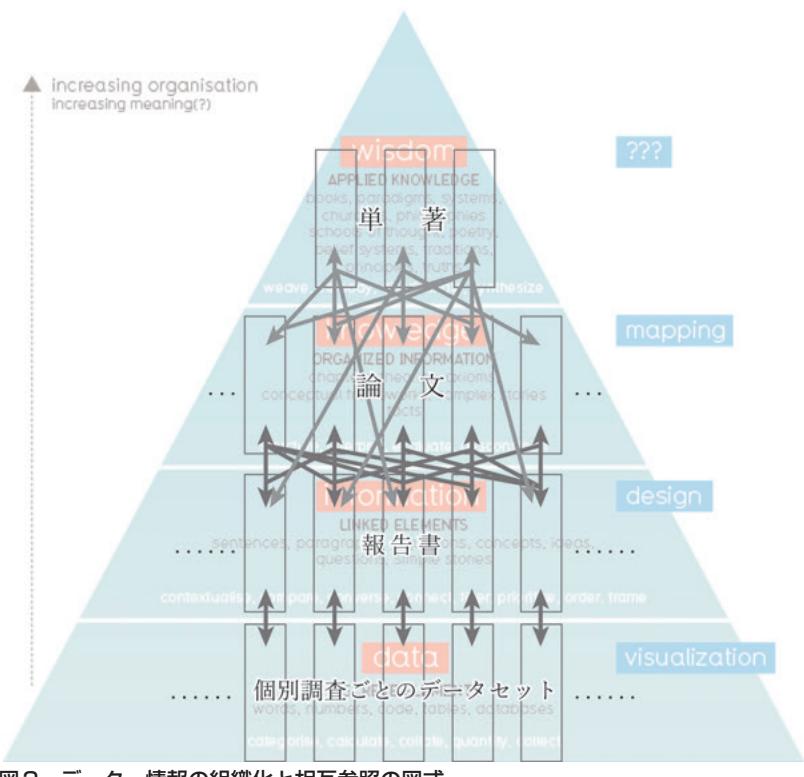


図2 データ・情報の組織化と相互参照の図式

対して、報告書では、報告者により整理され、記述・説明された調査記録だけが利用可能になっているからである。

このようにデータ・情報のフローの観点からは、調査記録から報告書をまとめる部分が狭窄部（ボトルネック）となっていると指摘できる。

5. 「記録保存」の理念と実情

同時に、現行の報告書はデータと分析過程の再現性に乏しく、調査記録が適切に整理・記述・説明されているかどうかの検証が事後的に困難であり「ブラックボックス化」しているという指摘（石井2019）にも、注意を払っておく必要がある³⁾。

図2のとおり、単著や論文は複数のものが、それぞれ同じ報告書を参照することが可能なため、解釈の差異はさておき、引用・参照しているデータ・情報に食い違いがある場合には、どこかに間違いがあるだろうと検証が可能である。同じような検証可能性を報告書に担保するためには、ある報告書の記載について、他の複数の報告書や論文が同じ調査記録を引用・参照できるような仕組みが必要である。しかし現状では、調査記録がそのまま公開・共有される仕組みはない。報告書では基本的に、調査記録のうち、選択・抽出され、報告者の視点・方法にもとづいて整理・編集されたものだけが記載・報告される。調査および記録の再現性はきわめて低い。

もちろん、調査者、報告者の手もとにはオリジナルの調査記録が残されているはずである。開発等に伴う事前調査はもちろんのこと、学術調査であっても、発掘調査は、遺跡の原状に対して不可逆的な変更を加えるものであるため、当初の検出状態から、調査の過程、そして結果にいたるまでの記録は、原状の変更すなわち破壊の代償・代替措置として保存されるべきである。しかし実際には、どこまで再利用・参照可能な状態で記録類が保存されているだろうか。

文化庁による2003年の報告⁴⁾では、図面・写真類あわせて全国で約1億437万点が保管されているが、

市町村レベルでは専用保管施設を持つところは1割未満であった。保管用控えの作成も図面・写真ともに2割程度である。また保管施設について「不明」とする回答も各設問に一定数（1.6～22.9%）含まれている（同報告図44～56）。そのすべてが、記録の所在自体が不明であることを示しているとは言い切れないが、実際に、報告書刊行が完了している調査記録をあらためて閲覧し、参照・利用することが困難な例が少なくないことは、多くが実感しているのではないだろうか。

6. 報告書というボトルネック2：再利用性

とは言え、一次データとしての調査記録は、多くの場合、参照・利用にあまり適さない形態・状態であることが多いだろう。ひとつには、それらの記録が、正規化され構造化された整然データ：tidy data（西原2017）ではないことがある。データの形（数値・テキスト）、単位・桁、書式、配列などが整っていない、雑然としたデータは、利用のために整理・調整・編集を行なわなければならない。再利用性という観点からは、このようなデータを、オリジナル（原データ）、一次データであるからといって公開・共有しても、再利用のためのコストが増大し効果が減ずる。場合によっては、障害の方が大きいかもしれない。

