

データの集成と全国遺跡報告総覧との連携利用 -『日本列島の旧石器時代遺跡』データベースの場合-

野口 淳 (奈良文化財研究所)

Collection of archaeological site data in collaboration with Comprehensive Database of
Archaeological Site Reports in Japan

NOGUCHI Atsushi (Nara National Research Institute for Cultural Properties)

・「日本列島の旧石器時代遺跡」データベース/JPRA Palaeolithic site database
・遺跡データ/Site data

はじめに

遺跡発掘調査報告書には重要な考古学情報が多数含まれているが、構造化されていないため、必要な情報を抽出するには検索し、読解しなければならない。考古学において広く行われてきた主題別(時代・遺跡種別・遺物等)の集成・データベースの整備は、実は構造化されていない報告書内の考古学情報について外部参照基準を与えることにはほかならない。

全国遺跡報告総覧(以下、「総覧」)は収録報告書の全文検索が可能である。しかし、たとえば「ナイフ形石器」で検索を行うと、実際にナイフ形石器が出土した遺跡だけでなく、何らかの形で言及している報告書がすべて検索結果に表示される。これに対して、旧石器時代遺跡のデータベースまたはナイフ形石器出土遺跡の集成が整備されているとき、先の全文検索結果から旧石器時代の遺跡、実際にナイフ形石器が出土した遺跡だけを抽出することが可能になる。

つまり個別の報告書、およびそれらを網羅するデータベース(「総覧」)が主題別データベースのソースとなると同時に、主題別データベースが報告書データベース内の情報のアクセス性、再利用性を高めることになる。ここでは日本旧石器学会による『日本列島の旧石器時代遺跡』データベースを事例として、DB構築と利活用における「総覧」との関

係についてまとめる。

1. 『日本列島の旧石器時代遺跡』データベース

『日本列島の旧石器時代遺跡』データベース(以下旧石器DB)は、日本旧石器学会の事業として2003年から足かけ7年、学会内外の179名の協力を得て2010年に書籍として刊行された(日本旧石器学会編2010)(図1)。「遺跡」より下位の地点・調査区・文化層を基準単位(レコード)として、遺跡名、所在地、位置情報(緯度経度)、主要な出土石器(点数ではなく有無)や遺構(石器集中部、礫群、その他)と



図1 データベース『日本列島の旧石器時代遺跡』書籍版表紙(日本旧石器学会編2010)

表1 旧石器DBの構成

データベース項目			
1	ID	識別情報	
2	遺跡群名		
3	遺跡名		
4	地点・地区名		
5	文化層・石器文化		
6	遺跡群名読み方		
7	遺跡名読み方		
8	地点・地区名読み方		
9	所在地		
10	緯度	位置情報	
11	経度		
12	測地系		
13	標高		
14	ナイフ形石器		付加情報
15	台形(様)石器		
16	斧形石器		
17	剥片尖頭器		
18	角錐状石器・三稜尖頭器		
19	槍先形尖頭器		
20	両面調整石器		
21	細石刃・細石核等		
22	神子柴型石斧		
23	有茎(舌)尖頭器		
24	搔器・削器		
25	彫器		
26	砥石		
27	叩石		
28	台石		
29	礫器		
30	その他の石器		
31	草創期土器		
32	ブロック・ユニット		
33	礫群・配石		
34	炭化物集中		
35	その他の遺構		
36	特記事項		
37	文献(書誌情報)	(その他)	
38	調査歴		
39	作成年月日		
40	作成者		
41	更新年月日		
42	更新者		

表2 レコード数の概要

データベース収録件数	
遺跡・地点・文化層	16,771
遺跡	12,582
文献書誌	9,137

いった遺跡内容と情報ソースとしての文献・報告書のレファレンスからなるレコード(表1)が、日本全国で16,771件収録されている(表2)。なお、書籍版には全レコードを掲載する余裕はなく、全国(主要石器別)および都道府県別の遺跡分布図と、ID・遺跡名・所在地・緯度経度・標高・主要石器・遺構・文献レファレンスおよび文献一覧が表形式で掲載されているほか、添付CD-ROMには全レコードを含む.xls、.txt(タブ区切り)および.csvデータが収録された。

また書籍版は刊行部数も限られ、再版の予定もないことから、刊行後6年を経た2016年には、書籍全体のPDF版と、全レコードの.xls、.txt(タブ区切り)および.csvデータが、学会ウェブサ



