

「トルキスタン集成」のデータベース化プロジェクトについて

帯谷 知可

「トルキスタン集成 Туркестанский сборник」（以下、見出しとプロジェクト・データベース名以外の部分では「集成」とは、ロシア帝政期の中央アジアに関するロシア語およびヨーロッパ諸語の文献・刊行物から成る一大コレクションである。初代トルキスタン総督カウフマン К. П. Фон Кауфман により、征服間もない新たな植民地中央アジアをよりよく知るための情報源として構想され、その命によって編纂が開始された。当時の新聞・雑誌の記事や論文、単行本、地図、統計資料などが広く収集され、同一の判型に製本し直され、第1巻から順に巻番号を付されて集積されていった。今日コレクション原本を所蔵しているウズベキスタン A. ナヴァーイー記念国立図書館 Alisher Navoiy nomidagi O'zbekiston Milliy kutubxonasi（以下、ナヴァーイー図書館）の希少本室には計 594 巻から成る「集成」が残されている。

コレクションとしては世界に1セットのみ存在するという、この貴重な資料集成は、2002～2004年にかけてタシュケントにおいて現地の出版社によりデジタル撮影され⁽¹⁾、現在では世界の数か所にその複製が所蔵されるに至っている。日本国内では、筆者の所属先である京都大学地域研究統合情報センター（以下、CIAS）⁽²⁾、北海道大学スラブ研究センターが図書館において利用に付しているほか、NIHU プログラム・イスラーム地域研究・東京大学拠点にも一部が所蔵されている。

CIAS ではここ数年、「集成」のデータベース化が進められており、試行錯誤を経て、最近ようやくある程度の展望が見えてきたところである。ここでは、その現状を紹介させていただきたい。

⁽¹⁾ デジタル化撮影の準備段階の作業には筆者も少なからず関与しており、当時の新プログラム方式による地域研究「現代イスラーム世界の動態的研究」（通称「イスラーム地域研究」、1997-2001年度）第3班「イスラームと民族・地域性」から一部支援を得て行われた。

⁽²⁾ 元々は国立民族学博物館に附置されていた地域研究企画交流センターが所蔵していたが、2006年の同センターの廃止に伴い、その所蔵資料が同年新設のCIASに移管されたという経緯がある。

1. 「トルキスタン集成」の概要

(1) 編纂の経緯と歴代の編者たち

まずКасымова [1984]⁽³⁾に従って、「集成」編纂の経緯を歴代の編者を中心に概観しておこう。

1867年、トルキスタン総督府が設置されたその年に、カウフマン総督は、中央アジアとりわけトルキスタン地方に関する書誌が必要であると思い立った。その意向を伝えられたペテルブルグの書誌学者メジョフ В. И. Межов (1830-1894) は、書誌だけでなく文献資料そのものを収集し手元に置くことが重要だと提案し、これはカウフマンの歓迎するところとなった。こうしてメジョフは中央アジア関連文献を集めては製本し、毎年数十巻をペテルブルグからタシュケントへ送ることとなった。1882年のカウフマンの没後もこの仕事は続けられたが、第416巻までが作成されたところで、1888年ローゼンバフ И. О. Розенбах 総督は財政的困難を理由に「集成」編纂の停止を命じた。

それから約20年の空白を経て、タシュケントにあるトルキスタン公共図書館の業務改善のための監視委員会のイニシアティブによって、1907年「集成」編纂は再開された。編纂の舞台はペテルブルグからタシュケントへと移った。当時の同図書館長や著名な東洋学者、民族学者らがこれに関与したが、中心となったのは中央アジア関連文献に精通した書誌学者ドミトロフスキー Н. В. Дмитровский (1841-1910) である。彼のもとで127巻、すなわち第417巻～第543巻が編纂された。

1910年には事業費不足のため編纂作業は小休止を余儀なくされ、それに続いてドミトロフスキーが死去した。これを受けて、図書館監視委員会のメンバーでもあった優れた中央アジア研究者セミョーノフ А. А. Семенов (1873-1958) に「集成」編纂が委ねられることとなった。彼によってロシア革命前夜までに48巻、すなわち第544巻～第591巻が作成された。

「集成」はロシア革命とその後の混乱を無事生き延び、ウズベク・ソヴィエト社会主義共和国（以下、ウズベク共和国）国立公共図書館と改称した図書館に保管され続けた。ウズベク共和国において「集成」を書誌学的な研究対象ともしていた書誌学者ベトゲル Е. К. Бетгер (1887-1956) が、1939年、チェレンチエフ著『中央アジア征服史』(М. А. Терентьев, *История завоеваний Средней Азии*, Санкт-Петербург, 1906.) 全3巻を「集成」第592巻～第594巻として追加した。こうして今日私たちが見る「集成」のコレクションとしての全体が完成した。

