大学図書館における人文社会系の専門書の所蔵状況

― CiNii Books と openBD を用いた予備的調査 ―

久保琢也（信州大学）

１．本発表の目的と課題

研究 IR では、研究活動に関する機関の意思決定や URA 等の研究支援者による日々の業務を支援するための調査・分析を行っている。また、分析に際しては、学内で整備しているデータベースだけでなく、学外の公開データベースや商用データベースが一般的に用いられている。特に、論文等の研究成果に関する情報については、国際的な学術文献データベースとそれに対応する分析ツールが普及しており、国内外のあらゆる研究機関の論文数や国際共著論文数・割合、被引用数 Top10%論文数・割合といった指標が利用できるようになっている。

その一方で、既存の学術文献データベースを用いた分析では、人文社会系の研究成果が見落とされるという問題が指摘されている。人文社会系の特徴の１つとしては図書が重視される点が挙げられ、雑誌論文だけでなく図書での論文発表も多くなされている（後藤, 2018）。また、引用される文献も雑誌論文より図書の方が多いという調査結果が報告されており（後藤, 2018; Zuccala & Guns, 2013）、学術研究の進展において図書は重要な役割を果たしていると言える。しかしながら、既存の学術文献データベースでは英語等の雑誌や図書が中心的に採録されており、日本の図書はほとんど採録されていない現状がある。そのため、既存の学術文献データベースだけに頼っていては、人文社会系の研究活動が過小に評価されてしまう恐れがある（稲石・神谷, 2018）。

このような背景において、発表者は人文社会系の図書の分析・可視化に取り組んでいる。本稿は、この取り組みの一環として行なった、大学図書館における人文社会系の専門書の所蔵状況に関する予備的調査の結果を報告する。ここで「専門書」とは、市販される図書に付与される C コード（図書分類コード）において、販売対象が「専門」である図書を指す。発表者の関心は学術研究の成果である学術図書にあり、必ずしも専門書と同一であるとは限らないが、現状では純粋に学術図書だけを抽出することが困難であるため、このような代替策をとっている。

本調査の具体的な目的は以下の２点である。まず、本調査では大学図書館に所蔵されている人文社会系の専門書の分野別、出版年別のタイトル数（異なり図書数）の特徴を明らかにする。大学図書館はその機能として学術研究活動を支える情報基盤の役割が求められていることから（文部科学省 科学技術・学術審議会 学術分科会 研究環境基盤部会 学術情報基盤作業部会, 2010）、大学図書館にはより網羅的に学術図書が所蔵されていると考えられる。したがって、大学図書館での所蔵状況を調査することで、毎年どの分野の図書がどの程度出版されているかといった全体的な動向を推察できるのではないかと思われる。

また、２点目の目的は、各図書が所蔵されている図書館の数（以下、所蔵館数）がどのような特徴を持つのかを明らかにすることである。所蔵館数は、論文の引用分析の考え方

を図書に応用した指標として提案されており、所蔵館数が多いほどインパクトが大きいことを示すと考えられている（White et al., 2009; Torres-Salinas & Moed, 2009）。図書館に所蔵されることの意味について議論はあるが、所蔵館数が多いほど、より多くの人の目に触れることになり、利用される可能性が高くなると考えられる（Zuccala & Guns, 2013; Torres-Salinas & Arroyo-Machado, 2020）。その利用形態の 1 つが論文等での引用であり、ある図書の所蔵館数とその図書の被引用数には弱いながらも正の相関関係が確認されている（Zuccala & Guns, 2013）。日本では研究 IR において所蔵館数の活用例は見られないが、第 2 期中期目標期間の国立大学法人評価において、社会・経済・文化面での卓越性を示す根拠データ例の 1 つとして挙げられている事例もある（林, 2018）。しかし、日本の専門書に関して所蔵館数がどのような特徴を持っているのかを明らかにした研究は見られない。

