

**第 61 回日本図書館情報学会**

# **研究大会発表論文集**

**東京大学**

**2013年10月12日（土）～13日（日）**

**日本図書館情報学会**

**Japan Society of Library and Information Science**

12:00～	受付(赤門総合研究棟 入口ロビー)
--------	-------------------

12:45～12:55	開会式(第7教室)	
	第1部会(第7教室)	第2部会(第8教室)
13:00～13:30	松戸宏予(佛教大学) イングランドの大学図書館における特別な支援の背景要素:ブロンフェンブレナー理論を援用して	伴周一(日本大学) 大久保尚紀(日本大学) 岡田悟志(日本大学) スマートフォンのカメラ機能を活用した箱形書籍・資料デジタルアーカイブ作成装置の研究開発
13:30～14:00	志茂淳子(和歌山県立医科大学図書館三葛館) 看護教育・研究をとりまく大学図書館の実態:「日本の図書館」を用いた分析	望月有希子(筑波大学大学院) 明治・大正期の図書館における資料保存
14:00～14:30	西野祐子(筑波大学大学院) 松野渉(筑波大学大学院) 池内有為(筑波大学大学院) 佐藤翔(同志社大学) 逸村裕(筑波大学) 大学図書館における未貸出図書の実態分析	茅野良太(東京大学大学院) ネットワーク構造として見るメディアおよび記号の編成:メディア環境記述の新たな枠組みの提案
14:30～15:00	武井千寿子(筑波大学大学院) 芳鐘冬樹(筑波大学) 逸村裕(筑波大学) 大学図書館における電子ジャーナルの引用と閲読のオペレーション分析	谷口祥一(慶應義塾大学) BSH/NDLSHとNDCにおける階層構造はどの程度一致するのか:件名と分類の統合的活用法に向けて
15:00～15:15	休憩	
15:15～15:45	宮田洋輔(常葉大学短期大学部) 安形輝(亜細亜大学) 池内淳(筑波大学) 石田栄美(九州大学) 上田修一(前慶應義塾大学) ウェブページの寿命:2001年に存在した1000万ページを対象にした調査	矢代寿寛(総合研究大学院大学/秋草学園短期大学) 小山守恵(秋草学園短期大学) 効果的な図書館見学方法の開発:ワークシートの試作と利用を通じて
15:45～16:15	上田修一(前慶應義塾大学) 石田栄美(九州大学) 宮田洋輔(常葉大学短期大学部) 南友紀子(慶應義塾大学) 倉田敬子(慶應義塾大学) レビューサイトにみる集合知の可能性:アマゾンカスタマーレビューを例として	坪井優美(筑波大学大学院) 柳田美樹(株式会社モバロブ) 鈴木佳苗(筑波大学) 乳幼児のいる家庭向けイベントの内容に関する検討:公共図書館での新たなサービスへの提案
16:15～16:45	佐藤翔(同志社大学) 川瀬直人(国立国会図書館) 塩崎亮(国立国会図書館) 北島颯正(国立国会図書館) 原田隆史(同志社大学) アクセスログに基づく国立国会図書館サーチの利用状況の分析	桂まに子(京都女子大学) 小規模図書館が市民・行政と共に地域の課題解決に取り組む手法:京都東山におけるまちづくり実践「図書館から始める文学まち歩き」をふまえて
16:45～17:15	阿辺川武(国立情報学研究所) 間下亜紀子(連想出版) 高野明彦(国立情報学研究所) Webcat Plus への問い合わせとその対応にみる名寄せ処理の課題	平山陽菜 小室里花(茨城県信用組合) 池内淳(筑波大学) 公立図書館に関する公開質問状の分析
17:15～17:45		磯部ゆき江(放送大学大学院) 三輪眞木子(放送大学) 公共図書館の電子書籍導入について

18:00～20:00	懇親会(東京大学医学部教育研究棟 13階 カポ・ペリカーノ 本郷店)
-------------	------------------------------------

9:00～	受付(赤門総合研究棟 入口ロビー)
-------	-------------------

	第1部会(第7教室)	第2部会(第8教室)
9:30～10:00	角田裕之(鶴見大学) 元木章博(鶴見大学) 天野晃(理化学研究所) タブレット端末用情報資源組織演習の学習教材アプリの開発	吉田右子(筑波大学) 白井哲哉(筑波大学) 大庭一郎(筑波大学) 原淳之(筑波大学) 平久江祐司(筑波大学) 毛利 るみこ(筑波大学) 図書館専門職養成史の批判的再検討:図書館情報学とアーカイブズ学の方法論的統合の可能性
10:00～10:30	村瀬愛美(鶴見大学) 元木章博(鶴見大学) 角田裕之(鶴見大学) 天野晃(理化学研究所) 農業生物資源研究所/筑波大学) 大学図書館の利用を促すシリアゲームの開発	松林正己(中部大学総合学術研究院出版室) D. C. ギルマンの青年期の図書館観:アメリカ研究図書館史におけるその役割
10:30～11:00	金丸早希(鶴見大学) 角田裕之(鶴見大学) 「BL」資料の図書館での扱い:小説における「腐女子」が好む属性の傾向の調査を通して	野口久美子(大妻女子大学) 高校教員の読書経験が生徒への読書指導に与える影響:質問紙調査をもとに
11:00～11:30	前田知子(政策研究大学院大学) 日本における科学技術情報政策:NIST 構想の実像と評価	松田めぐみ(東京大学大学院) 浅石卓真(東京大学大学院) 河村俊太郎(東京大学大学院) 高校の学校司書がもつ教科に関する学問の知識:物理学の基本文献の把握状況から
11:30～12:00	葉袋秀樹(前筑波大学) 「図書館の設置及び運営上の望ましい基準」の数値目標と日本図書館協会	中村百合子(立教大学) 橋詰秋子(国際子ども図書館) 学校図書館コレクション形成に関する2000年以降の議論の概観

12:00～13:00	昼食・理事会(教育学部 第一会議室)
13:00～14:00	臨時総会・学会賞授与式(第7教室)
14:00～14:30	日本図書館情報学会 60周年記念式典(第7教室)
14:40～17:00	シンポジウム(第7教室)
17:00～17:10	閉会式(第7教室)

○東京大学総合図書館の見学について

大会期間の開館時間中(9:00～19:00)であれば、総合図書館は自由に見学することができます。見学を希望される方は、研究大会に参加している旨を、受付にてお申し出ください。なお、13日(日)には総合図書館の見学会を企画しております。

# イングランドの大学図書館における特別な支援の背景要素： Bronfenbrenner理論を援用して

松戸宏予  
佛教大学

matsudo@bukkyo-u.ac.jp

## 抄録

特別な支援体制の背景要素を明らかにするため、分析ではBronfenbrennerの「人間発達の生態学」理論を援用した。大学図書館支援体制の背景要素は、①パートナーシップの精神を根底とする職員の意識。②大学図書館、障害アドバイザーサービス、IT関連部署、サポートワーカーとの連携による特別支援。③法的整備と組織としての体制づくり。④大学外での研修とネットワークの構築。これらの相互の影響が大学図書館の特別支援を支えていた。

## 1 はじめに

### 1.1 日本の状況：障害者差別解消法

日本では、2013年6月29日に「障害者差別解消法」が制定された。施行は2016年4月1日の予定である。この法律では「障害者の権利利益を侵害することとならないよう、(中略)必要かつ合理的な配慮をしなければならない」と定めており、国公立大学を含む行政機関等は法的義務を負い、私立大学を含む事業者は努力義務を負う。

合理的な配慮とは「障害者が他の者と平等にすべての人権及び基本的自由を共有し、または行使することを確保するための必要かつ適当な変更及び調整であって、特定の場合において必要とされるものであり、かつ、均衡を失したまたは過度の負担を課さないもの」である。支援の対象範囲に施設の利用として図書館も含まれる<sup>1</sup>。また、教材の充実のため大学等や図書館、出版社との連携の促進について検討することが望まれている<sup>2</sup>。

### 1.2 イギリスの状況

一方、イギリスでは1995年に障害者差別禁止法(Disability Discrimination Act)が制定され、2001年には1995年の障害者差別禁止法の改正として「特別な教育的ニーズ・障害法」(Special Educational Needs and Disability Act 2001)が制定された。この改正により、高等教育機関においても特別なニーズをもつ学生に対する支援が正式に開始された。高等教育機関に設置される図書館においても合理的配慮に基づく図書館サービスが行われている。

しかし、特別なニーズをもつ学生に対する図書館サービスは最初から順調におこなわれていたわけではない。特別なニーズをもつ学生にとっては入りにくい、利用しづらい大学図書館であった<sup>3</sup>。なお、現在では人的支援を中心としたサービスとなっている<sup>4</sup>。では、なぜ、イギリスの大学図書館は、特別な支援を業務の1つとして、提供できているのか。特別な支援体制の背景要素を明らかにする

ことを研究の目的とする。

## 2. 方法

### 2.1 Bronfenbrennerの「人間発達の生態学」<sup>5</sup>の援用

Bronfenbrennerは人間の発達を支えるうえで、環境との相互調整を重視した「人間発達の生態学」を唱えた。特徴として、①ダイナミックス：発達しつつある人間は、生活環境の中で、再構成を図るように成長しつつあること、②相互影響性：人と環境の二方向であること、③生態学的環境の4つのシステム：単一の直接的行動場面に限定されず、行動場面間の相互の連携がある。そして、より大きな状況から生ずる外的影響も含むことを挙げている。これらは、マイクロ、メゾ、エクソ、マクロのシステムが入れ子構造で示される。以下、4つのシステムについて説明する。

マイクロシステムは発達しつつある人にとって最も影響力を及ぼす環境である。発達しつつある人が家庭や学校、仲間集団など、具体的な行動場面において経験する活動、役割、対人関係のパターンを示すものである。なお、行動場面とは、人々が対面的相互作用を容易に行うことのできる場所を指す。活動、役割、対人関係は、マイクロシステムを構成する要素である。

メゾシステムとは発達しつつある人が積極的に参加しているマイクロシステムの間で起こる2つ以上の行動場面間の相互関係からなる。

エクソシステムとは発達しつつある人を積極的な参加者として含めていないが、発達しつつある人を含む行動場面で生起する事に影響を及ぼしたり、あるいは影響されたりするような事柄が生ずるような1つまたはそれ以上の行動場面である。例えば、地域が提供している福祉サービス、養育者の仲間関係、養育者の職場が挙げられる。

マクロシステムとは、ある状況における限局された特定のシステムではなく、その個人が含まれ

ている文化に固有な子ども観や育児観のような価値の枠組みや習慣、信念体系、思想などである。一貫性の背景にある信念体系やイデオロギーに対応するものである。分析では、この4つのシステムを援用して検討する。

## 2.2 方法の手順

### (1) インタビュー調査

第1回 (2012年2月:10大学とCLAUD事務局長)

第2回 (2012年9月:8大学とJISCTechDis・メールにてOpenUniv.)

(2) 大学図書館障害者サービスに関連した人々のインタビュー調査をもとに、質問項目ごとのリストを作成する。

(3) 障害者サービス担当者が特別な支援に関して、フォーマル、インフォーマルに連絡を取り合っている他部署、他機関を抽出する。

(4) (3)で抽出した他部署、他機関を特別なニーズをもつ学生の発達の観点の視点から、プロンプトブレインナーが提唱した4つのシステムを援用して検討する。

(4)-1 マイクロシステム、メゾシステム、エクソシステムについては、グループ(行動場面)ごとに抽出した他部署・他機関の①行動、②支援要素となる共通点、③必要条件(共通点で取り上げた要素を達成するための条件とは何か。その必要な条件をみいだすために、それぞれのサポートのつながりを探す)を、インタビューデータ、参考文献を通して検討する。検討結果を、表にまとめる。

(4)-2 マクロシステムについては、定義に沿って、価値観、価値観を表す考え、社会背景をインタビューデータ、参考文献を通して検討する。検討結果を、表にまとめる。

(5) (4)で検討した表をもとに、イングランドの大学図書館における特別な支援体制の背景を図にまとめる。

## 3 結果と考察

### 3.1 マイクロシステム

特別なニーズを持つ学生が学習支援を受けながら生涯学習者へと発達していく過程において、学生に直接影響を与えているのが、大学図書館、障害アドバイザリーサービス、IT関連部署、サポートワーカーである。

大学図書館では、主に特別なニーズをもつ学生に対して、特別支援として情報探索補助、資料取り置き、貸出延長、図書館利用のワークショップなど人的支援、支援テクノロジーソフトの提供、学習室の改装など学習環境の整備を行っている。

障害アドバイザリーサービスでは、主にディスレキシアなどのアセスメント、学習支援、インクルージョンに関連した啓発を行っている。

表1. マイクロ・メゾを区分とした分析結果

システム	グループ(行動場面)の 카테고리	行動場面での行動	共通点(支援要素)	必要条件
マイクロシステム	・大学図書館	・人的支援 ・支援テクノロジーソフトの提供 ・学習環境の整備	・合理的配慮 ・特別なニーズをもつ学生に対するミッション	連携意識
	・障害アドバイザリーサービス	・アセスメント ・学習支援 ・啓発		
	・IT関連部署	・IT講習 ・ATソフトの提供・更新		
	・サポートワーカー	*代理店: ・貸出・返却 ・IT補助		
メゾシステム	・障害アドバイザリーサービスと大学図書館	・特別なニーズをもつ学生の情報提供と個のニーズに応じた支援	・合理的配慮 ・特別なニーズをもつ学生に対するミッション ・職員意識	連携意識
	・IT関連部署と大学図書館	・図書館内での *IT講習の開催 *ATソフトの提供・更新		
	・サポートワーカー/ボランティアと大学図書館	・特別なニーズをもつ学生への補助⇔図書館利用の権限を与える。 *代理店貸出・返却 *ワークショップ受講		
	・障害アドバイザリーサービスとIT関連部署	・バリアフリーに向けた *ATソフトの提供・更新		
	・障害アドバイザリーサービスとサポートワーカー/ボランティア	・障害学生手当申込手続きのサポート⇔特別なニーズをもつ学生・大学による契約雇用		
	・IT関連部署とサポートワーカー	*IT講習開催と受講		

また、大学によっては、学生が障害学生手当を申請する際、申請手続きのサポートを行っている。

情報システムにかかわるIT関連部署では、IT講習や支援テクノロジーソフトの提供や更新を行っている。

サポートワーカーは特別なニーズを持つ学生にノートブック、図書館を利用するうえでの補助、身の回りの補助など、特別なニーズをもつ学生のニーズに応じて世話をを行っている。学生が障害学生手当を得た場合、個人で代理店を通じてサポートワーカーを雇用する場合、大学が代理店と契約してサポートワーカーを雇用する場合、あるいは、学習支援に主眼を置き、すでに研鑽を積んだサポートワーカーから研修を受けた大学院生が自校のサポートワーカーとなる場合もある<sup>6</sup>。大学図書館では、学生の代理として、貸出・返却やIT補助などを行う。

学生の学習支援にかかわる大学図書館、障害アドバイザリーサービス、IT関連部署、サポートワーカーの支援要素の共通点に、「合理的な配慮」「特別なニーズをもつ学生に対するミッション」が挙げられた。これは、「特別なニーズをもつ学生が、将来、社会人となるための準備がなされなければならない。そのために、特別なニーズをもつ学生が、(合理的な配慮のもと)健全学生と平等に扱われる必要がある」(Teesside大学)という大学図書館員の意見からも窺える。

### 3.2 メゾシステム

ただし、特別なニーズをもつ学生への特別支援は大学図書館が単独で行っているのではない。大学図書館は、①障害アドバイザリーサービス、②IT関連部署、③サポートワーカーらとも相互にかわりを持ちながら、特別支援を行っている。以下、相互の関わりについて述べる。

① 障害アドバイザリーサービスから、大学図書

館は、個別の支援が必要な学生についての情報提供を受ける。そして、大学図書館は、個のニーズに応じた支援を提供する。

② IT 関連部署に所属する職員が、大学図書館に常駐し、大学図書館内での IT 講習の開催、PC のトラブル対処、支援テクノロジーソフトの提供や更新に関わっている。

③ サポートワーカーは特別なニーズを持つ学生への補助を行う。ただし、サポートワーカーは、大学職員ではない。このため、大学図書館では、サポートワーカーらに図書館利用の権限を与え、代理貸出・返却や、図書館ワークショップの受講を認めている。図書館ワークショップを受けたサポートワーカーが、特別なニーズをもつ学生への補助をしやすくするためである。

### 3.3 エクソシステム

表 2 エクソを区分とした分析結果

システム	グループ(行動場面)の 카테고리	行動場面での行動	共通点(支援要素)	必要条件
エクソシステム	・政府	*法律作成 平等法 2010 *インクルージョンに関する施策 *DirectGo.	・物理的環境整備の根拠	法的整備
	・大学理事会	・特別なニーズをもつ学生受入表明 ・方針・受け入れ体制	・特別なニーズを持つ学生の受け入れ態勢作り	組織としての体制
	・SCONUL	・地域にかかわらず通信教育学生の大学図書館利用の受け入れ	大学図書館の後方支援	連携意識
	・CLAUD ・Open Rose	・障害関連知識の共有 ・特別な教育的ニーズを含めたインクルージョンに関する啓発	・サービス利便性のための研修 ・ネットワーク	
	・JISC TechDis	・障害者支援の展望 ・e-book など代替資料のアドバイス		
	・地域	・(地域) 雇用の機会⇄(大学) 社会に貢献できる人材の養成	・学生の自立	

また、特別なニーズをもつ学生は直接的な関わりを持たないが、特別な支援にかかる影響を間接的に受けているものとして、①政府、②大学理事会、③大学図書館研究会・大学図書館コンソーシアム・高等教育支援機関、④地域がある。

① 政府は、2001 年には 1995 年の障害者差別禁止法の改正として「特別な教育的ニーズ・障害法」(Special Educational Needs and Disability Act 2001)を制定し、高等教育機関においても特別なニーズをもつ学生の教育を義務づけた。

また、2006 年に障害者差別禁止法の改がなされ、2010 年には平等法が施行された。この物理的環境整備を支援要素として、大学において特別な支援の具体的な施策が講じられた。

② 例えば、大学の理事会では特別なニーズをもつ学生の受け入れ表明の作成が義務づけられた。方針・受け入れ体制・支援内容・将来計画は、サイトからも閲覧できる。この特別なニーズをもつ学生の受け入れ体制づくりを支援要素として、それまで、非公式に大学図書館利用者サービス担当が行っていた特別な支援も、公式な支援として正式に行われたのである。

③ 併せて、大学図書館後方支援として、図書館コンソーシアム、大学図書館研究会、高等教育支援機関が挙げられた。

### ③-1 図書館コンソーシアム SCONUL

UK とアイルランドの高等教育機関に関わる図書館、そして、国立図書館のコンソーシアムとして SCONUL (The Society of College, National and University Libraries 以下、SCONUL) は 1950 年に設立された。通信教育課程在籍の特別なニーズをもつ学生は、大学院に籍をおく社会人が多い。このため、受講生の居住先近隣にある大学図書館を利用できるように、各大学図書館では SCONUL の相互協力を利用していた。ただし、大学院生を対象としていること。試験期間中、職員が少ない夜などは、自校の学生が中心となるので、サービスを提供できないこと。また、海外への学習者に対しても、サービスを提供できないことが、支援の限界としてある。なお、SCONUL では、特別なニーズをもつ学生に対する支援の手引きを出版している。

### ③-2 大学図書館研究会 CLAUD・OpenRose

CLAUD (Creating Libraries Accessible to Users with Disabilities 以下、CLAUD)、OpenRose はともに、特別なニーズをもつ学生が大学図書館を利用する際、サービスの利便性の向上、問題の共有を目指して作られた研究会である。特に、障害関連知識の共有、特別な教育的ニーズを含めたインクルージョンに関する啓発、障害者支援の展望に関連したトピックに焦点を当てている。

CLAUD は 1996 年に南部・南西部の大学を中心に設立された。現在、22 の大学が加盟している。OpenRose は 2003 年に中北部を中心に設立された。現在、8 大学が加盟している。

### ③-3 高等教育支援機関 JISCTechDis

JISC (Joint Information Systems Committee 以下、JISC) は、教育と研究に電子技術を浸透させることを目的として登録された非営利団体である<sup>7)</sup>。そして、JISC の系列にある JISCTechDis は、高等教育機関で働く職員の他、学生や学習者、企業セクター、コミュニティセクターに対してインクルージョンと電子技術の浸透に関連した提言やアドバイスを提供している。

大学図書館との関連では、JISCTechDis のミッションとして、インクルージョンを前提に特別なニーズをもつ学生を主体的な学習者として育てること、代替資料の入手の利便性(アクセシビリティ)、インクルージョンに向けた大学内の連携組織づくりなどアドバイザーとしてのサポートが挙げられる。特に、JISCTechDis から提供する E-book platforms, Publisher Lookup Service を活用し、大学図書館では障害学生に対する代替資料として、出版社からデジタル資料を購読入手していた。なお、出版社との著作権の問題など交渉窓口にあっているのは JISCTechDis である。なお、③-1 であげた CLAUD や OpenRose とも連携をとっている。例えば、CLAUD は年 3 回の研修を行うが、JISC

TechDisの職員が2013年の夏研修にゲスト講師として参加していた<sup>8</sup>。このことから支援要素では、サービス利便性のための研修とネットワークの構築が挙げられる。また、SCONULで出版している特別支援の手引きをCLOUDのサイトからもダウンロードできるようにしている<sup>9</sup>。

④ 大学地域のコミュニティでは企業が就職の機会を提供し、大学では、社会に貢献できる人材の養成という相互に補助がなされている。支援要素は、学生の自立援助が挙げられる。

### 3.4 マクロシステム

社会全体の背景にあたるマクロシステムでは、イギリス社会を構成するうえで、パートナーシップが担当者の価値観として窺えた。パートナーシップには、社会の責任、平等の精神、パストラルケアに関わる発言がみられた。

表3. マクロを区分とした分析結果

システム	文化 (価値観)	価値観を表す考え (インタビューより)	背景
マクロシステム	パートナーシップ ・ 社会の責任  ・ 平等の精神 ・ パストラルケア	・ 同等の人間関係 ・ 支えあい (自分も障害者になるかもしれない) ・ 特別なニーズをもつ学生に対する配慮工夫は、全体の学生サービスの向上にもつながる。 ・ 社会全体として、障害者の自立した生活を保障したい → 平等とは、誰もが人間として他人に依存しないで生活できるようにしたい ・ 人への支援する心	社会背景：転職：人材の流動 経済的な要因：授業料の値上げ (2012～)

### 3.5 イングランドの大学図書館における特別な支援の背景要素

特別なニーズをもつ学生への直接影響にかかわるマイクロ・メゾシステムでは、大学図書館は、特別なニーズをもつ学生に対して、障害アドバイザーサービス、IT関連部署、サポートワーカーとの連携による支援を行っていた。各集団の共通要素では、「合理的配慮」、「特別なニーズをもつ学生に対するミッション」が挙げられた。学生への間接影響にかかわるエクソシステムでは、大学図書館担当者は、大学図書館地域別研究会 (CLAUD, OpenRose など)、大学図書館コンソーシアム (SCONUL)、高等教育支援機関 (JISCTechDis) とネットワークを結んでいた。支援要素では研修が挙げられた。この研修はインクルージョンを前提にしたサービス利便性の向上を目指している。また、大学図書館のインクルージョンに向けた物理的環境整備は、政府による法令や施策を根拠にしていた。社会全体の背景にあたるマクロシステムでは、障害者の自立した生活を保障していく、ともに支えあうパートナーシップが根底として挙げられた。

### 4. おわりに

イングランドの大学図書館における特別な支援

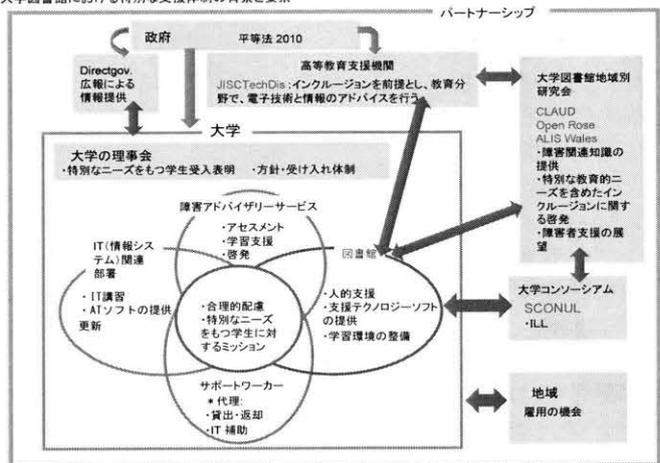


図1. イングランドの大学図書館における特別な支援の背景要素

の背景要素では、以下の結論が得られた。

- ① 職員の意識は、パートナーシップの精神を根底とする。
- ② 大学図書館支援は単独では行われていない。図書館、障害アドバイザーサービス、IT関連部署、サポートワーカーとの連携による。
- ③ 現場での相互の連携支援は、法的整備と組織としての体制づくりによる。
- ④ 大学外でサービス利便性のための研修とネットワークの構築がなされていた。これらの相互の影響が大学図書館の特別支援業務を支えていた。そして、別なニーズをもつ学生への支援へとつながっていた。

<sup>1</sup> 障がいのある学生の修学支援に関する検討会。障がいのある学生の修学支援に関する検討会報告 (第一次まとめ) について。文部科学省。2012。

<sup>2</sup> 同上。

<sup>3</sup> Fuller, Mary; Healey, Mick; Bradley, Andrew & Hall, Tim. Barriers to learning: a systematic study of the experience of disabled students in one university. *Studies in Higher Education*. v. 29, n. 3. 2004, p. 303-318.

<sup>4</sup> 松戸宏子。イングランドの大学図書館における特別な支援プロジェクトの形成過程。第60回日本図書館情報学会研究大会発表要綱。日本図書館情報学会。2012, p. 21-24.

