HOKKAIDO UNIVERSITY

Title	サハリン/樺太史研究DB（データベース）について ： 個人作成資料目録の統合と活用
Author(s)	中山, 大将
Citation	北方人文研究, 13, 171-173
Issue Date	2020-03-25
Doc URL	https://hdl.handle.net/2115/77253
Type	departmental bulletin paper
File Information	13_13_Nakayama.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

［研究会報告］

サハリン/樺太史研究 DB（データベース）について：個人作成資料目録の統合と活用

中山大将
（釧路公立大学）

1．サハリン/樺太史研究 DB（データベース）とは

サハリン樺太史研究者が個々人で蓄積した資料目録や収集資料の公開、研究に必要な基礎情報の公開を進めるための研究インフラであり、科研費「境界地域史への地域情報学活用：サハリン島ミクロ歴史情報データベースの構築と応用」（挑戦的萌芽研究、中山大将代表、2016–18 年度）の研究の一環として構築が開始された。

現在公開されているものとして、以下のものがある。

⑴樺太日日新聞 DB[1)]

⑵樺太地理情報 DB[2)]

⑶サハリン/樺太史研究文献 DB[3)]

2．DB 作成の動機

本 DB 作成の動機のひとつとなった事柄として、筆者が体験した、北海道大学附属図書館北方資料室所蔵の書誌情報不明資料であった『北樺太調査復命書』の書誌情報の確定過程を挙げることができる。

『北樺太調査復命書』は作成年も作成者も不明であったが、本文の記述からロシア内戦期に作成されたものであるとかわかり、唯一の人名として調査隊長「高清水清四郎」の名前が出てくる。他の資料から、高清水が北越製紙などでの勤務歴のある人物であることは判明していたが、この時期の勤務先や活動などは不明なままであった。

別件で『樺太日日新聞』の記事を確認していたところ「露領北樺太視察談」（1918 年 6 月 5、6、8、11、12、13 日）という記事を見つけた。この記事は、日本化学紙料株式会社が行なった 1918 年春の北樺太調査の講演記録であり、その調査日程・内容が『北樺太調査復命書』と一致することが明らかになった。

この偶然の発見までに 4 年間の時間が流れており、この一件は歴史研究における「経験」の重要性を物語っている。しかし、資料のデジタル化で、この部分の「効率化」を進めることができるのではないかと筆者には思えたのである。

1）https://nakayamataisho.wordpress.com/borderlandsdb/sakhalinkarafutodb/

2）http://app.cseas.kyoto-u.ac.jp/infolib/meta_pub/G0000311karafutoMAP2

3）http://app.cseas.kyoto-u.ac.jp/infolib/meta_pub/G0000311karafutoHIS2

3．DBの発想

従来の資料のDB化やデジタル化の形態として、各図書館のOPACや各種の資料画像公開サービスが挙げられる。前者は、資料名や書誌情報から資料を検索することができ、たいへんな便益を提供するものであるが、現状では、本文のようなそれよりも深いレベルの情報まではできない。後者は、国会図書館、アジア歴史資料センターを例として挙げることができ、一次史料をネット上で閲覧できることは、歴史研究の検証可能性の確保という意味できわめて有意義であるが、ここでも検索が可能であるのは資料名レベルにとどまり、より深いレベルの検索はできないという限界がある。

本DBでは「ミクロ歴史情報」という概念を設定する。これは、資料名レベルではなく、その中の「見出し」や「本文」がテキスト・データ化されたものである。もし、『樺太日日新聞』のミクロ歴史情報のデジタル化が完遂していたら『北樺太調査復命書』は数分で解決したと考えられる。

しかし、実際に『樺太日日新聞』全体にこの作業をするとなると極めて巨額の経費が要される。たとえば、『台湾日日新報』は、出版社によって本文まで完全デジタル化がなされているが、販売価格は約630万円にのぼる。この金額では、個人の購入は困難であり、予算規模の大きい図書館などが大型図書として購入するほかない。つまり、たとえデジタル化されても実質的に利用困難であれば、その意義は薄れてしまう。

そこで筆者は、研究者ひとりひとりは個人的に研究資料の目録を作っているはずであり、それを持ち寄れば、悉皆ではなくても低コストでデータベース化と公開が可能ではないかと考えるにいたった。