図2 「データベース『日本列島の旧石器時代遺跡』」(<http://paleolithic.jp/data/index.htm>) 日本旧石器学会(2018年11月6日閲覧)

イト上に利用条件付きで公開された(図2:日本旧石器学会「データベース『日本列島の旧石器時代遺跡』」<http://paleolithic.jp/data/index.htm>)。

旧石器DBの最大の成果は、旧石器時代遺跡の一覧が内容情報とともに整えられ、位置情報と紐づけられたことにある(野口2012a)。旧石器時代遺跡について時期別、主題別(特徴石器型式・石材・遺構等)の集成や分布図作成は旧石器DB以前にも繰り返し行われてきたが、原データを再利用可能な形=データベースとして公開提供するのはこれがはじめてであったし、日本列島全域をカバーすることや収録情報量など、日本考古学においても稀有な事例である。空間的な情報の要約や予測モデリング(近藤2012)、各種地理情報と連携しての分析(野口2012b)など活用の可能性が期待された。またウェブ版の公開後は、宮崎県による「ひなたGIS」(<https://hgis.pref.miyazaki.lg.jp/hinata/>)にも収録され、アクセシビリティがますます向上した(野口2018)(図3)。

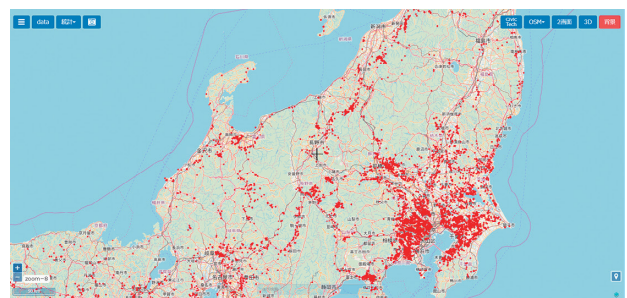


図3 「ひなたGIS」による旧石器DBの表示例(ズームレベル8、背景地図:オープンストリートマップ、<https://openstreetmap.jp/>)

2. 旧石器DBの更新改定

ウェブ版公開とともに学会内で検討されたのは、集成刊行後時間も経ち増加する追加情報を加えた更新版作成の必要性であった。ただし学会の体力－ボランティアベースで作業に従事する会員・協力者の負担と費用の両面－からみて、あらたな書籍版を編集刊行することは困難であると判断された。情報の追加は継続的に発生するものであり、書籍版では柔軟な対応が難しいという課題もある。そこでウェブ版をベースに、適宜、追加更新が可能な形式を模索することとした。

更新改定作業にあたっては、書籍版刊行後の追加情報の収録だけでなく、書籍版刊行時に積み残された課題の解決も必要とされた。提起された課題群は、**1) 位置情報の精度と確度、2) 遺跡内容情報の充実、3) 層序・年代など関連情報の充実、4) 文献書誌情報の参照性**である。しかしその多くは、旧石器DBそのものではなくデータソースとしての報告書に遡る問題である。そこで更新改定にあたり、「総覧」との積極的な連携を模索することとした。

旧石器DBをめぐる課題群のうち1～3は、報告書の所収情報とその記載が標準化されていないことに起因する。このうち**1) 位置情報**については後述する。**2) 遺跡内容情報**については、たとえば出土石器とその石材や検出遺構の内訳などの詳細なデータ化・数量記載への要望がある。しかし器種・石材の分類体系・用語および数量基準の揺らぎが大きい。そもそも統一的な基準策定は困難であり、将来的には類語整理・シソーラスの構築が望まれる（高田・国武2017）。**3) 関連情報**についても同様で、たとえば層序は地域的な基準によるものであり統一的な記載は難しい。広域火山灰や年代測定値については、それらの専門のDBが整備されているところでもある。そこで、更新改定にあたってこれらは書籍版の体裁・基準を大きく変更せず、将来的に外部DBを参照する方針とした。

4) 文献書誌情報の参照性の問題は、ひとえに遺

跡調査報告書が図書としての固有識別情報を持たないことに起因する。書籍版では、論文書籍等における引用参照に準じて編著者+刊行年を識別子としたが、入力担当者による表記揺れを解消できなかった。これについては、旧石器DB内での解決ではなく、外部DBを参照することとした。具体的には「総覧」の書誌ID、未収録書誌は国立情報学研究所学術情報ナビゲータ（CiNii）、国立国会図書館（NDL）による冊子・論文記事単位のIDをJPRA-IDに紐づけることとした。