このため、何らかのかたちで整理・編集を行ない、一覧表、集計表、グラフ、図面化し、組織化された情報として掲載・提供する報告書が望ましいという評価もあり得るだろう。実際、位置座標や計測数値等を含む遺構・遺物の一覧表など、一次記録に近い（あるいはそのもの）を収録している報告書も少なくない。

しかしここでも、印刷された冊子形態の報告書故の問題が指摘される。

たとえば発掘調査現場ではx, y, zの座標数値で記録されている遺構の形状や遺物の出土位置は、製図され清書され製版される過程で、輪郭形状や相対的

な位置関係、しかも平面か断面（立面）に次元数を減じられた情報に不可逆的に変換される。一覧表や集計表に関しても、たとえば表の枠（セル）を結合したり、誌面（版面）における体裁・表示のされ方を重視するため、再デジタル化の際に支障となる事例が少なくない。さらに言えば、編集・組版作業の遙か以前の段階から、管理に適した整然データとしてではなく、印刷出力状態を整えることを優先とした取り扱いが当然視されている状況もある。

もちろん、図面中に基準点や参照情報が提示されていれば、デジタル化し再度座標を付与することが不可能ではない。しかし、変換前のデータが提供されるならば再利用性は高まる。何より、デジタル・データとして取得・作成される一次記録や、その整理・編集済みデータが、印刷物、およびその形式に即して編集構成された製版データ（PDF）としてのみ公開され、流通することは、多くの手戻りを生じさせ再利用のコスト増加と利便性の低下を招いている。コンピューターを利用した情報処理が必須となっている中で、これは大きな不利益である。

現行の報告書の形態は、データ・情報をまとめて大量の複製を作り流通させるという、グーテンベルク革命以来の印刷物に最適化されたものである。媒体がそれしかなかった時代には当然、最善の選択肢だったし、それを利用することに慣れ親しんできた人間にとっては今でも最良の手段であると言える。しかしコンピューター、インターネットの利用が一般化した現代に、データ・情報の提供を、单一の、従来型の媒体のみとすることについては、あらためて費用便益を検討しなければならない。

7. 情報化時代の情報の流通と利用に向けて

これまでの経緯と蓄積もある報告書は、それ故に使いやすい部分があることも確かである。情報化時代において、印刷物としての報告書をベースに、検索可能性を高め情報の利用・流通を促進することや（高田 2019b）、付加的に調査記録・データを公開し

利用可能にすること（堀木 2019）は、従来の蓄積と新しい情報流通・利用を接ぎ木するアプローチとなる。

一方、2016年の官民データ活用推進基本法（平成28年法律第103号）は、国及び地方公共団体に「オープンデータ」への取り組みが義務付けられた。これを受けて政府省庁、地方公共団体では、行政データの公開化が進められている⁵⁾。考古学・埋蔵文化財データを含む文化財についての統一的な取り組みはまだないが、隣接分野として国土交通省関連の測量、地質等調査、設計・工事完成図書等についてはすでに要領・基準等が定められている⁶⁾。公共事業等の工事関連では、データの互換・交換のために対応を要求されることもあるだろう。さらに同法の謳う政策が大きく変更されない限り、考古学・埋蔵文化財データが対象とされることも将来的にはあり得る。

また文化財の利活用に関しても、台帳等の基盤情報やメタデータの整備と流通が重要になることが予測される。しかしながら現状では、発掘調査時に取得された記録類は、報告書の作成・刊行を区切りとする傾向が強く、そこから先、博物館等における利活用にまで連続的・シームレスに受け渡され共有利用されている事例は少ないのではないだろうか。出土品について収蔵・管理が変更になるたびに登録と台帳の作り直しを行なうのではなく、発掘調査の時点から一元的・継続的な管理がなされるならば、各種の付随的・背景的情報を含めた取り扱いが容易になるだろう。

そして最も重要なのは、コンピューター、インターネットの利用が広く一般化している現在、こうしたデータ・情報の取り扱いは大規模なシステムがなくても、小規模な組織や個人レベルでも可能になっているという点である。ひとたび、データ・情報の公開・流通へと舵が切られるならば、さまざまなアクターによるさらなる利活用、アウトプットが見込まれる。

おわりに

高度経済成長期以降、毎年のように増加する発掘調査件数に対して、調査を実施し、記録を残し、報告書を刊行することが当面のゴールとして設定され、膨大な報告書が蓄積されてきた。しかし調査件数は、1997年をピークにすでに減少局面に入っている。一方で、コンピューター、インターネットの急速な普及は、データ・情報の流通と利用の基盤を根本から変えつつある。

