

発掘調査報告書の頻出用語俯瞰図等の公開

考古学ビッグデータの定量的可視化：日本の発掘調査成果を1枚の画像で表現すると？

1、公開機能の概要

(1)機能公開

日 時：2017 年4 月27日（木）より公開

URL（全国遺跡報告総覧）：<http://sitereports.nabunken.go.jp/ja/visualization/term>

(2)機能の概要

- ・【発掘調査報告書の頻出用語俯瞰図】全国遺跡報告総覧に登録されている約 18,000 件の報告書（テキストデータ 15 億文字）に対し、考古学関係用語の出現回数を集計し、図化しました。
- ・【発掘調査報告書の頻出用語】よく使われる考古学関係用語を報告書ごとに集計し、多いもの（頻出用語）を報告書詳細ページに表示します。この頻出用語を絞込検索に使用できます。

2、経緯と期待される効果

【経緯（現状の課題）】

- ・日本では埋蔵文化財行政の着実な推進によって、膨大な発掘調査報告書（以下、発掘報告書）が発行され、長年にわたる日本考古学の研究蓄積があります。全国遺跡報告総覧においても 2017 年 4 月 23 日時点で、18,868 冊の資料、文字数 1,515,825,312 字、総ページ数 2,280,897 ページのデータ量があります。考古学は調査事例の蓄積によって研究が進展する蓄積型の学問です。「考古学ビッグデータ」というべき新たな可能性がある一方、情報過多によって処理不能となる「情報爆発」という一面があります。既に推定十数万冊の報告書が発行されており、さらに年間約 1,700 冊程度の発掘報告書が発行されているため、もはや人間がすべての発掘報告書を実際に閲覧し、内容を把握することは困難な状況となっています。

・全国遺跡報告総覧では、約 18,000 冊のテキストデータ約 15 億文字に対し、全文検索機能によって利用者が関心のある用語で検索することで、必要とする報告書を網羅的に探し出すことができます。しかし、発掘報告書では、遺跡の周辺環境の説明等で、当該遺跡とは直接関係のない用語も記述され、それらが全文検索時の一種のノイズとなっていました。ノイズを見極め、発掘報告書の概要を把握するには、すべてを閲覧するか、巻末の抄録を確認する方法があります。すべてを閲覧するには、膨大な手間がかかりますし、抄録は古い報告書にはついていません。文化財専門職員や研究者は、長年の研究活動で、ある程度発掘報告書を把握できている場合がありますが、一般市民の方が、手軽に発掘成果にアクセスするためにはハードルがありました。

【期待される効果と今後の発展可能性】

- ・【発掘調査報告書の頻出用語俯瞰図】
全国遺跡報告総覧に登録された発掘報告書の内容を可視化することで、日本考古学の成果の概略を把握することができます。今後、年代ごと、地域ごとに分析することで、学術的な動向や地域特性を客観的に分析できる可能性があります。

- ・【発掘調査報告書の頻出用語】

報告書の頻出用語のみを検索できるようになるため、直接的に関係のない検索結果(ノイズ)を除外し、高精度な検索が可能となります。検索結果で表示された報告書群に関連する頻出用語も自動提示されるため、それらを活用した絞込検索機能によって、体感的に必要なとする報告書を探すことができます。今後、頻出用語の構成を分析することで、内容が類似する報告書を利用者に自動提示することが可能になります。

- ・既に公開済みの英語自動変換機能（2016年8月24日公開）にて整備した和英対照の考古学用語およびシソーラスを連携させることで、今後、日本語の考古学用語を知らない場合でも、英語で日本の発掘報告書を体感的に調べることが可能になります。日本研究に興味がある海外の利用者に、より手軽に発掘報告書を閲覧できる環境となります。

【参考情報】

- ・全国遺跡報告総覧

「全国遺跡報告総覧」は、埋蔵文化財の発掘報告書を全文電子化して、インターネット上で検索・閲覧できるようにした“電子書庫”です。「総覧」は、全国遺跡資料リポジトリ・プロジェクトによって構築された遺跡資料リポジトリ・システムとコンテンツを国立文化財機構 奈良文化財研究所が引き継ぎ、運用しているものです。

貴重な学術資料でありながら、流通範囲が限られ一般に利用しづらい報告書をインターネット上で公開することで、必要とする人が誰でも手軽に調査・研究や教育に利用できる環境の構築を目指しています。

