

FRBR における「著作」実体としての日本の古典著作

FRBR 研究会の取り組み II

宮田洋輔 (慶應義塾大学) †, 上田修一 (慶應義塾大学)
谷口祥一 (筑波大学), 横山幸雄 (国立国会図書館),
鴫田拓哉 (つくば国際短期大学), 向當麻衣子 (慶應義塾大学)
†m@miyay.org

【抄録】『書誌レコードの機能要件』で提案された概念モデルに基づく OPAC の再構築 (FRBR 化) が現在の図書館目録の課題となっている。そこで、『日本十進分類法新訂 9 版』に掲載されているタイトルなどから選出した、日本の古典著作 159 点に対して、人手による著作同定を行った。この同定作業においては、判定結果の一貫性を高めるために、FRBR における著作の定義に依拠しつつ、必要な基準を順次設け、マニュアル化した。対象著作を、表現形、体現形を表わすように集計し、著作の成立年代、形式などでカテゴリ分けし、全体の特徴をみた。ここから、日本の図書館目録を FRBR 化する意義が認められた。

1. はじめに

「書誌レコードの機能要件 (FRBR)」は、図書館目録の概念モデルを提供した。OPAC の機能向上の方策の 1 つとして、FRBR に依拠した著作に基づく集中化とナビゲート機能の実装が試みられ、現在の図書館目録の課題となっている。こうした OPAC の FRBR 化を図るには、個々の書誌レコードに対して著作同定、すなわちいかなる著作に対応する体現形か判定することが不可欠である。これまでの図書館目録は、物理的な媒体、すなわち FRBR での体現形レベルを基盤として設計されており、著作の単位による検索または表示の機能を実現するには、これまで明示的に扱われてこなかった著作の同定が必要となる。

これまでの試みは、書誌レコードを対象とした、著作の機械的同定 (同一著作に属する候補の抽出とそれらに対する同一性の認定) によるものであった。主な手法として、機械可読式のレコードから、一部のフィールドを抽出し、著作を識別するための著作キーを構築、そのキーの照合によって、著作のクラスターを構築する方法がある。OCLC の FictionFinder や WorldCat, オーストラリア国立図書館によるプロトタイプシステムなどがその代表例である。一方、著作に関する情報の記録が限定的なわが国のレコード群を用いた試行例には、JAPAN/MARC 書誌レコードを対象とした宮田や谷口による機械的同定の実

験がある¹⁾²⁾。これらの試行結果によれば、機械的同定は有効ではあるが、大きく限界があるといわざるをえない。特に古典著作などについては機械的な著作同定は困難を極める。

そこで、発表者らは「FRBR 研究会」を組織し、既存書誌レコードに対する主に人手による著作同定作業とその作業結果の公開準備を進めている。これは、著作実体が機関に依存するものではなく、著作同定は一度行われれば繰り返し行う必要はなく、また同定結果は複数機関で共有できることに依拠している。著作同定の結果の一部については、既に報告している³⁾。本発表では、その継続研究として、より多くの著作群を対象とした大規模な同定作業から得られた日本の古典著作の特徴について報告する。

2. 古典著作の同定作業

以下では、1)同定作業の対象とした著作、2)同定作業の手順と 3)その際に用いている同定基準について記す。

2.1. 対象著作の選定

はじめに、対象とする古典著作のリストを作成した。対象とする古典著作のタイトルは、『日本十進分類法 新訂 9 版』の「本表編」に掲載されているタイトル、「関連索引」に掲載されているタイトル、国立国会図書館の「統一タイトル件名」中に登場しているタイトルを中心にリストを作成した。上記の資料に含

まれる著作は、日本の図書館目録で扱われうる、著名な著作群と考えられる。

次に、著作同定を行うためのレコードの集合を作成した。JAPAN/MARC 書誌レコード (J-BISC レコード:明治期～2009.03 収録分) に対して、主要な古典著作ごとに J-BISC の検索機能を用いて漏れがない包括的な検索を行い、ヒットした書誌レコード群を JAPAN/MARC フォーマットでダウンロードし同定作業用ファイル (候補レコード群) としている。J-BISC における検索は、例えば著作「紫式部. 源氏物語」の場合、「(タイトル: ゲンジモノガタリ or 源氏物語 or genji) or (件名: 源氏物語) or (著者名[著者標目+責任表示]: 紫式部)」という、ノイズを多く許容したものである。

前回の発表では、21 著作に対する同定作業の結果を報告した³⁾。本研究では、上記によって選定された 159 著作に関する同定結果を示す。

2.2. 著作同定の過程

書誌レコードで示された資料が特定の著作に含まれるかどうかの判定は、判定者が、書誌レコードに基づいて判定している。判定結果の例を図 1 に示した。

対象となる著作に含まれると判断した場合、その根拠となるフィールドとサブフィールドコードを図の下線部のように記録している (例では、「291\$A」と「377\$A」)。同一の著作ではないと判断した場合には、「D」として、その旨を記録している。また、著作の出現箇所を<w></w>によってタグ付けしている。