<sup>5</sup> ブロンフェンブレナー。人間発達と生態学：発達心理学への挑戦。磯貝芳郎、福富護 (訳)。川島書店、1996、355p.

<sup>6</sup> Bristol University. Communication and support services. Who are the support workers? <http://www.bristol.ac.uk/disability-services/currentstudents/services/cssinfo.htm> 1. (参照日 2013/08/28.)

<sup>7</sup> JISC. About us. <http://www.jisc.ac.uk/about> (参照日 2013/8/30)

<sup>8</sup> CLAUD. CLAUD Members' Meetings Publicity. <https://wiki.brookes.ac.uk/display/CLAUD/CLAUD+Members%27+Meetings> (参照日 2013/8/30)

<sup>9</sup> Robertson, Linda SCONUL Access Steering Group. Access for library users with disabilities. 2002. [http://www.sconul.ac.uk/sites/default/files/documents/access\\_disabilities\\_0.pdf](http://www.sconul.ac.uk/sites/default/files/documents/access_disabilities_0.pdf) (参照日 2013/8/30)

本研究は、平成25年度文部科学省科学研究費補助金 (課題番号 25330390) の研究成果の一部である。

# 看護教育・研究をとりまく大学図書館の実態：「日本の図書館」を用いた分析

志茂 淳子<sup>†</sup>

<sup>†</sup>和歌山県立医科大学図書館三葛館  
shimo-nc@wakayama-med.ac.jp

**抄録** 急激に増加した看護系学部を利用対象とする大学図書館の、国公立別図書館整備状況の実態を『日本の図書館』を用いて調査した。その結果、看護系図書館が取り扱う資料の主題が多様化していること、雑誌の電子ジャーナル化が進んでおり、各大学における文献入手率が高まったことがうかがえた。国公立大学いずれにおいても、看護基礎教育と大学院教育双方を視野に入れると、それぞれ蔵書構成が偏っていること、特に私立大学においてはインプットとアウトプットともに大学間の格差が存在すること、職員数はさらに危機的な状況であることが示唆された。

## 1 序論

1952年にわが国最初の看護の4年制大学として高知県立女子大学家政学部看護学科が開設された後、1990年までの約40年間で、保健師、助産師、看護師の国家試験受験資格を取得することのできる看護基礎教育（以下、看護教育という）を実施する学部あるいは学科（以下、看護系学部という）が開設されたのは、僅か11校であった。その後、看護系学部は急増し、2013年4月現在、211校となっている。これらの看護教育の高等教育化は看護学の学問としての確立とともに看護職の社会的位置づけを高めることにも寄与してきた一方で、大学間に格差をもたらしていることが指摘されている<sup>1)</sup>。背景には、大学の設備や運営に加えて学生の意識や学力水準の格差による卒後の離職率の高さや、教員確保の困難さに起因した教育の質の格差などの問題がある。

2006年3月末時点でのデータをもとに91機関を対象として看護に視点をのいた図書館整備状況について調査し（以下、2006年調査という）、実態を分析した<sup>2)</sup>が、看護系学部はますます増加しており、看護の大学院修士課程や博士課程の開設も相次いでいる。そのような中で、看護教育を行っている大学で組織する日本看護系大学協議会は、機関評価としての認証評価のみでは、看護学領域の教育・研究の実態が反映されにくい現状があるとして、看護学の専門領域に特化した評価体系を構築するための事業を実施している<sup>3)</sup>。報告書の中で、大学図書館は学術支援サービスの1つとして評価の対象になっており、本研究では、図書館レベルで、看護に関わる大学図書館の整備状況を把握し、2006年調査との比較を行いながら、現状と問題点について考察する。

## 2 対象と方法

大学図書館についての全国統計のうち、個票データが公開されている『日本の図書館』の最新版である『日本の図書館：統計と名簿2012』<sup>4)</sup>を用いて、

2012年3月末時点で学部開設の完成年度を既に迎えている大学図書館を調査対象として、職員数、蔵書冊数、受入資料数、利用状況、予算について現状分析を行った。分析には、原則として、看護系学部の所在するキャンパスに設置された図書館（以下、看護系図書館という）を対象として、看護系学部生の収容定員（以下、看護学生数という）をもとに、看護学生数を当該図書館利用対象学生数で除した係数（以下、看護係数という）から各項目における看護の占有率を計算した。なお、今回の看護係数の算定には、各大学ウェブサイトから、平成22年6月15日の学校教育法施行規則の改正により義務付けられた「教育情報の公表」を参考にして、当該図書館の利用対象者を当該キャンパスの学部生の入学定員に課程年数を乗じて当該図書館利用対象学生数とし、教職員や大学院生は除いた。

調査項目は、本研究の目的に合わせ、選択して用いた。インプットとしての職員数、蔵書数、受入資料数、図書館費、アウトプットとしての入館者数、貸出数、相互協力業務について、各項目に看護係数を乗じて算出し、分析した。

さらに、2006年調査との比較を行うとともに、それ以降に看護系学部が新設された図書館のデータと以前から設置されていた図書館のデータとの比較を行った。

## 3 結果

### 3.1 基本データ

『平成20年看護関係統計資料集』<sup>5)</sup>の看護大学一覧によると、対象となる機関は168機関であったが、未回答の機関と『日本の図書館：統計と名簿2012』<sup>4)</sup>の調査対象となっていない機関を除くと、分析対象は163機関であった。国公立別には、それぞれ42機関、44機関、77機関であった。

看護係数は、国公立別に見ると、中央値が、国立(n=42)は0.26(0.50-0.02)、公立(n=44)は0.43(1.00-0.11)、私立(n=77)は0.33(1.00-0.02)であった。

2006年調査では、中央値が、国立(n=35)は0.29(0.50-0.07)、公立(n=29)は0.57(1.00-0.10)、私立(n=27)は0.40(1.00-0.04)であった。

なお、本調査では、各項目の数値がそれぞれ正規分布をしていないことから、次項以降も全て中央値を用いることとする。

### 3.2 インプット

大学図書館のインプットを示す指標として、職員数、蔵書冊数、受入図書冊数、所蔵雑誌種数、受入雑誌種数、図書館費について分析した。

今回の『日本の図書館：統計と名簿2012』<sup>4)</sup>を用いた調査(以下、2012年調査という)では、対象となった163件を国公立別に分析するとともに、そのうち、2006年調査時にも対象となっていた群(以下、2006対象群という)と、2006年調査時には開設されていないあるいは完成年度に達していなかったために対象となっていなかった群(以下、2006対象外群という)についても分析した。また、2006年調査のデータも併せて比較した。

職員数については、専従、非常勤のデータを用いた。兼務職員は対象外とした(表1)。蔵書冊数については、総冊数と洋書内数を用いた(表2)。受入図書冊数については、受入総冊数と洋書内数を用いた(表3)。所蔵雑誌種数については、所蔵雑誌総数と洋雑誌内数を用いた(表4)。受入雑誌種数については、受入雑誌総数と洋雑誌内数を用いた(表5)。

インプットの最も重要な指標としての図書館費は、調査項目のうち、資料費、図書費、雑誌費を用いた。さらに、公表されている項目にはないが、資料費から図書費と雑誌費を減じて、その他の資料費を算出した。無回答のデータとその他の資料費については、数値が矛盾するデータを除いたものを有効回答とした(表6)。

		専従	臨時・非常勤
国立	2012 (n=42)	0.66(2.37-0.00)	0.94(2.53-0.00)
	2006対象(n=34)	0.73(2.37-0.23)	0.94(2.48-0.00)
	2006対象外(n=8)	0.49(1.69-0.00)	0.97(2.53-0.10)
	2006 (n=35)	1.07(2.86-0.38)	0.95(3.84-0.30)
公立	2012 (n=44)	0.64(2.55-0.00)	1.71(15.00-0.00)
	2006対象(n=30)	0.53(2.55-0.00)	1.95(7.00-0.00)
	2006対象外(n=14)	0.7(1.30-0.00)	1.52(15.00-0.00)
私立	2012 (n=77)	0.75(6.11-0.00)	0.37(4.45-0.00)
	2006対象(n=29)	1.14(6.11-0.00)	0.62(3.25-0.00)
	2006対象外(n=48)	0.44(3.34-0.00)	0.23(4.45-0.00)
	2006 (n=27)	1.64(7.60-0.00)	0.80(3.50-0.00)

		蔵書冊数	うち洋書
国立	2012(n=41)	36,545(89,796-5,106)	17,258(54,151-929)
	2006対象(n=33)	36,356(66,089-10,091)	17,154(46,520-1,578)
	2006対象外(n=8)	40,959(89,796-7,852)	21,449(54,151-193)
	2006(n=34)	40,357(99,174-14,400)	19,453(47,868-2,400)
公立	2012 (n=44)	49,714(97,533-20,135)	7,136(42,227-709)
	2006対象(n=30)	52,230(97,533-22,675)	9,449(42,227-3,351)
	2006対象外(n=14)	45,052(76,692-20,135)	6,078(30,541-709)
	2006(n=29)	45,500(90,000-21,414)	8,000(39,333-3,158)
私立	2012 (n=77)	27,151(1,438,390-5,106)	4,104(57,899-193)
	2006対象(n=29)	38,354(143,839-10,670)	6,293(57,899-1,209)
	2006対象外(n=48)	24,777(90,898-5,106)	3,056(49,040-193)
	2006 (n=27)	34,000(123,000-5,436)	5,000(41,200-615)

		受入図書冊数	うち洋書
国立	2012 (n=41)	539(1187-70)	101(594-7)
	2006対象(n=33)	549(1187-214)	107(594-15)
	2006対象外(n=8)	409(977-70)	40(199-7)
	2006 (n=33)	684(1412-247)	306(780-10)
公立	2012 (n=44)	1265(3908-467)	61(235-0)
	2006対象(n=30)	1311(3908-467)	55(235-0)
	2006対象外(n=14)	1111(2125-597)	79(187-6)
	2006 (n=28)	1554(5506-470)	148(455-14)
私立	2012 (n=77)	909(8142-214)	70(740-0)
	2006対象(n=29)	1036(4052-286)	76(740-7)
	2006対象外(n=48)	748(8142-214)	63(489-0)
	2006 (n=27)	1104(2770-236)	112(450-8)

		所蔵雑誌総数	うち洋雑誌
国立	2012 (n=41)	914(2850-189)	447(2114-29)
	2006対象(n=33)	904(2710-226)	426(2114-39)
	2006対象外(n=8)	1017(2850-189)	538(2044-29)
	2006 (n=34)	1041(2318-287)	527(1837-58)
公立	2012 (n=44)	670(3010-186)	120(1056-7)
	2006対象(n=30)	601(2741-186)	139(1056-35)
	2006対象外(n=14)	763(3010-206)	106(1035-7)
	2006 (n=29)	418(2872-148)	113(1163-40)
私立	2012 (n=75)	253(3736-9)	48(2741-0)
	2006対象(n=28)	316(3736-95)	94(2741-22)
	2006対象外(n=47)	201(3438-9)	33(1868-0)
	2006 (n=25)	343(1815-23)	83(883-10)

		受入雑誌総数	うち洋雑誌
国立	2012 (n=41)	203(565-36)	45(361-3)
	2006対象(n=33)	203(335-55)	45(178-9)
	2006対象外(n=8)	191(565-36)	41(361-3)
	2006 (n=34)	293(534-119)	108(371-10)
公立	2012 (n=43)	260(2225-76)	51(130-3)
	2006対象(n=29)	242(2225-76)	52(100-5)
	2006対象外(n=14)	268(928-82)	36(130-3)
	2006 (n=27)	223(747-107)	72(247-13)
私立	2012 (n=75)	139(3244-11)	20(1831-0)
	2006対象(n=28)	154(2074-30)	32(1508-4)
	2006対象外(n=47)	104(3244-11)	19(1831-0)
	2006 (n=26)	217(855-23)	49(275-10)

		図書費	雑誌費	その他資料費
国立	2012 (n=41)	1,881(7,377-0)	2,698(11,120-0)	1,627(21,266-0)
	2006対象(n=33)	1,892(7,377-632)	2,777(11,120-126)	2,469(21,266-0)
	2006対象外(n=8)	455(5,703-0)	660(4,781-36)	196(2,752-0)
	2006 (n=34)	1,878(8,465-0)	9,555(17,731-0)	1,554(14,291-0)
公立	2012 (n=43)	2,825(9,989-270)	3,757(11,021-0)	2,467(17,667-0)
	2006対象(n=29)	3,231(9,989-270)	3,842(11,021-0)	2,399(17,652-0)
	2006対象外(n=14)	2,478(5,914-1,343)	2,231(6,473-694)	4,704(17,677-393)
	2006 (n=27)	4,900(16,900-307)	6,438(19,910-1,651)	1,052(6,207-0)
私立	2012 (n=75)	2,305(28,548-0)	2,133(19,992-29)	1,712(50,199-0)
	2006対象(n=28)	3,050(9,915-0)	2985(19,992-254)	3,226(50,199-0)
	2006対象外(n=47)	2,143(28,548-501)	1,843(7,610-29)	1,444(43,675-0)
	2006 (n=26)	2,679(10,092-270)	4,000(37,983-998)	938(16,359-0)

### 3.3 アウトプット

大学図書館のアウトプットを示す指標としては、入館者数と貸出数、相互協力業務について分析した。

年間入館者数と年間貸出数については、表7と表8の通りである。相互協力業務では、文献複写に注目し、複写受付件数と複写依頼件数をとり上げ、結果は表9の通りとなった。

		入館者数
国立	2012 (n=41)	34,083(66,451-3,648)
	2006対象 (n=34)	34,867(66,451-8,267)
	2006対象外 (n=7)	26,577(47,475-3,648)
公立	2012 (n=35)	32,000(92,000-1,000)
	2012 (n=44)	31,860(120,982-5,752)
	2006対象 (n=30)	30,574(120,982-5,752)
私立	2006対象外 (n=14)	33,602(89,937-13,599)
	2006 (n=27)	39,000(80,000-6,000)
	2012 (n=75)	21,302(75,893-8,878)
私立	2006対象 (n=28)	36,053(75,893-8,878)
	2006対象外 (n=47)	16,275(57,509-5,418)
	2006 (n=26)	34,000(92,000-1,000)

		貸出数
国立	2012 (n=42)	3,541(8,153-415)
	2006対象 (n=34)	3,702(8,153-1,128)
	2006対象外 (n=8)	2,656(6,363-415)
公立	2006 (n=27)	4,000(8,000-1,000)
	2012 (n=44)	10,507(40,114-2,281)
	2006対象 (n=30)	9,799(40,114-2,281)
私立	2006対象外 (n=14)	11,475(21,557-3,076)
	2006 (n=29)	9,000(30,000-2,000)
	2012 (n=77)	5,028(22,560-166)
私立	2006対象 (n=29)	6,900(22,560-1,027)
	2006対象外 (n=48)	3,986(12,391-166)
	2006 (n=27)	8,000(28,000-1,000)

		複写受付	複写申込
国立	2012 (n=41)	569(3,623-0)	468(1,829-0)
	2006対象 (n=34)	579(3,623-0)	477(1,829-0)
	2006対象外 (n=7)	444(1,879-131)	367(1,199-75)
公立	2006(n=34)	1,006(4,540-196)	961(1,998-328)
	2012(n=43)	423(9,252-0)	564(2,196-134)
	2006対象 (n=30)	425(9,252-0)	496(2,196-134)
私立	2006対象外 (n=13)	413(9,252-4)	533(969-169)
	2006(n=28)	487(4,012-19)	927(2,409-122)
	2012 (n=74)	108(3,671-0)	234(1,927-21)
私立	2006対象 (n=27)	367(3,671-79)	444(1,927-74)
	2006対象外 (n=47)	63(2,199-0)	157(1,109-21)
	2006 (n=24)	474(4,674-10)	656(3,677-102)

### 4 考察

「看護図書館」は、『最新図書館用語大辞典』<sup>6)</sup>によると、専門図書館の一種である「医学図書館」に含まれる概念と位置付けることができる。そのことを裏付けるように、本研究の対象とした看護系図書館は、医師や薬剤師、理学療法士をはじめとした医療技術職を養成する学部、さらには社会福祉系の学部などとの共用館が大部分を占めていた。特に、看護係数が2006年調査よりも低下している要因の一つに、これまで抑制していた医学部の入学定員が、「経済改革の基本方針 2008」閣議決定により、

2009年度以降増員されたことが挙げられる。また、看護係数が「1」を示す、純粹な意味での「看護図書館」は、国立大学で0%、公立大学で32%、私立大学では18%であった。このことから、昨今は、医学部との共用館である「医学図書館」も狭義の医学には収まっておらず、従来よりも広範囲の主題を扱っており、欧米における「Health science library」の様を呈していると言える。また、2006年調査対象外群においては、医療系として初めて看護系学部が設置された大学が見受けられ、看護系図書館が取り扱う資料の主題が多様化していることも示唆された。

蔵書冊数について、国立大学において、2006年調査と比較して減少しているのは、医学部定員増加による影響が考えられるが、蔵書冊数に対する洋書占有率は依然として5割前後を維持しているのに対し、公立大学と私立大学では15%前後、公立大学の2006対象外群においては僅か8%であった。これらのことは、国立大学が主に研究支援に重点を置いている所以であろうが、看護教育を支援するという意味では資料が偏っていて十分ではないと推測できる。一方で、公私立大学においても修士課程や博士課程が開設されている昨今、外国語資料の充実にも目を向ける必要があろう。公立大学においては、特に2006年対象群において蔵書冊数が確実に増加しており、2006年調査においても見られたように、蔵書数が確保されていると言えよう。私立大学においては、蔵書冊数、洋書占有率ともに、2006年調査と同様、大学間の格差が存在することが示唆された。

一方、受入図書冊数と図書費はいずれも2006年調査よりも減少傾向にあった。洋書については国立大学においても冊数が大幅に減少しているものの図書費は維持していることから、洋書の電子ブックへの移行が示唆された。また、図書については、2006年調査と同様、冊数、資料費ともに公立大学が最も充実しており、国立大学との差は明らかであった。受入雑誌種数と雑誌費については、国公立いずれにおいても洋雑誌の占有率が減少していることが顕著であり、雑誌費についても低下している一方で、その他資料費が全体的に増額していることから、看護系図書館においても、雑誌の電子ジャーナル化が進んでいる現状が窺えた。

電子ジャーナル化が進んでいる状況は相互協力業務にも表れている。米田<sup>7)</sup>が指摘したとおり、2006年調査と比較すると、ビッグディール後のILLにおいては看護系図書館においても文献の流通量が減少し、各大学における文献の入手率が高まったことが推測できる。佐藤<sup>8)</sup>が指摘したILLにおける看護文献の流通動向のその後の研究が期待されるが、「メディカルオンライン」や「Pier Online」などに見られる看護の商業雑誌

あった公立大学の2006対象群が2012年調査においては最も低値であった。その反面、貸出数は公立大学のみ増加した。反対に2006対象群の国立及び私立大学は入館者数が増加し、貸出数は低下した。この背景には、特に国立大学や大規模私立大学で実践しているラーニングコモンズの設置や学習支援への取り組みなどの影響が考えられる。

データを細かく見ると、私立大学において、2006対象外群の複写受付件数や入館者数、貸出数が低調なことが私立大学全体の数値の低下に影響を及ぼしていた。このことから、新設看護系学部におけるジャーナルコレクションの貧弱さや看護の学習・研究活動への支援が発展途上であることがうかがえ、大学間に格差が存在することが示唆された。

これらのインプットとアウトプットを支える職員数は、2006年調査と比較して国公立いずれにおいても減少した。さらに、公立大学において職員の非常勤率が約71%と、依然として高かった。私立大学においては非常勤職員率が全体と比較して低値であることからさらなる委託・派遣職員の増加が推測される。いずれにしても、求められる大学図書館像<sup>9)</sup>を視野に入れると、看護系図書館の危機的な状況に拍車をかけていると言えよう。

## 5 研究の限界

本研究は、大学図書館の調査・研究への使用が妥当とされる<sup>10)</sup>『日本の図書館』を用いて実施したが、電子情報資源へのアクセスを保証することが大学図書館の基本的な課題である<sup>9)</sup>現在、いくつかの限界が明らかになった。まず1点目は、職員数における、委託や派遣によるスタッフ数についての調査項目が存在しないことにより、相対的な比較ができないことである。2点目は、所蔵・受入資料数について、電子情報源への対応がされていないことで、電子情報化時代の大学図書館の実態をありのまま評価できないことである。3点目に、電子資料の利用に関する指標が設けられていないことである。利用者の属性によっては来館型から非来館型の利用へシフトしている現在、図書館のアウトプットを客観的に評価するための調査項目の設定と、電子資料の利用について全ての図書館が簡便かつ効率よく算出できるような環境やツールの整備、研修機会の提供が望まれる。

ますます進んでおり、ILLにおいても各大学における文献の入手率が高まったことがうかがえた。国立大学においては、2006年調査から依然として図書、雑誌ともに洋書占有率が高いため、看護基礎教育を支援するという意味では十分ではない反面、公私立大学においては洋書占有率が低いことで大学院教育を視野に入れると十分とは言えず、看護系図書館としての蔵書の偏りが明らかになった。また、特に私立大学において、インプットとアウトプットともに大学間の格差が存在することが示唆された。さらに、看護系図書館を支える職員数は危機的な状況に拍車をかけていた。

## 【引用文献】

- 1) 日本看護歴史学会編集『日本の看護120年：歴史をつくるあなたへ』日本看護協会出版会、2008。
- 2) 志茂淳子「看護教育・研修を支援する大学図書館の実態」『看護と情報』16, 2009, 74-79。
- 3) 東京大学大学院医学系研究科地域看護学分野編『看護学専門領域の評価基準・評価体制の開発研究事業報告書：看護系大学・大学院の質向上システムの構築を目指して：平成19年度文部科学省大学評価研究委託事業』日本看護系大学協議会、2008。
- 4) 日本図書館協会編『日本の図書館：統計と名簿2012』日本図書館協会、2013。
- 5) 日本看護協会出版会編『看護関係統計資料集：平成20年』日本看護協会出版会、2008。
- 6) 図書館用語辞典編集委員会編『最新図書館用語大辞典』柏書房、2004。
- 7) 米田奈穂ほか「ビッグディール後のILL：千葉大学附属図書館玄鼻分館における調査」『大学図書館研究』76, 2006, 74-81。
- 8) 佐藤義則「近年のNACSIS-ILLにおける看護文献の需要と供給」『看護と情報』14, 2007, 69-76。
- 9) 文部科学省『大学図書館の整備について（審議のまとめ）—変革する大学にあって求められる大学図書館像—』オンライン、  
[http://www.mext.go.jp/b\\_menu/shingi/gijyutu/gijyutu4/toushin/1301602.htm](http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/toushin/1301602.htm), [accessed 2013-09-01]。
- 10) 立石亜紀子ほか「日本の大学図書館統計の現状と課題」『現代の図書館』47(1), 2009, 54-65。

# 大学図書館における未貸出図書の分析

\*†西野祐子 \*松野渉 \*池内有為 \*\*佐藤翔 \*\*\*逸村裕

\*筑波大学大学院

\*\*同志社大学社会学部

\*\*\*筑波大学図書館情報メディア系

図書館情報メディア研究科

†s0911636@klis.tsukuba.ac.jp

## 抄録

本研究は、大学図書館における図書の利用実態の解明を目的とし、2006年度に受入された貸出可能な図書10,401件を対象に、2006年4月から2011年3月までに一度も貸出されなかった図書、すなわち未貸出図書と、同期間に貸出された図書とを比較し、その属性の違いを分析した。集計の結果未貸出図書の割合は34%となった。さらに言語別に貸出された図書と未貸出図書のクロス集計を行い、 $\chi^2$ 乗検定を行った結果有意差が見られた。

## 1. 研究背景・目的

大学図書館において蔵書は重要な経営資源である<sup>1)</sup>。図書館蔵書の利用のされ方に関する研究は、図書館情報学における古典的問題意識であり、大学図書館においても多くの研究がなされてきた。近年では松野ら(2012)の研究によって、大学図書館の蔵書の3%で貸出の80%を充足することが明らかになった<sup>2)</sup>。また、2010年にCornell大学図書館で行われた調査では、1990年受入図書のうち2010年までに一度も貸出されなかった図書、すなわち未貸出図書の割合が55%であったと報告されている<sup>3)</sup>。

未貸出図書の分析に関する先行研究は以下のものがある。まず、公共図書館を対象に、伊藤ら(1984)の研究が先駆的に行われている。これは未貸出図書の価格や分野、重さなどの図書の形態で分析している<sup>4) 5)</sup>。

大学図書館の蔵書を対象にしたものは、鬼頭ら(1986)の研究がある。この研究では、開館した年度である1953年に受け入れた図書に未貸出図書が多いこと、NDCの200番台の図書の未貸出割合が高いことが明らかにされた<sup>6)</sup>。

未貸出図書が蔵書に占める割合の、経年変化に特化した研究としては岸田ら(1994)

がある。この研究では、蔵書に占める未貸出図書の割合は毎年一定であり、とくに0類や2類の図書の未貸出図書の割合が高いことが述べられている<sup>7)</sup>。

先行研究における未貸出図書の分析では、未貸出図書だけを対象としており、言及された未貸出図書の特徴が、貸出された図書(以下貸出図書)と統計的に有意な差をもつかどうか明らかになっていない。

本研究では、大学図書館の蔵書における未貸出図書を、貸出図書と比較することで、大学図書館における蔵書の利用実態を把握することを目的とする。未貸出図書の発生原因が解明されることで、貸出される図書の特性と合わせて、図書館経営における意思決定に貢献することができよう。

## 2. 研究方法

### 2.1 分析手法

本研究では、まずKent et al (1979)の研究<sup>8)</sup>で用いられた手法を参考に、2006年度に受入を行った図書の、その後の貸出回数を調査した。次に言語、大きさ、ページ数、複本の有無、主題について、測定の結果貸出回数1回以上の貸出図書と、0回の未貸出図書をそれぞれ集計した。そして