以上を踏まえ、本調査では、大学図書館に所蔵されている人文社会系の専門書の分野別、出版年別のタイトル数、及び、これらの図書の所蔵館数の特徴を明らかにする。また、この調査結果から得られた知見をもとに、研究 IR において図書を分析する際の留意点や、今後の課題等について述べる。

2．方法

大学図書館における人文社会系の専門書の所蔵状況の調査を行うにあたり、主に CiNii Books（https://ci.nii.ac.jp/books/）を用いた。CiNii Books は全国の大学図書館等が所蔵する図書の情報を検索することができるサービスであり、各図書が所蔵されている図書館の情報も取得することができる。参加図書館数は 2020 年 3 月 31 日時点で 1341 機関となっている（https://www.nii.ac.jp/CAT-ILL/archive/stats/cat/org.html）。

大学図書館では広く一般書も所蔵されているため、openBD も併用することで各図書の C コードを取得し、販売対象が「専門」である図書を抽出した。openBD とはカーリルと版元ドットコムにより開発された API システムであり、書誌情報や書影が提供されている。C コードは図書分類コードであり、C に続く 4 桁によって、販売対象（1 桁目）、形態（2 桁目）、内容（3、4 桁目）を表す。なお、CiNii Books 及び openBD の利用に当たり、それぞれの管理機関から事前に許諾を得ている。

具体的な方法はとしては、まず、CiNii Books API により 2015 年から 2019 年に出版された全ての図書（日本語）のレコードを取得した（NCID : 262568 件）。また、これらのレコードからユニークな ISBN を抽出したところ 224408 件であった。次に、openBD を用いてこれらの ISBN から C コードの取得を試みたが、取得できたのは約半数の 114543 件であった。このうち、販売対象が「専門」であるデータは 28270 件であった。なお、CiNii Books では 1 つの NCID に複数の ISBN が紐づいていたり、1 つの ISBN が複数の NCID に紐づいている場合があるため、集計は NCID と ISBN のユニークなペアを単位とした。この結果、最終的な分析対象は 28463 件となった。分析では分野ごとの集計を行うが、分野分類については右表に示す C コードの内容コードを用いた。

表 1．C コードによる分野分類

C コード（3, 4 桁目）	内容
00～04	総記
10～16	哲学・宗教・心理
20～26	歴史・地理
30～39	社会科学
40～47	自然科学
50～58	工学・工業
60～65	産業
70～79	芸術・生活
80～87	語学
90～98	文学

３.結果

３.１．タイトル数

まず、分野別のタイトル数の年次推移を図１に示す。各分野の粒度が異なるため、分野間の比較には慎重を要するが、全体的には「社会科学」の図書が最も多く、また、増加傾向にあることが分かる。また、「哲学・宗教・倫理」、「歴史・地理」、「語学」、「文学」、「芸能・生活」といった人文学系の分野では一定数の図書が確認できるが、全体的な年次推移は横ばい傾向であると言える。なお、「自然科学」の図書が２番目に多いが、これは「自然科学」の下位分野である「医学･歯学･薬学」の寄与が大きく、５年間の通算では約７割を占めている（5209/7317）。

