

親・子・孫書誌機能の利用と活用 ——データのグループ構築による階層化

堀池 理生（立命館大学アート・リサーチセンター 研究員）
E-mail rhoriike@fc.ritsumeai.ac.jp

1. はじめに

これまでのテクニカルサポート通信でも紹介してきたように、立命館大学アート・リサーチセンター（以下 ARC）のデータベース（以下 DB）には、データの蓄積のみでなく、活用のための様々な機能がある。本稿ではその一つとして、リソース DB（古典籍 DB、近代書籍 DB、文化資源 DB など）での親・子・孫書誌機能の応用的活用方法を紹介する。

本機能を活用することで、利用者は上下巻などに別れた書籍の管理だけでなく、複数の内容が入ったノートのような個別に扱うと煩雑になるデータを階層的に整理できる。

2. 親・子・孫書誌機能の概要

2-1. 概要

ARC のリソース DB で、検索した際に通常表示されるデータを親データとする。その中で、シリーズとして出版された書籍など階層的な管理が必要なデータについて、親（統一のタイトル）、子（巻または号）、孫（号または各記事）のように分けて管理できる。【図 1】のように「A1」というタイトルの雑誌の場合、「A1」が親データとなり、その下に年または月ごとに刊行される巻を「A1-1」、「A1-2」のような子データとして扱う。「A1-1」内の記事は、「A1-1-1」、「A1-1-2」のように孫データとして登録できる。

「A1」で検索すれば親データの「A1」のみが表示され、A1 に「子データ有」などの表示が付く。「子データ有」をクリックすれば子データ「A1-1」「A1-2」が表示され、「A1-1」に「孫データ有」などの表示が付く。同様に子データからは、「孫データ有」をクリックすれば「A1-1-1」などが表示される。

具体例で確認する。「ARC 近代書籍ポータルデータベース」¹⁾にて資料名「演芸画報」で検索すると資料番号「AN00271184」の「演芸画報」の横に「子書誌有」というボタンがある【図 2】。ここをクリックすると演芸画報の各巻が一覧で表示される【図 3】。さらにその各巻、例えば「第 1 巻第 1 号」の横に「子書誌有」のボタンが

あり、これをクリックすると記事のデータが一覧で表示される【図 4】。

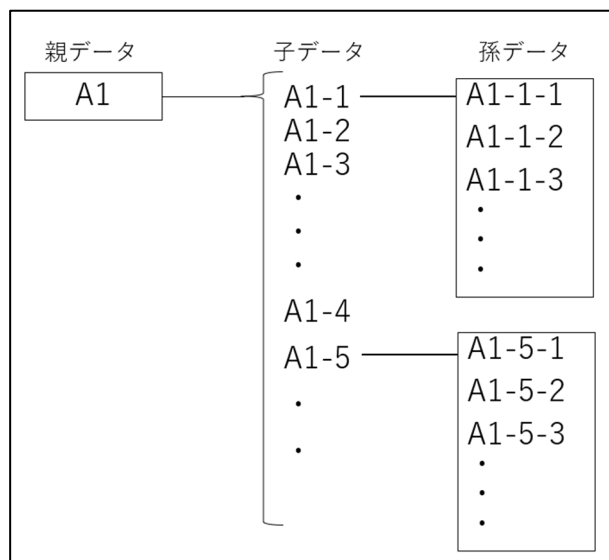


図 1 親・子・孫データの構造図

資料番号	詳細	形態	資料名
AN00271184	詳細書誌	雑誌	えんげいがほう 演芸画報

図 2 『演芸画報』横の「子書誌有」ボタン

詳細	形態	資料名
詳細書誌	雑誌	えんげいがほう 演芸画報 第1巻第1号
詳細書誌	雑誌	えんげいがほう 演芸画報 第1巻第2号
詳細書誌	雑誌	えんげいがほう 演芸画報 第1巻第3号
詳細書誌	雑誌	えんげいがほう 演芸画報 第1巻第4号
詳細書誌	雑誌	えんげいがほう 演芸画報 第1巻第5号

図 3 『演芸画報』巻号(子データ)と「子書誌有」ボタン

図 8 検索結果(親データ、子データ含む)

基本	URL	解説
Cinii No.		ARCX最新番号
DB管理No.	arc-demo	資料No arc-demo
*SET番号	SET番号 arc-demo	孫書誌 - v 子書誌有
*統一資料名	arc親子書誌機能デモ用データ	

図9 図5の「SET 番号」拡大図

検索時に「A1」で検索した場合に他の「A1」が付いたデータが埋もれてしまうのを防ぐため、検索すると基本的に親データのみ表示されるように設定されている。子データを検索したい場合は、検索時に「子書誌を含める」と書かれたチェックボックスにチェックを入れれば、「A1-1」「A1-2」のみが検索でき、「親書誌を含める」と「子書誌を含める」の両方にチェックを入れれば親データも子データもまとめて検索が可能になる。