そして、試行として『樺太日日新聞DB』の作成を開始した。データベースのシステムとしては、「Myデータベース」（京都大学地域研究統合情報センター[4]）を利用しており、費用はかかっていない。「Myデータベース」の特徴として地理情報との関連付けも可能であり、これは後述する『樺太地理情報DB』と連携することで検索方法の多様性の向上を試みている。こうして作成されるDBには研究で生まれた副産物の有効利用、社会還元という側面があることは見落とさないでおいていただきたい。

4．サハリン/樺太史研究DBの意義

以下、サハリン/樺太史研究DBの具体的な意義について述べておく。

⑴ 『樺太日日新聞DB』

現在公開中なのは、農業編、漁業編、観光編、アイヌ犬橇編、中川小十郎巡視編である。

類似した先行DBとして、『神戸大新聞記事DB』があるが、同DBには、収録記事・範囲の限定性がある一方で、本DBには拡張性があり、現在も記事目録の提供者を募り新規の公開準備を進めている。また、同DBの収集基準が不明瞭である一方で、本DBでは、基となる記事目録の作成者とその収集基準（研究テーマ）が明確である。

自分の関連テーマでも目視では漏れていた記事や未見の記事を発見することが可能であり、「他人の目を借りる」研究が可能となる。また、本文までテキスト・データ化されている場合には、見出しからはわからない本文中に登場する人物や事項まで検索することができる。

特定トピックの検索の実例としては、「現在のブッセ（遠淵）湖の牡蠣は日本時代の養殖用の牡蠣

4）2017年1月に「東南アジア地域研究研究所」へと統合・改称。

が定着したものなのか」という問い合わせがあった際に、本 DB を用いて関連記事を発見することができたことが挙げられる。こうした素朴な興味に答え、それを新たな研究テーマにつなげるためのツールとしても利用することが可能であろう。

また、上述の通り、後述の樺太地理情報 DB との連携によって、記事に関連する地点の表示や地図からの検索（指定範囲に関係する記事の検索）も可能となる。

⑵　樺太地理情報 DB

本 DB は、集落レベルの地理情報（地名、緯度経度、ロシア語地名など）を入力したものであり、それらのデータから地点を検索できる。また、「My データベース」の特徴として、地図から指定範囲に含まれる地名を検索するということも可能である。

地点の確定にあたっては、戦前の陸海軍地図などと Google Map を照合して各地点を特定するという作業を行なった。また、現在のロシア語地名から日本語地名を検索できるようにもし、ローマ字表記にも対応できるようにしてある。

Google Map などの一般的な地図サービスでも、樺太の地名の一部は検索可能であるが、実はそれらで検索・表示可能な日本語地名はそもそも数に限りがあり、その位置関係は不正確な場合がある。たとえば、Google Map 上の「新問」は、現在の Новое 駅のことである。しかし、現在の Новое 駅は、実際には旧南新問駅であり、この駅は当時、新問の市街地とは離れたところにあり、当時の「新問」の位置を知るためには不正確と言わざるを得ない。

⑶　サハリン/樺太史研究文献 DB

サハリン樺太史研究会の『年度活動報告書』を利用し、主に 2008 年度以降のサハリン樺太史研究の書誌情報を掲載している。

一般的な学術文献検索サイトである Cinii との違いとして、漏れやすい論集内の個別論文を収録していること、Cinii には再録されないような小規模な研究会の会誌に掲載された論稿なども把握できる限り掲載していることが挙げられる。英語等の外国語書誌情報や要旨があるものはそれらの情報も収録し、異言語圏間の研究状況の相互把握の促進を図っている。

本シンポジウム関連論考でとりあげられた諸文献の書誌情報も順次本 DB に収録していく予定である。

5．境界地域史研究資料統合活用計画

「境界地域史研究資料統合活用計画」とは、サハリンに限らず他の地域も対象にしてこうした DB 化を進めるための研究プロジェクトであり、「境界地域史研究資料統合活用計画推進協議会」はそこで生まれた DB のうち閲覧制限のある資料（新聞画像など）を閲覧するための ID とパスワードを管理・提供するための団体であり、筆者によって運営されている。

6．今後の展望

『樺太日日新聞』DB は、提供者を募り DB 化する目録を拡大していく。樺太地理情報 DB では、アイヌ語地名、旧露領時代地名、山岳や河川、施設なども今後収録していくことを予定している。サハリン/樺太史研究文献 DB は、情報の更新は当然ながら、地理情報 DB との連携、要旨・本文中の地名の抽出を行ない、地名や地図から関連研究を検索できるような方法の実現を試みる。