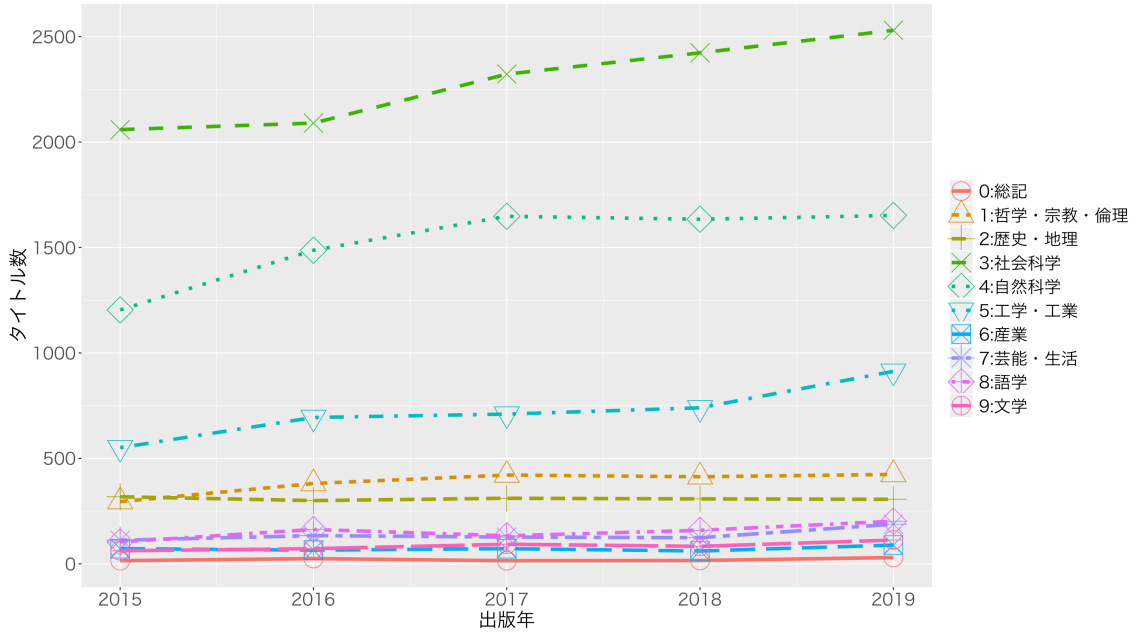


図１. 各分野でのタイトルの推移（2015～2019）

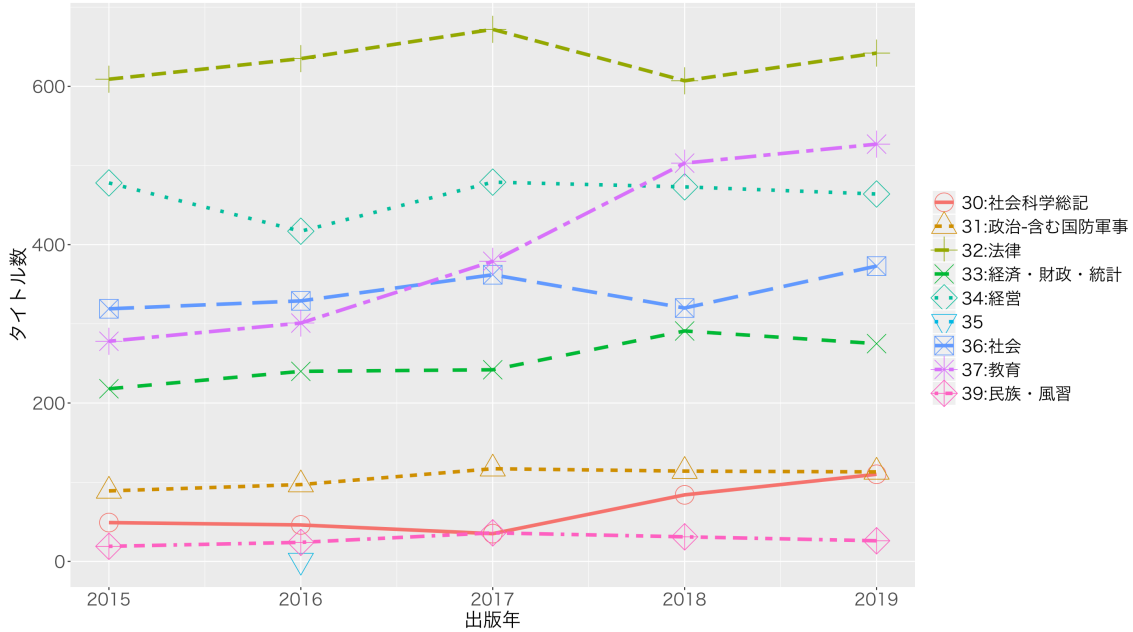


図１. 社会科学系分野でのタイトル数の推移（2015～2019）

また、「社会科学」の下位分野におけるタイトル数の推移を示したのが図２である。図２より、「社会科学」のそれぞれの下位分野でも全体的にタイトル数に差があることがわかる。年次推移については全体的に横ばい、あるいは微増している分野が多いが、「教育」分野での増加傾向が顕著に見て取れる。

