

学校図書館の蔵書分析:13校の全蔵書データを対象に

安形 輝(亜細亜大学) agata@asia-u.ac.jp
橋詰秋子(国立国会図書館国際子ども図書館)

1. 目的と背景

1.1 研究目的

国際子ども図書館では「中高生向け調べものの部屋の準備調査プロジェクト」¹⁾として、国際子ども図書館における中高生向け調べものの部屋の開設準備を行っている。そのプロジェクトのために、積極的なサービスを展開している中学校中心とする学校図書館 13 館からほぼ目録データ全ての提供を受けた。今回はそれらの目録データの蔵書分析を行った結果について報告する。

蔵書分析の目的は、第一に学校図書館における中核となる資料群はあるのか、そのような資料群がある場合、それはどのようなものかを明らかにすることである。さらには蔵書分析を通じて、調査対象の学校図書館の蔵書の特徴や、学校図書館蔵書同士の近さなども明らかにする。

1.2 今回の蔵書分析の独自性

従来の所蔵調査は、多くの場合、チェックリストを対象とした小規模なサンプリング調査であった。しかし、どの図書館に対しても有効な適切なチェックリストを作成することは非常に難しい。また、小規模なサンプリングでは、蔵書の特徴づける重要

な資料がカバーできない恐れもある。

一方で、近年、図書館の多くが OPAC をウェブ上で公開するようになり、蔵書全体を対象とした大規模な所蔵調査が実施できるようになった。ただし、OPAC での検索を前提とした所蔵調査には、検索対象となる網羅的な資料リストが必要である。全出版物を網羅したリストはそれに近いもの、例えば日本全国書誌はあるが、完全なものには存在しない。また、OPAC に対する負荷を考慮したときに、全出版物を対象とした調査の実施は難しい。そのため、一定期間に刊行された出版物を対象とした所蔵調査²⁾や、特定ジャンルを対象とした所蔵調査³⁾が行われてきた。

今回は蔵書全体を対象としている点で、蔵書の一部しか対象としていなかった従来の所蔵調査とは異なる。

2. 蔵書分析の手法

2.1 調査対象図書館

調査対象となった学校図書館の概要は表1に示す⁴⁾。なお、A に関しては同じ自治体の他の中学校 3 校についても目録データの提供を受けた。そのため、以下では、A は A-1 とし、

表 1 調査対象図書館の概要

学校	生徒数	図書館の対象	コレクション数	1人あたり冊数	貸出数
	【概数】		(図書のみ)	【概数】	(/年) 【概数】
公立共学中学 A	450	中	13,000	29	6,000
公立共学中学 B	500	中	17,000	34	20,000
国立共学中学 C	480	中	22,000	46	12,200
私立共学中学 D	600	中	55,000	92	20,000
私立男子中学 E	720	中	32,000	44	23,000
私立女子中学 F	350	中	28,000	80	10,000
私立共学IS・中高一貫 G	450/700	幼～高	69,000	99	6,000
私立共学中高一貫 H	2,000	中高	41,000	21	40,000
私立男子中高一貫 I	1,800	中高	82,000	46	11,000
私立男子中高一貫 J	1,100	中高	79,000	72	14,000
2012年度学校図書館調査 平均値	-	中	11,138	28.3	-
	-	高	25,280	34.5	-

同自治体の他の学校図書館については A-2 から A-4 として参照する。また、G は、提供データ以外の洋書を約 20,000 冊所蔵している。

2.2 目録データの整理

調査対象館の目録データは書誌データの精粗や形式がさまざまであったため、以下のように目録データの整理を行った。

- ・CSV 形式への統一
- ・文字コードの統一
- ・区切り記号の統一とデータ項目の識別
- ・第一著者の抽出
- ・巻号数のアラビア数字への変換
- ・複数目録の統合
- ・データの修復
- ・可能な場合、視聴覚資料の削除

なお、A-1～4 と同自治体の中学校は 5 校あり、すべてから目録データの提供を受けていた。しかし、1 校の目録データについては欠損データが多く、調査対象から除外した。