「集成」はメジョフ、ドミトロフスキー、セミョーノフという3人の編者が編纂を担当した部分ごとに異なる様相を呈している。例えば、セミョーノフはそれ以前には収録されてい

⁽³⁾ 本書のうち「集成」の編纂史について重要な基本情報が記述されている12～48ページは帯谷 [2008: 6-25] に邦訳を掲載している。

た新聞記事・論文をあえて除外して、編纂にあたった。そうした相違は、当時の編纂をめぐる周囲の事情や、3人各々のアプローチや、あるいは能力を反映したものでもあったと言えるが、結果としてコレクションとしての「集成」の体系性や一貫性という点でマイナスに作用した面もあることは、利用にあたって念頭に置いてよいだろう。

(2) 冊子体インデクス

「集成」の各巻の冒頭には一部のものを除いて目次が付けられているが、これとは別に、冊子体のインデクス（索引）が複数存在する。このことは、編纂史を検討する上でも、「集成」を利用する上でも注目すべきことである。

筆者が入手しえたインデクスとして以下の4点があり、ここでは便宜的に各々インデクス I、II、III、IV と呼ぶことにする。

インデクス I: *Туркестанский сборник сочинений и статей относящихся до Средней Азии вообще и туркестанского края в особенности, составляемый по поручению г. Туркестанского генерал-губернатора К. П. фон Кауфмана В. И. Межовым. Томы 1-150. Систематический и азбучные указатели сочинений и статей на русском и иностранном языках*, Санктпетербург, 1878.

インデクス II: *Туркестанский сборник сочинений и статей относящихся до Средней Азии вообще и туркестанского края в особенности, составляемый по поручению г. Туркестанского генерал-губернатора М. Г. Черняева В. И. Межовым. Томы 151-300. Систематический и азбучные указатели сочинений и статей на русском и иностранном языках*, Санктпетербург, 1884.

インデクス III: *Туркестанский сборник сочинений и статей относящихся до Средней Азии вообще и туркестанского края в особенности, составляемый по поручению г. Туркестанского генерал-губернатора И. О. Розенбаха В. И. Межовым. Томы 300-416. Систематический и азбучные указатели сочинений и статей на русском и иностранном языках*, Санктпетербург, 1888.⁽⁴⁾

インデクス IV: О. В. Маслова, *Систематический указатель к тт. 417-591 Туркестанского сборника. Ч. 1. (1-2411). Ч. 2. (2412-3630)*, Ташкент, 1940. (タイプ打ち、未刊行)

基本的にはこれらは「集成」の第1巻～第591巻に所収された資料を巻順に4段階に分けてカバーしている⁽⁵⁾。第592巻～第594巻は、上述のように、1点の書籍から成るので、この

⁽⁴⁾ インデクス III には対象とする巻が「300～416」とあるが、「301～416」の誤りであろう。

⁽⁵⁾ インデクス IV は第417巻～第591巻に所収された資料のうち、ロシア領トルキスタンに関連する資料のみを採録したものである。Касимова [1984]によれば、インデクス IV が対象としているのと同じ巻

最後の3巻がインデクスに含まれていないことは大きな問題ではなく、これらのインデクスがあれば、「集成」のロシア領トルキスタン関連の所収資料については検索が可能ということになる。(ただし、後述するように、これらのインデクスには問題もあり、実際には容易には検索不能なものもある。)

インデクスⅠ～Ⅲは「集成」コレクションの最初の編者であったメジョフ自身が作成したものである。インデクスⅣはドミトロフスキー、セミョーノフが編纂した部分について、ソ連時代になってからウズベク共和国の書誌学者マスロヴァが作成した。いずれも当該巻の所収資料をカテゴリー分類した上で書誌情報が採録されており、インデクス毎に書誌情報に通し番号がふられている。この通し番号を単純に合計すると、書誌件数は第1巻～第591巻の合計で8343となる⁶⁾。また、各インデクスには人名・事項等のアルファベット順索引も付いている。

インデクスⅠ～Ⅲのカテゴリー分類は、学術分野や地域などによるもので、相互に若干の相違はあるもののおおむね共通である。インデクスⅣのカテゴリー構成はそれらとは異なっており、このインデクス作成当時のソ連のイデオロギーや歴史観を反映したと思われるようなカテゴリーが追加されている点が特徴的である。また、特にインデクスⅣにおいて数多く目につくが、一部の資料についてインデクスの編者による資料の内容紹介や解説が付されている。