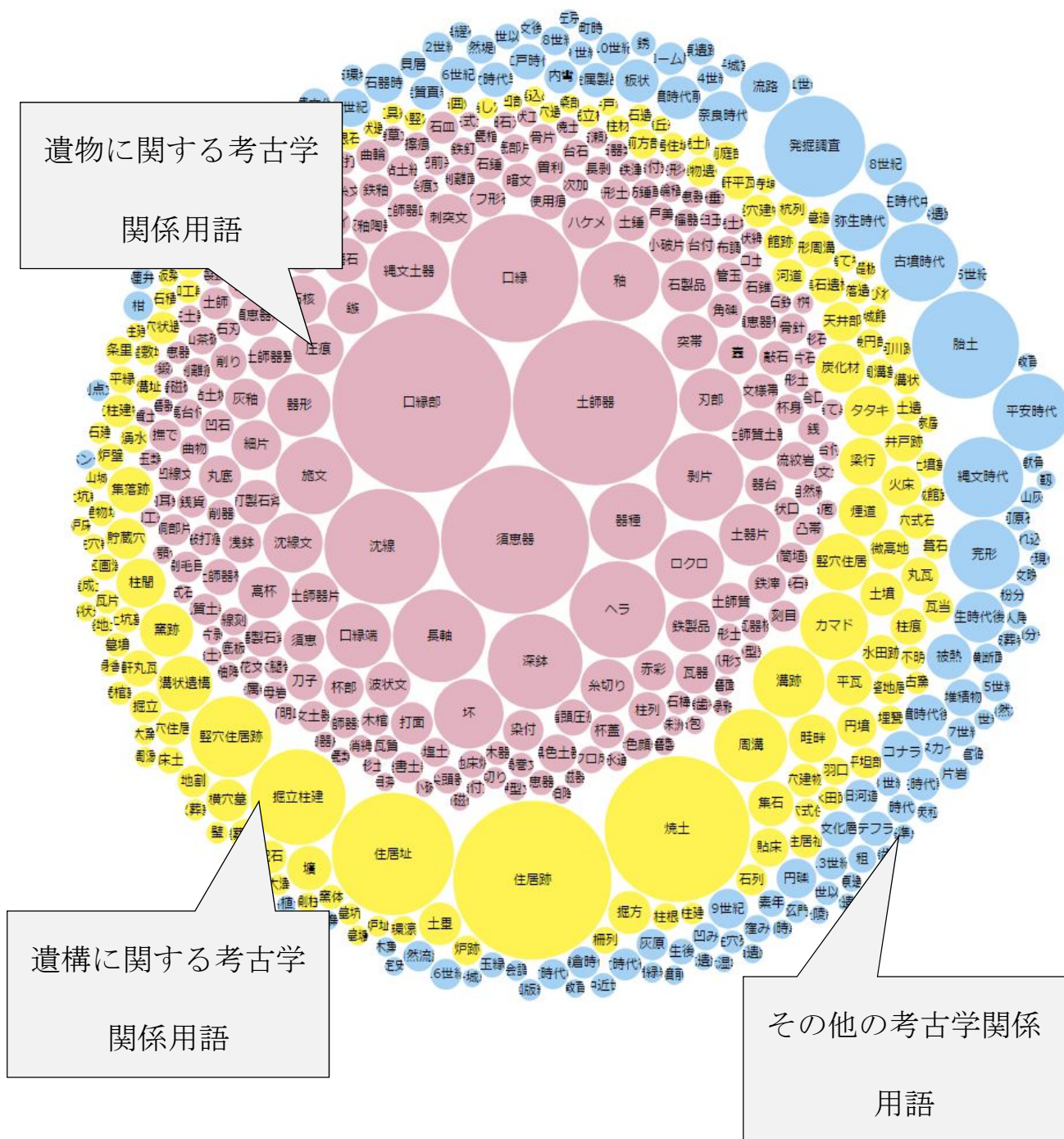
現在、全国の大学や自治体等の368機関の報告書類18,868件を収録しています（2016年4月23日時点）。PDF ページ数 2,280,897、収録文字数 1,515,825,312 と膨大なデータを保持しています。

お問い合わせ先 奈良文化財研究所企画調整部 高田祐一 Tel 0742-30-6711 Mail ytakata@nabunken.go.jp

画面サンプル ※公開前に変更になる可能性があります。

【発掘調査報告書の頻出用語俯瞰図】

発掘報告書では、遺物に関する記述が多いことが分かります。特に土器関係で「口縁」に関して多くの記述があることが判明します。



画面サンプル ※公開前に変更になる可能性があります。

【発掘調査報告書の頻出用語】

簡易検索

フリーワード

絞込

PDF全文を対象に する しない

検索

検索条件を追加

絞込

本文頻出用語

- 土器 7
- 穴住居 6
- 口縁部 5
- 暗褐色土 5
- 焼土 5
- カマド 4
- 口縁 4
- 坯 4
- 掘立柱建物 4
- 縄文時代中期 4
- 胎土 4
- 弥生時代 3
- 柳沢遺跡 3

検索結果 76 件

適合度順 並び順

1 2 3 4 5 ...