2.3 目録データの統合

複数の図書館の蔵書を比較するためには、同一資料に対する目録データを同定することが必要となる。既往研究の手法等を参考にしながら⁹⁾、同定における ISBN を単独で用いることの問題を検討した上で、以下のような手順で同定作業を行った。

- 1) ISBN-13 は ISBN-10 に変換
- 2) ISBN-10 が同一であれば同一とみなす
- 3) 書名が同じで、著者、出版社、出版年のいずれも同一の場合、同一とみなす
- 4) 統合された目録データに足りない項目があった場合には追加

目録提供を受けた図書館のうち、ISBN が付与されている書誌データ数が最も多かった I を基準とし、他の図書館の目録データを追加する形で統合を行った。

上記の手順で機械的に統合された目録データは全体として異なりタイトル数は 237,512 点、延べ資料数 425,484 点であった。

2.4 人手による書誌分かれの同定

各館の目録データの作成や入力方法の違いから、機械的な同定には限界がある。

特に、多巻ものの参考資料は、書名や副書名、巻号の入力の仕方がさまざまであった。そこで、参考資料を中心として人手での書誌分かれの統合を行った。機械的に統合したデータを、人手で 500 タイトル以上を統合し、最終的に異なりタイトル数 237,001 点となった。以下の分析ではこのデータを用いる。

2.5 国会図書館からの書誌データの検索

各学校図書館では、その館独自の事情に基づき、目録データを作成している。そのため、同じタイトルであっても異なる分類が付与されることがある。また、精粗の異なる目録データに基づくため、最小限の書誌情報しかわからないデータが数多くあった。

そこで、共通した NDC と詳細な書誌データ入手するため、国立国会図書館の目録データを検索した。国会図書館サーチへの負荷を考慮し、2004 年以前は J-BISC のデータを、2005 年以降は国会図書館サーチの検索 API を用いた。

具体的な書誌データの検索手順は以下のとおりである。

- 1) ISBN があるデータは ISBN から検索し、検索結果が返ってくれば取得
- 2) 書名と出版年(不明な場合は省略)を検索キーとして検索し、著者名が一致する検索結果の 1 件目を取得

統合データ 237,001 点のうち、J-BISC ないし国会図書館サーチで検索でき書誌データを付与できたのは 183,000 点である。検索できなかったもの 54,001 点のうち、45,800 点は所蔵館数が 1 館のみの資料であった。検索できなかったものは、多巻資料、視聴覚資料、独自作成資料などであった。

なお、データ分析においては、国会図書館で書誌データが検索できた資料については国会図書館のデータを、それ以外の資料については図書館の書誌データを用いる。

2.6 重複率の算出とクラスター分析

図書館の所蔵の重なりを示すために以下のように重複率を定義し算出している。

重複率は図書館同士の重なりを示す指標であり、戸田⁶⁾と同様に、図書館 i と図書館 j の和集合に対する積集合の形で算出している。

$$\text{重複率}(\text{所蔵}_i, \text{所蔵}_j) = \frac{\text{所蔵}_i \cap \text{所蔵}_j}{\text{所蔵}_i \cup \text{所蔵}_j}$$

重複率は所蔵に対して Jaccard 係数と同じ意味を持つ。また、この重複率を 1 から減算したものを図書館蔵書の距離として、Ward 法にてクラスター分析も行った。

2.7 チェックリストとの照合

『学校図書館基本図書目録 2011 年版』に掲載された中学校分 1,031 点や青少年読書感想文全国コンクール第 8 回(1962 年)～第 59 回(2013 年)における課題図書リスト 743 点についても識別できる範囲で照合している。ただし、課題図書には著者とタイトルしか書誌情報がないため、それらでの照合である。

3. 蔵書分析の結果

3.1 所蔵の多い資料(全体)

目録データ全体で所蔵館数が多い、タイトル、著者、出版社について表 2-4 に示す。

表2のうち、「13 歳のハローワーク」は学校図書館職員に対するインタビューでは寄贈を受けた資料という言葉があった。表2からは漏れているが「ニッポンの嵐」も寄贈を受けたそうで、9 館が所蔵していた。課題図書として取り上げられることが多い、「沈黙の春」などのタイトルは一館での所蔵冊数も多い。