筆者はこれらのインデクスを眺めながら、まずはこれらに掲載されている書誌情報をひとつのフォーマットで統合し、「集成」全体の書誌情報として整備する必要があるのではないかと思ひ至り、科学研究費補助金・基盤研究(C)「地域研究資料としての『トルキスタン集成』に関する総合的書誌研究」(平成2005年度～平成2007年度)においてその作業に着手した。その成果として、インデクスⅠ～Ⅳの書誌情報をすべて、ロシア語の書誌については現代ロシア語の正書法に修正した上で、エクセル・ファイルに入力し、暫定的な性格のものではあるが書誌一覧リストを作成した[帯谷:2008]。

について、ロシア領トルキスタン以外の、隣接する諸地域に関連する資料を採録した、ベトゲル編のインデクス(タシュケント、1948年、タイプ打ち、未刊行)がナヴァーイー図書館に存在する。マスロヴァ編のインデクスⅣと、このベトゲル編のインデクスは、第417巻～第591巻について補完関係にあることになるが、後者については筆者はまだ現物を確認できていない。

⁶⁾ ただし、番号の抜けやひとつの番号に対する複数書誌の登録などが見られ、また後述するように書誌のカウント方法にも問題があるケースがあるので、「集成」全体としての書誌数はまだ正確につかめていない。

2. 「トルキスタン集成」データベース

筆者の所属先 CIAS では、地域情報学 Area Informatics の構築が活動の柱のひとつとなっており、その一環として地域研究のために有用なさまざまなデータベースの開発とそれらを統合するシステムの構築が行われている（CIAS データベース一覧 <http://www.cias.kyoto-u.ac.jp/database/>）。「集成」についても、地域研究資料の基盤整備と情報化という観点から、収奪的でない現地資料収集のあり方として現地との協働により希少資料をデジタル化し、さらにそれを国際的にも共有し、現代にみあった活用方法を検討していく一例として、そのデータベース化が推進されることとなった。その第一段階は、書誌情報検索データベースの構築であり、第二段階は書誌情報と資料原本の画像をリンクさせ、資料そのものの閲覧まで PC 上で可能な形へ展開させることである。

(1) 書誌情報検索データベースの構築

インデクス I～IV 掲載の書誌情報を統合したエクセル・ファイルが得られたことを受けて、それを CIAS のデータベース・システムにのせる実験的な試みが行われ、2009 年度に『「トルキスタン集成」データベース暫定版』として公開した。

このデータベースでは、キーワードを入力することによって、簡易検索、詳細検索により「集成」所収資料の書誌情報と、当該資料が「集成」の第何巻の何ページにあるかを検索することができる。キーワードとして、資料タイトル、編著者・訳者名、出版地、出版社、出版年、掲載紙・誌名のほか、インデクス I～IV における分類カテゴリーとインデクス編者による内容紹介・解説に含まれる語がヒットする。また、詳細検索画面では、検索したい資料の言語、種別（単行本か、新聞・雑誌の記事・論文か、その他か）、「集成」の特定の巻を指定することも可能である。

2012 年 1 月末現在、エクセル・ファイル上の元データの複数回のチェックとその結果のデータベースへの反映を経て、現行の「集成」データベースはインデクス I～IV の記載をほぼ忠実に反映したものとなった。従来、冊子体の各インデクスを個々に開いて書誌情報を探さなければならなかったものが、一度に PC 上で検索できるようになったわけである。

以下に述べるようにまだいくつもの課題を抱えた形ではあるが、「集成」書誌情報検索データベースは次のアドレスに置かれており、インターネット上で利用することができる。

http://area.net.cias.kyoto-u.ac.jp/infolib/meta_pub/G0000003Turkestan

(2) 書誌情報検索＋資料閲覧データベースへ

「集成」の利用にあたって、PC 上で書誌情報検索をした後に、書誌情報からそのまま資

料画像にジャンプし、読むことができたとしても便利になる — しごく当たり前の発想であるが、これを実現に移すにあたって、数々の問題に直面することになった。技術的には、CIAS が所蔵する「集成」CD122 枚の JPEG または TIFF 画像を PDF 化し、書誌情報に 1 件ずつリンクする、というだけの作業である。しかし、2010 年度に上述のエクセル・ファイルをもとにこの作業を開始してみたところ、リンクできないものが山のように出てきたのである。