3. 遺跡位置情報をめぐる課題

報告書を参照して取得されたはずの位置情報に多くの「誤り」や不統一があることは、書籍版編集時に遺跡分布図を作成する段階で確認された。これらはいくつかの異なる要因およびその組み合わせによる。

- a. **報告書自体の誤り**：収録記載された位置情報に、校正ミス、測地系の誤表記がある
- b. **情報基準の不一致**：記載位置情報は報告書収録調査地点の代表点か、周知の遺跡範囲の代表点か、または範囲を示す複数の位置情報の併記か、等
- c. **DB入力時のミス**：DB入力作業者のミス

このうちaは、報告書等の記載情報を地図化することではじめて判明する。いったん刊行された報告書誌を修正することは事実上不可能なので、誤情報の流通が続く。こうした誤りは決して少ないとは言えず、考古学・埋蔵文化財行政全般にかかわる重大な問題である。

bは位置情報の記載方法が標準されていないことによる。これも地図化によりはじめて判明するものでもある。誤情報とまでは言えないが、記載情報が目的とする精度・解像度を有しているかどうか分からないという点で、報告書誌の確からしさ・信頼性を損なっている。

cは人間が作業に関与する以上、避けられない課題である。旧石器DBでは、作業の指針・標準の明示

化・共有が不十分だったこともあり、たとえばa・bの問題に気付くなどした作業者が報告書記載情報の代替値を取得・入力した際に、学会データベースの書式(60進法表記規則)に、ウェブ地図の10進法の数値をそのまま入力することが多数生じた。これは緯度経度の分秒の桁に59以上の数値が入力されていることで初めて判明する。逆に0~59の入力値については地図化し検証しない限り誤りかどうかは分からない。ただし原データが「標準化」されていれば、チェック・修正・再入力等の人間の作業関与の度合いが低下するので問題解決につながるものもある。

ところで書籍版刊行時には、旧石器DBの収録位置情報にこれらの問題が含まれていることが認識されていたが、すべての位置情報を学会において検証する余裕はなく、極端な外れ値-都道府県境を越える、陸地から外れるなど-以外は入力値をそのまま掲載せざるを得なかった。そこで更新改定にあたっては、JPRA-IDに対応する地点・調査区レベルで、地理院地図などウェブ地図サービスを利用し、基本的にはズームレベル15またはそれより大きな縮尺で位置座標を確認、必要に応じて追加修整することとした。その後「ひなたGIS」に.csv形式で一覧情報の読み込み・書き出しの機能が追加されたこともあり、より効率的に作業を進めるための方策をさらに検討している。

なお位置情報については、報告書抄録DB、各自治体が整備する遺跡地図・台帳なども参照可能である。しかしそれらの位置情報が範囲なのか代表点なのか、周知の遺跡範囲なのか個別調査地点なのか、測地系は何か、等々、情報が標準化されていない点では報告書誌の記載情報と同じ問題を抱えている。繰り返しになるが、これは旧石器DBに限ったものではない。また、いったん刊行・公表されると内容が固定され変更できない印刷媒体の報告書や静的なDBの形式では解決できないものでもある。そこで旧石器DBは、「旧石器時代の遺跡」の位置に関する情報基盤を提供するものとして、独自に抽出確認し

た情報を保持し、遺跡・地点・調査区・文化層単位のレコードを基本単位とする識別子(JPRA-ID)に紐づけることとした。

4. 新しい旧石器DBの設計方針と考古遺跡DBの最適化

以上を踏まえ更新改定版では、旧石器DBの収録情報を、**A) 識別情報**、**B) 位置情報**、**C) 書誌情報**、**D) 付加情報**に位置づけ直すこととした。**A**は一義的にはJPRA-IDだが、人間が可読な要素として遺跡名、地点名、文化層名を含む。**B**は前述のとおり、旧石器DBが独自で取得記載する。**C**は基本的に外部DBを参照連携する。**D**は、報告書誌(**C**)より読み取り可能な情報であるが、利便性を高めるため付加される(野口ほか2017)。

その基本設計は、単一のDBに可能な限りの情報を集約・収録するのではなく、基本情報(**A+B**)を整備した上で、関連するDBと連携することで拡張可能にすることである。つまり最小の設定としては、**D**に位置つけた項目はなくてもよい。それらは**C**の文献書誌情報を通じて紐づけられた報告書誌から取得可能だからである(図4)。