表 2 所蔵が多いタイトル

書名	出版年	所蔵館数	所蔵冊数
13歳のハローワーク	2003	13	25
きみの友だち	1987	13	21
素数ゼミの謎	2005	13	14
いのちの食べかた	2004	12	28
沈黙の春	1974	12	24
世界を信じるためのメソッド	2006	12	22
ぼくを探しに	1977	12	18
化石	1991	12	16
魔法使いハウルと火の悪魔	1997	12	14
鬼の橋	1998	12	13

表3は所蔵が多い著者である。この表で所蔵館数はこの著者のタイトル群を延べで所蔵

している館数を意味する。人気のある著者では、星新一は表3の他の著者と比較すれば、著作は少ないが、多くの図書館で所蔵されている。また、手塚治虫の著作は複本が多くあることがわかる。

表 3 所蔵が多い著者

著者	カバー数	所蔵館数	所蔵冊数	平均館数	平均複本数
赤川次郎	476	775	872	1.63	1.13
学研	329	745	871	2.26	1.17
司馬遼太郎	420	732	1,002	1.74	1.37
NHK取材班	387	710	773	1.83	1.09
講談社	347	654	723	1.88	1.11
手塚治虫	277	535	876	1.93	1.64
旺文社	414	531	689	1.28	1.30
星新一	142	494	713	3.48	1.44
あさのあつこ	177	457	545	2.58	1.19
朝日新聞社	376	448	573	1.19	1.28

表4は所蔵が多い出版社を示している。この目録データ群で最も所蔵数が多いのは 2004 年である。2004 年の出版新刊点数上位の 1 位講談社と 7 位岩波書店の順位は逆である。また、データ群がカバーしているタイトル数は講談社が多いにも関わらず、所蔵冊数は岩波書店が多い。つまり、学校図書館は講談社よりも岩波書店から出版された資料を購入する傾向があり、複本も多く購入している。

表 4 所蔵が多い出版社

出版社	所蔵冊数	カバー数
岩波書店	34,419	15,241
講談社	29,756	16,389
新潮社	13,561	6,414
筑摩書房	10,708	5,924
ポプラ社	9,106	4,015
小学館	8,681	4,374
角川書店	8,428	4,550
平凡社	7,845	4,706
集英社	8,041	4,473
中央公論社	6,408	3,434

類別に分析したときには 13 館中 11 館が所蔵していた 65 タイトルのうち、9 類(文学)が 46 タイトルであった。9 類の次に多かったのは 4 類で 5 タイトルと極端に減る。ほぼ共通して所蔵される資料群は 9 類に集中している。

3.2 所蔵の多い参考図書(著作別)

参考図書に関しては版次をまとめ著作別にまとめたときに、所蔵が多いタイトルを集計したものが表5である。表5において著作単位では少なくとも参考図書に関しては「広辞苑」

等よく知られたタイトルが所蔵の上位となった。旧版を所蔵し、予算に余裕がない場合、最新版ではなく他を優先することを示唆している。

表 5 所蔵の多い参考図書(著作別)

書名	出版社	館数
ギネス世界記録	ポプラ社 など	13
日本国勢図会:日本がわかるデータブック	矢野恒太記念会	13
世界国勢図会:世界がわかるデータブック	矢野恒太記念会	13
朝日ジュニア学習年鑑	朝日新聞出版	13
広辞苑	岩波書店	13
日本のすがた:表とグラフでみる社会科資料集	矢野恒太記念会	12
教え方の辞典	小学館	11
世界大百科事典	平凡社	11
教科書にでる人物学習事典	学習研究社	11

3.2 各図書館の蔵書の特徴

学校図書館の蔵書の特徴を示すために学校図書館の NDC の類ごとのタイトル数、構成比、出版年の中央値を示したのが表6である。紙幅の関係で、BとIの結果だけを示す。Bは所蔵資料の出版年が比較的新しく規模の小さな公立共学中学で、Iは規模が大きな私立男子中高一貫校である。

