

著者名典拠データの比較分析

酒井清彦 京藤貫

国立情報学研究所開発・事業部コンテンツ課

1. 調査結果概要
2. 日本名における特徴
3. 中国名における特徴
4. 韓国名における特徴

前回のワークショップの議論の中で、各国各機関がどのように著者名典拠データを作成しているかを明らかにするため、日本名、中国名及び韓国名それぞれについて、共有している可能性が高い著者 10 名程度をピックアップして、各機関でどのように記録されているかを紹介し比較分析を行う、という結論になった。以下にその結果を紹介する。

1. 調査結果概要

ピックアップされた著者名は、以下の各 20 名ずつ計 60 名であった。

	日本名	中国名	韓国名
1	紫式部	艾青	高銀
2	清少納言	巴金	金九
3	夏目漱石	杜甫	金大中
4	芥川竜之介	費孝通	金マリア
5	谷崎潤一郎	郭沫若	金素雲
6	川端康成	胡風	金泳三
7	三島由紀夫	胡適	金宇中
8	山崎豊子	金庸	金日成
9	井上靖	李登輝	金芝河
10	井上清	李自珍	朴正熙
11	松下幸之助	魯迅	朴趾源
12	石原慎太郎	毛沢東	石潤基
13	大江健三郎	任繼愈	尹東柱
14	多湖輝	孫文	李光洙
15	湯川秀樹	王力	李文烈
16	黒柳徹子	呉昌碩	李箱
17	乙武洋匡	笑笑生	李昇基
18	梅原猛	葉昌熾	李御寧
19	瀬戸内晴美	袁牧	崔南善
20	佐藤一郎	張愛玲	許浚

このそれぞれについて、NACSIS-CAT に登録されているかどうか、検索を行った。

4 名のみ、NACSIS-CAT に登録されていなかった。その 4 名とは、「袁牧」（中国名 19）、「金マリア」（韓国名 4）、「李箱」（韓国名 16）及び「許浚」（韓国名 20）であった。

2. 日本名における特徴

NACSIS-CAT での日本名著者名典拠レコードの登録件数は、280,000 件であり、全体の 24%である。

(1) 全体一語で扱う名称

役職、官職等と姓もしくは名が一緒になって用いられ、全体で姓名形のように慣用されている名称は、役職、官職等と姓もしくは名とは切離さずに全体で一語とみなす。

（例）紫式部（日本名 1）、清少納言（日本名 2）

いずれも、式部、少納言が役職を表すため、全体で一語と見なしている。

(2) 同名異人の区別

付記事項には、生没年、職業、専門分野、世系等の種類があるが、今回の例では、生没年、職業における区別が現れている。

（例）佐藤一郎（日本名 20）

NACSIS-CAT では「佐藤一郎」のレコードは 4 レコード存在している。それぞれのレコードは出生年で区別されている。

通常は生没年で区別することになるが、生没年が不明の場合、職業、専門分野が付記事項の第一要素になる場合もある。

（例）井上清（日本名 10）

「井上清」は NACSIS-CAT に 5 レコードある。1 レコードは付記事項がなく、3 レコードは出生年での区別がされ、1 レコードは職業（「造園家」）で区別されている。

(3) SF における表記の多様性

SF は、後の検索利用の便宜を考えて、目録担当者が作成するものである。それには、ローマ字による表記形、読み方の違い、読みの区切り方の違い、本名等がある。

ローマ字による表記形では、読みをローマ字化したもののほか、AACR2 に基づいた形（例：夏目漱石（日本名 3）、芥川竜之介（日本名 4）、石原慎太郎（日本名 12））、他言語によるもの（例：紫式部（日本名 1）、清少納言（日本名 2）、芥川竜之介（日本名 4）、石原慎太郎（日本名 12））がある。

(4) 注記

標目として決定した形がどの資料のどの部分に表示されていた形であったのか、を本来記録しておくところであるが、どの部分に表示されていたか（例えば「タイトルページから」「カバーから」など）についての情報が記録されて

いるレコードはあまり多くない。

標目形の中で表現しなかった個人に関する情報、例えば具体的な職業又は身分、経歴等の記録については、かなり多くのレコードに記録されている。

3. 中国名における特徴

NACSIS-CAT での中国名著者名典拠レコードの登録件数は、40,000 件であり、全体の 3.4%である。

(1) 全体一語で扱う名称

日本名と同様に、姓名を分離することができない場合、名前全体を一語として扱う。

(例) 魯迅 (中国名 11)

(2) ローマ字表記の多様性

SF に記録されているローマ字表記が日本名及び韓国名に比較して多い。大部分は、ピンインの表記方法に多種類あることによっている。

(例) 費孝通 (中国名 4), 魯迅 (中国名 11), 毛沢東 (中国名 12), 王力 (中国名 15)

(3) 日本語読みの多様性

SF に記録されている漢字表記に対応するかな読みが、一般的な日本語読みよりは、ピンインをかなで表記したものの方が多い。別名を日本語読みした場合、検索できない可能性が高いと思われる。

(例) 毛沢東 (中国名 12), 孫文 (中国名 14)

「毛沢東」は、日本語読みとして「マオ、ツェトン」がから見よ参照フィールドに記録されているが、より一般的なヨミである「マオ、ツートン」も記録されていた方が望ましい。同様に「孫文」も「ソン、イッセン」が記録されていることが望ましい。

4. 韓国名における特徴

NACSIS-CAT での韓国名著者名典拠レコードの登録件数は、3,000 件であり、全体の 0.3%である。

(1) 統一標目形の読み

統一標目形の読みは、母国語読みがかなで標記されていなくても、母国語読みがよく知られており、その発音を基にした日本語のかな標記が参考資料等で確立している場合、当該かな標記を読みとしている。この「よく知られている」という言葉の表す範囲が目録担当者によって変動することがあり、検索用インデックスの作成に影響を及ぼしている。

(例) 高銀 (韓国名 1), 尹東柱 (韓国名 13), 李文烈 (韓国名 15)

(2) 注記

注記自体が記録されていないレコードが 17 件のうち 5 件（約 3 割）存在している。原因は、注記が必須ではなかった比較的早い時期にレコードが作成されてから以後、その著者名典拠レコード修正が加わっていないためであると考えられる。リンクされ、又は確認される機会が少ないことが影響していると推測される。

（例）金大中（韓国名 3）、朴正熙（韓国名 10）

(3) 情報量の少なさ

全体的に記録されている情報量が他の言語と比較して少ない。これには、韓国・朝鮮語資料の取扱い自体が未だ暫定的な運用であることが影響していると考えられる。

（例）朴正熙（韓国名 10）、石潤基（韓国名 12）、李光洙（韓国名 14）