貸出図書と未貸出図書との間の統計的有意差の有無を検証した。なお分析には予め図書館が加工したデータを使用した。

## 2.2 分析対象

本研究では、学生数 15,000 人規模の国立 A 大学の、蔵書数 300 万冊程度の中央図書館において、2006 年度に受入された貸出可能な図書を分析対象とした。貸出履歴データは、2006 年 4 月から 2012 年 3 月までの、中央図書館の貸出可能な図書の貸出履歴データを使用した。利用者は限定せず、学生、教職員、学外利用者も含む。全貸出履歴データから本研究の分析対象とする貸出履歴データを抽出した結果を表 1 に示す。

未貸出図書の抽出は以下のように行った。その結果を表 2 に示す。2006 年度に受入された資料 31,142 冊から、雑誌や参考図書など貸出不可能な資料を除く。そしてその中から中央図書館の図書 10,401 件のみ抽出する。この書誌データと貸出履歴データとを照合し、書誌ベースで貸出回数を集計する。複本がある場合は、各複本の貸出回数を合算した。そして貸出回数が 0 となった 3,860 件を未貸出図書とした。

表 1 2006 年 4 月—2012 年 3 月までの  
貸出履歴データ延べ件数

貸出履歴データ延べ件数	2,242,497
うち対象とした件数	1,531,853

表 2 対象データタイトル件数

2006 年度受入資料冊数	31,142
うち貸出可能な図書	21,924
うち中央図書館の図書	11,401
うち貸出図書	7,541
未貸出図書	3,860

## 3. 分析結果

未貸出図書が、貸出可能な 2006 年度受入

図書数全体に占める割合は 34%となった。

### 3.1 言語

貸出可能な受入図書の言語は全部で 25 種類ある。日本語 (62.8%)、英語 (26.3%)、中国語 (5.6%)、ドイツ語 (1.7%)、フランス語 (2%)、韓国語 (0.4%) の 6 言語で、受入図書の 97%を占めている。外国語図書の割合は 40.7%である。言語を日本語と日本語以外つまり外国語図書に分けてクロス集計を行い、 $\chi^2$  二乗検定を用いて 1%両側検定を行った結果、有意差が見られた。

また、貸出の有無と、図書本文の言語でクロス集計を行い、 $\chi^2$  二乗検定を用いて 1%両側検定を行った結果、有意差が見られた。中国語は受入図書件数の 82%が未貸出図書となっている。これは中国語図書の大半が『俗文学叢刊』や『胡適全集』等の史料集や全集で、その多くが未貸出図書となっているためである。表 3 は言語別クロス表と日本語か否かの別クロス表を一つにまとめたものである。

表 3 言語と貸出の有無別図書件数

	未貸出 図書	貸出図書	合計	未貸出 図書割合
日本語	1,394	5,770	7,164	19%
外国語	2,466	1,771	4,237	58%
英語	1,594	1,407	3,001	53%
中国語	528	115	643	82%
フランス語	117	108	225	52%
ドイツ語	108	81	189	57%
韓国語	34	16	50	68%
その他言語	85	44	129	66%

### 3.2 大きさおよびページ数

大きさは、図書の縦または横の長さのうちの長いほうとした。これを判型に基づいて表 4 のように四つに区分し、貸出の有無

別に図書件数を集計した。そして $\chi^2$ 乗検定を用いて両側 1%の検定を行った結果有意差がみられた。大きさが大きくなるにつれて未貸出図書の割合が高くなっている。

表4 大きさと貸出の有無別図書件数

大きさ	未貸出	貸出有	合計	割合
15-19cm	95	687	782	12%
20-24cm	715	3,647	4,362	16%
25-29cm	215	571	786	27%
30-34cm	78	101	179	44%

ページ数については、極端にページ数の少ないまたは多い図書を除き、t検定を用いて、貸出の有無別にページ数の平均値の有意差を両側 1%で検定したが、有意差はみられなかった。図1は、日本語図書貸出の有無別の日本語図書のページ数の分布を示す箱ひげ図である。また貸出の有無別の日本語図書のページ数の平均値および中央値を表5に示す。

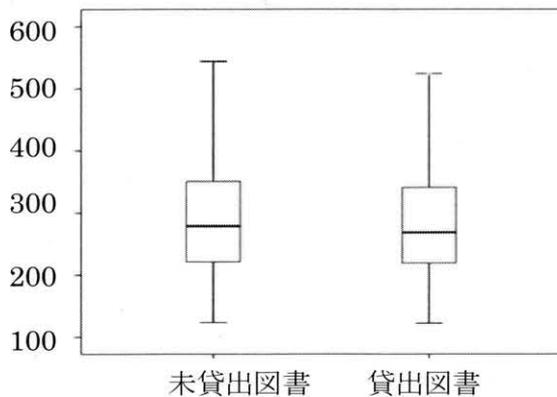


図1 日本語図書のページ数分布

表5 日本語図書ページ数平均値と中央値

	平均値	中央値
未貸出図書	292.88	278.50
貸出図書	285.73	268.00

### 3.3 複本の有無

貸出の有無と複本の有無別の図書件数をクロス集計し、 $\chi^2$ 乗検定を用いて両側 1%の検定を行った結果有意差がみられた。クロス表を表6に示す。この結果から、複本のある図書は未貸出図書となりにくいことがわかった。

表6 貸出の有無と複本の有無別図書件数

	未貸出	貸出有	合計
複本なし	3,788	7,159	10,947
複本あり	72	382	454
合計	3,860	7,541	11,401

### 3.4 主題

NDCの一次区分別に貸出の有無をクロス集計した。その結果、2類、4類、9類の未貸出図書の割合が40%を越えた。詳細を分析するため、二次区分、三次区分別に同様の集計を行った結果、2類は「222 中国史」の未貸出図書の割合が高くなっていた。4類は「41 数学」、9類は「92 中国文学」と「93 英米文学」の未貸出図書の割合が高い。未貸出図書のほとんどは、2類は『嘉慶帝起居註』22冊のような中国語の史料集であった。また4類は数学と物理のテキスト・シリーズになっている英語図書であった。そして9類は英米文学の英語図書および『俗文学叢刊』100巻のような中国文学の中国語全集となっていた。

### 4. 考察

本研究の分析結果から外国語図書、判型が大きい図書は未貸出図書となりやすいと考えられる。そして外国語図書の未貸出図書が多いことが、判型、出版年等の他の要素に影響を与えていると考えられる。とりわけ中国語の史料集、文学全集といった多巻ものにおいて未貸出図書の割合が高い傾向が見られた。これは、多巻ものの性質上、全巻をそろって所蔵していなければ意味を

なさない一方で、利用が一部の巻に留まってしまうためであると考えられる。

しかし外国語で書かれていることも要因として存在すると考えられる。例えば英語図書の入数が多き英米文学については、未貸出図書となった図書 104 件すべてが英語図書となっており、英米文学の図書全体に占める割合は 68% になっていた。

また、数学や物理学の英語で書かれたシリーズものの図書は、大学院の各研究科の予算で購入され、研究科所有の資料室に貸出されているものが大半であったが、複本として図書館に排架されている図書は貸出が一度もなかった。これは大学院生や教員が研究用に読むための内容であり、大学院生や教員は手近な資料室を利用しているため貸出がなかったと考えられる。このことから、図書の購入主体や購入、寄贈の別も未貸出図書の特徴にかかわってくることも示唆される。

以上結果から考えられる要因について考察した。多巻ものや研究向けの図書の所蔵が大学図書館の役割であるとするならば、これらが未貸出図書となっても問題にはならない。つまりそれらを要因から除けば、図書館の選書がうまく機能しているとも捉えられる。今後は、図書館利用をさらに活発にするためにどのような蔵書構築を行うかを含め、いっそうの検討を続けたい。

## 引用文献

- 1) 岸田和明. 蔵書評価に関する調査研究 : 蔵書評価とその手法. 図書館調査研究レポート. 2006, No. 7, p. 5-14.
- 2) 松野渉, 本田咲美, 池内有為, 佐藤翔, 逸村裕. “大学附属図書館における貸出履歴の分析”. 日本図書館情報学会春季研究集会発表要綱. 三重, 2012-5-12, 日本図書館情報学会. 日本図書館情報学会, 2012, p. 57-60.
- 3) The Collection Development Execu-

tive Committee. “Report of the Collection Development Executive Committee Task Force on Print Collection Usage Cornell University Library”. Cornell University Library Staff Web. [http://staffweb.library.cornell.edu/system/files/CollectionUsageTF\\_ReportFinal11-22-10.pdf](http://staffweb.library.cornell.edu/system/files/CollectionUsageTF_ReportFinal11-22-10.pdf). (参照日:2013/09/05)

- 4) 伊藤昭治, 芝勝徳. 公立図書館における大規模開架と貸出図書の分析. 図書館界. 1983, Vol. 35, No. 4, p. 171-184.
- 5) 伊藤昭治, 竹島昭雄, 芝勝徳, 柏悦子, 川崎良孝ほか (読書調査研究グループ). 大規模開架の公立図書館において、一度も借りられなかった本の分析. 図書館界. 1984, Vol. 36, No. 3, p. 166-175.
- 6) 鬼頭當子, 黒沢公人. 全面開架制度の図書館に於ける未貸出図書の分析. 私立大学図書館協会会報. 1986, No. 87, p. 67-99.
- 7) 岸田和明, 逸村裕, 高山正也. 大学図書館における館外貸出データの分析手法・オブソレッセンスと貸出頻度の分析を中心として. 図書館研究シリーズ. 1994, No. 31, p. 79-127.
- 8) Kent, Allen; Cohen, Jacob; Montgomery, K. Leon; Williams, James G.; Bulick, Stephen; Flynn, Roger R.; Sabor, William N.; Mansfield, Una. Use of library materials : the University of Pittsburgh study. New York, M. Dekker, 1979, 272p, (Books in library and information science, v. 26).

# 大学図書館における電子ジャーナルの引用と閲読のオブソレッセンス分析

武井千寿子 (筑波大学大学院図書館情報メディア研究科: naoe.chizuko@ynu.ac.jp)

芳鐘冬樹 (筑波大学図書館情報メディア系)

逸村裕 (筑波大学図書館情報メディア系)

## 抄録

SpringerLink の全 11 分野と ScienceDirect の 20 分野から分野毎に無作為抽出した学術雑誌約 1,200 誌を対象に、分野毎の引用と閲読のオブソレッセンスの相関を調査した。Cited Half-life, Download Half-life などの指標を用いて調査した結果、物理学をはじめとする医学以外の理系分野において、長期におけるオブソレッセンスに関する指標の相関が比較的高かった。また、観察年によって、心理学などは相関の程度が著しく変動しうることが示唆された。

## 1. はじめに

図書館における蔵書評価方法の一つに、オブソレッセンス分析がある。オブソレッセンス分析とは、経年変化に伴う文献利用の廃れを調査し評価を行う手法である。

現在、大学図書館では、Big Deal 離脱時のサーバーティネットとして、バックファイル整備が急務となっている。効果的なバックファイル整備方法の検討については、武井らが指摘するように、既往研究でも注目されているところである<sup>1)2)</sup>。しかし、既往研究の多くは、少数の分野を対象にした調査に留まっている<sup>3)4)5)6)</sup>。武井らの研究は、Springer の SpringerLink 全分野について引用と閲読のオブソレッセンスの関係を調査し、前者から後者の予測可能性を示唆するものとなっているが、標本数が少なく、その結果を分野全体に適用するには不十分なものであった。

そこで本研究では、より多くの標本を得るため、SpringerLink のほか、Elsevier の ScienceDirect も調査対象とする。SpringerLink は STM 中心の全 11 分野からなるコレクション、ScienceDirect は STM のほか、社会科学分野も含む全 23 分野のコレクションである。両者を分析することで、さらに広範な分野について調査および比較できるほか、重複する分野においては、より多くの標本をもとに分析可能となる。そして、分野によってあてはまりの良い指標、すなわち、バックファイル導入後の効果を予測するのに有効なオブソレッセンスの指標が異なるという想定のもと調査を行い、分野毎の引用と閲読のオブソレッセンスの差異を明らかにする。

## 2. 研究方法

本研究では、横浜国立大学を調査対象とする。用いるデータは、引用データとして 2009 年～2012 年版の Journal Citation Reports (以下、JCR)、閲読データとして SpringerLink (2010 年～2012 年) と ScienceDirect (2001 年～2012 年) の出版年別フルテキスト利用統計 (COUNTER Journal Report 5) を用いる。また、SpringerLink は全 11 分野、ScienceDirect については、全 23 分野のうち、分析対象誌数が 10 未満であった 3 分野を除く 20 分野を対象とする。分析にあたり、各分野から以下の手順で分析対象誌を無作為抽出した。

①SpringerLink の全 2,782 誌、ScienceDirect の全 1,792 誌から、一度も閲読がない雑誌を除き、さらに出版社のタイトルリストから分野が判明した雑誌のみ抽出した<sup>7)</sup>。その結果、SpringerLink は 1,567 誌、ScienceDirect は 1,657 誌となった。

②①の雑誌のうち、JCR 当該年版に指標の値が掲載されたものを抽出し、分野別に閲読数(アクセス数)の降順に並びかえた。分野全体の傾向を調査するため、閲読数の累積比率に応じて、70%未満、70%以上～90%未満、90%以上の 3 つに分け、各区分から 15 誌ずつ無作為に抽出し、15 誌以下の区分についてはその全数を調査対象とした。また、抽出にあたり、用いる指標の算出を可能にするための以下の条件を加え、合致するものをそれぞれ抽出した。

(a) 長期におけるオブソレッセンス分析のため、当該年の閲読数が 1 以上であるもの。

(b) 短期におけるオブソレッセンス分析のため、2010～2012 年にコレクションに含まれるもの。

(c) 各指標の関係を調査するため、(a)(b)ともに満たすもの。

この結果、最終的な分析対象タイトル数は、SpringerLink は (a) 450 誌, (b) 475 誌, (c) 432

表1 SpringerLink 分野別タイトル数

分野	a	b	c
Behavioral Science (BS)	18	31	18
Biomedical and Life Sciences (BL)	45	45	45
Business and Economics (BE)	38	42	33
Chemistry and Materials Science (CM)	45	45	45
Computer Science (CS)	45	45	44
Earth and Environmental Science (EE)	45	45	45
Engineering (EG)	42	42	42
Humanities, Social Sciences and Law (HS)	38	45	34
Mathematics and Statistics (MS)	45	45	45
Medicine (MD)	44	45	36
Physics and Astronomy (PA)	45	45	45
合計	450	475	432

表2 ScienceDirect 分野別タイトル数

分野	a	b	c
Agricultural and Biological Sciences (AB)	41	41	41
Biochemistry, Genetics and Molecular Biology (BG)	45	45	45
Business, Management and Accounting (BM)	36	34	34
Chemical Engineering (CE)	40	40	40
Chemistry (CH)	36	35	35
Computer Science (CS)	45	45	45
Earth and Planetary Sciences (EP)	45	45	45
Economics, Econometrics and Finance (EP)	45	45	45
Energy (EN)	22	21	21
Engineering (EG)	45	45	45
Environmental Science (ES)	36	36	36
Health Sciences (HE)	45	43	42
Immunology and Microbiology (IM)	37	37	36
Materials Science (MT)	43	42	42
Mathematics (MA)	36	36	36
Neuroscience (NS)	38	34	34
Pharmacology, Toxicology and Pharmaceutical Science (PT)	30	29	29
Physics and Astronomy (PA)	33	33	33
Psychology (PC)	36	29	29
Social Sciences (SS)	39	37	37
合計	773	752	750

誌, ScienceDirect:は (a) 773 誌, (b) 752 誌, (c) 750 誌となった。分野別タイトル数は表1, 表2に示すとおりである。

また、本研究では、以下の指標をオブソレッセンスの尺度として用いて調査を行う。

(1) 引用のオブソレッセンスに関する指標:

(1A) Cited Half-life (CHL)

(1B) Immediacy Index / Impact Factor (II/IF)

(2) 閲読のオブソレッセンスに関する指標:

(2A) Download Half-life (DHL)

(2B) Download Immediacy Index / Download Impact Factor (DII/DIF)

CHL と DHL は値が大きいほど廃れが遅く、II/IF と DII/DIF は値が大きいほど廃れが早いことを表す。また、CHL と DHL は長期における廃れを表すのに対し、II/IF と DII/DIF は出版後数年の短期における廃れに関する指標である。これらの指標について、まず2012年時点の値を求めた。CHL, II, IF はJCR2012年版からデータを入手し、DHL, DII, DIF はそれぞれCHL, II, IF の定義を閲読に準用して算出した。ただし、

(b)の短期におけるオブソレッセンス分析では、ゼロ除算回避のためDIIとDIF各々の値に1を加えて用いた。さらにScienceDirectについては、2011年と2012年、2つの観察基準年における結果を比較し、直近の文献利用の変化を観察するため、2011年時点の値も同様の方法で入手した。

具体的な調査手順としては、まず長期(CHLとDHL)、短期(II/IFとDII/DIF)それぞれに関して、引用と閲読のオブソレッセンスの一致、つまり相関の程度を分野毎に調査する。そして、SpringerLinkとScienceDirectの結果を合わせ、分野間の差異を明らかにする。さらに、ScienceDirectについては、2011年と2012年の結果を対照し、その差異を分析する。

### 3. 結果

引用と閲読のオブソレッセンスがどの程度一致しているかみるため、まず(A)CHL対DHL、(B)II/IF対DII/DIF、それぞれについて相関を観察した。II/IFとDII/DIFは2.77~13.27と分布の歪度が高く、またJCRにおけるCHLは最大値が10であり、実際にはそれ以上の値でも10として扱われることから、正規分布に従ってい

ないと判断し、スピアマンの順位相関係数  $\rho$  を算出した。(A) CHL 対 DHL, (B) II/IF 対 DII/DIF, それぞれの相関係数を分野毎に示したのが表 3 である。SpringerLink と ScienceDirect では分野数も収録対象分野も異なることから、両者の結果を比較しやすくするため、各々の分野を人文・社会科学系, 医学系, 化学・工学系, 数学・コンピュータ科学系, 農学・環境科学系, 物理学系の6分野へ再分類し, 併せて表3に示した。

SpringerLink の 11 分野全体における相関係数は, (A) は  $\rho = 0.50$  ( $p < 0.05$ ), (B) は  $\rho = 0.12$  ( $p < 0.05$ ) であり, (A) については中程度の相関が観察されたが, (B) の相関はほとんどみ

表 3 両コレクションにおける分野の傾向

分野	2012A	2012B	2011A	2011B	
人文・社会科学系	BS(S)	0.21	-0.08		
	BE(S)	<b>0.55</b> *	0.31 *		
	HS(S)	0.35 *	-0.27		
	BM(E)	0.09	0.07	-0.31	0.02
	EF(E)	0.26	0.18	0.13	-0.02
	PC(E)	0.16	-0.16	-0.04	<b>-0.52</b> *
	SS(E)	0.05	0.34 *	0.36 *	0.22
医学系	BL(S)	<b>0.51</b> *	0.16		
	MD(S)	0.28	-0.14		
	HE(E)	0.09	0.28	0.22	-0.27
	IM(E)	0.05	-0.30	0.18	0.13
	NS(E)	0.30	-0.29	0.18	<b>-0.43</b> *
	PT(E)	0.08	0.39 *	0.27	-0.31
化学・工学系	CM(S)	<b>0.57</b> *	0.05		
	EG(S)	<b>0.50</b> *	0.37 *		
	BG(E)	0.26	0.03	<b>0.50</b> *	0.03
	CE(E)	<b>0.60</b> *	0.03	<b>0.57</b> *	-0.01
	CH(E)	0.30 *	-0.34 *	<b>0.66</b> *	0.36 *
	EG(E)	0.34 *	0.15	<b>0.42</b> *	0.22
	MT(E)	<b>0.56</b> *	-0.11	<b>0.56</b> *	-0.20
数学・コンピュータ科学系	CS(S)	<b>0.46</b> *	0.04		
	MS(S)	<b>0.43</b> *	-0.03		
	CS(E)	<b>0.25</b> *	0.14	0.23	0.01
	MA(E)	0.36 *	-0.27	<b>0.41</b> *	-0.06
農学・環境科学系	EE(S)	<b>0.47</b> *	0.16		
	AB(E)	0.15	-0.08	0.36 *	-0.21
	ES(E)	<b>0.46</b> *	0.15	0.39 *	0.13
物理学系	PA(S)	<b>0.59</b> *	0.16		
	EP(E)	0.32 *	0.09	0.32 *	-0.09
	EN(E)	<b>0.62</b> *	0.29	<b>0.73</b> *	-0.08
	PA(E)	0.35 *	0.33	0.33	0.16
分野全体	(S)	<b>0.50</b> *	0.12 *		
	(E)	0.30 *	0.07	0.37	0.03

S: SpringerLink の分野, E: ScienceDirect の分野

\* $p < 0.05$

られなかった。分野個別では, (A) について最も強く有意な相関がみられたのは Physics and Astronomy (0.59) で, (B) では Engineering (0.37) で比較的強い有意な相関がみられた ( $p < 0.05$ )。一方, ScienceDirect の 20 分野全体の相関係数は, (A) は 2012 年:  $\rho = 0.30$  ( $p < 0.05$ ), 2011 年:  $\rho = 0.37$ , (B) は 2012 年:  $\rho = 0.07$ , 2011 年:  $\rho = 0.03$  であった。分野別にみると, 観察基準年によって変動があるが, (A) は Energy で 2012 年:  $\rho = 0.62$ , 2011 年:  $\rho = 0.73$  と最も強く有意な相関が観察されたほか, 2012 年は Chemical Engineering (0.60), 2011 年は Chemistry (0.66) でも比較的強い有意な相関がみられた ( $p < 0.05$ )。(B) の指標については, 2011 年の Psychology で  $\rho = -0.52$  ( $p < 0.05$ ), Neuroscience で  $\rho = -0.43$  ( $p < 0.05$ ) と中程度の相関が観察された以外は, SpringerLink 同様にほとんどの分野で弱い相関しかみられなかった。また, 再分類後の 6 分野の傾向から, 医学系以外の理系分野, 特に物理学系と化学・工学系分野において (A) の CHL 対 DHL の指標のあてはまりの良さが読みとれる。

さらに, 両コレクションの分野全体について, (c) の抽出条件に合致した標本をもとに, (A), (B) 以外の指標も含めて, 指標同士の相関を調査した。既往研究の結果との対比と, 元々の値の強さを考慮するため, ピアソンの積率相関も順位相関と併せて調査した。積率相関を算出する際, 分布の歪度を抑えるために対数変換を行った。表 4, 5 は SpringerLink のそれぞれの結果であるが, ScienceDirect の 2011 年, 2012 年についても類似の結果が得られた。なお, 表のグレー部分は引用と読書の指標の相関, 枠で囲った部分はオブソレッセンスに関する指標の

表 4 SpringerLink の 11 分野全体における各指標の相関の程度 (順位相関)

	IF	DII	DIF	CHL	DHL	II/IF	DII/DIF
II	0.78 *	-0.10 *	-0.01	-0.14 *	-0.06	0.59 *	-0.05
IF	1	-0.12 *	0.04	-0.18 *	-0.05	0.00	-0.08
DII		1	-0.73 *	-0.20 *	-0.06	-0.01	0.92 *
DIF			1	0.18 *	0.09	-0.07	-0.93 *
CHL				1	<b>0.49</b> *	0.03	-0.21 *
DHL					1	-0.02	-0.08
II/IF						1	0.03
DII/DIF							1

\* $p < 0.05$

表5 SpringerLinkの11分野全体における各指標の相関の程度(対数変換後の積率相関)

	IF	DII	DIF	CHL	DHL	II/IF	DII/DIF
II	0.76 *	-0.07	-0.03	-0.16 *	-0.06	0.62 *	-0.02
IF	1	-0.07	0.01	-0.19 *	-0.03	-0.04	-0.04
DII		1	-0.76 *	-0.19 *	-0.08	-0.03	0.93 *
DIF			1	0.19 *	0.12 *	-0.06	-0.94 *
CHL				1	0.47 *	-0.01	-0.20 *
DHL					1	-0.05	-0.11 *
II/IF						1	0.02
DII/DIF							1

\* $p < 0.05$

相関を示している。順位相関、積率相関の結果に大きな差はみられず<sup>8)</sup>、オブソレッセンスに関する指標間では、(A)の指標にあたるCHL対DHLで約0.5( $p < 0.05$ )と最も強い有意な相関が観察されたが、それ以外では弱い相関しかみられなかった。

#### 4. 考察と今後の展望

引用と読書のオブソレッセンス分析結果より、SpringerLinkの8分野とScienceDirectの一部の分野において、長期におけるオブソレッセンスの指標であるCHL対DHL間で0.4以上の有意な相関が観察された。また、全分野を6分野に再分類したところ、物理学など医学以外の理系分野でCHL対DHLの指標で比較的強い有意な相関が観察された。これらより、本研究でも引用のオブソレッセンスの指標から読書のオブソレッセンスを予測する、すなわち、JCRのCHLの値に基づいて、バックファイル導入後の効果がある程度予測することが可能であると言える。

また、分野全体について順位相関に加え、積率相関も用いて指標同士の相関を調査した結果、両者に大きな差異はみられなかった。既往研究の一例として、中国の全文データベースを対象に調査したWanらの調査結果では、農学・林学分野でDII対IIは $\rho = 0.24$  ( $p = 0.0964$ )、DII対IFは $\rho = 0.41$  ( $p = 0.0034$ )、II対IFは $\rho = 0.59$  ( $p < 0.0001$ )、心理学分野ではDII対IIは $r = 0.8$ であった<sup>9)</sup>。本研究ではDII対IIおよびDII対IFはほとんど相関が見られなかったが、II対IFは強い有意な相関が観察された。これは、対象分野および対象データベースの違いが影響していると考えられる。

また、ScienceDirectの2011年と2012年の結

果から、年によってPsychologyなどのいくつかの分野は相関の程度が著しく変動する可能性があることと、より強い相関を示す指標が異なる可能性があることが観察された。II/IF対DII/DIFの変動については、出版後数年間における読書は、論文数や刊行頻度、特集などの影響を受けやすいことから、ある程度予測できる。また、CHL対DHLの変動についても、雑誌の他社への移管やタイトル変遷、廃刊等の影響が考えられる。しかし、今回は2つの基準年しか調査していないため、その要因についてはさらなる調査が必要である。

#### 注・引用文献

- 1) 武井千寿子, 芳鐘冬樹, 逸村裕. “大学図書館における電子ジャーナルの利用状況: 効果的なバックファイル整備のための読書と引用のオブソレッセンス分析”. 第60回日本図書館情報学会研究大会発表要綱. 福岡, 九州大学, 2012年11月, p. 53-56.
- 2) Chizuko, T.; Yoshikane, F. and Itsumura, H. “Use of electronic journals in university libraries: an analysis of obsolescence regarding citations and access,” In *Proceedings of the ISSI conference*, Vienna, Austria, 15-19 July, 2013, Vol. 2, p. 1772-1783.
- 3) Schloegl, C. and Gorraiz, J. “Global usage versus global citation, metrics: the case of Pharmacology journals,” *Journal of the American Society for Information Science and Technology*. Vol. 62, No. 1, 2011, p. 161-170.
- 4) Schloegl, C. and Gorraiz, J. “Comparison of citation and usage indicators: the case of Oncology journals,” *Scientometrics*. Vol. 82, No. 3, 2010, p. 567-580.
- 5) Gorraiz, J.; Gumpenberger, C. and Schloegl, C. “Differences and similarities in usage versus citation behaviours observed for five subject areas,” In *Proceedings of the ISSI conference*, Vienna, Austria, 15-19 July, 2013, Vol. 1, p. 519-535.
- 6) Guerrero-Bote, V. P. and Moya-Anegon, F. “Relationship between downloads and citation and the influence of language,” In *Proceedings of the ISSI conference*, Vienna, Austria, 15-19 July, 2013, Vol. 2, p. 1469-1484.
- 7) ScienceDirectの雑誌は、複数の分野に分類されているものが多い。SpringerLinkと分析条件をそろえるため、本研究ではWeb of Scienceに最初に記載されている分野名を採用した。
- 8) 対数変換せず求めた積率相関もほぼ同様だった。
- 9) Wan, J. -K.; Hua, P. -H.; Rousseau, R. and Sun, X. -K. “The journal download immediacy index (DII): Experiences using a Chinese full-text database,” *Scientometrics*, 2010, vol.82, no.3, p. 555-566.