表 6 図書館 B と I の類ごとの蔵書構成

類	公立共学中学B			私立男子中高一貫校I		
	タイトル数	構成比	出版年	タイトル数	構成比	出版年
0	417	2.8%	2002	1,473	2.9%	1989
1	374	2.5%	2003	2,698	5.4%	1989
2	1,494	10.1%	2001	7,057	14.0%	1990
3	1,648	11.1%	2002	7,702	15.3%	1993
4	1,421	9.6%	1999	6,665	13.2%	1990
5	905	6.1%	2002	2,601	5.2%	1992
6	430	2.9%	2003	1,358	2.7%	1991
7	1,533	10.3%	2000	5,777	11.5%	1990
8	394	2.7%	2003	1,323	2.6%	1994
9	6,196	41.8%	2005	13,667	27.2%	1985

表5からは、規模が小さく資料購入費が少ない公立共学中学 B は読み物中心の蔵書、私立男子中高一貫校 I は 3 類の割合が 9 類に続いて高い。また、B は 9 類に他の類に比べ、新しい資料を、逆に B は 9 類に他

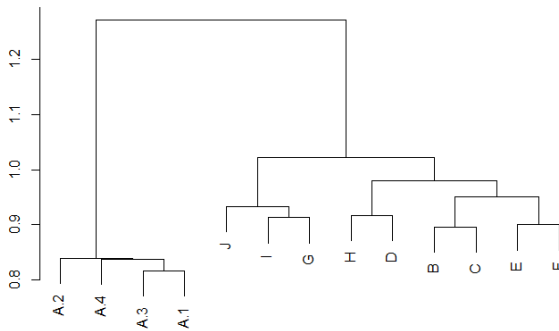


図 1 クラスタ分析

の類よりも古い資料を所蔵している。

3.3 各図書館の重複分析

13 校の学校図書館の所蔵の重複率を算出した。重複率は最も高い館同士でも 16%程度であり、1%未満の館同士もあった。重複率に基づきクラスター分析を行った結果を図1に示す。図1から同じ自治体の館同士の類似性が高いことがわかる。また、国公立同士、私立同士で固まる傾向にある。

4. まとめ

学校図書館 13 館のほぼ全目録データを対象とした蔵書分析の結果、どこの図書館も所蔵する資料群はタイトル単位では 9 類以外では少なかった。ただ、参考図書については著作単位で見たときに所蔵が多い資料群があった。蔵書の類似性は情報交流がある自治体内の図書館同士では確認された。

本分析の限界と課題は、目録データの同定の難しさと 13 校という調査対象が学校図書館全体をどの程度代表するかにある。

【注・引用文献】

- 1) 中高生向け調べものの部屋の準備調査プロジェクト. <http://www.kodomo.go.jp/promote/school/room.html>
- 2) 大場博幸; 安形輝; 池内淳; 大谷康晴. 図書館はどのような本を所蔵しているか : 2006 年上半期総刊行書籍を対象とした包括的所蔵調査. 日本図書館情報学会誌, vol. 58, no.3, 2012, p.139-154
- 3) 小山信弥ほか 5 名. 日本の大学図書館におけるマンガの所蔵状況. 三田図書館・情報学会研究大会発表論文集. 三田図書館・情報学会研究大会発表論文集, 2012, p.33-36
- 4) 「中高生向け調べものの部屋の準備調査プロジェクト」事例調査に基づく数値。また、学校図書館調査は国 SLA 研究調査部「2012 年度学校図書館調査報告」『学校図書館』2012 年 11 月号による。
- 5) 谷口祥一. 総合目録データに対する機械的書誌同定と著作同定の試み : ゆにかねっとレコードによる実験. 日本図書館情報学会誌. Vol.57, no.4, p.124-140(2011)等を参考にした。
- 6) 戸田あきら. 「蔵書の重複分析による公共図書館ネットワーク効果の研究」『図書館界』vol.47, no.1, 1995, p.2-12