このように、大学図書館に所蔵される図書のタイトル数は分野別、出版年別で異なることが分かる。その理由としては、各分野での研究者数の差や図書の重要性の差等が考えられるが、本件については、今後より詳細な検討が求められる。

３.２. 所蔵関数

図３は各分野における所蔵館数の分布を確率密度曲線で表したグラフである。図の視認性を確保するため、出版年については 2015 年、2017 年、2019 年に限定している。図３より、大学図書館での所蔵館数は全体的に右に裾が長い分布をしており、一部の図書がより多くの大学図書館において所蔵されている状況にあることがわかる。また、分野による差はあるが、この分布は出版年が古くなるにつれて右側に移行している様子が窺える。表２は分野別、出版年別の所蔵館数の中央値を示しているが、「社会科学」を除く全ての分野において、出版年が古くなるほど所蔵館数の中央値が高くなる傾向にあることがわかる。この傾向は出版から大学図書館に所蔵されるまでのラグを反映しているものと思われる。

表２. 人文社会系図書の図書館所蔵館数の中央値

分野	出版年		
	2015	2017	2019
哲学・宗教・倫理	83.0	73.0	64.0
歴史・地理	82.5	76.0	65.0
社会科学	56.0	59.0	48.0
芸能・生活	74.0	63.0	45.5
語学	80.0	78.0	53.0
文学	76.5	72.0	59.0