## スマートフォンのカメラ機能を活用した

### 箱形書籍・資料デジタルアーカイブ作成装置の研究開発

伴 周一<sup>†</sup> 大久保 尚紀<sup>†</sup> 岡田 悟志<sup>†</sup>

<sup>†</sup> 日本大学理工学部

van@phys.ge.cst.nihon-u.ac.jp

#### 抄録

誰もが既存の書籍・試料のデジタルアーカイブ・電子書籍を作成できる装置があれば、図書館などの大量の書籍を所蔵する施設にとって大変有益で、コストの問題からデジタル化のすすめにくい地域独自の資料を有する地方の小規模図書館・資料館に特に役立つ。そこで手軽で小型軽量・優れたカメラ機能・データ処理機能をもつスマートフォンを活用した書籍撮影装置について研究開発を行っている。その装置の基本設計・試作・基本実験の結果について報告する。

#### 1. はじめに

電子書籍端末といえば、いわゆる書籍リーダー機能に特化したものからタブレット PC まで多種多様で幅広い価格の商品が市場に存在する。ハード面からもデジタル書籍・デジタルアーカイブを手軽に使える環境になってきている。またスマートフォンの大画面化により電子書籍端末として十分に機能している。結果として利用者も増加し電子書籍・電子ブックの活用は拡大し、多種多様の書籍のデジタル化が進んでいる。

元々デジタルデータの無い書籍・資料などは簡単にデジタル化を進めにくく、デジタル化の恩恵が多いはずの図書館・企業等の資料室などで意思通りにデジタル化が進められない。特に地域独自の資料を有する地方の小規模図書館・資料館においては、図書館としての独自性のアピールできる素材を生かすことがなかなか難しい。そこで誰でも安全・低価格な撮影装置があれば多くの問題が解決すると考え、書籍を半開き状態で撮影する箱形書籍撮影装置の研究開発を行っている。

さらに撮影と同時にデータ処理を自動化出来れば、誰もが既存の書籍・試料のデジタルアーカイブ・電子書籍を作成できる。そこで前回までにタブレット PC である iPad-mini を活用した撮影装置の設計・試作・撮影実験について報告した。また撮影と同時にデータ処理を行うアプリケーションの設計コンセプトとその試作 iOS アプリケーションについて報告し、画質を

含めた問題点についても報告した。今回はさらなる操作性の向上と画質の改善のために、タブレット PC より小さく、操作性が優れ、カメラ機能としても優れ、汎用性が高い、ハイスペックスマートフォンを撮影用カメラとして使用した箱形書籍撮影装置を研究開発することを目的とする。

#### 2. 書籍撮影装置の基本概念と基本構造

書籍を分解すること無しに、撮影を専門外の人間が行うことは大変難しく、高い次元で操作性や安全性が確保されなければならない。また開いた書籍の平面化を安全かつ安定に行う必要があり、そのうえで書籍を安定な撮影条件（配置・固定・光量・光質・カメラ等）で撮影する。さらに撮影者の肉体的・精神的な負担が小さく無ければ数百・数千頁の撮影を行うことはできない。つまり一枚一枚の撮影での調整が無い（ほとんど無い）事が重要となる。

そこで図 1 のような内部に LED 照明を有し下面に撮影面をもつ箱型の撮影装置に、書籍を半開きの状態で添え、上からデジタルカメラで撮影を行う装置を開発した。これは書籍を半開きで箱に添えることで、撮影面が容易に平面化でき、カメラとの安定した位置関係が保持され、箱内が小さなスタジオとなって安定に撮影を行い続けることが出来る。またこの構造は、頁が変わってもカメラと撮影面の位置関係は変化せず同じ撮影条件が常に維持される。

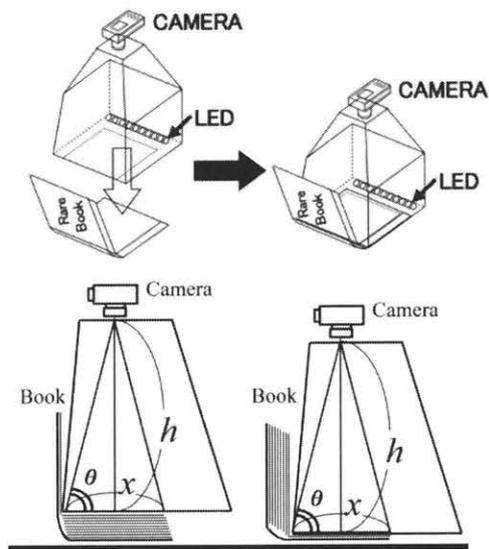


図1 箱型撮影装置の基本原理と特徴  
ページが変わっても無調整でカメラと  
書籍の位置関係が維持される。

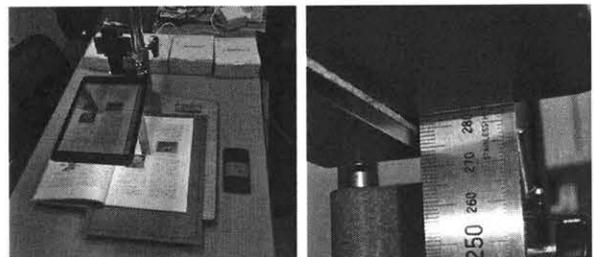
この装置では奇数頁を連続撮影し、その後偶数頁を連続撮影することから、撮影後に頁の90度回転・頁の再ナンバリング・表紙・裏表紙・背表紙の対応などのパソコンでのデータ処理が必要である。これらも普通の人を使うという観点から考えると、これは煩わしい操作であり、書籍デジタル化の障害要素と考えている。つまりタブレットPCやスマートフォンのカメラ機能が利用出来て、これらのデータ処理アプリケーションができれば大幅に問題解決すると考えている。

### 3. スマートフォンのカメラ機能

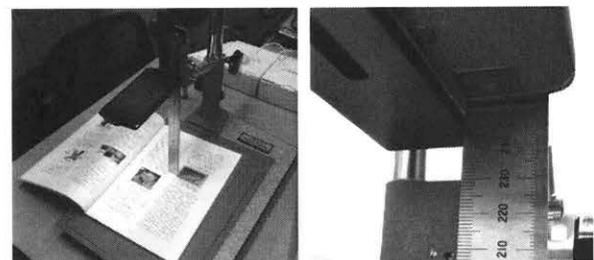
これまでの装置ではコンパクトデジタルカメラ・小型一眼レフデジタルカメラなどで極めて良好な書籍の撮影を行うことが出来、モニターテストでも良好な結果が得られていた。しかしタブレットPCのカメラ機能を用いた場合、問題点はカメラの性能であった。

タブレットPCのカメラ機能の性能は現在のコンパクトデジタルカメラに及ぶことは無く、特に解像度に関しては200~300万画素程度で書籍撮影に応えられる性能とは言えなかった。2012年頃からようやくApple社iPad-miniや

SONY Xperia S-seriesのように500万画素以上のCMOSカメラを搭載したものが市場に出回るようになり現実的に利用が可能になってきた。A4サイズを200dpiでスキャンした場合約600万画素、300dpiで約1000万画素であるので、500万画素では満足いくものではない。さらにデジタルズームのみ持ち・光学ズームを持たないタブレットPCでは、最大画素数を生かすためには撮影面との位置関係は一意的に決まる。図2はApple社iPad-miniと富士通スマートフォンARROWS F-02EでA4サイズをフルサイズ撮影する場合の撮影面までの相対距離の測定例である。



Apple 社 iPad-mini



富士通スマートフォン ARROWS F-02E

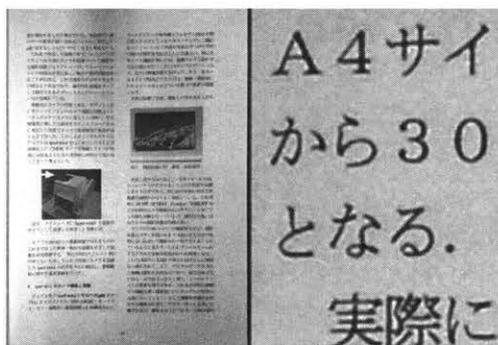
図2 タブレットPC(上)とスマートフォン(下)のカメラと撮影対象との相対距離測定例  
A4サイズをフルサイズ撮影した場合

Apple 社iPad-miniでは相対距離283mmとなった。同様の測定でSONY Xperia S-series TABLETでは302mm、REGZAタブレットでは292mmであり、相対距離約30cmがタブレットPCでのA4限界値のだいたいの相場となる。しかしこの限界値のままでは撮影画像角付近のゆがみを考えると好ましいとはいえ、わずかにでも距離を多くとり、周辺部のゆがみをデジタルズームにより排除する対応をとる方が望ましい。同時にこれは最大画素数では撮影できないことを意味する。それに対して富士通ス

スマートフォン ARROWS F-02E の場合この距離は 252mm であった。これは仮に撮影面までの距離 300mm の撮影装置を想定すると、

$$\left(\frac{252}{300}\right)^2 \times 1630 \text{ 万画素} \approx 1150 \text{ 万画素} \quad (1)$$

となる。この画素数は A4 を約 310dpi でスキャンしたものと同程度の画素数となる。つまりこれはこれまで作成したコンパクトタイプ箱形撮影装置<sup>4, 5)</sup>にそのまま転用できる(またこの距離を 28cm にした場合 1320 万画素 1320 万画素約 340dpi となる)。以上のようにタブレット PC を活用するより、コンパクトで、画素数も多く、ゆがみ対応も改善できる装置が作成できることになる。撮影例を示す。図 3 は撮影面までの距離を A4 フルサイズに調整した場合の撮影例である。拡大図をみるとハイスペックカメラ機能を持つスマートフォンならば文字のエッジがきれいに撮影できていることがわかる。



(A) Apple 社 iPad-mini (500 万画素)



(B) スマートフォン F-02E (1630 万画素)

図 3 A4 フルサイズ撮影画像比較

#### 4. スマートフォンカメラ機能を活用した書籍撮影装置の撮影実験と比較

これまで基礎実験から A4 サイズ+αの書籍に対応した撮影面とカメラの相対距離 30cm の装置で、カメラをスマートフォン F-02E と iPad-mini の場合について実験した。図 4 は実験に使用した装置で、基本構造はコンパクトタイプ箱形書籍撮影装置であり<sup>4)</sup>、撮影面は低反射透明アクリル板で半開きの書籍を軽く押さえる構造を持ち、カメラは上部にエラストマシート(粘着シート)により固定する。

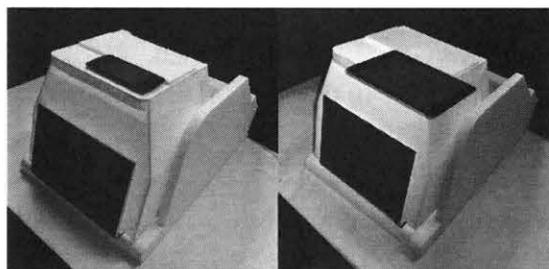


図 4 箱形書籍撮影装置比較

(右) iPad-mini 装着

(左) スマートフォン F-02E 装着



図 5 操作画面比較

(左) デジタルカメラ SONY Cyber-shot DSC-WX170 (中央) F-02E (右) iPad-mini

デバイス内のカメラの位置とパネルの操作の関係で、書籍のサイズによってカメラのセンター位置が手前に来るとき、iPad-mini では操作側に飛び出す配置になることがある。この問題もスマートフォンにより簡単に解決した。5 インチある F-02E の画面ならば撮影上困ることは無く、軽快な操作が可能であった。重量は iPad-mini 312g で F-02E 157g である。撮影実験した書籍と撮影時間を表 1 にまとめ、図 6 は撮影結果の一例を示す。

表1 iPad-mini と F-02E の撮影結果比較

覚園寺 大塚巧藝社 覚園寺 1975年 146枚 (素手) iPad-mini: 20分 8.2秒/頁 F-02E: 25分 10.3秒/頁
廣重 岩波書店 内田實 1978年二刷 862枚 (白手袋着用) iPad-mini: 204分 14.2秒/頁 F-02E: 214分 14.9秒/頁
PROGRESS IN METAL PHYSICS BUTTERWORTHS Sci. PUB. Ed. BRUCE CHALMERS 1949年初版 418枚 (素手) iPad-mini: 102分 14.6秒/頁 F-02E: 88分 12.6秒/頁
解析概論 岩波書店 高木貞治 1978年22刷 574枚 (白手袋着用) iPad-mini: 130分 13.6秒/頁 F-02E: 125分 13.1秒/頁

撮影速度は F-02E と iPad-mini ではほとんど差は無く、14~15秒/頁程度 (500頁の書籍を約2時間で撮影) である。画質は格段に改善され文字のエッジがきれいに撮影できた。なお自動輝度調整で全体に灰色がかった画像になる問題は残る結果となった。

以上からスマートフォンのカメラ機能活用は本装置にとって非常に有効な方法であった。さらにカメラ固定部を大幅に小さく出来た。

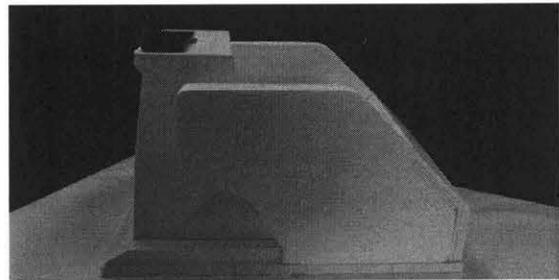


図7 スマートフォン用撮影装置外観

これは書籍を半開きのまま2つのカメラで同時に見開きを撮影する装置が作成可能になる。撮影専用アプリケーション開発とともに今後の課題である。

5. まとめ

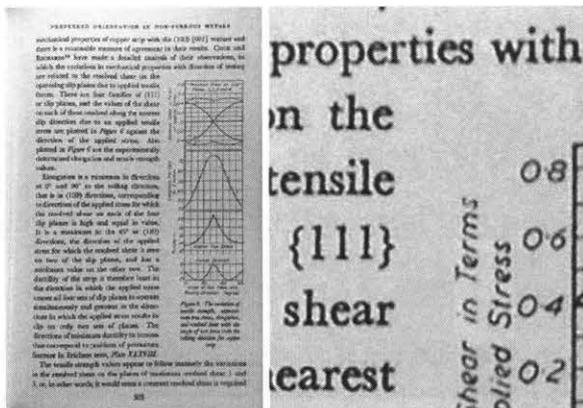
スマートフォンのカメラ機能を活用した箱形書籍撮影装置の基礎実験の結果、極めて良好な結果が得られた。また見開き頁を同時に撮影する装置の可能性がでてきた。

謝辞

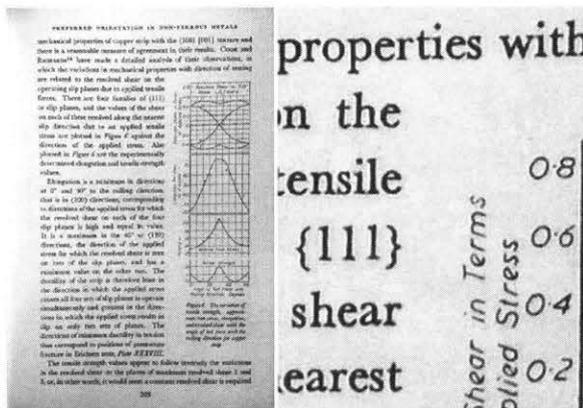
本研究は JSPS 科研費 23650124 の助成を受けたものです。

引用文献

- (1) 伴 周一他, 日本図書館情報学会春季研究集会発表要綱 (2010) P.95-98
- (2) 伴 周一他, 日本図書館情報学会研究大会発表要綱 (2011) P.77-80
- (3) 伴 周一他, 日本図書館情報学会春季研究集会発表要綱 (2012) P.133-136
- (4) 伴 周一他, 日本図書館情報学会研究大会発表要綱 (2012) P.93-96
- (5) 伴 周一他, 日本図書館情報学会春季研究集会発表論文集 (2013) P.17-20



(A) Apple 社 iPad-mini (500 万画素)



(B) スマートフォン F-02E (1630 万画素)

図6 撮影画像比較

## 明治・大正期の図書館における資料保存

望月有希子（筑波大学大学院図書館情報メディア研究科）wangyue@slis.tsukuba.ac.jp

### 抄録

本研究では、近代図書館創成期に行われていた資料の保存管理方法を解明することを目的とし、明治・大正期に出版された図書館管理業務と資料保存に関する文献から、保存管理方法の詳細、傾向、変化を分析した。分析の結果、明治・大正期の図書館における主な資料の保存管理方法は①曝書、虫害対応、②蔵書点検、③製本、装丁、④消毒、廃棄であった。そして図書館の発展と社会背景から、保存管理業務の変化と新しい方法が加わる過程を確認した。

### 1. 研究の背景と目的

日本では古代から書籍の保存管理方法として曝書が行われてきた。曝書とは書籍を日干しすることで、書籍に付着した虫を払い、蓄積された湿気の除去を行うことである。古代にはこれと同時に、修復作業のための書籍の破損状態の調査、蔵書点検、書庫清掃も行われており、これらの作業を含めて曝書と呼んでいた<sup>1)</sup>。

日本における曝書の最古の実施記録は、『正倉院御物出納文書』の中にある奈良時代の延暦6(787)年6月26日の「珍財帳」、同12(793)年6月26日の「曝涼目録」にみられる<sup>1)</sup>。それ以降、社寺、公家、武家を中心に曝書が行われてきた。

しかし、現在の図書館において曝書とは、蔵書点検を表す言葉としてわずかに残っているだけで、書籍を日干しする作業はほとんど行われていない。そして、近代図書館創成期における資料の保存管理の方法も明らかとなっていない。そのため、本研究では明治・大正時代に出版された、図書館管理業務と資料保存に関する文献から、近代図書館創成期に行われていた資料の保存管理を歴史的に解明することを目的とする。

### 2. 対象・方法

#### 2.1 調査対象

明治25年刊行の西村竹間著作の『図書館管理法』<sup>2)</sup>、明治33年刊行の文部省編纂の『図書館管理法』<sup>3)</sup>、明治45年刊行の文部省編纂の『図書館管理法(改訂版)』<sup>4)</sup>、大正4年刊行の日本図書館協会編集の『図書館小識』<sup>5)</sup>の4つの文献を調査対象とした。これらは、図書館の管理業務に関する教本的文献であるため対象に選んだ。なお、大正11年刊行の『図書館小識』の増訂版もあるが、資料の保存管理の内容は大正4年版と同じであるため、ここでは取り上げない。

#### 2.2 調査方法

上記の4つの文献から、明治・大正時代の図書館における資料の保存管理方法の詳細、傾向、変化を分析した。そして、ここから得られた結果を、東京図書館・帝国図書館の年報、目録、資料保存関連の資料から検証した。

### 3. 結果

#### 3.1 明治・大正期における資料の保存管理方法

分析の結果、明治・大正期における主な資料の保存管理方法は、①曝書、虫害対応、②蔵書点検、③製本、装丁、④消毒、廃棄であったことが明らかとなった。

### 3.1.1 曝書と虫害対応

虫害対応は、曝書を行うこととしていた。曝書の方法は、実施回数は年に1回、実施時期は空気が乾燥した10、11月頃、対象資料は和紙、唐紙が使用された和漢書としていた。曝書後の対応は、日干しした書籍を、樟脳、殺虫菊、害虫駆除剤と一緒に書函に保管し防虫処置を行うこととしていた。そして、明治後期になると、曝書による日干し作業に加え、ホルマリンガスや二硫化炭素ガスを使用した燻蒸による殺虫処理が取り入れられていた。

### 3.1.2 蔵書点検

書籍の紛失を確認するため、蔵書点検を行うこととしていた。蔵書点検の実施回数は年に1回、曝書と同時に行うか、または業務に支障がでない時期に行うこととしていた。調査の範囲は、明治25年刊行の『図書館管理法』では全蔵書を行うとしていたが、明治33年刊行の『図書館管理法』以降は、業務の兼ね合いを考え、蔵書の部分的調査も可としていた。調査方法は、函架目録と照合し、紛失だけでなく、破損内容も調査することとしていた。そして、紛失が確認されたものを記録する帳簿を作り、紛失書籍の貸出状況の確認や、館内の搜索などを行うこととしていた。また、紛失書籍の対応として、複本を入れることを提案していた。ここから、明治時代に現在の蔵書点検作業の原型ができていたことを確認した。

### 3.1.3 製本、装丁

製本、装丁を行うことは、紛失や破損を予防するために有効であると考えられていた。装丁の方法は、紛失時の搜索に対する利便性から、表紙を一様にするとしていた。そ

して、良好な材料、強固な縫製、適宜な技術で製本したものは、利用者も丁寧に扱う傾向があるため、破損や劣化を回避することができるとしていた。

また、製本方法以外に、製本の手順、発注、支払い、統計・記録まで詳しく業務手順が示されていた。

### 3.1.4 消毒、廃棄

大正4年刊行の『図書館小識』では、伝染病患者が利用した書籍や汚れが目立つ書籍の対応について記載されていた。その対応方法は以下の4点であった。

①病菌に触れた図書は、速やかに消毒をすること。館内にその設備がない場合には、専門の消毒所に送ること。

②重い伝染病患者が利用した図書は、焼却廃棄する手段を考えること。

③腸チフス、猩紅熱等のウイルスは数年後に汚染書籍から伝播した例があるので、注意すること。

④閲覧が頻繁で手沢が甚だしい図書は適宜廃却し新本を補充すること。

このように、伝染病による汚染が疑われる書籍や、頻繁な利用により汚れが目立つ書籍は、消毒や廃棄が考えられていた。

## 3.2 時代経過による変化

明治25年刊行の『図書館管理法』から大正4年刊行の『図書館小識』の流れの中で、4つの変化が見られた。第一に、図書館における資料保存の業務内容の変化である。図1に4つの文献における、資料保存管理の業務内容の記載の割合を示した。明治25年刊行の『図書館管理法』による資料保存業務は、曝書が主であり、約半分が曝書の記載となっていた。次いで点検が33.3%、装丁が15.7%の割合であった。しかし、明治33年刊行の『図書館管理法』になると、書籍の製本・装丁が

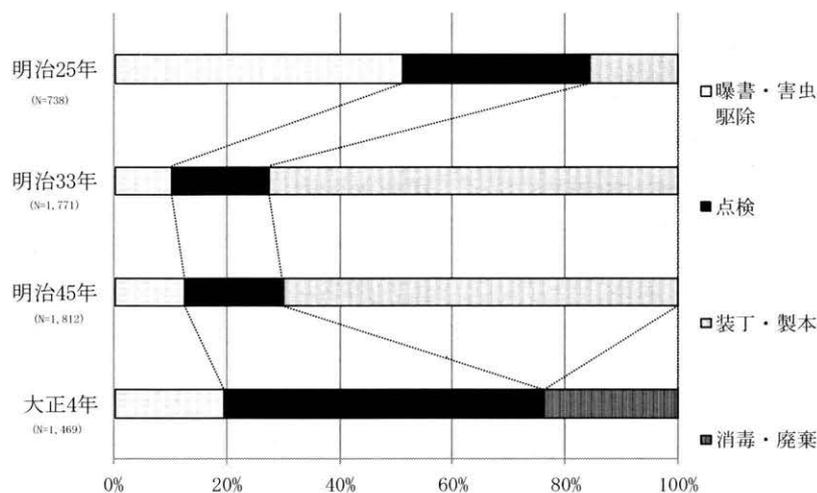


図1 明治・大正期の図書館における資料保存業務の割合

出所：『図書館管理法』（明治25年<sup>2)</sup>、明治33年<sup>3)</sup>、明治45年改訂版<sup>4)</sup>）、『図書館小識』<sup>5)</sup>より作成。