一番の問題は、インデクス I ～ IV に記載されている、資料の「集成」中の巻番号、ページ番号の不備にあった。それらが間違っている、あるいはページ番号の記載がないものがあるにも多かったのである。そこで、あらためてエクセル・ファイル上の書誌情報を逐一、資料画像と突き合わせる作業を行うこととなった。エクセル・ファイルはインデクス I ～ IV の掲載順に入力されているので、まずはその順に従って頭から照合を行っていった。これについては 2011 年度に終了することができたが、まだなお行方不明のものがかかなりある状況だった。次に、エクセル・ファイル上の書誌情報を「集成」の巻順に並べ替え、今度は資料画像を第 1 巻の冒頭から順に見ながらエクセル・ファイルと照合することとした。この過程で、予想はしていたのだが、書誌情報がインデクスに採録されていない資料がある、インデクスに 1 件の書誌として採録されているものが実際には複数の資料から成る、資料のタイトルとインデクスのタイトルが一致しない、といったケースがかかなりあることが明らかとなり、それらをエクセル・ファイルに追加入力する作業も同時進行で進めることとなった。

この照合・追加入力作業は現在も継続中なのだが、近日中にも照合済みの分について、5000 件程度は現行データベースに資料画像をプラスし、限定つきながら公開することができるのではないかと見通しを持っているところである。これによって当初構想した「トルキスタン集成データベース」の原型が目に見える形となるわけである。

3. 課題と展望

以上に述べたように、「集成」のデータベース化は、データベース暫定版の公開以降も、書誌情報の点検・整備と問題点の発見・解決が同時進行しているような形で進んでいる。その課題として、筆者は現在大きく分けて二つのことを念頭に置いている。

第一に、書誌情報検索+資料閲覧データベースとしての「集成」データベースのさらなる改良である。このうち、CIAS のデータベースの共通の仕組みとして解消されうる問題として、多言語間の横断検索機能（現行ではロシア語資料はキリル文字で、英語資料はローマ字でキーワードを入力しないとヒットしない）、ヒットしたキーワードがどこにあるのかを示す機能、インデクス I ～ IV についてもそのオリジナルの画像を PC 上でばらばらとブラウ

ジングする機能の導入などがあげられる。また中央アジア研究者として、資料の利用者としての視点から改良しうる点としては、より便利で魅力的なインターフェースの検討、現代の視点からキーワードを追加することによる検索効率の向上などがある。キーワード追加については、2010年度にスタートさせた CIAS 個別共同研究ユニット『『トルキスタン集成』のデータベース化とその現代的活用の諸相』(<http://www.cias.kyoto-u.ac.jp/project/kyodo2010-63>) においてメンバーの皆さんのご協力のもとに具体化させているところである。

第二に、インデクス I～IV の問題点の解消である。期せずしてデータベース化の過程で数々の問題点が明らかになってきたが、まずはそれをしかるべく整理・提示する必要があるだろう。その上で、すべての書誌情報を正確に網羅した、いわばインデクス改訂版と、巻別目次を作成してはどうかと考えている。「集成」の利用者や上述の共同研究メンバーの間でもこのような要望はあるのだが、国際的な資料共有の結果として、資料原本が存在する現地への還元という観点からもインデクス改訂版と巻別目次の作成は大きな意義を持つだろう。

「集成」に関連する先行研究や上述の共同研究における議論の中でも指摘されてきたように、「集成」には、必ずしもロシア帝政期の中央アジア関連の重要文献・資料が「網羅的」「体系的」に収集されたわけではない、ペテルブルグとタシュケントで入手可能だった文献・資料に限定されているためカザフ草原についてはかなり手薄である、現地語文献にはほとんど目が向けられていないなど、種々の欠点や制約は確かにある。しかし一方で、ロシア帝政期の中央アジアについて、特定のテーマについて何かしらの情報を得たい場合に、まずは紐解いてみる価値のある、玉石混交の膨大なコレクションであることは間違いない。そのための有効なツールとして、ぜひ「集成」データベースをご活用いただければ幸いである。また、ご活用いただくことでそこからフィードバックを得て、さらなる改良のための工夫を今後も引き続き検討していきたいと考えている。

参考文献

帯谷知可 2008 『地域研究資料としての『トルキスタン集成』に関する総合的書誌研究』(課題番号 17510215、平成 17 年度～平成 19 年度科学研究費補助金基盤研究 (C) 研究成果報告書) 京都：京都大学。

Касымова, А. Г. 1984. *Туркестанский сборник*, Ташкент: Узбекистан.

(京都大学地域研究統合情報センター)