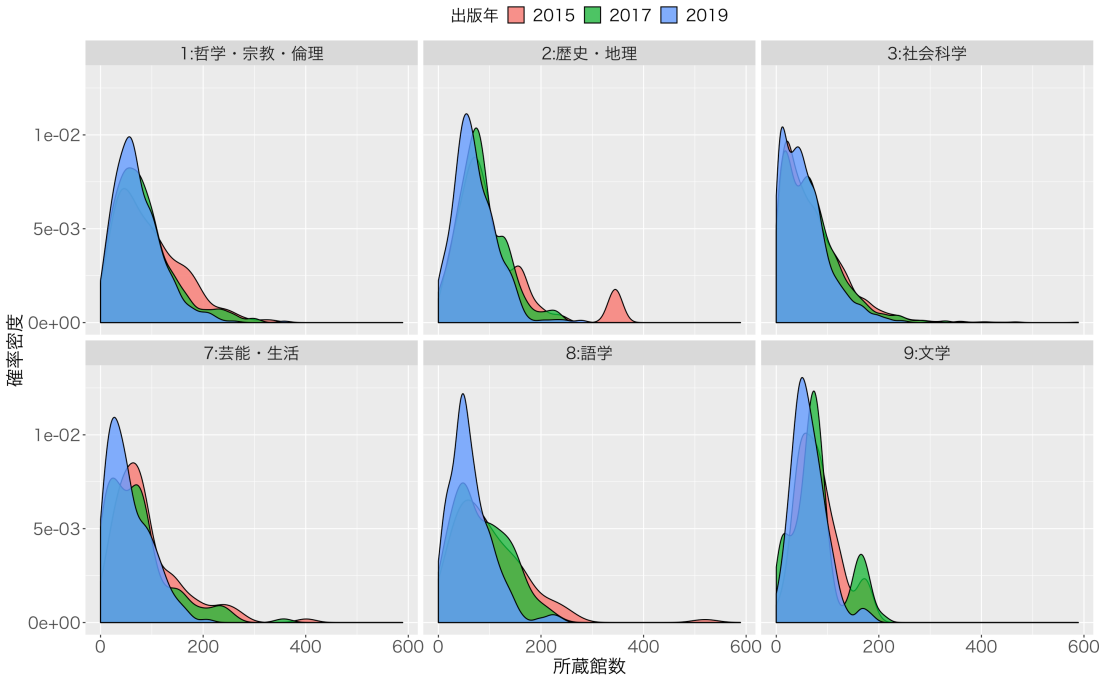


図３. 人文社会系図書の大学図書館での所蔵館数の分布

図４は「社会科学」の下位分野において同様の可視化を行った図である。図３と同様に全体的に右に裾が伸びた分布をしているが、分野間の差異がより明確に見て取れる。例えば、「法律」や「経営」においては所蔵館数が比較的に低い値で分布が集中しているのに対し、「経済・財政・統計」、「社会」では、所蔵館数が「法律」や「経営」よりも高い値において分布のピークがあり、また、集中度もそれほど高くない。表３は各分野の出版年別の中央値を示しているが、この表からも分野によって中央値に大きな差があることが分かる。一方、年次推移については、概ね出版年が古いほうが所蔵館数が多い傾向にあるが、「法律」、「経済・財政・統計」、「経営」においては、2015 年の図書の方が 2017 年の図書よりも所蔵館数が少ない。

このように、図書館の所蔵館数についても、その分布や中央値は分野や出版年によって異なることが窺える。

表３. 社会科学系図書の図書館所蔵館数の中央値

分野	出版年		
	2015	2017	2019
政治	103.0	92.0	82.0
法律	43.0	46.0	39.0
経済・財政・統計	80.0	91.5	73.0
経営	36.0	43.0	39.0
社会	93.0	79.0	72.0
教育	59.0	56.0	48.0
民族・風習	83.0	79.5	67.0

