

# 考古学ビッグデータの可能性と課題

野口 淳 (奈良文化財研究所)

高田祐一 (奈良文化財研究所)

Archaeological Big Data: Potential and Challenges

Noguchi Atsushi (Nara National Research Institute for Cultural Properties)

Takata Yuichi (Nara National Research Institute for Cultural Properties)

・考古学ビッグデータ / Archaeological big data ・電子公開 / Internet publishing

## 1. 考古学ビッグデータ

考古学は、(おもに) 過去の物質資料にもとづき、人類の行動、生活、社会、文化および取り巻く環境などを解明する科学である。考古学資料は、生きた文脈 (systemic context) から離れ、埋没、遺存する過程で断片化し、散逸化するため、考古学者は、発掘調査などを通じて得られる様々なデータを記録し、整理・分析し、過去の復元・再構成を目指す。ここでは、限られた特定のサンプルではなく、できる限り多くのデータの取得が求められる。日本では、統計データの整備されている1975年以降、2016年までの間に、312,522件の発掘調査が行なわれている。(届出件数：文化庁『埋蔵文化財関係統計資料－平成30年度－』)そして戦前を含めて約125,000冊の発掘調査報告書等が公刊されていると推測されている(高田2019)。これは世界的に見ても他に類を見ない規模であり、そこに掲載されている図表、文章記載、一覧表等の数値情報は、まさに考古学ビッグデータといえる。

近年、コンピュータ技術の発展とともに、そうしたビッグデータを統計的に解析し、いままで知られていなかったあらたな知見が得られるようになった。国民共有の財産である文化財でもある考古学資料の、長年の蓄積を有効活用し社会に還元することが、考古学研究者に求められている。

## 2. 何が課題なのか？

報告書に代表される公刊された記録類は、分析対象となるデータの宝庫である。しかしそれらを効果的に活用するためには課題がある。

一つは、ほとんどの場合、報告書上のデータがそのままではコンピュータなどで処理をするのに適した状態にないことである。人間が読みやすい、理解しやすいように書式、体裁を整えられた文章や図、表は、情報処理に適しているとは限らないため、機械可読性を向上させる必要がある。

もう一つは、報告書が一次記録ではなく、報告者が編集した抜粋や要約であることである。発掘調査で得られた多量の記録を、整理作業を通じてまとめることは、報告書作成において必須である。一方で、その過程で取捨選択される情報があることも否めない。しかし報告書にまとめられる内容が最適のものであったとしても、異なる視点や分析方法による検証の可能性が開かれている方が、将来にわたってより豊かな成果が得られることは間違いない。

## 3. オープンデータ・オープンサイエンスへ

近年、科学、および学術的営みに全般ついて、透明性、公開性、そして手続き的再現性の要求機運が高まっている。考古学も例外ではない。オープンデータ、オープンサイエンスの潮流である。報告者

により取捨選択され、報告書として刊行される発掘調査記録について、そのオリジナルを、アクセス可能かつ再利用可能な形で公開することは、この潮流に沿った取り組みとなる。

発掘調査現場における記録から、整理・分析のほぼすべて過程がデジタル化、コンピュータ処理となっている現状では、公開化のための追加コストはそれほど嵩まず、公開化による利得が上回るが見込まれる。発掘調査～整理作業～報告書編集・刊行というワークフローに、データの構造化・整然化を埋め込むだけで、オープンデータへの準備は整う。

また、整理作業や分析の過程が記録され、報告されることが少ないという課題もある。仮に報告書をまとめるのに利用したデータがすべて利用可能な状態にあったとしても、整理・分析の手順や方法、過程が示されていないければ、第三者が報告書と同じ結果を得られるかどうかは定かでない。

この点について、報告書における整理・分析の手順、使用した方法・プログラム・機器と設定・条件等の明示を徹底することが望まれる。同時に、分析の方法そのものを、使用データセットとあわせて公開・共有することも重要である。

ビッグデータに取り組む情報化時代の考古学にとって、今後、データと手続きの開示と共有は必須の要件になると推測される。今回、このような観点から、考古学におけるオープンデータ・オープンサイエンスの取り組み推進の世界的第一人者である Ben Marwick 氏を招聘して研究集会を開催した。直接参加できなかった方を含めて、その意義と成果を共有するために、(本報告)が活用されることを期待する。

#### 【註】

高田祐一 2019「報告書のデータ量を推計する」『文化財の壺』7号