資料保存業務の主な業務となり、明治25年刊行の『図書館管理法』の記載が15.7%であったのに対し、72.6%に激増していた。そして、曝書の記述は10.2%に減少していた。

この要因は、図書館の閲覧人数、貸付図書数、蔵書数の変化が影響していると考えられる。帝国図書館における明治25年の閲覧人数は68,056人、貸付図書数は427,794冊、蔵書数は305,437冊であった<sup>6)</sup>。それに対し明治33年の閲覧人数は倍増の135,713人、貸付図書数は47%増の620,971冊、蔵書数は31%増の401,419冊であった<sup>6)</sup>。ここから、閲覧人数、貸付図書数、蔵書数の急激な増加が、書籍の破損や紛失を増大させたと推測され、図書館における資料保存業務の重点が、曝書から製本や装丁に変化したと考えられる。

第二に、明治45年の『図書館管理法』になると、害虫駆除の方法は曝書だけではなく燻蒸が加わったことである。

明治39年の帝国図書館の年報には、内務省衛生試験所で発明された燻蒸設備を設置したと記載がある<sup>6)</sup>。そして、大正元年、2年にはホルマリンガス燻蒸を行った記載もある<sup>6)</sup>。ここから、明治末期から曝書に加え、燻蒸に

よる害虫の対応も行われ始めたことを確認した。

第三に、大正4年の『図書館小識』には製本・装丁の記載がなくなり、資料の保存管理の主な業務が点検となったことである。

図1の資料保存業務の割合からわかるように、明治33年と明治45年刊行の『図書館管理法』には、装丁・製本の記載が約7割を占めていた。しかし、大正4年刊行の『図書館小識』には

記載がなくなり、点検が56.7%を占め、資料保存管理の主な業務となっていた。

明治時代、書籍の装丁は和装から洋装に変化し、図書館の蔵書も和装本が少なくなり、洋装本が主流となっていった。

図2に、帝国図書館（東京図書館）の蔵書における、明治時代の刊行年による和装本と洋装本の割合を示した。ここから、和装本の割合が、明治元年は96%を占めていたのに対し、明治10年には84%に減少し、明治19年になると全体の半数を下回り42%となったことがわかる。そして、明治25年になるとさらに半分の20%となり、明治最終年の45年には5%までに低下し、明治末期には、大部分が洋装本となったことがわかる<sup>7)</sup>。

このため、大正時代に入ると、図書館の蔵書の多くが洋装本となったため、書籍の製本や装丁の必要が少なくなったと推定される。

第四に、大正4年刊行の『図書館小識』には、汚染書籍の対応として、消毒、廃棄による処理が新しく加わったことである。

図1の資料保存業務の割合に示されているが、大正4年の『図書館小識』において、消毒、廃棄は、新しく加わった上、全体の23.8%

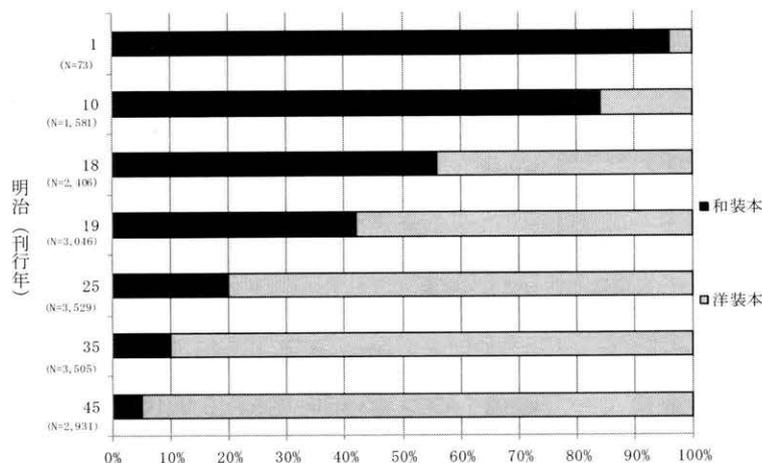


図2 明治期の和装本と洋装本の割合

出所：大沼宜規「明治期における和装・洋装本の比率調査—帝国図書館蔵書を中心に」より作成

を占めている。

当時は伝染病の病原菌が書籍から伝染すると考えられていた。そのため、海外においても、伝染病患者が利用した書籍に対する法律が作られていた。1893年、イギリスでは伝染病患者が発生した場合、公共図書館長に報告し、伝染病患者の家族に対し図書帯出を停止し、既に帯出されている図書が返納された場合は、消毒、焼却、伝染病病院に引き渡すことが義務付けられた。1903年、ドイツでは伝染病患者が使用した学校教科書は焼却する規定が出された。1904年、オーストリアでは結核患者が使用した図書は消毒する規定が出された<sup>8)</sup>。

日本においては、大正11(1922)年の伝染病予防施行規則において、図書類の消毒はホルムアルデヒドを使用することとされた<sup>8)</sup>。このため、伝染病を拡大させないために、書籍の消毒や廃棄が重要と考えられたと思われる。

#### 4. まとめ

明治・大正時代の図書館における主な資料保存管理の方法は、①曝書、虫害対応、②蔵

書点検、③製本、装丁、④消毒、廃棄であった。そして、図書館の発展により、資料の保存管理業務における主たる業務内容が、曝書から製本・装丁、点検へ変化していったことが明らかとなった。また、科学技術の進歩により害虫対応の方法として、従来の曝書に加え、燻蒸が取り入れられ、伝染病の流行という社会背景から、汚染書籍を消毒、廃棄する新しい対応が加わったことが明らかとなった。

#### 引用・参考文献

- 1) 杳掛伊左吉. 曝書史稿 : 書籍保存の歴史. 二宮山房, 1970, 87p.
- 2) 西村竹間著. 図書館管理法. 日本図書館協会, 1978, 66p., (復刻図書館学古典資料集).
- 3) 文部省編. 図書館管理法. 日本図書館協会, 1978, 139p., (復刻図書館学古典資料集).
- 4) 文部省編. 図書館管理法. 改訂版, 日本図書館協会, 1978, 124p., (復刻図書館学古典資料集).
- 5) 日本図書館協会編. 図書館小識. 日本図書館協会, 1978, 202p., (復刻図書館学古典資料集).
- 6) 国立国会図書館支部上野図書館編. 帝国図書館年報. 国立国会図書館, 1974, 428p.
- 7) 大沼宜規. 明治期における和装・洋装本の比率調査: 帝国図書館蔵書を中心に. 日本出版史料. 2003, 8, p.126-153.
- 8) 今澤慈海. 図書館経営の理論及實際. 叢文閣, 1926, 719p.

# ネットワーク構造として見るメディアおよび記号の編成

## ～メディア環境記述の新たな枠組みの提案～

茅野良太<sup>†</sup>

<sup>†</sup> 東京大学大学院教育学研究科

<sup>†</sup> mafumafu0202@yahoo.co.jp

### 抄録

本稿の目的は、主体の認識が成立する条件としてのメディアの構成を明らかにする方法を提案することである。つまり、図書館等の環境において主体がメディアを認識することを考えるとき、主体が意味づけをしてしまったのではない形でメディアの構成を捉えるための方法である。以上の目的のために本稿では、メディアの編成を、メディアがパッケージする記号と形態の共通要素からなるネットワーク構造として見なす方法を提案する。

## 1. はじめに

図書館情報学と教育学の接点の1つに学校図書館というテーマがあるものの、学校図書館という領域に限定するまでもなくそもそも図書館情報学は教育学的である。なぜなら「学ぶ」ということは現行の学習指導要領における「主体的に学習に取り組む態度」<sup>1)</sup>によってではなくむしろ学習する主体を取り巻く環境によって引き起こされると考えることもできるのであり、図書館情報学はそのような環境を構築する技術についての学であると言えるからである。たしかに言語やメディアは社会的なものであり、図書館はその社会性に立脚した制度である。しかし言語やメディアの社会性を問題にせず、主体を取り巻き主体の認識を成立させる条件としての環境を考えることもできる。この考え方に従えば主体は主体的であるよりむしろ環境に対して受動的な存在であると言える。

しかしながら、主体の認識や学習の条件となるような環境あるいは対象・モノについて十分に注目されてきたとは言い難い。たとえばカント (Kant, I) はその認識論において「対象」を認識の1つの条件としているが<sup>2)</sup>、カント自身もその後の認識論哲学も、その議論の中心は主体の認識の側であり、カントの言う「対象」すなわち環境やモノの問題は取り残されてしまっている。また教育学の学習理論においても対象・モノの問題が不問にされがちであり、佐藤によれば現在の学習理論では「学習者の思考が構成する対象の現実性を追求しえても、学習者

の身体が感受する実在性を追求することはできず、学習を「モノや人や状況に媒介された実践として認識し組織し直す」必要があるという<sup>3)</sup>。本稿はこの課題についてメディアの構成の側からアプローチする。つまり、主体にとって未だ社会的な意味づけのなされていない形で実在的に感受されるようなメディア (図書) の構成を明らかにするための手法を提案するものである。

## 2. 先行研究

以上のような本稿の関心は、認識する主体の自発性よりもむしろ主体の認識の様態に影響を及ぼす環境やモノに注目しているという点で、ブルデュー (Bourdieu, P) の文化資本論とマクルーハン (McLuhan, M) のメディア論に通じるものがある。ブルデューは各階層が保有している文化資本が世代間で伝達されることが階層の再生産に大きな役割を果たすことを指摘している<sup>4)</sup>。またマクルーハンは「メディアはメッセージである」や「身体の拡張」といった主張を行い、人間の思考の様式や習慣を変容させるものとしてのメディアに注目した<sup>5)</sup>。

この両者を一括りにして論じるのはやや強引であるものの、両者とも、認識する主体を環境やモノに対して受動的な存在として捉えていると言うことができる。しかしこれらはいずれも本稿の目的に十分に応えるものではない。というのも、これらの議論は文化資本およびメディアの特徴を社会的に意味があると認められる形

で取り出してしまっているからである。つまりこれらにおいては、文化資本およびメディアの特徴を社会現象との関わりにおいて見出した上でそのような文化資本およびメディアが社会に対して与える影響を論じるという同語反復的な議論になっており、モノとしての文化資本やメディアの姿は明らかにされていない。このような議論では、本稿の目的である、主体が実在的に感受するところの対象・モノへのアプローチには十分でない。

以上の2つの議論は思弁的なものであるが、メディアそのものをより実体的に扱う研究が図書館情報学で行われている。メディア環境の構成について扱っている研究の代表的なものとして、図書館の蔵書構成の調査・分析がある。これらの研究はある環境のメディア構成を記述するという点において本稿と関心を共有している。しかし蔵書分析の研究は分類番号や日本図書コードのような予め技術的・社会的に設定された項目を用いてメディアを記述しているため<sup>6)</sup>、本稿の目的のためにはこれらと異なる方法を採用必要がある。他に、図書館のみに限定せずにメディアの社会的な流通全体を視野に収めた計量的な研究がいくつかある<sup>7)</sup>。これらの研究は広く社会のメディア環境を記述するものであり、その点では本稿と関心を共有する。ただしこれらの研究はそれぞれの時代や社会のメディア流通の動向を明らかにしているものの、そのメディアが主体によって認識されるレベルにおける考察はまだなされていない。

### 3. 基本的な世界構成および方法の概観

では、主体の認識を成立させる条件としてのメディア、言い換えれば主体にとって実在的に感受されるメディアの構成を明らかにすることは、いかにして可能になるだろうか。本稿の目的からすると、社会化された意味づけを先取りすることなく観察点を定めることが重要である。本稿では、近代の知識流通に大きな役割を果たしたメディアとして「図書」を取り上げ、次のような世界を想定してみることにする。

(1)メディアを「記号」と「メディアの形態」の組み合わせとして記述する。(2)記号として「タイトル」と「著者名」に注目し、形態とし

て「シリーズ」に注目する。(3)記号およびメディアが人間にとって直接認識可能な形で現実に存在している場所として「環境世界」を、その背後にあるメディアの可能的集合として「可能世界」を想定する。(4)メディアにパッケージされる記号である「タイトル」および「著者名」とメディアの形態である「シリーズ」は環境世界内において主体によって認識される。(5)メディアにパッケージされる記号である「タイトル」および「著者名」とメディアの形態である「シリーズ」のうち、共通のものをリンクづけし、ネットワークを描く。

#### 3.1. メディアの外表面に表象される異同

本稿ではメディアの定義を「記号をパッケージするもの」という簡便なものにとどめておく。そうするとメディアを、メディアによってパッケージされる「記号」とメディアそのものの「形態」という、大きく分けて2つの要素の組み合わせによって記述することができる。記号として「タイトル」と「著者名」に、メディアの形態として「シリーズ」に注目する理由を述べる。

主体が環境世界内において図書や言葉について十分に意味づけされていない形でそれらを認識することを想定するなら、メディアおよび記号で構成される環境世界とは主体にとって、ただ単にモノとしての記号の異同と、ただ単にモノとしてのメディアの形態の異同とが、ただ単にあるパターンとして広がっている世界であると想像できる。また環境世界内において主体によって第一次的に感受されてしまうような記号やメディアの形態とは、メディアの内部ではなく外表面に表示されていなくてはならない。図書の場合を考えてみると、まず記号としてはタイトルと著者名がメディアの外表面に記されているので、この2つを、環境世界内において主体によって一時的に感受される記号として採用することができる。同様にメディアの形態としては、図書のサイズや版形、表紙の色や絵柄、文字のフォントや配置などの諸特徴が考えられる。ただしこれらを1つ1つ調査するのは困難であることもあり、ある程度形態上の同型性を記述することができると考えられるシリーズという項目を採用する。またシリーズを採用する

のには、後述するように、描かれたネットワークを知識の構造として見なすことを許すためという積極的な意義もある。

このようにして本稿では、図書に「タイトル」「著者名」「シリーズ」という3つの要素を付与することとする。

### 3.2. パラテキストとしてのリンクづけ

先に、主体によって未だ意味づけされずに環境において認識されるメディアの要素として「タイトル」「著者名」「シリーズ」という3つの項目を採用したが、これらの持つ意味は、単にメディアの外表面に表示され主体によって第一次的に感受されることのみにとどまらない。これらの外表面に表示された記号あるいは形態的表象は、ジュネット (Genette, G) のパラテキスト概念に相当するものである。パラテキストとはすなわち、複数の「タイトル」や「著者名」あるいは「シリーズ」の中に何か共通の要素が認められるならば、それらの複数のメディアの間には何らかの関係づけがなされるということである<sup>8)</sup>。ここにおいて、複数点の図書において3つの要素のうちで共通のものが認められる場合に、その複数点の図書をリンクづけることの正当性を示すことができた。つまり、主体にとって未だ意味づけはなされていないものの、環境世界の中に複数の同じ記号や形態が認められる程度には秩序のあるメディア構成が知覚される時、記号とメディアの形態の共通要素からなる秩序を記述するのがネットワークである。リンク付けは図1のようになされる。まず主なパターンを挙げるだけでも、ある図書に対して著者名を共有する図書、タイトルを共有する図書、シリーズを共有する図書の間リンク付けをすることが可能であろう。

ところで、著者名については同じ複数の著者名があればリンクづけをすることとなるが、タイトルについてはタイトル中の言語表現の一部の一致があればリンクづけをすることにする。具体的には、タイトル中に見られる固有名(人名、地名、書名など)を要素として取り出す。

### 3.3. 環境世界と可能世界

本稿で設定している2つの世界について確認

しておく。1つは、主体が生活する具体的な環境でありまた同時に図書を有する世界であり、これを「環境世界」と呼ぶこととする。環境世界はたとえば家庭、学校、図書館、書店などの具体的な場所が想定され、ここには現実態としての図書の集合が存在している。もう1つは、社会に流通している図書の点数としての集合全体であり、これを「可能世界」と呼ぶこととする。可能世界には、個々の環境世界においてありえたかもしれない全ての種類の図書すなわち可能態としての完全な図書の集合が存在していると考えられる。可能世界を完全に記述することは難しいが、たとえば国会図書館のNDL-OPACに掲載されている全ての図書といった形で近似できると考えられる。2つの世界を合わせた世界構成は図2のようになる。

### 4. ネットワーク分析と予想される結果

ここまでで、図書の「タイトル」「著者名」「シリーズ」の中で共通のものをリンクづけするという手法について述べた。では、このようにして描かれるネットワークは何を意味するのだろうか。環境世界の中のみ限定して考えるなら、得られたネットワークはまず環境世界において主体に感受される記号とメディアの形態という表象の分布の様態を示す。言い換えれば、メディアによってパッケージされた記号的表象の異同に基づいたメディアの編成パターンである。次に、「タイトル」「著者名」「シリーズ」を一段階社会化して意味づけされた形で、知識を何らかの形で表現するユニットであると捉えるならば<sup>9)</sup>、得られたネットワークは、環境世界における知識の分布の様態を示すものであるとも言える<sup>10)</sup>。さらにこの環境世界におけるネットワークを可能世界に拡張してみると、そこで得られたネットワークは、環境世界のメディアの集合によって支えられる知識が、可能世界全体の知識に対して関わるその仕方であると言える。

### 5. まとめと課題

本稿では「タイトル」「著者名」「シリーズ」というメディアの外表面上の表象を出発点にしてそれらをネットワークとして描くことにより、

主体の現実的な認識に先立って社会的な意味づけを仮定してしまうことを極力回避した形で、メディアの編成の様態を明らかにすること、さらには知識の編成の様態を明らかにする可能性を示すことができた。ただし以上の議論には不十分な部分も多い。たとえば、著者名やシリーズなどの社会的な表徴をあえて現実的に仮定していることや、そもそも記号や形態が環境世界の中で特権的に知覚されてしまうことを仮定していることなどである。これらについてはさらに議論が必要であるが、とりあえず本稿では試論としてアプローチの方策を提案することを行った。

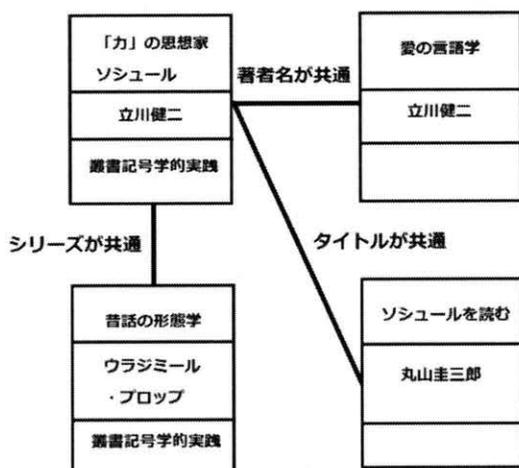


図1 メディア間のリンク付けの仕方

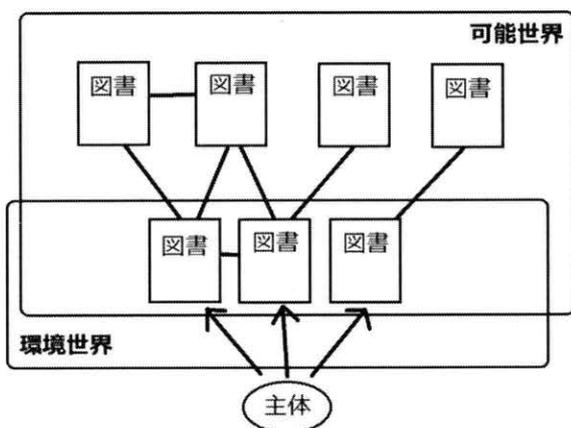


図2 環境世界と可能世界

注および引用文献

1) たとえば「高等学校学習指導要領」第1章第1款1を参照。小学校と中学校も同様。

[http://www.mext.go.jp/a\\_menu/shotou/new-cs/youryou/](http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/) (参照 2013-9-3)

2) Kant, I. 『純粋理性批判〈上〉』 [*Kritik der reinen Vernunft*. 1787,] 篠田英雄, 岩波書店, 1961, p.92-93.

3) 佐藤学 『教育方法学』 岩波書店, 1996, p.72-73.

4) 文化資本には嗜癖や習慣も含まれるが、特に本稿と直接関わるのは「物質的遺産」である。Bourdieu, P. 『ディスタンクシオン—社会的判断力批判〈1〉』 [*La Distinction. Critique sociale du jugement*. Les Editions de Minuit, 1979,] 石井洋二郎, 藤原書店, 1990, p.121.

5) McLuhan, M. 『メディア論—人間の拡張の諸相』 [*Understanding Media: the Extensions of Man*. McGraw-Hill, 1964,] 栗原裕, 河本仲聖, みすず書房, 1987, p.7.

6) Cコードを分析項目にした蔵書分析は以下を参照。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.

7) たとえば以下を参照。海野敏, 影浦峯, 戸田慎一 「戦後日本における印刷メディア受容量変化の数量的検証」 『日本図書館情報学会誌』 Vol.58, No.1, 2012, p.1-17.

8) Genette, G. 『スィューテキストから書物へ』 [*Seuils*. editions du Seuil, 1987,] 和泉涼一, 水声社, 2001, p.11-18.

9) 固有名を知識であると見なすことの考察は以下を参照。蓮實重彦 『反=日本語論』 筑摩書房, 2009, p.68-86.

10) 文献ネットワークを知識の構造と見なす方法は科学計量学に見られる。Lydesdorff, L. 『科学計量学の挑戦—コミュニケーションの自己組織化』 [*The Challenge of Scientometrics—The Development, Measurement, and Self-Organization of Scientific Communications*. DSWO Press, 1995,] 藤垣裕子, 林隆之, 富澤宏之, 平川秀幸, 調麻佐志, 牧野淳一郎, 玉川大学出版部, 2001, p.151-153.