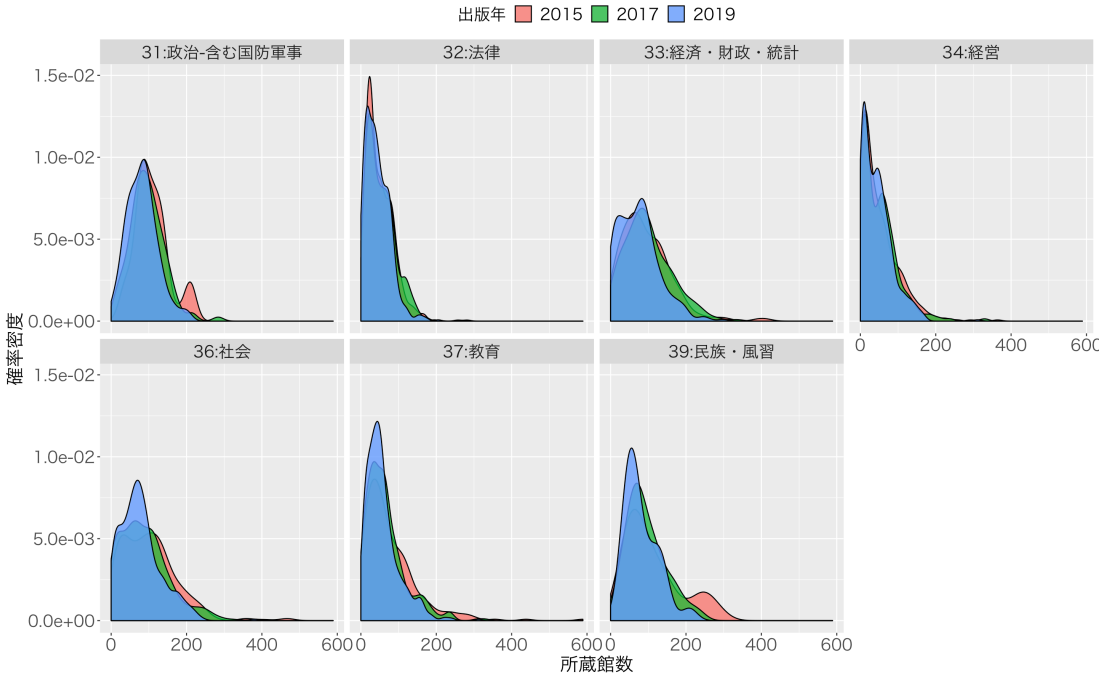


図４．社会科学系図書の大学図書館での所蔵館数の分布

４．まとめと今後の展望

以上、簡単ではあるが、本稿では CiNii Books に登録される人文社会系の専門書のタイトル数および大学図書館での所蔵館数の特徴を概観した。あくまで限られたサンプルを対象とした調査ではあるが、この調査により、分野や出版年によってタイトル数や所蔵館数の分布が異なる可能性を示した。このような差異は、研究 IR において図書を分析対象とする際に重要であり、分野や出版年を超えた比較の際には注意が必要であることを示して

いる。例えば、出版された年や分野の異なる図書の所蔵館数の多寡を判断する場合、出版年や分野の差を考慮して所蔵館数の標準化を行ったり、同じ出版年、分野の図書内でのパーセンタイルを用いるといった方法が必要であろう。

今後の課題としては分析対象データの拡充及び精選が挙げられる。今回の調査方法では CiNii Books から取得した ISBN に対して半数程度しか C コードを取得できなかったが、今後は他のデータベースを参照するなど分析対象となる図書数を増やす必要がある。また、C コードによって専門書に分類される図書の中には、資格試験の参考書といった、必ずしも学術研究の成果の発表を意図したものではない図書も散見される。そのため、何らかの基準を設けた上で、これらの図書を削除した上で分析することが望まれる。

また、本調査では CiNii Books に登録されている全ての大学図書館を対象として所蔵館数の調査を行ったが、所蔵館数が表す意味はどのような図書館を対象とするかに依存する（Torres-Salinas & Arroyo-Machado, 2020）ことに留意する必要がある。大学によっては附属図書館に特定の分野の図書が所蔵されているものの、当該分野の研究者が所属していないこともありうる。こういった事例は所蔵館数の解釈を難しくする 1 つの要因であるため、例えば、学術的インパクトを特に重視する場合には、特定の研究分野に強い大学の図書館に限定して所蔵館数を調査することも方法の 1 つとして挙げられるだろう。

以上、本稿では大学図書館における人文社会系の専門書の所蔵状況に関する調査結果を報告したが、今後はデータの拡充・精選と合わせて、多様な角度から人文社会系の専門書の所蔵状況に関する特徴を明らかにする必要があるだろう。

【参考文献】

[1] 稲石奈津子, 神谷俊朗（2018）. 人文・社会科学系の研究力可視化に向けた取り組み –研究推進・研究支援の観点から, 学術の動向, 23 巻 10 号, pp. 58-61.

[2] 文部科学省 科学技術・学術審議会 学術分科会 研究環境基盤部会 学術情報基盤作業部会（2010）.「大学図書館の整備について（審議まとめ）-変革する大学にあって求められる大学図書館像-」

[3] 後藤真(2018), 研究の量的評価は人文学に対して可能なのか –人間文化研究機構の試み, 学術の動向, 23 巻 10 号, pp. 42-49.

[4] Torres-Salinas, D., & Arroyo-Machado, W. (2020). Library Catalog Analysis and Library Holdings Counts: origins, methodological issues and application to the field of Informetrics. Evaluative Informetrics: The Art of Metrics-Based Research Assessment, pp.287-308.

[5] Torres-Salinas, D., & Moed, H. F. (2009). Library Catalog Analysis as a tool in studies of social sciences and humanities: An exploratory study of published book titles in Economics. Journal of Informetrics, 3(1), 9–26.

[6] White, H. D., Boell, S. K., Yu, H., Davis, M., Wilson, C. S., & Cole, F. T. H. (2009). Libcitations: A measure for comparative assessment of book publications in the humanities and social sciences. Journal of the American Society for Information Science and Technology, 60(6), pp. 1083-1096.

[7] Zuccala, A. & Guns, R. (2013). Comparing book citations in humanities journals to library holdings: Scholarly use versus 'perceived cultural benefit, Proceedings of the 14th International Society for Scientometrics and Informetrics Conference, 1, pp. 353-360.