基本的に 1 つの著作に対しては、一人の判定者が著作同定を実施し、問題のあった事例について、研究会メンバー内での合意を図った。合意は次に述べる著作同定の基準に反映させ、判定結果は適宜、見直しを行い、一貫したものとなるように努めている。

また、本研究の過程で、シリーズに対する判定を行った。資料が属する特定のシリーズによって、編集方針等の観点から同じ著作であるか否かの判断を行えるように、シリー

00155002075-291A	
020\$AJPSB55002075	
100\$A19911219 1953	0JPN 1312
101\$AJPN	
102\$AJP	
251\$A 日本古典文学全集\$B 現代語訳\$D[第 11 卷]	
270\$A 東京\$B 河出書房\$D1953	
275\$A251p 図版\$B19cm	
291\$A<w>枕草子</w>\$F 池田亀鑑 // 訳	
551\$A ニホン コテン ブンガク ゼンシユ\$XNihon koten bungaku zensyuu\$B251A1\$D11\$A ケ ン タ イ コ ヲ ヤ ク	
\$XGendaigoyaku\$B251B1	
591\$A マクラノサシ\$XMakuranosousi\$B291A1	
677\$A918\$V6	
801\$AJPSBNational Diet Library,JAPAN	
\$C20031004\$GNCRT\$2jpnmarc	
905\$A918-G292	
00155006562-377A	
020\$AJPSB55006562	
100\$A19920124 1965	0JPN 1312
101\$AJPN	
102\$AJP	
251\$A 日本文学全集\$D 第 3	
270\$A 東京\$B 河出書房新社\$D1965	
275\$A475p 図版\$B20cm	
291\$A 王朝日記随筆集	
350\$A 監修者: 谷崎潤一郎等	
377\$A 内容: 土佐日記(池田弥三郎訳) 蜻蛉日記(室生犀星訳) 更級日記(井上靖訳) <w>枕草子</w> (田中澄江訳) 方丈記(佐藤春夫訳) 徒然草(佐藤春夫訳) 注釈(池田弥三郎) 年譜(阿部秋生) 解説(中村真一郎) (以下略)	

図 1 同定結果の例

ズの調査を行った。判断に際し、シリーズの現物を確認し、それぞれに収録された内容から、同一の著作と判断できるかどうかを、調査した。その結果を合議によって再度検討し、シリーズに基づく判断用のリストを作成した。

2.3. 著作の同定基準

著作同定作業を進めるに当たって、FRBR における著作の定義に依拠しつつ、必要な基準を順次設け、マニュアル化によって著作同定のための規則を構築している。

まず、下記の基本的方針を採用している。

- ①FRBR の示す基準に整合させる。「著作と著作の関連」に列記されている後継、補遺、追補、要約、改作、変形、模造の場合には、相互に異なる著作とする。他方、「表現形と表現形の関連」に列記されている縮約、改訂、翻訳、編曲の場合には、同一著作とする。
- ②できるだけ既存の目録規則、その他基準類に整合させる。たとえば、古典著作の場合には、『国書総目録』『古典籍総合目録』を尊重

し、できるだけ著作の単位を一致させる。

③原則として書誌レコードのみを見て判定する。つまり、資料現物は基本的に参照しない。ただし、Web上で参照可能なページ群（Yahooオークション・近代デジタルライブラリーなど）の参照は可とする。

④国立国会図書館による作業方針とその作業結果をできるだけ活用して判定する。請求記号（分類記号）、件名標目、著者標目の付与方針などを踏まえ、書誌レコードに示されている付与結果を尊重しつつ判定する。なお、同館における作業方針自体の変遷などにも配慮する。

基本の方針の上に、著作のタイプやカテゴリに合わせた基準を設けている。古典著作に関する基準として、前回の発表³⁾の段階では、下記のものを採用していた。

①同一著作とするもの：

「校注書」「現代語訳（部分訳を含む）」、「影印本」、「縮約（abridgement）」、「要約（condensation）」、「抜粋」

②異なる著作とするもの：

「評釈書（原テキストが記載されている場合も含む）」、「学習参考書（原テキストが記載されている場合も含む）」、「児童書」、「漫画」、「索引」、「書（原著作の本文を、ある書法をもって書き表したもの）」、「ダイジェスト（digest）」、「抄録（abstract）」、「梗概・あらすじ」

その後の同定作業によって、下記のものそれぞれ追加した。

①同一著作とするもの：

「一部」、「部分」、「抜書」、「残簡」
「部分訳」
「音訳」、「朗読」
「マイクロブック」
「抄」
「諸本対照」
「オンデマンド版」
「本文付き索引」
「選」

②異なる著作とするもの：

「抄訳」

「速記」

「絵本」

「暗誦」

「加工」

「平かな絵入り」

「絵巻物」、「滑稽本」

この判断には、著作の成立に対する、著者の貢献の大きさを基準としている。例えば、「絵本」や「書」に関しては、オリジナルの著者よりも、その後に関与した者による貢献がより大きいと考えられるため、異なる著作と判断した。また、「暗誦」と「速記」に関しては、「同一性」を追跡することが出来ないことから、同じ著作とは判断できないと考えた。