# BSH/NDLSH と NDC における階層構造はどの程度一致するのか 件名と分類の統合的活用法に向けて

谷口 祥一 (慶應義塾大学文学部)  
taniguchi@z2.keio.jp

[抄録] BSH 第4版と NDLSH (2013年8月時点データ) の個々の件名に付与されている NDC 新訂9版の分類記号を手がかりに、BSH と NDC の組み合わせ、NDLSH と NDC の組み合わせにおいて、どの程度上位下位関係の階層構造が一致するのか定量的に調査した。BSH と NDLSH をベースにした照合、NDC をベースにした照合をそれぞれ実施し、階層構造が一致しない箇所をプログラムで数えた。NDC の縮約・不均衡項目の正規化、関連項目の活用、上位分類項目からの件名の限定的継承などを組み入れた場合についても実験した。

## 1. はじめに

分類法は体系性と網羅性を重視し、関連主題項目が並置され集中される特質をもつ。一方、件名法(統制語彙による主題表現法)は、自然言語をベースにして、その直接性や柔軟さに力点を置くため、上位語を複数有する構造もみられる。こうした異なる特性を有する主題表現方式およびその表現結果であるが、統合的に活用する方策を検討する意義はあると考える。

そこで、統合的な活用法の検討の一環として、分類法、件名法の身近な例である BSH (基本件名標目表) 第4版、NDLSH (国立国会図書館件名標目表)、NDC (日本十進分類法) 新訂9版を取り上げ、それぞれの階層構造がどの程度照応するのか確認する。NDC は十進分類法として分類項目の上位下位関係が明示されており、全体として木構造をなす。BSH/NDLSH も上位語・下位語の関係による階層構造をもつが、複数の異なる上位語をもつ場合などがあり、全体としてはグラフ構造である。BSH と NDLSH の個々の件名に付与されている NDC 分類記号を手がかりに、BSH と NDC の組み合わせ、NDLSH と NDC の組み合わせにおいて、どの程度上位下位関係の階層構造が一致するのか定量的に調査した。

本研究と同様なアプローチを採用した研究はないが、和田らは個々の書誌レコードに付与されている分類記号と件名標目の組み合わせが、件名標目表において指示されている対応づけとどの程度一致するのか調査している<sup>1)</sup>。

## 2. 階層構造の照合処理

それぞれの機械可読形式データを入手 (NDLSH は Web NDL Authorities から 2013年8月7日時点のデータを取得) し<sup>2)4)</sup>、BSH と NDLSH については、参照や細目のみのものを除いた。その結果、BSH は 8,034 件名、

NDLSH は 19,184 件名であった。なお、これらすべての件名が、NDC9 版の対応する分類記号をもっている (BSH では平均 1.34 個、標準偏差 0.64、NDLSH では 1.37 個、標準偏差 0.62)。また、NDLSH はこのうち 4,615 件名 (24.1%) は、上位語も下位語ももたない孤立した件名であった (BSH では 433 件、5.4%)。

NDC については、本表データから二者択一項目や不使用項目などを除いた分類項目 9,432 を採用した。加えて、BSH 件名には付与されているが、NDC 本表にはない 435 項目を追加した (計 9,867 項目)。同様に、NDLSH 件名には付与されているが、NDC 本表に設定されていない 1,255 項目を追加した (計 10,687 項目)。

これら前処理済みデータを用いて、① BSH/NDLSH をベースにして、件名の上位語・下位語関係に対して、それぞれの件名に対応づけられた NDC 分類項目間が同じく上位下位関係にあるのかプログラムで照合し、不一致の箇所を検出した。同様に、② NDC 分類項目間の上位下位関係に対して、それぞれ対応づけられた BSH/NDLSH の件名間で上位語・下位語関係にあるのか照合した。

照合処理においては、以下のような点に考慮した。

1) BSH/NDLSH 件名と NDC 分類項目の間には多対多の対応関係がある。そのため、対応づけられた複数ある件名 (または分類記号) のうち、いずれか一つが上位下位関係にあれば合致したものとする。

2) すべての NDC 分類項目が BSH/NDLSH 件名に対応づけられているわけではない。これは NDC をベースにした照合を行う際に問題となるが、a) 件名をもたない箇所では不一致は発生しない (いかなる件名とも合致する) とみなす、あるいは b) 上位 NDC 項目に対応づけられてい

る件名を継承させて照合することとした。

3)NDCには分類記号が体系をそのまま表現できていない縮約項目(215項目)と不均衡項目(350項目)がある。縮約項目は、有力な下位の区分肢を昇格させ、分類記号の短縮を行った項目、不均衡項目は逆に分類記号上、下位に位置づけられているが、本来は同位にある項目を指す。これらは機械可読データでは印が付されているため、照合実験では、NDCの分類記号をそのまま使用する方式とは別に、上記の縮約項目・不均衡項目について本来の階層構造を表現できる形式に機械的に変換したものを使用した(以下、正規化NDCと呼ぶ)。具体的には、階層の各レベルを3桁数字で表す方式を採用し、これに機械的に変換した。

4)NDCには、分類記号上で階層が飛躍する(漸近的ではない)箇所がある。例えば、007.6(データ処理、情報処理)に続けて007.609(データ管理)が出現し、中間階層の007.60がない。階層構造の照合においては問題となるため、機械的に中間階層の項目を生成した。前処理段階において、BSH/NDLSHには出現するが、NDC本表には出現しない分類項目をそれぞれ追加しているため、生成された中間階層項目の数もそれぞれ異なる。BSHとの照合用では336項目、NDLSHでは409項目(正規化NDCの場合419項目)が生成され追加された。

また、照合に以下の処理を追加した場合についても実験した。

1)NDCにおける関連項目の活用:NDC本表において各分類項目の下に関連する分類項目への参照が指示されている。このうち、「をも見よ参照」(→:)を抽出し、実験では当該分類項目と擬似的に等価の項目として照合に用いる方式をも試みた。1,170の項目において、延べ1,564の関連項目が追加された。

2)上位分類項目からの件名の限定的継承:対応するBSH/NDLSH件名をもたないNDC分類項目について、上位分類項目がもつ件名を継承して擬似的に付与する方式をも試みた。

### 3. 照合実験結果

#### 3.1 BSHとNDCの照合

BSHとNDCの階層構造の照合実験結果を表1に掲載する。

まず、BSHをベースとした照合(表1のA)では、最上位語(248件名)ごとに照合を行った結果、複数の上位語をもつケースがあるため総出現件数名10,421となる。それぞれの件名の上位語・下位語関係に対して、個々の件名に

対応づけられたNDC分類項目間が同じく上位下位関係(または同一項目への対応づけ)にあるのか照合し、上位下位関係にない不一致の箇所を検出した。その結果、不一致の箇所は2,902、全体の27.9%であった(表1のA.①)。その不一致をNDC分類記号の桁数で分けて集計した結果を、表2の「照合A.①」列に示した。NDCの先頭1桁すなわち10区分の段階で既に異なるケースが542(件名出現総数10,421に対して5.2%)、2桁すなわち100区分の段階で異なるケースが620(同5.9%)あった。なお、こうした集計は、件名に複数の分類記号が対応づけられている場合には、その最も大きな桁数での不一致箇所(分類階層上での最も近い箇所)を採用し行っている。ちなみに、この実験においては個々の最上位語から順に階層を降りていく照合法をとっているため、上位語・下位語関係をもたない孤立した件名は結果的に除外されている。

照合の範囲を直近の上位下位という2階層から、3階層に広げ照合したときには、不一致は2,084(20.0%)に大幅に減少した(表1のA.②)。これは、上位下位関係にある件名間では分類記号が上位下位関係(または同一項目)にないが、そのさらに上位語に付与されている分類記号とは上位下位関係(または同一項目)にある場合を合致するとみなした照合である。不一致をNDC分類記号の桁数で分けた集計結果を、表2の「照合A.②」列に示した。

次に、不均衡・縮約項目を本来の体系を表すよう変換した正規化NDCを用いたとき、階層関係が一致しない箇所は2,896となった(表1のA.③)。基本的な照合(表1のA.①)と比べると全体で6件不一致が減少しているが、実際には18件減少し、12件新たな不一致が発生している。不均衡・縮約項目への対処が階層構造の照合において必ずしも効果的とはいえない結果となった。正規化NDCを用い、かつ照合範囲を3階層に拡張したときには、不一致箇所は2,085となった(表1のA.④)。

同様に、BSHが対応づけているNDC分類項目に加えて、NDCによる関連項目を元の項目と擬似的に等価な項目として加えた照合実験を行い、得られた結果を表1のBに示した。不一致の箇所が2,672(表1のB.①)となり、関連項目を加えないときに比べて、230の不一致が減少している。照合の範囲を3階層に拡張した場合、あるいは正規化NDCを用いた場合、表1のAと同じような減少傾向を示した。

次に、NDCをベースにして、NDC分類項

表 1. BSH と NDC の階層構造の照合結果

BSH をベースにした照合: BSH → NDC				
件名の出現総数: 10,421 (件名の異なり数: 7,601; 階層関係をもたない孤立件名は除外)				
	①基本的照合	②3階層の範囲での照合	③正規化NDCを用いた照合	④正規化NDC+3階層の範囲での照合
<b>A: 最上位語(248語)ごとに照合</b>				
上位下位関係の不一致	2,902 27.9%	2,084 20.0%	2,896 27.8%	2,085 20.0%
		A.①との相違 -818 (内訳: -818, +0)	A.①との相違 -6 (内訳: -18, +12)	A.②との相違 +1 (内訳: -2, +3)
				A.③との相違 -811 (内訳: -811, +0)
<b>B: 最上位語(248語)ごとに照合+NDCの関連項目を追加した照合</b>				
上位下位関係の不一致	2,672 25.6%	1,866 17.9%	2,659 25.5%	1,863 17.9%
	A.①との相違 -230	A.②との相違 -218	A.③との相違 -237	A.④との相違 -222
NDC をベースにした照合: NDC → BSH				
分類項目数: 10,203 (9,867+機械的に生成した中間項目 336)				
そのうち、対応する件名をもつ項目数: 5,632				
	①基本的照合	②3階層の範囲での照合	③正規化NDCを用いた照合	④正規化NDC+3階層の範囲での照合
<b>C: 網目表(100区分)の単位で照合</b>				
上位下位関係の不一致	2,258 22.1%	2,085 20.4%	2,198 21.5%	2,069 20.3%
		C.①との相違 -173 (内訳: -662, +489)	C.①との相違 -60 (内訳: -111, +51)	C.②との相違 +16 (内訳: -44, +28)
				C.③との相違 -129 (内訳: -614, +485)
<b>D: 網目表(100区分)の単位で照合+上位分類項目からの件名の限定的継承</b>				
上位下位関係の不一致	2,861 28.0%	2,141 21.0%	2,808 27.5%	2,129 20.9%
	C.①との相違 +603 (内訳: -0, +603)	C.②との相違 +56 (内訳: -58, +114)	C.③との相違 +610 (内訳: -0, +610)	C.④との相違 +60 (内訳: -65, +125)

表 2. 不一致の発生レベル (BSH をベースとした照合結果)

不一致のレベル	照合 A.①	照合 A.②	照合 A.③	照合 A.④
NDC 先頭 1 桁	542	489	542	489
先頭 2 桁	620	586	620	586
先頭 3 桁	768	516	782	525
先頭 4 桁	757	408	746	405
先頭 5 桁	209	83	196	79
先頭 6 桁以降	6	2	10	1
不一致総数	2,902	2,084	2,896	2,085

目間の上位下位関係に対して、それぞれ対応づけられた BSH の件名間で上位語・下位語関係にあるのか照合した。(表 1 の C.①)。

全分類項目数 10,203 のうち、2,258 項目 (22.1%) において不一致となっている。なお、前述の通り、件名をもたない箇所では不一致は発生しない (いかなる件名とも合致する) とみなしている。併せて、対応する件名をもたないときには、上位 NDC 項目に対応づけられている件名を継承させる方式を試みた。その結果が表 1 の D であり、すべての NDC 項目で対応する件名をもつことになり、そのため不一致の箇所が大幅に増加している (2,861 項目、28.0%) (表 1 の D.①)。

表3. NDLSHとNDCの階層構造の照合結果

NDLSHをベースにした照合:NDLSH → NDC				
件名の出現総数:19,643		(件名の異なり数:14,569; 階層関係をもたない孤立件名は除外)		
	①基本的照合	②3階層の範囲での照合	③正規化NDCを用いた照合	④正規化NDC+3階層の範囲での照合
<b>A: 最上位語(1,533語)ごとに照合</b>				
上位下位関係の不一致	6,343 32.3%	5,559 28.3%	6,303 32.1%	5,530 28.2%
<b>B: 最上位語(1,533語)ごとに照合+NDCの関連項目を追加した照合</b>				
上位下位関係の不一致	6,054 30.8%	5,219 26.6%	5,984 30.5%	5,156 26.2%
NDCをベースにした照合:NDC → NDLSH				
分類項目数:11,096		(10,687+機械的に生成した中間項目 409)		
分類項目数(正規化NDC):11,100		(10,687+機械的に生成した中間項目 413)		
そのうち、対応する件名をもつ項目数:7,717				
	①基本的照合	②3階層の範囲での照合	③正規化NDCを用いた照合	④正規化NDC+3階層の範囲での照合
<b>C: 網目表(100区分)の単位で照合</b>				
上位下位関係の不一致	4,300 38.8%	4,730 42.6%	4,252 38.3%	4,702 42.4%
孤立件名を除いた場合	980 8.8%	1,410 12.7%	932 8.4%	1,382 12.5%
<b>D: 網目表(100区分)の単位で照合+上位分類項目からの件名の限定的継承</b>				
上位下位関係の不一致	5,360 48.3%	4,891 44.1%	5,310 47.8%	4,866 43.8%

### 3.2 NDLSHとNDCの照合

NDLSHとNDCの階層構造の照合実験、すなわちNDLSHをベースとした照合、NDCをベースにNDLSHを用いた照合のそれぞれを実行し結果を得た。その結果は、表3に掲載してある。ここで採用した照合方式は、すべてBSHとNDCの照合に用いたものと同じである。

NDLSHをベースとした照合では、BSHをベースとした前述の照合と類似する結果となった。件名の総出現数が19,643とBSHに比べてほぼ倍の数であるが、NDCへの対応づけによる上位下位関係の照合では、不一致の箇所6,343と同じく倍の数となった。全体の件名数に対する不一致の割合は32.3%である。この結果を基点にして、3階層に照合範囲を拡張したとき、正規化NDCを用いたとき、さらにはNDCの関連項目を追加したときなど、それぞれ不一致が減少する傾向は、BSHの場合と相似する。

他方、NDCをベースにした照合では、BSHの場合に比べて、不一致の割合が増加する。

BSHの場合、全体の22.1%(表1のC.①)が不一致であったのに対し、NDLSHでは全体の38.8%(表3のC.①)が不一致となっている。さらに、対応する件名をもたない項目について、上位項目から件名を継承させて照合したとき、不一致が大幅に増加している。これらの大半は、孤立した件名(4,615件)が引き起こした不一致である。孤立した件名を除いて集計した場合を、表3のCの下に併せて示してある。例えば、C.①基本的照合では、不一致数980(全体の8.8%)であった。

#### 注

- 1) 和田匡路, 川向直樹「件名標目表の代表分類記号と書誌レコードの一致する割合」第58回日本図書館情報学会研究大会発表要綱. 2010, p.61-64.
- 2) 日本図書館協会. BSH4-Computer File (基本件名標目表第4版機械可読データファイル). 2002.
- 3) 国立国会図書館. Web NDL Authorities.一括ダウンロード用ファイル. <http://iss.ndl.go.jp/ndla/download/> (参照 2013-08-06)
- 4) 日本図書館協会. NDC・MRDF9 (日本十進分類法新訂9版) 機械可読ファイル, 1996.

# ウェブページの寿命: 2001年に存在した1000万ページを対象にした調査

宮田洋輔(常葉大学短期大学部) m@miyay.org

安形輝(亜細亜大学), 池内淳(筑波大学)

石田栄美(九州大学), 上田修一(前慶應義塾大学)

## 抄録

ウェブページの寿命を明らかにするため、2001年に収集された1,000万件のウェブページを対象に、長期的な生存調査を実施するとともに、Internet Archiveを利用して、ウェブページがいつ誕生し、いつ消失したのかに関する調査を行った。生存調査の結果、9割以上のページが収集から12年を経た現在、消失していることが明らかになった。寿命調査の結果、2001年に収集され現在消失しているウェブページの平均寿命は、1,108.2日であることが明らかになった。

## 1. はじめに

ウェブは、公開後も容易に更新でき、削除できるなど、従来のメディアと異なる動的な特性を持っている。その関心のもとに、これまで、ウェブがどのように変化するか、あるいは、どのように消えていくかといった研究が行われてきた。

Internet Archiveの創設者であるKahleは、1997年時点で、ウェブページの寿命は44日であると述べている<sup>1)</sup>。ウェブページがどの程度の期間でアクセスできるか/できなくなるかについても実証的な研究が行われている。Koehlerは1996年に収集した361件のURLについて、4年にわたり、ウェブページの生存や変化について調査を行い、2年間で約半数のウェブページが消失することを明らかにした<sup>2)</sup>。最近の研究では、2007年2月に存在していた、白人の権利擁護団体のウェブサイト163件を対象とした調査でも、2.4年で半減期を迎えることが明らかになっている<sup>3)</sup>。

本研究グループでは、2001年に収集された1000万件のウェブページを対象として、2003年に生存調査(以下2003年調査)を行なった<sup>4)</sup>。その結果、Koehlerによる研究と同程度に、2年を経ると47%のページがアクセスできなくなるという結果が得られた。さらにその結果を用いて、2003年当時のjpドメインのウェブページ数を推計した。

ウェブの消失に関する異なるアプローチとして、Ainsworthらは、どれだけウェブアーカイブに保存されているかを調査した<sup>5)</sup>。AinsworthらはDMOZ, Delicious, Bitly, 検索エンジンのインデックスから取得したURL計4,000件を用いて、2010年11月から2011年1月にかけて調査を行った。その結果、サンプル集合によって傾向は異なるものの、ウェブの35%~90%程度はいずれか1つのウェブアーカイブにアーカイブされていることを明らかにした。

ウェブページの生存に関する既往研究では、ウェブ

ページを取得した日付を基盤として、ウェブページの消失までの期間が調査されてきた。ウェブページが「誕生」した日を正確に取得することは困難であり、取得日を「誕生日」に近似して扱うほかない。一方で、取得日をウェブページの誕生日と設定することによって、実際の誕生から取得までの期間が除外されてしまい、消失までの期間が実際よりも短く評価されてしまう可能性がある。

さて、Ainsworthらの結果を踏まえると、ウェブページの多くはウェブアーカイブ中に存在している。ウェブアーカイブは、調査のための取得日より誕生に近い日時にウェブページを取得している可能性がある。そこで、ウェブアーカイブが最初に取得した日付を利用することでよりウェブページの誕生日に近い日付を得ることができる可能性がある。この関係を図1に示した。

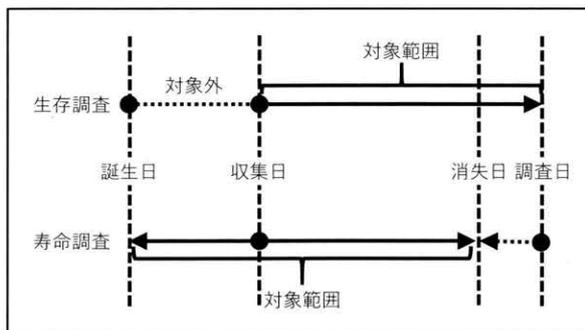


図1 生存調査と寿命調査

そこで、本研究では、2003年調査と同じデータ集合を対象として、前回の調査から10年後という長期的視点から大規模な生存調査を実施するとともに、Internet Archiveを利用して、ウェブページがいつ誕生し、いつ消失したのかに関する調査を実施した。本研究の目的は、2つの調査を組み合わせることで

ウェブページの寿命を明らかにすることである。

## 2. 調査方法

ウェブページの寿命を明らかにするために、調査対象のウェブページが、1)現在アクセスできる状態にあるのか、2)Internet Archive でいつからいつまでアーカイブされているかを調査した。

### 2.1. 調査対象

今回の調査にも、2003年調査と同じく、NTCIR-3 Webタスクのために収集・構築された100GBの文書集合(NW100G-01)を用いた。NW100G-01は、NIIのホームページを起点として、2001年8月29日から11月12日にかけて(46日間)、収集された文書群の部分集合であり、1,100万件以上のウェブページによって構成されている<sup>6)</sup>。NW100G-01から、jpドメインのページのみを対象に無作為抽出した1,000万件のウェブページを調査対象とした。

2001年頃のjpドメインのページ数は、6,507万ページと推定されており<sup>7)</sup>、1,000万件のウェブページは、当時のjpドメイン全体の約1/6を占める文書集合であった。

### 2.2. 生存調査

調査対象となる1,000万件のウェブページについて、生存調査を行った。既に述べたように、筆者らは2003年に同一のウェブページ群を対象とした生存調査を行っていることから、結果の比較を行うことができるよう、調査方法は2003年調査に合わせた。

具体的な手順としては、ウェブページのURLを対象として、設定したタイムアウト時間内にアクセスが可能かどうかを調査した。アクセスできた場合を、「生存」とみなした。アクセスできなかった場合を「消失」とし、消失していた場合には、その理由を記録した。タイムアウトした場合には「タイムアウト」とし、サーバそのものが見つからない場合には「ホストエラー」とした。サーバにアクセスできたが該当ページへのアクセスできなかった場合にはレスポンスコードに基づき、それぞれのエラーとして扱っている。調査期間は2013年7月22日から7月28日の7日間である。

### 2.3. 寿命調査

生存調査と並行して、同じ1,000万件のウェブページを対象として、Internet Archiveの登録データの取得を行った。

ウェブページがいつ誕生しいつ消失していったかに関する情報を正確に取得することは困難である。しかしながら、ウェブアーカイブがページを取得したデータを用いることで、生存分析のために取得した

日時を利用するよりも、より正確な日時に近いデータを取得できる可能性がある。そこで、本研究では、現在3540億ページを保存しているInternet Archiveからデータを取得した。

Internet Archiveからのデータの取得には、Memento API<sup>8)</sup>を用いた。Memento APIは、様々なウェブアーカイブ上に蓄積されたデータを透過的に取得できるように設計されたAPIである。

2013年7月に行った生存調査のデータとInternet Archiveから取得したデータとを組み合わせることでウェブページの寿命の近似値を得た。ウェブページの誕生と消失については、以下の手順によって設定した。

ウェブページの誕生については、Timemapでの「first」(最初にアーカイブによって取得された日時)を用いた。その際にInternet Archiveから当該ウェブページのデータが取得できなかった場合には、寿命調査の対象から除外した。Internet Archiveの取得開始が、NTCIRによる取得日より遅い場合は、NTCIRのクロール期間の最終日である2001年11月12日00時をウェブページの誕生の近似値として用いた。

次にウェブページの消失は、生存調査の結果、消失していることが分かった場合は、Timemapの「last」(アーカイブによって直近で最後に取得された日時)を用いた。生存調査で、現在も生存していたページは寿命の集計から除外した。

上記の方法で判定したウェブページの誕生日時と消失日時の日数をウェブページの寿命とした。

## 3. 調査結果

### 3.1. 生存調査

2003年調査と比較して、生存調査の結果を表1に示した。ウェブページの状態は次のように判別した。

- タイムアウト: 接続時間内に応答がなかった
- ホストエラー: サーバに接続できなかった場合
- サーバエラー: ステータス・コード 500~505
- 認証失敗: ステータス・コード 401~403, 407
- ファイル移動: ステータス・コード 301, 303, 307
- ファイルなし: ステータス・コード 404
- 他のエラー: その他のエラーコード
- 生存: ステータス・コード 200

2003年の列が前回の調査の結果で、2013年の列が今回の調査の結果を示している。調査の結果、

2001年に存在した1000万のウェブページのうち、2013年7月現在でアクセスできたのは、779,176件(7.79%)で、9割以上のウェブページが収集から12年の間にアクセスできなくなっていた。2003年調査の5,336,099件(53.36%)よりもアクセスできるページの比率が減少していた。

アクセスできなかった原因は、2003年調査では、「ファイルなし」がアクセスできなかった理由の多くを占めていた。本調査では、「ホストエラー」が全体の46.5%で最も多くなっていた。ホストエラーに含まれるのは、ウェブサーバがなくなっている場合、DNSによってドメイン名が解決できない場合などが含まれる。2003年調査で最も多かった「ファイルなし」は35.02%で、2番目に多かったが、2003年調査と占める比率は同程度であった。

表1 ページ集合の生存状況

	2003年調査		2013年調査	
	n	%	n	%
生存	5,336,099	53.36%	779,176	7.79%
消失				
タイムアウト	84,255	0.84%	349,896	3.50%
ホストエラー	878,467	8.78%	4,647,225	46.47%
サーバエラー	17,254	0.17%	4,942	0.05%
認証失敗	74,172	0.74%	66,795	0.67%
ファイル移動	386,620	3.87%	240,721	2.41%
ファイルなし	3,219,881	32.20%	3,501,616	35.02%
他のエラー	3,252	0.03%	409,629	4.10%
合計	10,000,000	100%	10,000,000	100%

つぎに、主要なSLD(セカンドレベルドメイン)別での集計を表2に示した。なおTLD(トップレベルドメイン)は全てjpドメインである。これらの中では、goが最も生存率が低く、4.5%である。goドメインのページにアクセスできなかった理由で最も多かったのは、「ホストエラー」で70.7%であった。

### 3.2. 寿命調査

Internet Archiveに登録されたデータを用いて、寿命の分析を行った。寿命調査は、Internet Archiveの

みに対してアクセスを行うこととなるため、サイトへの負荷を考慮しながら調査を行った。約400万件のデータが取得できた時点で、分析を行った。

はじめに、Internet Archiveへの登録と現在の生存の関係について表3に示した。466万件のうち、4,006,158件(86.0%)がInternet Archiveに登録されていた。Internet Archiveに登録されたページのうち89.2%が現在はアクセスできなくなっているページであった。一方、Internet Archiveに登録されていなかったウェブページは約65万件(14.0%)であった。そのうち、95.5%がアクセスできなくなっているページであった。フィッシャーの正確確率検定の結果、Internet Archiveへの登録とウェブページの生存との間に有意な関連が見られた( $p < 0.01$ )。

表3 Internet Archiveへの登録と生存状況

IAへの登録	消失		生存		合計	
	n	%	n	%	n	%
あり	3,572,998	89.2%	433,160	10.8%	4,006,158	86.0%
なし	624,331	95.5%	29,511	4.5%	653,842	14.0%
					4,660,000	100%

つぎに、2001年に収集され、現在は消失していたページ3,572,998件に対して、ウェブページの寿命の分析を行った(表4)。ウェブページ全体の平均寿命は1,108.2日であった。標準偏差は、879.5日で大きくばらつきがあった。2001年に収集されたウェブサイト中で最も早くにInternet Archiveに登録されていたのは、1996年4月4日からであった。

表4 対象ページ全体の寿命

平均	標準偏差	中央値
1108.2	879.5	948.9

表5に、寿命を1年毎にグループ化した集計を示した。1年程度の寿命を持つページが87万件で最多であった。6年以降は概ね年とともに減少傾向に

表2 SLDごとの生存状況

SLD	生存		消失											小計			
	n	%	タイムアウト	ホストエラー	サーバエラー	認証失敗	ファイル移動	ファイルなし	その他のエラー	n	%	n	%				
co	346312	7.5%	116060	2.5%	1949518	42.0%	3089	0.1%	21271	0.5%	141002	3.0%	1845171	39.7%	223812	4.8%	4646235
ac	130344	8.8%	125295	8.5%	684984	46.4%	1336	0.1%	23588	1.6%	23345	1.6%	443689	30.1%	43392	2.9%	1475973
or	125490	10.3%	31283	2.6%	498731	41.1%	101	0.0%	5739	0.5%	22198	1.8%	491830	40.5%	39120	3.2%	1214492
ne	66203	6.6%	29977	3.0%	579100	58.1%	126	0.0%	8098	0.8%	17675	1.8%	233068	23.4%	61907	6.2%	996154
gr	32746	10.4%	5530	1.7%	153649	48.6%	37	0.0%	1625	0.5%	2609	0.8%	114667	36.3%	5262	1.7%	316125
go	14054	4.5%	8595	2.8%	219731	70.7%	2	0.0%	1201	0.4%	7807	2.5%	52616	16.9%	6784	2.2%	310790
ed	12620	8.5%	2059	1.4%	74461	50.4%	12	0.0%	390	0.3%	4770	3.2%	50279	34.0%	3206	2.2%	147797

あった。一方で17年間継続していたページも僅かではあるが存在した(111件)。累積比率で見ると、2年から3年の間に、存在しているウェブページは半減していた。2年程度で半減期を迎えるという結果は、先行研究での結果と大きく異なることはなかった。

表5 寿命1年ごとの集計

年数	n	%	累積%
0	300792	8.4%	8.4%
1	850287	23.8%	32.2%
2	592358	16.6%	48.8%
3	622572	17.4%	66.2%
4	464425	13.0%	79.2%
5	210479	5.9%	85.1%
6	207094	5.8%	90.9%
7	127516	3.6%	94.5%
8	73474	2.1%	96.5%
9	41475	1.2%	97.7%
10	30853	0.9%	98.6%
>10	51,673	1.4%	100.0%
合計	3,572,998	100.0%	

