

データベースの自由検索が不自由なとき ——標本資料の検索を変える一試み

文・写真
太田心平

基幹研究 ● 朝鮮半島関連の資料データベースの強化と国際的な接合に関する日米共同研究(2017-2019年度)

検索窓は自由の入口か

読者はデータベースを日々どう活用しているだろうか。もっとも使用頻度が高いのは、ウェブサイトを紹介したインターネット上の情報の検索だろう。1990年代後半に登場した検索エンジンは、人の手によりカテゴリー分けされたディレクトリ型だった。だが今日では、インターネット上のサイトを無差別に自動収録して検索対象とするロボット型が圧倒的に利用されている。フリーワードの(どんな入力文字でも受け入れる)検索窓に検索ワードを入力して検索すると、プログラムによって収集、解析、編集されたデータベースから、えたい情報が、しかも膨大に表示される。

こうしたフリーワード検索は、検索結果がきわめて合理的なものと思われがちである。たしかに、今日のフリーワード検索は、検索者にとって利用価値が高いと予測される順に情報が表示されるなど、利便性や公平性が向上してきた。ただ、これが本当に合理的なのかは、意見がわかるところだろう。各ウェブサイトの運営者たちも、検索サイトの運営者側も、広告や履歴収集などを通じて検索者の誘導を試みているからだ。誘導といえば悪いことのように聞こえようが、これは検索窓の技術を補うためのものでもあり、しばしば「キュレーション」とも呼ばれる。

キュレーションが必要になるひとつの理由は、すべての検索者が的確なフリーワードを使うとは限らないことである。世の中、検索に長けた人ばかりではない。適切な検索語彙が思い浮かばないということも、じつはたいへん多い。自分が詳しくない分野の情報を調べるのは誰でも苦労するものであるうえ、物事を調べるという習慣自体があまりない人びとにとっては検索窓が悩みのタネにすらなる。「どんなワードで検索しろと?」といったところだ。検索能力の格差を補うためにも、ウェブサイトや検索サイトの提供元がキュレーションすることは、意義をもつ。

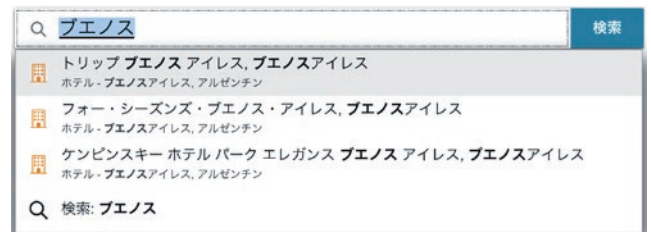
もちろん、このことばは、展示を企画することを意味するキュレーションから転用されたものだ。博物館は字義どおり「広範囲の物を擁する館」で、多種の資料が多量にある場所には、案内や解説が必要である。利用者のニーズや、その博物館の社会的な役割を反映させ、資料を配置し、解釈する作業が必要になる。外観からは何を展示している博物館かわからなかったり、展示場の解説パネルがデータベースの検索窓のような空欄で、自分で検索しろというものだったなら、さて、利用者はどうするだろうか。

人とモノ、人どうし、博物館と社会をつなげる

現代の検索サイトの進化を侮ってはいけない。多くの検索サイトでは、検索ワードを入力しはじめるだけで、他の検索者がよく検索している語彙を欄外に表示する技術など、利便性の向上がみられる。検索者たちがつながっていくのである。

あるいは、入力途中から候補を表示し、誘導していくものも

ある。検索者と運営者側のキュレーションがつながっているのだ。このキュレーションがわかりやすい一例が、ホテル検索サイトだ。検索者は自分が行き先の地名や泊まりたいホテルの名称を検索窓に入力するが、「見つかりませんでした」という結果が表示されることも多い。そうならぬよう検索窓は、検索者が入力中に、サイト側が提示できる検索候補を提示する。このことは、じつはそのサイトに登録されている地名やホテル名からしか、検索しえない仕組みになっていることを意味する。こうした検索は、フリーワード検索の外観をもちながら、実際はあらかじめ用意された経路に検索者を誘導する。となると、1990年代さながらのディレクトリ検索と大きく違わないわけだ。だが、2010年代の我われは、サイト運営者の手でキュレーションされたもの、つなげられたもの、つまりはディレクトリを、単なるフリーワード検索よりも画期的に便利だと思うようになった。



入力途中から候補を表示する検索の例(trivago.com)。

また、メタ検索サイトが広がり続けている。いちど検索をかけるだけで、複数の検索サイトを横断的に検索し、それらの結果を一覧表示してくれる検索サイトだ。インターネット通販の価格比較などもそれで、検索サイトどうしをつなげる取り組みである。これにより、各社のウェブサイトのひとつひとつ検索していくという煩雑な作業を、検索者は大きく省けるようになった。

一方、博物館のキュレーションはどうか。利用者が寄せる展示の感想や、データベースを使った履歴は、博物館に収集され、それぞれの改善のために活用される。こうやって、目に見えない形で利用者はつながっている。ただ、上記でホテル検索サイトや通販サイトの例で挙げたようなキュレーション、運営側による検索者の補助は、博物館のデータベースでどこまでおこなうのだろうか。第1に、ディレクトリ検索をどこまで親切に提供しえるだろうか。第2に、個別のデータベースの枠を越えて、横断的に同じディレクトリ構造を有するポータルは、用意されているだろうか。第3に、以前にデータベースを利用した検索者の体験が、以後の検索者に役立つというような、検索者をつなげていくキュレーションは、博物館のデータベースにおこなわれているだろうか。