考古学・埋蔵文化財保護について「1997年以降に最適化したシステム」への移行が要請されている（横山 2018）。その根幹は、情報化時代においては、データ・情報の公開と流通にあるだろう。蓄積された報告書にもとづいて考古学・埋蔵文化財保護の未来を拓くためにも、いっそうのデータ・情報の公開・流通の促進が望まれる。

【註】

- 1) 原図：Data, Information, Knowledge, Wisdom? *Information is beautiful* (<https://informationisbeautiful.net/2010/data-information-knowledge-wisdom/> : 2019年12月18日閲覧) に加筆
- 2) 埋蔵文化財発掘調査体制等の整備充実に関する調査研究委員会「行政目的で行う埋蔵文化財の調査についての標準（報告）」2004.10.29 https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/hokoku_06.pdf 文化庁ウェブサイト「埋蔵文化財」ページで公開 (<https://www.bunka.go.jp/seisaku/bunkazai/shokai/maizo.html>、文化庁文化財部記念物課監修 2010『発掘調査のびき－整理・報告書編－』pp.210-238に収録)
- 3) 印刷物では出力結果を容易に改変できない＝真正性を保持できるとする文化庁の見解（文化庁文化財第二課埋蔵文化財部門 2019: 表2）に対する石井の指摘は重要である。また紙面に固定されることで改変が困難になることは、真正性の確保という側面だけでなく、適切な対処がなされない限り修正を要する

誤情報が流通し続けることも意味する。この点については、『日本列島の旧石器時代遺跡データベース』（日本旧石器学会 2010）の地図化に際して、報告書記載の位置情報に含まれるミス・誤情報によって生じる問題を指摘している（野口 2019a）。

- 4) 埋蔵文化財発掘調査体制等の整備充実に関する調査研究委員会「出土品の保管について（報告）」2003.10.20 https://www.bunka.go.jp/seisaku/bunkazai/shokai/pdf/hokoku_06.pdf 文化庁ウェブサイト「埋蔵文化財」ページで公開 (<https://www.bunka.go.jp/seisaku/bunkazai/shokai/maizo.html>、文化庁文化財部記念物課監修 2010『発掘調査のびき－整理・報告書編－』pp.271-275に本文のみ収録)
- 5) 政府 CIO ポータル「オープンデータ」<https://cio.go.jp/policy-opendata>、2019年12月18日閲覧
- 6) 国土交通省「電子納品に関する要領・基準」<http://www.cals-ed.go.jp/>、2019年12月18日閲覧

【引用文献】

- 石井淳平 2019 「考古学情報の再現可能性－バージョン管理システム Git を利用した調査データの管理と公開－」『日本考古学協会第85回総会研究発表要旨』 pp.162-163
- 高田祐一 2019a 「デジタル技術を活用した発掘調査報告書のアクセス性向上の試行」『日本考古学協会第85回総会研究発表要旨』 pp.164-165
- 高田祐一 2019b 「発掘調査報告書の電子公開による情報発信とその新たな可能性」『奈良文化財研究所研究報告21：デジタル技術による文化財情報の記録と利活用』 pp.73-78
- 高田祐一 2019c 「報告書のデータ量を推計する」『文化財の壺』7: 4-5
- 西原史暁 2017 「整然データとは何か」『情報の科学と技術』67 (9): 448-453, https://doi.org/10.18919/jkg.67.9_448
- 野口 淳 2019a 「データの集成と全国遺跡法億総覧との連携利用－『日本列島の旧石器時代遺跡』データベースの場合－」『奈良文化財研究所研究報告21：デジ

タル技術による文化財情報の記録と利活用』 pp.86-90
野口 淳 2019b 「考古学・埋蔵文化財行政と情報処理－
　　ストックとフローの観点から－」『日本考古学協会第
　　85回総会研究発表要旨』 pp.156-157
文化庁文化財第二課埋蔵文化財部門 2019 「埋蔵文化財
　　保護行政におけるデジタル技術の導入について」『奈
　　良文化財研究所研究報告 21：デジタル技術による文
　　化財情報の記録と利活用』 pp.1-6
文化庁文化財部記念物課監修 2010 「第 VI 章 報告書の
　　記載事項」『発掘調査のてびき－整理・報告書編－』
　　同成社, pp.160-182

堀木真美子 2019 「調査データの活用－整理と公開－」
　　『奈良文化財研究所研究報告 21：デジタル技術によ
　　る文化財情報の記録と利活用』 pp.66-69
横山 真 2018 「三次元技術を考古資料の記録に用いる
　　ことの意義」『國史學』 226: 77-97
Childe, V. G. 1956 *Piecing together the Past: the
interpretation of archaeological data*. Routledge.
<https://doi.org/10.4324/9781315748597> (2014, Taylor
& Francis) ／邦訳：近藤義郎訳1964『考古学の方法』
　　河出書房