SLD ごとの寿命の平均値と標準偏差、件数を表6に示した。表から、acが1,445.5日と最も長く、4年近い寿命があったことが分かる。一方で、grやedは、比較的寿命が短く、誕生から3年程度で消失していることが分かる。

表6 SLDごとの平均寿命

SLD	平均	標準偏差	n
ac	1445.5	1075.5	507,177
co	1408.9	1259.5	125,488
go	1329.4	1018.4	547,215
ne	1231.7	1017.0	124,806
or	1206.6	1017.0	1,921,908
gr	1113.8	978.4	384,173
ed	1057.0	875.5	56,586

次に、カプラン・マイヤー法を用いて生存関数を推定した。現在も生存していたページは打ち切りデータとした。分析には、R2.15.2のsurvivalパッケージを用いた。SLDごとの中央値、95%信頼上限値、95%信頼下限値は表7の通りである。最も寿命が長いorの場合、1,263から1,269日で半数が消失に至る。最も短いneの場合、895から903日で半数が消失する。次に、ログ・ランク検定を用いて、SLDごとの生存時間に有意差があるか否かを確認したところ、

有意水準1%で有意差が検出された。

表7 SLDごとの寿命の要約

SLD	n	イベント数	中央値	信頼区間	
				下側95%	上側95%
or	507,177	436,095	1,266	1,263	1,269
ac	547,215	475,263	1,155	1,151	1,158
その他	338,805	311,674	1,088	1,085	1,092
gr	124,806	108,374	1,083	1,077	1,087
go	125,488	114,992	1,044	1,037	1,053
co	1,921,908	1,730,975	1,023	1,022	1,024
ed	56,586	49,555	928	917	937
ne	384,173	346,070	899	895	903

#### 4. まとめ

本調査では、収集から12年を経過したウェブページ群の生存分析と、Internet Archiveを用いた寿命の調査を行った。その結果、2001年に収集されたウェブページの9割以上が消失していることがわかり、またウェブページの平均寿命は、1,108.2日であった。

#### 引用文献

- 1) Kahle B. Preserving the Internet. Scientific American. 1997, vol. 276, no. 3, p82-83, <http://web.archive.org/web/19970215093036/http://www.sciam.com/0397issue/0397kahle.html>, (accessed: 2013-09-04)
- 2) Koehler W. Web page change and persistence: a four-year longitudinal study. JASIST. 2002, vol. 53, no. 2, p. 162-171.
- 3) McCluskey, M. Website content persistence and change: Longitudinal analysis of pro-white group identity. Journal of Information Science. 2013, vol. 39, no. 2 188-197.
- 4) 池内淳ら. "ウェブの動的変化に関する調査". 2003年度三田図書館・情報学会研究大会発表論文集. 慶應義塾大学, 2003-11-8. 三田図書館・情報学会, 2003, p. 19-22
- 5) Ainsworth S G. How Much of the Web Is Archived?. JCDL '11. 2011, p 133-136, <http://arxiv.org/abs/1212.6177>(accessed: 2013-09-04)
- 6) 江口浩二ら. NTCIR-3 WEB: Web 検索のための評価ワークショップ. NII journal. 2003, no. 6, p. 31-56.
- 7) 内田斉. メディアとしてのWebの成長を測る: サーチロボットを使ったWebコンテンツ統計調査の試み. <http://www.a-brain.com/result/report/015/>
- 8) Memento: Adding Time to the Web. <http://mementoweb.org/>

# レビューサイトにみる集合知の可能性

アマゾンカスタマーレビューを例として

上田修一† 石田栄美 宮田洋輔 南友紀子 倉田敬子  
前慶應義塾大学 九州大学 常葉大学短期大学部 慶應義塾大学大学院 慶應義塾大学  
† ueda@z5.keio.jp

## 抄録

新しく生まれたメディアとして商品やサービスに対する消費者の評価、意見を掲載するレビューサイトを取り上げ、「集合知」という観点から検討した。これまで集合知と見なされている事例や集合知の特性を述べた言説から要素を得て、アマゾンのカスタマーレビューサイトを例とし、29の質問項目を作り、全国の18歳以上の男女1,240人を対象にパネル利用調査を行った。レビューサイトには集合知としての特性が認められた。

## 1 はじめに

デジタル化とネットワーク化が急速に進展し、メディアの状況は流動的であり、常に枠組みが問い直されている。こうした中で、新興のメディアの実態と特性を明らかにし、メディア全体の中に位置づけることは、現在、メディア研究において活発な領域となっている。

ここで取り上げるメディアは、日常生活において頻繁に利用されるようになった商品やサービスに対する消費者の評価や意見を掲載するウェブサイトやサイト中の機能である。これをここでは「レビューサイト」と呼ぶ。

こうしたレビューサイトは、「クチコミ」サイトとも呼ばれ、ソーシャルメディアや質問回答サイトなどとともに消費者生成メディア (CGM: Consumer Generated Media, または CGC: Consumer Generated Content) の一つとみなされている。しかし、消費者生成メディアという概念には、単に消費者の意見を採り上げているという意味しかない。レビューサイト独自の特性を示しているわけではない。

権威があるわけでもなく、知り合いでもない他人の意見が公表され、利用されることはどのような意味を持つのか、レビューサイトはどのようなメディアとしての特性を持っているのかを「集合知」という観点から検討する。

## 2 集合知

個を超えた集団の持つ知性や知能は、「集合的

知性 collective intelligence」として捉えられてきた。動物は「群知能 Swarm Intelligence」を使って、集団として餌を探したり、捕食者を避けたりしている<sup>1)</sup>。一方、人々の使う集合的知性は、スロウィッキー<sup>2)</sup>やページ<sup>3)</sup>によって、専門家ではない素人の意見の統合によってもたらされる「集合知 wisdom of crowds」として新しく光が当てられるようになった。

集合知の働いている例としてよく取り上げられるのは、一つは、見本市での牛の重量をあてるコンテストや映画のアカデミー賞受賞者の予想であり、もう一つは、ウィキペディアやオープンソースソフトウェアである。

重量当てや賞の予想の場合、回答の平均値か最も票数を集めた選択肢が正解（もしくは正解に近い）となる事象が観察されてきた。スロウィッキーは、集合知の要件として、1) 独立性：各自は独立して意見を出す、2) 多様性：色々な人たちが参加する、3) 分散性：各自がそれぞれの情報源を持つ、4) 集約：みんなの意見を集約するメカニズムがある、をあげている。

他方で、ウィキペディアの場合は正解がなく、多くの素人が構築するものではあるが、他人のまとめたものに修正を加えていくので独立性はない。さらにオープンソースソフトウェアは、素人ではなく多数の専門家による協同作業の成果物といえる。

このように、集合知の概念には多様な現象や状況が含まれており、包括する考え方が示され

ているとはいえない。しかし、ウェブでは多数の一般の人々がCGMを用いた意見表明ができるようになり、集合知が実現される新しい社会的な場や仕組みが整いつつあると言える。

これまでの論考から、集合知であるかどうかの最も基本的な共通する要件は(1)「多様性」といえる。それ以外の要素に関しては、その有無によって集合知のパターン、種類が異なってくると考えられる。種別を判定するための要素としては、その他に以下が挙げられる。

- (2) 素人の意見であるかどうか
- (3) 意見の独立性が担保されているかどうか
- (4) 各自が情報源を持つ(分散性)かどうか
- (5) 意見集約、統合化がなされるかどうか。
- (6) 新たな知がもたらされているかどうか。

映画賞の予想は、(1)(2)(3)は当てはまるが、(4)は不明で、(5)は単純な集計で統合とはいえず、(6)はない。オープンソースソフトウェアは、(2)(3)は当てはまらず、残りが当てはまり、特に(5)(6)で大きな成果をあげているといえる。つまり、これら6要素のうちのどれが当てはまるかによって、集合知の種類が検討できる。

### 3 レビューサイト

現在のレビューサイトの原型は、1997年にアマゾンから提供されはじめた「カスタマーレビュー」であると考えられる。「カスタマーレビュー」は、本や雑誌以外の商品も対象とし始め、やがて、他のオンラインショッピングサイトもレビュー機能を標準的に備えるようになり、商品ばかりでなくレストランやホテルのようなサービス業も対象となった。

レビューサイトでは、個々の商品やサービスについて、利用者の意見、星の数を用いた評価を集積することが基本であり、さらにそれらの意見についての評価やコメントを付与できる場合もある。

レビューサイトは、これまで情報処理分野では意見や評判の自動抽出のための情報源として、言語学では談話分析の材料として使われてきた。また、商品購入者の評価だけでなく購入体験についての情報を入手できることが、第三者の購入、消費を促していることが指摘され、マーケティングで注目されるようになった<sup>4)5)</sup>。一方、商品の販売側やサービスの提供側がレビューサ

イトを不正にコントロールしようとする事例もあり、書評について、感心するようなレビューもいる反面、悪意が働いている場合や浅い意見もあると評されている<sup>6)</sup>ようにレビューサイト特有の問題がある。

## 4 調査目的と方法

### (1) 調査目的

本調査では、レビューサイトが2で述べた集合知という観点から見て、どのように位置づけられるかをレビューサイト利用者の認識から明らかにすることを目的とする。

その際レビューサイトについて漠然とたずねるのではなく、具体的なサイトを対象にその特性に関する認識を調査することとした。対象としたレビューサイトは、アマゾンの「本」に関するカスタマーレビューサイトである。

### (2) 調査項目

2で示した「集合知」の種別を判定する6要素を「本」のカスタマーレビューに当てはめ、(1)多様な意見、(2)素人の意見、(3)意見の独立性、(5)統合化、(6)新たな知の5要素について、複数の質問項目に具体化した。(4)の各自が情報源を持つかどうかは、「本」のカスタマーレビューとして対象となる「本」以外の情報源というものは想定しにくいため質問項目とはしなかった。また、(1)の多様な意見を明確にするため、多様であること(1a)と多数であること(1b)とを分けて質問項目を作成した。

同時に対照群として、集合知の特性とは考えられない「購入」、「情報源」、「共感」としての利用にあたる質問項目も含めた。全体で29項目となった。

### (3) 調査手順

2013年8月上旬に、全国の18歳以上の男女1,240人を対象として、パネル利用によるインターネット調査を実施した。29の質問項目に対し5段階のリッカート尺度により回答を求めた。

なお、附帯調査として、アマゾンの本に関するカスタマーレビューを使う際に利用する機器、本やその他のレビューサイトの利用状況、新聞・雑誌上での紹介や書評を、読む本を選ぶときの参考にしているか否かも併せて質問した。

事前に下記の質問で回答者を限定した。

- 1) 1 か月以内に、「専門書、教養書」「小説」「マンガ」のいずれかを1冊以上読んだ。
- 2) アマゾンで「専門書、教養書」「小説」「マンガ」のいずれかのカスタマーレビューを読んだことがある。

性別と年代(20代以下, 30代, 40代, 50代, 60代)は均等に割り付けた。

## 5 調査結果

### (1) 回答者属性, 読書量, 利用方法

回答者の属性には, やや偏りがみられた。

表1 回答者の職業

職業	人数	比率
公務員	49	4.0
会社員(含む経営者)	469	36.0
自営業	85	6.9
自由業	38	3.1
専業主婦(主夫)	215	17.3
パート・アルバイト	139	11.2
学生	81	6.5
その他, 無職	164	13.2
全体	1,240	100%

小説がよく読まれているが, 読書冊数はマンガが多い。専門書と教養書もよく読まれている。

表2 1か月以内に読んだ冊数

本のタイプ	4冊以上		1~3冊		0	
	人数	比率	人数	比率	人数	比率
専門書, 教養書	119	9.6	624	50.3	497	40.1
小説	207	16.7	603	48.6	430	34.7
マンガ	214	17.3	388	31.3	638	51.5
雑誌	158	12.7	628	50.6	454	36.6

レビューを読む人々の中で, 投稿するのは約3割である。

表3 カスタマーレビューの利用

	レビューを読む		レビューを投稿		「参考になりましたか?」に回答	
	人数	比率	人数	比率	人数	比率
よく	334	26.9	28	2.3	52	4.2
時々	467	37.7	76	6.1	148	11.9
数回	396	31.9	163	13.1	248	20.0
一度	43	3.5	80	6.5	52	4.2
ない	0	0.0	893	72.0	740	59.7

なお, 回答者の94.1%がパソコンから利用していた。

### (2) 集合知としてのカスタマーレビュー

質問項目29項目の回答を点数に換算して分析した。具体的には, 表4のように「どちらでもない」を0として, 「とてもそう思う」を2, 「少しそう思う」を1, 「あまりそう思わない」を-1, 「全くそう思わない」を-2として計算した。

表4 点数換算方式

著者について豊富な情報を得ることができる					
回答数	151	633	236	192	28
配分点数	2	1	0	-1	-2
点数	302	633	0	-192	-56
総点数	687				

この結果を次ページの表5に示した。次に, カテゴリごとの平均点数を算出した。

表6 カテゴリ別の点数

カテゴリ	平均
(1a)多様な意見	810.0
(1b)多数の意見	995.0
集合知 (2)素人の意見	519.0
(3)意見の独立性	260.0
(5)統合化	539.0
(6)新たな知	629.7
購入	99.0
情報源	628.5
共感	37.3

「集合知」の中では, 「いろいろな意見を知ることができる」(1308), 「学者, 作家, 評論家といった人が書いた書評とレビューは違う」(1019)のように「(1a)多様な意見」や「(2)素人の意見」のカテゴリの項目に支持が多い。

「一冊の本に数多くのレビューがある」(995)に示される「(1b)多数の意見」にもかなりの支持がある。

「その本のレビューに対して他のレビュアーが意見や評価をしている」(260)のような「(3)意見の独立性」はあまり特徴として認識されていない。

「(6)新たな知」のカテゴリの質問も平均点が630点と, 支持が多かった。

「購入」カテゴリの支持は, 全体的に点数が

低く(平均99), アマゾンカスタマーレビューは, レビューの数や星の数によって本を購入するサイトと必ずしも認識されていないと言える。

「情報源」カテゴリの点数は, 平均629点と比較的高い傾向にある。最も支持が多いのは「その本の雰囲気わかる」(1253)である。「出版社の宣伝に使われている」(103), 「星の数でその本の評価がわかる」(189)は, 相対的に支持が少ない。また, 「購入」カテゴリの結果と併せると, 星の数に示される評価よりも, レビューの内容を重視していると考えられる。

「共感」カテゴリは, 「その人の意見なら信用できるというレビューアがいる」(-890), 「レビューの内容はどれも納得できる」(-339) 「レビューを評価したり応援できる(「参考になった」やコメントをつける)」(-188)と支持の低い項目がある反面, 「自分が読んだ本について共感できる意見が見つかる」(743), 「自分と似たような感想, 意見を持つ人が見つかる」(755)は, 支持が多かった。また, 「自分の感想が多数

の人と同じであることを確かめることができる」(143)への支持は分かれていた。

#### 引用文献

- 1) レン・フィッシャー. 松浦俊輔訳. 群れはなぜ同じ方向を目指すのか?. 白揚社, 2012. 309p,
- 2) ジェームズ・スロウィッキー. 小高尚子訳. 「みんなの意見」は案外正しい. 角川書店, 2006. 286p.
- 3) スコット・ペイジ. 水谷淳訳. 「多様な意見」はなぜ正しいのか. 日経BP社, 2009. 486p.
- 4) 澁谷覚. クチコミの有用性: 購入シミュレーションとしてのネットクチコミ. 日経広告研究所報. Vol. 45, No. 3, p. 11-18. 2011.
- 5) 佐々木裕一. 商品購入における評価サイトの有効度. 情報メディア研究. Vol. 3, No. 1, p. 29-42. 2005.
- 6) 豊崎由美. ニッポンの書評. 光文社, 2011. 230p.

表5 質問項目と点数

カテゴリ		質問項目	点数
集合知	(1a)多様な意見	いろいろな意見を知ることができる	1308
		その本に対する星の数やレビューの内容にばらつきがある	864
		自分が読んだ本について自分とは異なる意見が見つかる	897
		その本に否定的な意見を知ることができる	995
		専門的な意見を知ることができる	-14
	(1b)多数の意見	一冊の本に数多くのレビューがある	995
	(2)素人の意見	自分の感覚に近い人たちが書いている	19
		学者, 作家, 評論家といった人が書いた書評とレビューは違う	1019
	(3)意見の独立性	その本のレビューに対して他のレビューアが意見や評価をしている	260
	(5)統合化	その本に対するレビューのそれぞれが一つの方向(意見)にまとまっていない	539
(6)新たな知	否定的なレビューが多くても, 自分には合いそうな本だと思える	351	
	星の数とは関係なく自分の意見を持つことができる	726	
	レビューを読むことで考えてもいなかった「認識」を得ることができる	812	
購入	レビューの数が多いことで本を買うかどうか判断している	-76	
	星の数で本を買うかどうか判断している	-61	
	買うつもりでいた本でもレビューの内容が悪くて買うのをやめることがある	194	
	レビューの内容を読んで本を買うかどうか判断している	339	
情報源	出版社の宣伝に使われている	103	
	星の数でその本の評価がわかる	189	
	著者について豊富な情報を得ることができる	687	
	レビューであらすじなど大まかな内容がわかる	748	
	その本の問題点(翻訳がひどい, 間違いが多いなど)がわかる	791	
	その本の雰囲気がわかる	1253	
共感	その人の意見なら信用できるというレビューアがいる	-890	
	レビューの内容はどれも納得できる	-339	
	レビューを評価したり応援できる(「参考になった」やコメントをつける)	-188	
	自分の感想が多数の人と同じであることを確かめることができる	143	
	自分が読んだ本について共感できる意見が見つかる	743	
	自分と似たような感想, 意見を持つ人が見つかる	755	

# アクセスログに基づく国立国会図書館サーチの利用状況の分析

佐藤翔 (同志社大学社会学部) min2fly@slis.doshisha.ac.jp  
川瀬直人 (国立国会図書館) 塩崎亮 (国立国会図書館)  
北島顕正 (国立国会図書館) 原田隆史 (同志社大学社会学部)

## 抄録

本研究では国立国会図書館が提供する検索サービス、国立国会図書館サーチ (NDL サーチ) のアクセスログを分析し、提供する各検索機能の利用状況を明らかにした。分析の結果、検索利用全体の4割は簡易検索の利用だが、詳細検索など他の機能も利用されていること、検索結果を絞り込む機能が範囲を広げる機能よりよく使われていること、複数のデータベースをまとめて検索し、適合度順で結果を表示できるNDL サーチの機能は、そのまま受容されている場合が多いこと等がわかった。

## 1. はじめに

本研究では国立国会図書館 (NDL) が提供する検索サービス、国立国会図書館サーチ (NDL サーチ) のアクセスログを分析した。分析結果から、利用者の行動実態を明らかにすることが研究の目的である。

NDL サーチは「国内の各機関が持つ豊富な「知」を活用するためのアクセスポイント」となることを目指し、2012年1月から正式に稼働した<sup>2)</sup>。NDL が所蔵する図書や雑誌記事、デジタル化したコンテンツ等のNDL 自身が保有する情報だけではなく、都道府県立図書館や大学図書館、国立情報学研究所等の学術機関、出版社の近刊情報など、約200のデータベースと連携し、約7,300万件の文献情報等を検索することができるサービスである<sup>3)</sup>。連携機関から収集したメタデータをただまとめて提供するのではなく、書誌・著作同等等のメタデータマネジメントを行なう点、多様なコンテンツをまとめて適合度順の検索結果表示を実現している点、検索結果の絞り込み・拡張など、利用者が自身の検索結果をマネジメントする機能を提供している点等に特徴がある。今後の図書館システムの一つのモデルともなりうる、いわゆるディスカバリインタフェースとしての仕組みを有するサービスとしても注目を集めている<sup>4)</sup>。

本研究ではこのNDL サーチの利用実態について、特に各検索機能の利用状況を中心に、アクセスログに基づいて分析した結果を報告する。その結果はNDL サーチ自身の評価や機能改修に活かすことができるのはもちろんのこと、今後の図書館における検索サービスのあり方を検討する上でも重要な示唆を与えるものと考えられる。

## 2. 調査方法

### 2.1 分析対象データ

#### (1) 分析期間

本研究ではNDL サーチが正式稼働を開始した

2012年1月から、2012年9月までのアクセスログを分析対象とする。全期間分のログを対象とするとデータ量が膨大になるため、本稿ではそのうち2012年4月の1ヶ月分のログの分析結果を報告する。4月を選んだのは、他の月に比べてもっとも平均的なデータ量の月であったためである。

#### (2) 分析から除外するログ

アクセスログ中には利用者自身が検索クエリを入力した場合のほか、NDL サーチが提供するAPI に対するアクセスや検索ロボットと考えられるアクセス、NDL サーチのシステム自身が利用者行動に伴って自動でクエリを発する場合など、機械による動作の記録も含まれている。これらの記録は利用者自身の行動を知る上ではノイズとなりうるため、分析対象から排除した。API に対するアクセスや自動的に発行されるクエリについてはURL から特定し、検索ロボットについては佐藤が機関リポジトリのアクセスログ分析時に作成したロボットリストに基づき排除した。

また、NDL 職員によるアクセスは機能検証を目的とするもの等が含まれるため、IP アドレスから特定し、分析から排除した。

加えて、本研究の目的はNDL サーチの中でも検索機能の利用状況を明らかにすることであるため、検索に関係しない利用の記録も分析から除外した。

#### (3) セッション情報の付与

本研究の目的である各検索機能の利用状況の分析のためには、個々の利用者の一連の行動 (セッション) を特定できる必要がある。NDL サーチのアクセスログの中にはセッションを示すID 等は含まれていないため、本研究ではIP アドレスとアクセス日時に基づいて個々のアクセスをセッション単位にまとめ、セッションごとに利用時間や検索回数等の分析を行なうことができるように

した。この方法ではプロキシを利用している機関からのアクセス等でセッションの誤同定が発生する(異なる利用者からのアクセスをまとめてしまう)ことがあるが、極端に利用時間が長いセッションや検索回数の多いセッションはセッションごと分析から除くことで対応した。具体的には、利用時間が全セッションのうち上位1%に含まれるセッション、検索回数上位1%のセッション、書誌データ閲覧回数上位1%のセッションを分析から除外した。

最後に、検索行動の分析に直接関係しない、1セッション中で1度も検索を行っていないセッション(トップページを閲覧したのみのものや、書誌データのみを閲覧し続けているもの等)は分析から除外した。

以上の処理の結果、2012年4月の全アクセスログ26,569,178件中、分析対象となったログは1,193,222件(約4.5%)で、総セッション数は148,941件であった。

## 2.2 分析内容



図1. NDLサーチの検索結果画面の一例

図1はNDLサーチの検索結果画面の一例を示したものである。画面中央部に表示されているのが検索クエリに適合したレコードの情報である。

画面上部には「簡易検索」、「詳細検索」、「障害者向け資料検索」の3つのタブが表示されている(図1は検索結果ページであるが、これらのタブはトップページにも表示される)。このうち簡易検索を選択すると検索ボックスが1つだけ、詳細検索と障害者向け資料検索ではフィールドごとに分かれた複数の検索ボックスと、検索対象データベース等を選択するチェックボックスが表示される。これらの検索機能を用いる場合、利用者は自身でクエリを入力する必要がある。

図1の画面左側には「検索結果の絞り込み」(絞

り込み)、右側には「著者名キーワード」「連想キーワード」等のキーワードが表示されている。中央の検索結果を中心とし、左側は資料種別や出版年等で結果を絞り込むための機能(「絞り込み」)、右側は検索結果中の語から連想されるキーワードや関連語を表示し、新たな検索のヒントを提供する機能(「拡張」)である。また、検索結果一覧の上部には検索結果を「適合度順」「出版年が新しい順」等、利用者が望む順序に並び替えるための「並び替え」が表示されている。そのほか、図1には含まれていないが、個々の書誌閲覧画面では著者標目やキーワードから新たな検索を行なうためのリンク(「書誌からのリンク」)も表示される。これらの機能においては、利用者は新たに検索クエリを入力せずに、自身の検索結果画面からのリンクをクリックするだけで新たなクエリを作成することができる。

本研究ではこれら「簡易検索」、「詳細検索」、「障害者向け資料検索」、「絞り込み」、「拡張」、「並び替え」、「書誌からのリンク」の7つの検索機能について、利用者がどの程度用いているか、さらにそれぞれの機能を用いた後、検索結果の中から書誌詳細へのリンクをクリックする回数ほどの程度かを分析する。また、「並び替え」、「詳細検索」、「絞り込み」についてはどのような観点からの検索や並び替えを行なう場合が多いか等の詳細も分析した。

## 3. 結果

### 3.1 セッションごとの行動時間、行動回数

表1は1セッションあたりの時間(最初のログと最後のログの日時の差分)、セッション中に含まれる検索回数、書誌閲覧数の状況をまとめたものである。セッション中に一度も検索が含まれていないログは分析から除いているので、検索回数の最小値は1になる。また、セッション時間については1セッション中に一度しか行動が含まれていない場合については集計から除いている。平均値も示してあるが、データの偏りが大きいため代表値としては中央値を見るほうがふさわしいと考えられる。

表1. セッションごとの行動時間、行動回数

	セッション時間(秒) (N=133,809)	検索回数 (N=148,941)	書誌閲覧 (N=148,941)
平均値	831.5	4.7	2.3
中央値	148	2	1
最頻値	8	1	0
最小値	0	1	0
最大値	11,229	56	32

一度以上、なんらかの検索行動を行なっている利用者に限ると、1セッションにおける行動時間の中央値は148秒(2分28秒)であり、その中で2回、なんらかの検索を行ない、書誌詳細を1件は閲覧する、という傾向が見られる。