日米の2つの博物館で

国立民族学博物館(以下「本館」)のフォーラム型情報ミュージアムという研究プロジェクトは、本館が所蔵する資料の収集元

の人びとや、その資料に関する研究をおこなっている館外者と協働して、そのデータベースを開発ないし強化するものである。その過程では、資料がしめす意味や象徴がさらに明らかとなるだけでなく、資料のデータベースのあり方や、所蔵と利用の方法、そして民族学的な博物館の未来像が探求されていく。成果として期待されるのは、データベースそのものの再構築、公開、運営に加え、それらの作業の過程からえられた知見の発信である。

このなかの1つのプロジェクトとして、本館の朝鮮半島関連資料のデータベースの再構築が進んでいる。このプロジェクトの目的は、データベースに多言語情報やコメント機能を追加することなどにより、関係者とともに資料情報を補強する作業であるが、それにとどまるものではない。

まず、ディレクトリ検索を導入することとした。これは、本館の展示場に配置されているように、すべての資料をどれかのカテゴリーに仕分けするという作業ではない。たとえば、既存のデータベースで「帽子」と登録されている標本番号「H0018277」の資料を例にとろう。これは男性用なので、「衣類と装飾具 > 男性用 > 帽子」というカテゴリーに入れるとともに、特権階層の衣装なので、「衣類と装飾具 > 特権階層用」というカテゴリーにも入れるべきだろう。また、この材質はウマの毛なので、「動物と植物 > 哺乳類 > ウマ」にも入れることとなる。なお、本館の既存のデータベースの検索窓で「朝鮮半島 特権階層」や「馬 朝鮮半島」を検索しても、この資料は出てこない。また、この資料の現地語名(ハングルで「갓」、アルファベット表示「kat」)には、それらの限定条件をもつ帽子という意味があるため、検索者が現地語を理解するかどうかで、伝わる情報量が変わることになる。



標本番号「H0018277」。朝鮮半島の特権階層の男性用に、ウマ毛などで作られた帽子。

こうしたカテゴリー分けは、残念ながら本館の標本資料データベースになかった。他の博物館では、資料をカテゴリー分けしてデータベースに登録しているところもあるが、各博物館が使うカテゴリーは独自のもので、国際的な統一がない。国際的な統一を優先してHRAF(Human Relations Area Files)のOCM(Outline of Cultural Materials)という既存のカテゴリーを使う手もありえるだろう。しかし、もともとOCMは学術論文等をカテゴリー分けし、研究者が必要な先行研究を参照しやすいように作られたものなので、標本資料を一般の人びとが検索するデータベースに対して使うのには不便も多い。あえて使うとしても、概念的で難解なOCMの用語を、研究者以外の人びととどう共有

するか模索する必要もあり、やはり挑戦的な取り組みになる。本プロジェクトでは、独自のディレクトリの構築に取り組んでいる。

つぎに、本プロジェクトは、本館のデータベースだけを再構築するものでないという特徴ももつ。本館とアメリカ自然史博物館(以下「AMNH」)の朝鮮半島関連の民族学資料データベースを、同時に再構築している。これにより第1に、もしどちらかの館が所蔵する資料になんらかの偏り(たとえば、陶器が多い)などがあっても、朝鮮半島の民族誌資料をより俯瞰するようなカテゴリー分けができよう。第2に、本館にとっては資料の英語名を決めるのにも参考となる。第3に、いちどに両館の資料が検索できれば、検索者の利便性もあがる。第4に、このデータベースが完成すれば、資料の収集元である韓国人びとにとっても意義がある。データベースを介して自文化に関する情報をやりとりする国際的な場ができるからだ。第5に、AMNHもまだカテゴリー分けをおこなっていないので、AMNH側、ひいては米国にもよい機会となる。

本プロジェクトのもうひとつの特徴は、標本資料だけではなく、写真や映像という資料もいっしょにデータベース化する点だ。上記の帽子も、実際に被っている人物の写真や映像とともに提供できれば、検索者はより多くの情報(どんな服と組み合わせているかなど)をえられるだろう。

こうしてキュレーションされたデータベースができ上がり、公開されつつ修正され、コメントも加わっていくことを願う。このシステムに参加してくれる第3、第4の博物館が現れてくれるようにと、準備もしている。日本語や英語やハングルを使っていない博物館の朝鮮半島の民族誌資料のデータベースとも、ひとつひとつの資料情報を翻訳するという手作業なしに、追加でつながりあえるだろう。本プロジェクトで作ったカテゴリーをその博物館の資料に適用してもらえば、各館の類似する資料が一覧表示できるからだ。



アメリカ自然史博物館で写真資料にタグ付けをする作業(2018年9月)。

おおた しんぺい

国立民族学博物館超域フィールド科学研究部准教授。アメリカ自然史博物館上級研究員。専攻は社会文化人類学、北東アジア研究。共編著書に『東アジアで学ぶ文化人類学』(昭和堂 2017年)、『한민족 해외동포의 현주소: 당사자와 일본 연구자의 목소리』(學研文化社 2012年)がある。