広島大学東広島キャンパス埋蔵文化財発掘調査報告書

副書名: 農場地区の調査

巻次: 1

ダウンロード

発行機一覧 > 全国 > 藤ノ越遺跡

藤ノ越遺跡

URL	http://earmas.enut.jp/sitereports/14575
ファイル	14575_1_藤ノ越遺跡.pdf (14.3 MB)
書名	藤ノ越遺跡
書名かな	ふじのこしいせき
副書名	県営畑地帯総合整備事業に伴う埋蔵文化財発掘調査報告書
巻次	
シリーズ名	静岡県埋蔵文化財調査研究所調査報告
シリーズ番号	171
編者名	足立 順司
編集機関	財団法人静岡県埋蔵文化財調査研究所
発行機関	財団法人静岡県埋蔵文化財調査研究所
発行年月日	20070110
作成機関ID	22000
郵便番号	4228002
電話番号	0542624261
住所	静岡県静岡市駿河区谷田23-20
報告書種別	報告書・概報

B!ブックマーク 0

G+1 0

G+ 共有 0

ツイート

いいね! 0 シェア

詳細ページ表示回数: 0

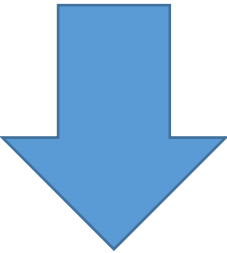
ファイルダウンロード数: 0

本文頻出用語

藤ノ越遺跡 古墳時代前期 足立順司 堅穴住居 溝状遺構 凹石 堅果類 縄文時代後期 縄文時代 弥生時代 向畑遺跡 弥生時代後期後葉 安養寺遺跡 掘立柱建物跡 土留め 焼土 炭化材 平坦部 地業 板状 台付 貼り床 土器片 建築材 凹地 貯蔵穴 土器細片 流れ込み 暗褐色土 堆積土層 丘陵部 テラス状 ハケメ 口縁 杯部 高杯 剥片 菖蒲ヶ谷遺跡 篠原修二 磨石

本文頻出用語

藤ノ越遺跡 古墳時代前期 足立順司 **堅穴住居** 溝状遺構 凹石 堅果類 縄文時代後期 縄文時代 弥生時代 向畑遺跡 弥生時代後期後葉 安養寺遺跡 掘立柱建物跡 土留め 焼土 炭化材 平坦部 地業 板状 台付 貼り床 土器片 建築材 凹地 貯蔵穴 土器細片 流れ込み 暗褐色土 堆積土層 丘陵部 テラス状 ハケメ 口縁 杯部 高杯 剥片 菖蒲ヶ谷遺跡 篠原修二 磨石



関心のある用語を選択することで絞込検索

簡易検索

絞込

詳細項目を表示

フリーワード

絞込 **本文頻出用語** 堅穴住居 × **本文頻出用語** カマド ×

PDF全文を対象に する しない

検索

検索結果 4 件

適合度順 並び順

千歳市オサツ2遺跡(1)・オサツ14遺跡 **ダウンロード**

副書名：都地区道當畑地帯総合土地改良事業用地内埋蔵文化財発掘調査報告書

巻次：

シリーズ名：[財団法人北海道埋蔵文化財センター調査報告書](#)

シリーズ番号：96

編著者名：[鈴木 信](#) | [三浦 正人](#) | [鎌田 望](#) | [千葉 英一](#)

発行(管理)機関：[公益財団法人北海道埋蔵文化財センター](#) - 北海道

本文頻出用語

- カマド 4
- 焼土 4
- 堅穴住居 4
- 須恵器 4
- 5c 3
- 口縁部 3
- 周溝 3
- 坯 3
- 掘立柱建物 3
- ロクロ 2
- 口縁 2
- 土器片 2
- 土師器 2
- 土坑 2
- 平安時代 2
- 暗褐色土 2
- 煙道 2
- 石錐 2
- 糸切り 2
- 縄文土器 2