### 3.2 各検索機能の利用状況

#### (1) 全体の傾向

表2は2.2で説明した各検索機能の利用状況と、それらの機能を利用した者がその後、検索結果に表示された書誌詳細を閲覧した回数をまとめたものである。

対象データ中のなんらかの検索機能の利用を示すログは605,437件であり、うち最も多いのは簡易検索の利用である。簡易検索の利用が多いのは事前の予想通りであるが、その割合は全体の4割程度にとどまっている。その他の利用者自身がクエリを入力する検索のうち、障害者向け資料検索の割合はごくわずかであるが、詳細検索の利用は14.3%で、簡易検索の3分の1程度の数に至っている。OPACの利用行動における先行研究では簡易検索の利用がほとんどであるとの指摘もあるが<sup>6)</sup>、NDLサーチは異なる傾向を示している。

利用者自身がクエリを入力しない機能の中で、最も使われているのは並び替え機能(全体の約20%)である。内訳を見るとデフォルト設定である適合度順の並び替え指定が最も多く(96,906件)、なんらかの理由で適合度順以外の表示に変更していた者が、適合度順表示に戻す場合が多いものと考えられる。他には出版年の降順(新しい順、13,183件)、出版年昇順(古い順、9,468件)等への並び替えが多い。

並び替えに次いでよく使われるのは絞り込み機能で、全体の15.2%に至っている。一方で拡張機能の利用は全体の1.9%にとどまっている。書誌からのリンクの利用は全体の6.8%で、一定数存在する。

表2. 各検索機能の利用状況

検索の種類	A: 検索数	検索中の割合	B: 書誌閲覧	B/A
簡易検索	248,281	41.0%	115,187	46.4%
詳細検索	86,531	14.3%	32,951	38.1%
障害者向け	4,155	0.7%	1,779	42.8%
絞り込み	92,172	15.2%	22,663	24.6%
拡張	11,339	1.9%	2,063	18.2%
並び替え	121,614	20.1%	32,730	26.9%
書誌リンク	41,345	6.8%	13,966	33.8%
合計	605,437	100.0%	221,339	36.6%

#### (2) 詳細検索機能の内訳

NDLサーチでは前述のとおり多くのデータベースを統合して検索することが可能であるが、詳

細検索時には対象とするデータベースを指定することができる。詳細検索の利用が多い理由として、実際には利用者は複数のデータベースの検索を求めている(NDLの情報のみ検索したい)場合が多く、詳細検索機能を使って余計なデータベースを検索から除いている可能性も考えられる。

そこで利用者が詳細検索時にどの項目を指定していたかを分析してみた結果が表3である。なお、詳細検索時には同時に複数の項目を指定することができるため、表3の値の合計は詳細検索利用数の合計とは一致しない。

表から、もっとも多いのはタイトルでの検索であり、2番目に多いデータベースを指定しての検索の2倍以上、利用されている。表3のうちデータベースや資料種別、所蔵館についてはそれらだけの検索は不可能(他の項目の値を指定する必要がある)ため単純な比較は不可能であるが、少なくとも対象データベースを絞って検索するためだけに詳細検索が使われているわけではないことが見て取れる。

さらに、利用者が同時にいくつの項目を指定する場合が多いかを見てみると、もっとも多いのは1項目のみの場合で44,384件と過半数を占める。その約半数(21,115件)はタイトルでの検索である。2項目以上指定する場合もタイトルと著者、タイトルと出版者等の組み合わせでの検索が多い。詳細検索利用者はデータベースを絞り込むためではなく、項目を指定した検索のために詳細検索を利用していると言える。

表3. 詳細検索時に指定する項目(N=86,531)

指定項目	件数
タイトル	51,336
データベース	22,631
ISBN	19,687
著者名	18,538
資料種別	15,455
出版年	8,721
出版者	8,134
所蔵館	2,951
件名	2,480
分類	2,320

#### (3) 絞り込み機能の内訳

表4は絞り込み機能の利用者が、どの項目を指定して検索結果の絞り込みを行なったかをまとめたものである。こちらも同時に複数項目を指定するので合計は表2の値と一致しない。

表から、絞り込み時にもっとも指定されるのは資料種別であり、次いでデータベース、所蔵館と続く。絞り込み機能は検索結果の件数が多すぎる

と利用者が感じた場合に用いられるものと考えられるが、その場合でもデータベースでの絞り込み以上に資料種別での絞り込みが用いられている。

表 4. 絞り込み時に指定する項目 (N=92,172)

指定項目	件数
資料種別	44,062
データベース	30,205
所蔵館	18,678
出版年	10,779
分類(記号)	4,310
特徴語	2,929
国・地域	153
分野(語)	102

#### 4. 考察

本研究の分析結果のうち、特徴的な点をまとめると以下のとおりである。

- a) NDL サーチでもっとも利用頻度が多いのは簡易検索であるが、詳細検索や他の検索機能の利用も多い
- b) 検索結果のマネジメント機能では結果を絞り込む機能の方が検索を拡張していく機能よりもよく使われる
- c) 並び替え機能もよく使われているが、適合度順のままにする、もしくは適合度順に戻す場合が多い
- d) 詳細検索、絞り込み機能等を用いる際、利用者はデータベースを指定するよりも他の項目を指定することの方が多い

これらの結果をどう解釈するかは複数の可能性がありうる。その中でも一つ言えるのは、利用者の多くは NDL サーチを複数のデータベースを統合して検索できる状態のまま、かつ検索結果は適合度順に表示する設定のままに利用していることである。もちろんデフォルトの設定を変更する利用者がそう多くはないであろうことに留意する必要はあるが、NDL が所蔵しない情報も含めた検索サービスを提供する、NDL サーチの狙いは利用者からも受け入れられていると考えられる。

検索結果の絞り込み機能、並び替え機能等、利用者が自身の検索結果をマネジメントしながら検索を行なえる機能がよく使われている点も注目に値する。このような機能はいわゆるディスカバリインタフェースにおいて重要な要素となりうるとの指摘があるが<sup>7)</sup>、本研究の結果からも利用者の需要の存在が示されたと言えよう。ただし、これらの機能が利用されるのは利用者が最初の検索結果表示に満足できなかったためとも考えられる。

表 2 に示したとおり、検索結果のマネジメントを行なった場合の方が、その後いずれかの書誌情報を閲覧する割合が低いこともその傍証となる(適当な情報が表示されなかったのでクリックしなかった可能性がある)。利用者の満足度等はアクセスログからは明らかにできないが、インタビューや実験等を組み合わせることでこの点についても今後、検証できる可能性はある。

同様に解釈が分かれないのが、拡張機能の利用の少なさである。この点については現在の NDL サーチの拡張機能の性能が低いために利用されないのか、そもそも検索結果の絞り込みに比べて拡張に対する需要が低いのか、という可能性が考えられるが、これもログのみから明らかにすることはできない。ただし、拡張機能が使われる場合にどのようなキーワードであればクリックされるか等を見ることは可能であり、今後の課題の一つと言える。

本研究はまだ分析の端緒段階にある。今後は対象とするログの範囲を拡大するとともに、よく指定される資料種別やデータベースの分析や、書誌閲覧状況との関係等のさらに踏み込んだ分析も行なっていく。また、前述のとおりログ分析以外の手法も組み合わせつつ、NDL サーチの機能要件、ひいては図書館等における検索システムに求められる機能について検討を続けていく予定である。

#### 引用文献

- 1) 国立国会図書館サーチ. <http://iss.ndl.go.jp/>, (2013-08-28 参照).
- 2) 中山正樹. 国立国会図書館におけるデジタルアーカイブ構築: 知の共有を目指して. 情報管理. 2012, vol.54, no.11, p.715-724.
- 3) “機能概要”. 国立国会図書館サーチ. <http://iss.ndl.go.jp/information/function/>, (2013-08-28 参照).
- 4) 原田隆史. 国立国会図書館サーチとディスカバリインタフェース. カレントアウェアネス. 2012, no.311, p.2-4.
- 5) 佐藤翔. コンテンツ入手元として機関リポジトリが果たしている役割. 筑波大学, 2013, 287p, 博士論文.
- 6) 種市淳子, 逸村裕. 短期大学図書館における情報探索行動: 目次を付与した OPAC のログ分析と検索実験をもとにして. 名古屋大学附属図書館研究年報. 2007, vol.5, p.57-68.
- 7) 宇陀則彦. ディスカバリサービスに関する少し長いつぶやき. デジタル図書館. 2012, no.43, p.11-14.

# Webcat Plus への問い合わせとその対応にみる 名寄せ処理の課題

阿辺川 武<sup>†</sup> 間下 亜紀子<sup>†‡</sup> 高野 明彦<sup>†‡‡</sup>

<sup>†</sup> 国立情報学研究所 <sup>†‡</sup> 特定非営利活動法人連想出版 <sup>†‡‡</sup> 国立情報学研究所

<sup>†</sup> abekawa@nii.ac.jp

国立情報学研究所が運営している Webcat Plus では、複数の大規模な書誌データベースを統合し、それらを検索する機能を提供しているが、2012年12月に内部の名寄せ処理を一新した。本稿では新しい名寄せ処理の仕組みと、Webcat Plus に寄せられる問い合わせから判明したシステムの不具合および課題について報告する。

## 1 はじめに

国立情報学研究所では、書誌情報の他に目次・概要情報を利用することで実現した連想検索機能や本、人、作品という3つの軸で検索できる機能を特徴とする Webcat Plus<sup>1</sup> を2002年より提供している。Webcat Plus では、全国の大学図書館や国立国会図書館(以下、NDL)の書誌、全国のお古書店から提供される書誌など延べ3000万件以上の本の情報を収集し、日本最大規模の本の情報や各書誌情報へのリンクリゾルバの提供を目的としている。2010年6月にデータベースおよびインタフェースを大幅に刷新し、2012年12月に名寄せ処理について抜本的な見直しを行った。

Webcat Plus では複数のデータソースで同一の情報を統合する名寄せ処理をおこなっているが、膨大な書誌情報を機械的に高精度で名寄せすることは難しく、誤名寄せが一定量存在する。運営側で発見できる誤りには限りがあり、一般ユーザから誤名寄せを指摘されることも少なくない。その都度、誤りの原因を調査し、システムの改善や、1次データソース情報に修正を入れている。

本稿では、最初に Webcat Plus で開発した名寄せ処理の概要説明をおこない、名寄せに関する統計情報を提示する。次に Webcat Plus に寄せられる問い合わせから誤名寄せを指摘している例を選択し、その原因および対応策について説明する。

## 2 Webcat Plus の名寄せ処理

### 2.1 名寄せ処理の方針

2010年に稼働したシステム(旧システム)での問題点を踏まえ、2012年12月から稼働しはじめたシステム(新システム)では、次の方針をたてた。

- 旧システムでは集合書誌での登録が多い NACSIS-CAT の情報を最優先ソースとしてマスターデータを構成していた。新システムでは各巻ごとの情報を提示したいため、NDL の JAPAN/MARC(以降、JP/MARCあるいは MARC) を最優先で扱う
- 新システムでは「別の情報である可能性がある場合は別の情報として扱う」ことを大前提とし、データの類似度が高くても、同一であると確定できない場合は無理に名寄せしない
- 名寄せの結果が誤りである、と判断した場合に情報を改めて分割、あるいは訂正する必要があるが、旧システムではデータの分割が困難であったため、これを容易にする

### 2.2 名寄せの処理の流れ

- 共通フォーマット変換: 各データソースからタイトル・著者・出版年など名寄せに必要なフィールドを抽出
- データ正規化: Unicode 正規化(NFKC)、ハイフン除去、異体字の統一などを文字列のマッチングをおこなうために表記を統一
- 指定した条件による名寄せ: 表3、表4に示す条件を優先度順に照合し、条件を満たすレコードを同一とみなす
- フィルタリング: 誤って名寄せしてしまった情報を除外する条件を指定

### 2.3 図書名寄せ

各データベースによって書誌項目が異なり、情報の粒度、書誌単位の方針などが大きく異なるため、図書名寄せ処理は大変難しい問題であり、現在までに様々な手法が提案されている<sup>2</sup>。新シス

表 1: 主なデータソースの全国書誌番号, ISBN 設定率と典拠情報保有率および名寄せ件数 (2013/8 現在)

データソース名	登録件数	全国書誌番号あり	ISBN あり	典拠あり	他と名寄せされた件数
NACSIS-CAT 図書	10,205,495	985,391 (10%)	4,123,347 (40%)	6,712,058 (66%)	2,987,315 (29%)
JP/MARC 図書	4,654,367	4,654,363 (100%)	1,999,080 (43%)	3,767,688 (81%)	3,089,293 (64%)
NDL デジタル化資料	1,180,951	907,348 (77%)	783 (0%)	0 (0%)	891,860 (76%)
Book データベース	1,387,160	0 (0%)	1,361,919 (98%)	0 (0%)	1,358,727 (97%)
Nielsen Book Data	11,998,976	0 (0%)	11,998,976 (100%)	0 (0%)	741,502 (6%)

表 2: 人物名寄せ対象となるデータ件数と名寄せ件数 (2013/8 現在)

データソース名	登録件数 (A)	生年のある件数 (B)	書誌由来名寄せ対象となった件数 (C)	名寄せ処理対象となった件数 (B+C/A)	他と名寄せされた件数 (D)
CAT 著者名典拠	1,622,432	593,757	151,193	749,254 (46%)	365,686 (22%)
JP/MARC 著者名典拠	1,023,320	491,064	111,284	610,002 (60%)	371,994 (36%)
青空文庫人物	785	727	0	727 (88%)	590 (72%)
Wikipedia 人物	193,039	186,890	0	186,890 (95%)	39,920 (20%)

テムでは、名寄せ処理条件に厳密で子細な定義を設けることはせず、誤った名寄せはをなるべく生じさせないという方針のもと、表 3 に示す比較的簡易な名寄せ条件を設定した。

最優先条件である「名寄せキー」とは、機械的な同定では同一としないレコードであっても、有識者が確認すれば同一の書誌、人物等を指し示すレコード群を強制的に名寄せするために、管理者が人手で設定する情報である。また逆に他と名寄せしたくないレコードに対しては、一意な名寄せキーを設定すると、他の情報と名寄せされない。

名寄せ後、マスターデータには、優先度の高いデータソースからの情報を登録する。新システムでは JP/MARC であると述べたが、JP/MARC では官報が年単位で登録されているのに対し、NDL デジタル化資料では日単位で登録されているなど、より細かい単位となっているため NDL デジタル化資料を最優先のデータソースとしている。

2013 年 8 月現在の主なデータソースの件数および図書名寄せの結果を表 1 に示す。

## 2.4 人物名寄せ

人物については表 3 の条件を使用して名寄せを行なうが、図書のように優先度は設定していない。項目 3 の条件では、一方の人物に生年が記述され、もう一方で生年が存在しない場合、名寄せが行われない。図書名寄せで人物を含む書誌レコード同士が名寄せされた場合、それぞれの人物も同一で

表 3: 図書の名寄せ条件

優先度	条件
1	名寄せキーの値が一致
2	ISBN のいずれかが一致
3	全国書誌番号が一致
4	タイトル, 別タイトルのいずれかが一致 著者名のいずれかが一致 出版者名のいずれかが一致 シリーズ同士が一致するか未定義 巻号 (版) が一致するか未定義

あるとみなすことができる。そこで書誌レコードが名寄せされたとき、名寄せされた書誌が著者 ID を持っており、著者 ID 同士がまだ名寄せされていない場合、これらの著者は同一の人物であるとして「書誌由来名寄せキー」を設定する。

2013 年 8 月現在の主な典拠データの件数および人物名寄せの結果を表 2 に示す。

表 4: 人物の名寄せ条件

項目	条件
1	名寄せキーの値が一致
2	書誌由来名寄せキーが一致
3	上記名寄せキーが不一致 著者名が一致 生年が一致 没年が一致するか未定義

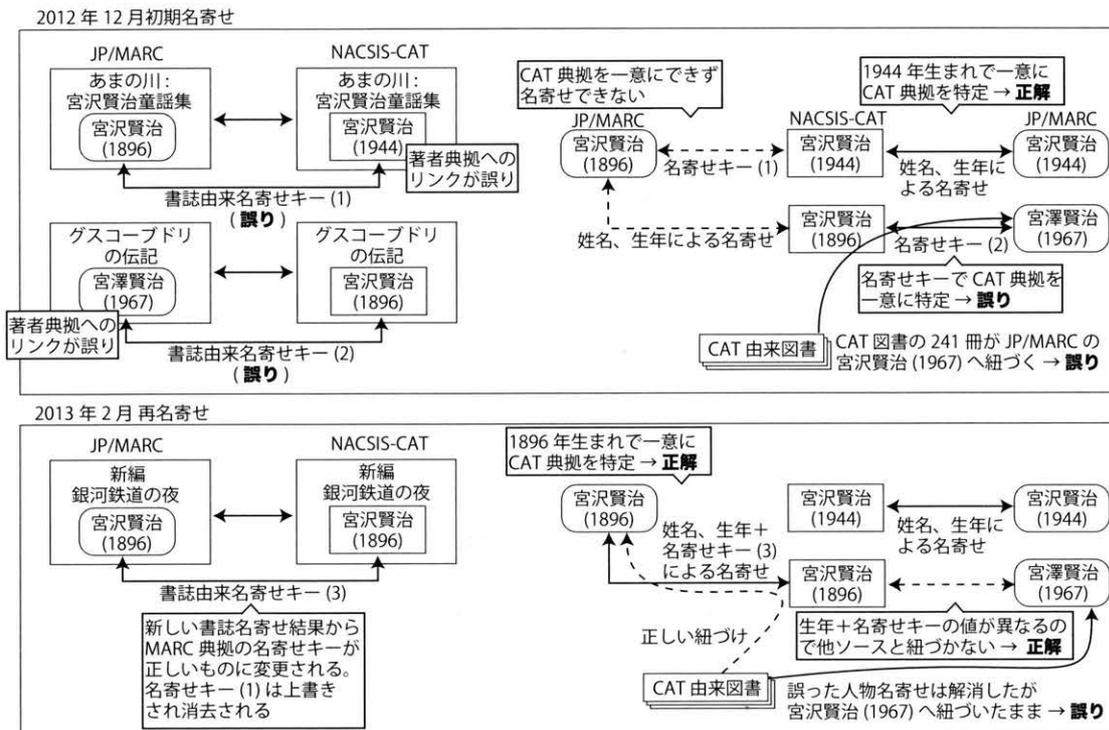


図 1: 宮沢賢治の著者と図書の紐づけ誤りの原因

### 3 問い合わせ例とその原因および対応

本節では Weecat Plus に寄せられる問い合わせと、どのような対応を行なっているかについて事例とともに紹介する。

#### 問い合わせ例 1

『競争と景気循環 第2版』(NACSIS-CAT: BA43829547)を他のデータ(初版)と分けてほしい

[原因] 名寄せ条件では、双方の書誌の巻号(版)に値があり、かつそれらが不一致のとき名寄せしないが、通常初版では巻号のフィールドは未定義であり、初版と第2版で名寄せが行われていた。

[対応] 表5のように分割したい書誌に別々の名寄せキーを付与する。

[課題] 刊行年が異なる場合に名寄せしないというフィルタリング条件を追加すれば、この例は解決できる。しかし、この条件により本来同一書誌であると考えられる書誌が名寄せされない可能性があるため、全体への影響を検討し、追加するかを決定する必要がある。

表 5: 問い合わせ例 1 の対応処理

CAT: BA43829547	名寄せキー“A.”を設定
CAT: BN08213350	
JP/MARC: 2205659	名寄せキー“B.”を設定
BOOK: B9225704	

#### 問い合わせ例 2

著者・編者に「宮澤賢治」と入力し、出版元を「新潮」「新潮社」と入力して検索すると、「検索結果 0 件目」となります。しかし、新潮社から宮澤賢治の本は出ています

[原因] リニューアル時のプログラムのミスで出版元を指定した検索が機能していなかった。

[対応] プログラムを修正し、出版元を指定した検索を有効にした。

このとき調査の過程で、本来「宮沢賢治 (1896)」が著者である 241 冊の図書が、誤って「宮澤賢治 (1967)」の著者となっていることが判明した。

[原因] 元データソースの著者 ID の選択ミス、および人物名寄せ変更後の紐づけの再登録ミス。

[詳細] 図 1 を参照。2012 年 12 月の初期名寄せ時に MARC, CAT それぞれの書誌で誤った著者 ID が登録され、これにより 2 つの誤った書誌由来名寄せキーが生成された。この名寄せキーの影響で、「MARC 宮沢賢治 (1896)」と「CAT 宮沢賢治 (1896)」が正しく名寄せされず、「MARC 宮澤賢治 (1967)」と「CAT 宮沢賢治 (1896)」が誤って名寄せされた。この人物名寄せを基に「CAT 宮沢賢治 (1896)」に紐づいている CAT 由来の図書 241 冊が、優先度の高い「MARC 宮澤賢治 (1967)」に紐づけられてしまった。

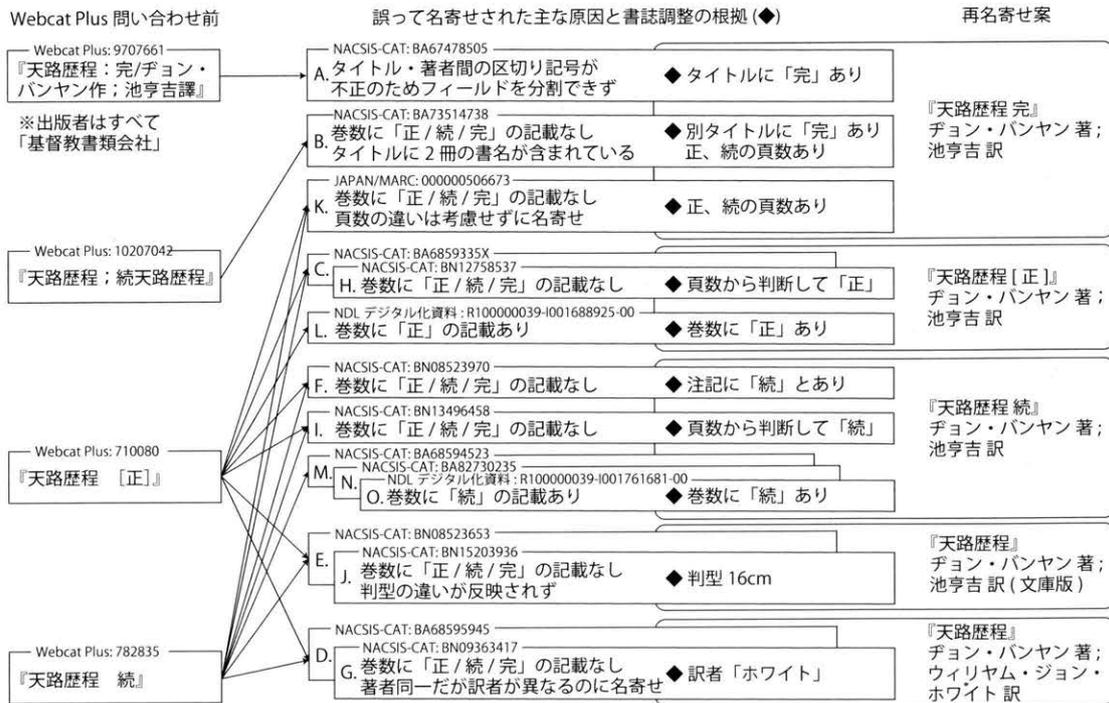


図 2: 『天路歷程』の名寄せ誤りの原因と修正方針

次いで 2013 年 2 月の再名寄せ時に、新しい書誌名寄せから正しい書誌由来名寄せキーが生成され、以前の誤った名寄せキーが上書消去された。この名寄せキーから「MARC 宮沢賢治 (1896)」と「CAT 宮沢賢治 (1896)」が正しく名寄せされ、「MARC 宮澤賢治 (1967)」と「CAT 宮沢賢治 (1896)」の誤名寄せは解消されたが、「MARC 宮澤賢治 (1967)」に誤って紐づけされた 241 冊については、紐づけがそのまま変更されていなかった。  
 [課題] 人物名寄せに変更が合った場合には、それを検知し、変更前の名寄せで紐付けられている図書がある場合、新しい名寄せに合わせて紐付けを修正する処理を追加する。なお MARC 内の「グスコブドリの伝記」の誤った典拠 ID は、2013 年 3 月に MARC 側で修正されている。

問い合わせ例 3

CiNii Books から Webcat Plus へのリンクの不整合を発見しました。  
<http://ci.nii.ac.jp/ncid/BN09363417> から画面内のリンクで Webcat Plus へ遷移すると BA6859335X の書誌が表示される

[原因] 誤名寄せの原因は図 2 を見ていただきたい。元データの入力ミス、名寄せに必要なデータの不足、現状の Webcat Plus の処理、からなる。  
 [対応] 本件については、名寄せキーによる強制

名寄せをするか、名寄せ処理自体の改良を行なうか、まだ正式な対応策は決定しておらず、ここでは再名寄せ案を示す。

[課題] 図 2 の書誌調整の根拠を用いて機械的な名寄せをすると、頁数や判型の違いの考慮、注記から必要な情報を読み取るなど、大幅なシステムの変更となる可能性が高く、慎重な検討を要する。

4 おわりに

Webcat Plus では、なるべく誤った名寄せを行わないという方針のため、人手であれば同一とみなせる情報が名寄せされずに大量に残っている。今後はこれらの情報を効率的に抽出、処理する仕組みを導入する必要がある。

問い合わせ例 2 に見るように現在の名寄せ処理は、誤った情報に対して脆弱であるため、今までの情報と矛盾するときにアラートを出すなど頑健なシステムを目指す。

注

- 1) 国立情報学研究所. Webcat Plus. <http://webcatplus.nii.ac.jp/>. 2010.
- 2) 谷口祥一. FRBR OPAC 構築に向けた著作の機械的同定法の検証-JAPAN/MARC 書誌レコードによる実験. Library and information science, Vol.61, pp.119-151, 2009.