3. 著作同定結果の集計

著作同定結果の全体的傾向を、「著作の成立日付（年代）」、「著作の形式」、「著者の有無」によってカテゴリ分けし集計した。

a) 著作の成立日付（年代）

著作の年代は、NDC「9類 文学」での年代の区分やその他の情報源に基づいて、著作が成立した年代を設定した。

b) 著作の形式

著作の形式・種類は、NDCでの「文学共通区分」で設定されている文学形式を基準とした。その他のものは「その他」にまとめた。

c) 著者の有無

著作に対して特定の著者が認識されている場合は「著者有り」、いない場合は「無著者」とした。

カテゴリごとに、著作が持つ「表現形数」と「体現形数」を集計した。「表現形数」と「体現形数」は、それぞれ以下によって近似した。

表現形数：一著作の中で<w></w>でタグ付けされたタイトルの種類数

体現形数：一著作の中で同定されたレコード数と複数冊からなる場合の冊数を足した数

表 1 同定結果の集計

カテゴリ	著作数	同定対象数	表現形				体現形			
			平均	標準偏差	最大値	最小値	平均	標準偏差	最大値	最小値
成立年代										
奈良	10	689.4	4.1	6.4	22	1	64.2	84.4	247	5
平安	67	205.0	3.9	6.0	28	1	57.5	147.4	1205	2
鎌倉	54	142.7	2.9	4.8	28	1	37.1	57.5	342	4
南北朝	7	254.1	2.0	1.4	4	1	40.4	40.6	112	4
室町	18	91.2	2.9	2.1	9	1	15.9	13.8	57	2
江戸	3	461.0	8.3	8.5	18	2	67.0	40.7	92	20
形式										
詩歌	37	219.0	4.3	6.3	24	1	33.7	47.6	247	2
戯曲	1	37.0	5.0	-	5	5	20.0	-	20	20
小説. 物語	72	240.5	2.9	4.0	28	1	55.8	146.7	1205	3
評論. エッセイ. 随筆	4	530.0	14.3	15.3	28	1	128.3	70.0	188	27
日記. 書簡. 紀行	14	130.2	3.3	3.8	14	1	37.4	33.2	119	4
漢詩文. 日本漢文学	7	57.3	1.1	0.4	2	1	13.7	11.1	37	2
その他	24	138.9	3.0	3.3	13	1	35.2	38.1	143	2
著者の有無										
著者有り	68	213.9	3.8	6.0	28	1	52.5	147.1	1205	2
無著者	91	204.3	3.2	4.5	24	1	40.6	54.6	342	2
全体	159	208.4	3.5	5.2	28	1	45.7	104.5	1205	2

同一タイトルで異なる表現形の存在あるいは異なるタイトルで同一の表現形の存在や、書誌レコードの記述の単位の相違による違いなどによって、どちらも正確ではないが、大局的な傾向を示すための近似値としては十分であると考え。カテゴリごとの「著作数」と、カテゴリに属する著作ごとの同定作業対象となったレコード数の平均(「同定対象数」)、「表現形数」、「体現形数」の平均等を表 1 に示した。

この結果から、これら著作は全体として一定の「体現形」と「表現形」を持っており、日本の図書館目録を FRBR 化する意義が認められるだろう。また、著名な著作の中でも、同定される実体数にはかなりののばらつきがあることがわかった。一方、「成立年代」や「著作の形式」のカテゴリによる際立った特徴は見られなかった。

4. 考察と将来の課題

本研究では、日本の古典著作に対する同定作業の結果に基づいて、FRBR の著作実体として見た日本の古典著作の傾向を分析した。先にも述べたように、FRBR 化における著作同定の作業は一度行われれば繰り返し行う必

要はなく、また同定結果は複数機関で共有可能である。本研究で対象とした著作は、図書館目録中で扱われる日本の著名な古典著作の大部分を含んでいる。今後は、同定作業と規則の精緻化とともに、これらの成果の公開と共有が課題となってくるだろう。本研究会では、利用者向けに「著作」実体に基づいて資料を提示するページや、機械処理を想定した API の作成・公開の方法を検討している。

引用文献

- 1)宮田洋輔. "JAPAN/MARC レコードから自動構築可能な著作識別子の提案". 三田図書館・情報学会研究大会発表論文集 2008 年度. 慶應義塾大学, 2008-9-27. 三田図書館・情報学会, 2008, p. 69-72.
- 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. 2009, no. 61, p.119-151.
- 3)谷口ら. "OPAC の FRBR 化を目指した人手による著作同定作業: FRBR 研究会の取り組み". 2010 年日本図書館情報学会春季研究集会発表要綱. 同志社大学, 2010-5-29. 日本図書館情報学会, 2010, p. 75-78.