

ビッグデータ時代における中国人留学生史研究

郭 夢 珪（神奈川大学）

本会議では、地方留学生研究が重要な論点の一つとして位置づけられていた。筆者は「清末における江蘇省崇明県の留日学生」と題して報告を行った。江蘇省と浙江省は、留学生の派遣時期と人数の両面において、清末期の留日学生ブームを先導した地域であり、地方留学生研究の意義を考えるうえで重要な研究対象であることが、改めて確認された。

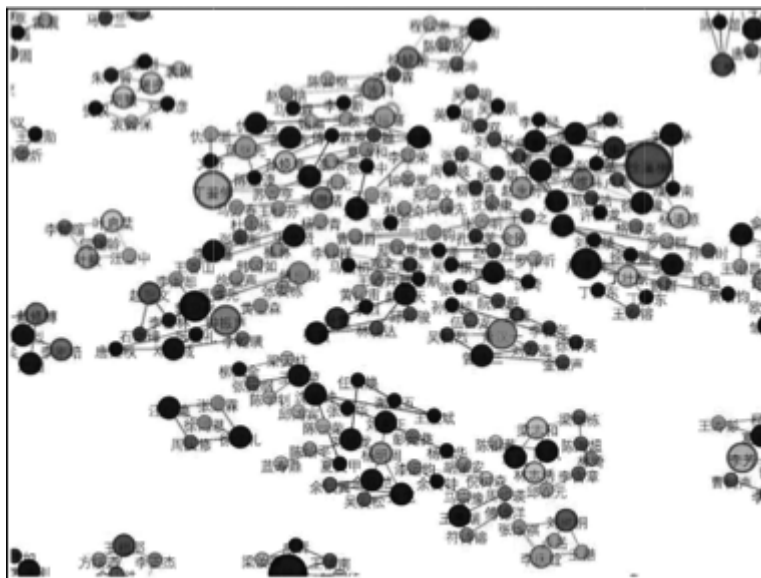
本会議を通じて筆者にとって特に印象的であったのは、近年中国で大きな進展を見せている、ビッグデータを活用した歴史研究である。今回の報告の中でも、留学生名簿を基礎史料として構築されたデータベースを用いる研究は、今後の留学生史研究の方向性を示すものとして、とりわけ注目された。具体的には、浙江工商大学・王侃良氏による「ビッグデータ時代における近代中国の日本留学生研究——浙江工商大学『近代中国留日学生データベース』」と題する報告である。

研究の目的は、19世紀末から20世紀20年代にかけて日本へ留学した約2万9千人の中国人留学生を対象とする大規模な歴史データベースを構築し、その分析を通じて留学生像を再検討することにある。データ総数は約9万7千件に及び、氏名、出身地、留学費用の区分（官費・自費）、渡日時期といった基本情報に加え、借款・担保関係など、留学生間の社会的ネットワークに関する情報も含まれている。

研究チームは、表記や記載内容が不統一な膨大な史料を整理・統合し、例えば314種に及ぶ出身地表記を27の省に整理するなど、データの精度と分析可能性を大きく向上させた。こうした基盤整備により、従来の「文献読解」を中心とする近代史研究から、「データ分析」を重視する研究手法を導入することが可能となっている。

分析の結果からは、いくつかの特徴的な傾向が示された。第一に、1905年に留日学生数が急増している点である。これは、清末の「新政」のもとで日本留学が奨励されたことと密接に関係している。第二に、官費留学生が51.4%、自費留学生が48.6%とほぼ拮抗しており、留学事業が中央政府や地方政府の主導だけでなく、民間の主體的な関与によっても支えられていたことが明らかになった。さらに地域別にみると、湖南・湖北・浙江の三省から多くの留学生が輩出されており、経済的・文化的条件が留学機会に影響していた可能性が示唆された。これらの点は、従来の理解を見直すうえでも示唆に富む。

また、この研究の特徴として、留学生同士の「借款担保」関係に着目した社会ネットワーク分析が挙げられる（図1）。調査の結果、2923人の個人と2738件の関係が確認され、全体として367の小規模なグループからなる、断片的なネットワーク構造が明らかになった。統計的検定により、地縁（同郷）関係の影響は学縁（同校）関係よりも強く、経済的な相互扶助は主として同郷者間で行われていたことが示されている。このことから、当時の留学生社会は、全国的なネットワークというよりも、地域共同体を基盤として形成されていたことが理解できる。



(図1) 留学生の「借款担保」関係ネットワーク

以上のように、浙江工商大学における「近代中国留日学生データベース」の構築と分析は、「データ駆動型の近代史研究」という新たな研究パラダイムの可能性を示すものであった。ただし、この新たな研究手法には、なお検討すべき課題が残されている。例えば、基礎データとなる留学生名簿については、データベース化に先立って行うべき史料批判の方法が十分に確立されているとはいえない。また、データ入力の手による作業に依存する以上、入力者の知識や判断によって誤差が生じる可能性も否定できない。さらに、データ分析の「俯瞰的」な特性を、個別の人物史研究にいかに関接するかについても、今後さらに検討する必要がある。

王侃良氏はこれまで一貫してビッグデータを活用した研究に取り組んでおり、データベースの構築にとどまらず、その理論的枠組みにも関心を寄せてきた。また、同氏は浙江工商大学デジタル・ヒューマニティーズ重点実験室の執行主任も務めている。近年では、「从典籍数字化到数字弘学：数字人文与弘学数字化的未来」（王侃良・江静ほか、『浙江工商大学学报』2024年第6期）、「数据驱动：当代日本史学的数字人文研究范式」（王侃良・戦林澤、『日本学研究』第39輯、2025年12月）などの論文を発表し、この分野で活発に研究活動を展開している。こうした点から、同氏はデータ駆動型歴史研究を牽引する研究者の一人に位置づけられ、今後のさらなる研究の発展が期待される。

コロナ禍以降、中国の歴史学界ではビッグデータやAI技術の活用が新たな潮流となっている。留学生史研究の分野においても、こうした方法論がすでに導入され、実証的研究に応用されている点は注目に値する。ビッグデータの活用により、従来の研究では捉えきれなかった集団の動向や社会関係の構造が可視化されつつあり、留学生像の再構成に新たな視座を提供しているといえよう。

一方、こうしたデータ駆動型研究は、従来の史料読解に基づく質的研究と対立するものではなく、むしろ両者を相補的に組み合わせることによって発展させるべきものである。すなわち、マクロな傾向把握とミクロな個別分析とをいかに有機的に結びつけるかが、今後の重要な課題となる。その意味で、本研究は方法論的にも実証的にも多くの示唆を与えるものであり、今後の展開を引き続き注視したい。