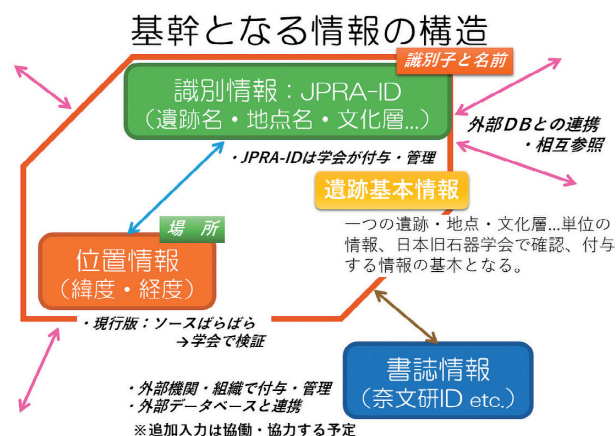


図4 旧石器DB(更新改定版)の構造

なおCiNiiやNDLの書誌DBには所蔵情報が含まれる。**C**の文献書誌情報にはこれら所蔵情報等へのリンクを含むため、アクセシビリティが良好ではない考古学文献類の検索性が向上する。さらに「総覧」をはじめ全文表示が可能な参照先が増えると、ネッ

ト (WWW) 環境下では、旧石器DBから直接文献を閲覧できるようになる。すでに「ひなたGIS」では、地図・分布図と内容情報が連動して表示される。更新改定版にデータが差し替えられると、地図・分布図から文献の閲覧までワンストップでの検索表示が可能になる。

このように新しい旧石器DBでは、収録情報と相互の関連付け、およびその表示を一体化したパッケージ型のDBではなく、鍵となる基本情報を整備し外部DBとの連携による拡張性を有したDBを目指している。文献書誌情報を「総覧」等外部DBに依拠し、地図表示については「ひなたGIS」と連携することで、書誌情報の収集確認検証や地図作成等の負担を軽減することができる。一方で、旧石器時代遺跡情報という専門性の高い情報については積極的に基盤を提供する。活動主体やその背景にかかわらず、可能な範囲でできることを積み上げ相互に参照・利用できるようにすることは、それぞれの持つリソースを最大化することにつながるのではないだろうか。

なお旧石器DBの更新改定作業はまだ緒についたばかりであるが、書籍版編集時の教訓を踏まえ、特定少数の作業者に過度な負担がかからないよう、分担して共同作業する方法も検討した。具体的にはクラウドサービス (Google ドライブ) を利用し、同時並行的に分散して作業を行なえるようにした (近藤ほか 2017)。作業内容の共有と手順の確認のためマニュアルも作成してワークショップ (マッピングパーティー) を各地で開催、順次、更新改定作業を進めている (神田ほか2017、野口2018)。

【註】

- 1) 日本旧石器学会「データベース『日本列島の旧石器時代遺跡』」<http://palaeolithic.jp/data/index.htm> (2018/11/6閲覧) ※全国版遺跡データのみ.xlsx

【引用文献】

- 神田和彦ほか 2017「日本旧石器学会データベース委員会の取組み (3)」『日本旧石器学会第15回研究発表シンポジウム予稿集』
- 近藤康久ほか 2017「日本旧石器学会データベース委員会の取組み (2)」『日本旧石器学会第15回研究発表シンポジウム予稿集』
- 高田祐一・国武貞克 2017「全国遺跡報告総覧における旧石器関係シソーラスの構築」『日本旧石器学会第15回研究発表シンポジウム予稿集』
- 日本旧石器学会編 2010『日本列島の旧石器時代遺跡—日本旧石器 (先土器・岩宿) 時代遺跡のデータベース—』
- 野口 淳 2012a「趣旨説明：シンポジウム「旧石器時代遺跡・立地・分布論の新展開」開催にあたって」『日本旧石器学会第10回研究発表シンポジウム予稿集』
- 野口 淳 2012b「地形・地質・考古遺跡情報の連携と旧石器時代遺跡の立地・構成について」『日本旧石器学会第10回研究発表シンポジウム予稿集』
- 野口 淳 2018「新しい『日本列島の旧石器時代遺跡』データベース—オープンデータ・オープンサイエンス時代の考古学研究を目指して—」『日本旧石器学会ニュースレター』38
- 野口 淳ほか 2017「日本旧石器学会データベース委員会の取組み (1)」『日本旧石器学会第15回研究発表シンポジウム予稿集』