デモデータの場合、検索欄上部の「資料名」欄で「arc 親子書誌機能デモ用データ」と検索すると親データのみが表示される【図7】。検索欄にある「子書誌を含める」のみにチェックを付けると、「arc 親子書誌機能デモ用データ 子データ」が表示され、「arc 親子書誌機能デモ用データ」の親データは表示されなくなる。検索欄に戻り、「親書誌を含める」にもチェックを付け、親子ともにチェックが付いた状態で検索すると、親データも子データも検索に表示される【図8】。

以上が親・子・孫書誌機能の概要である。この機能を利用すれば複数冊ある作品・資料の分冊管理以外にも、親データの内容を子データとして分割するといった使い方も可能である。雑誌のように雑誌そのものの書名や巻号だけでなく内容の見出しもデータとして登録すれば、資料の閲覧が容易になる。

3. 親・子・孫書誌機能の活用『歌舞伎』を例に

親・子・孫書誌機能の利用と、この機能を利用することでどのような応用ができるか、具体的に資料を用いて紹介する。

明治末期に刊行されていた雑誌『歌舞伎(第一次)』⁴⁾は統一データとしての親データとその各巻の子データが登録されている。この子データの下に孫データとして各記事のデータを登録することで、各巻に収録されている記事の見出題で検索可能にする。これにより DB 上で見出題に含まれた語から、巻をまたいだ検索や同一内容での絞り込みができるようになる。

DB 上で実際に資料を確認する。『歌舞伎(第一次)』の1巻の目次から各記事の見出題の情報がわかる。これを新規作成のデータとして「統一資料名」の欄に見出題を記入し、孫データとしての必要な項目も記入する。管理番号は巻号の管理番号の後ろに「_001」などの連番で、桁数が揃うよう調整する。

ここで「親書名」の欄にタイトルと巻号を入れ、「始頁」欄に子データの該当するページ(DB 上の画像番号)【図10】を入れておく【図11】。すると、記事から該当ページを直接閲覧することが可能になる。

これを登録すると孫データとして追加される⁵⁾。「始頁」欄に記入したことでデータ右側に「1024 画面」ボタンが付き、ここから元資料の該当画像を確認できるようになる。



図10 画像ページ

*統一資料名	常盤座句評		
*資料名(旧)	常盤座句評		
*別称			
親書名	歌舞伎 (第一次)	親巻号	001
子書名	常盤座句評	始頁	2
子書名(旧)		子書巻号	001
子書巻号(旧)		子書頁	2

図11 データ編集画面一部

データとして登録されると、データベースの機能を利用することができるようになる。例えば特定のキーワードや人名で他の巻に書かれた記事を同時に検索できる。また他の雑誌の記事索引も作成していれば合わせて横断的に検索も可能となる。

【図12】は親書名「歌舞伎(第一次)」と件名「市村座」で孫データを検索した結果である。各巻から「市村座」を含むデータが表示されており、効率的に求める情報を得ることができている。

1024画面 1280画面 小画 一覧	AN0040756X-001_011 arcID: AN0040756X-001_011 子孫両書誌を表示	詳細 書誌	市村座劇評 (歌舞伎 (第一次) 001)
1024画面 1280画面 小画 一覧	AN0040756X-017_018 arcID: AN0040756X-017_018 子孫両書誌を表示	詳細 書誌	市村座初日評 (歌舞伎 (第一次) 017)
1024画面 1280画面 小画 一覧	AN0040756X-019_013 arcID: AN0040756X-019_013 子孫両書誌を表示	詳細 書誌	市村座メチャクチャ評 (歌舞伎 (第一次) 019)
1024画面 1280画面 小画 一覧	AN0040756X-022_011 arcID: AN0040756X-022_011 子孫両書誌を表示	詳細 書誌	市村座芝居断 (歌舞伎 (第一次) 022)
1024画面 1280画面 小画 一覧	AN0040756X-027_005 arcID: AN0040756X-027_005 子孫両書誌を表示	詳細 書誌	市村座の臨狂言 (歌舞伎 (第一次) 027)
1024画面 1280画面 小画 一覧	AN0040756X-033_008 arcID: AN0040756X-033_008 子孫両書誌を表示	詳細 書誌	市村座の浦島を見て (歌舞伎 (第一次) 033)
1024画面 1280画面 小画 一覧	AN0040756X-038_008 arcID: AN0040756X-038_008 子孫両書誌を表示	詳細 書誌	市村座の伊勢物語

親・子・孫書誌機能の利用と活用
— データのグループ構築による階層化

図 12 件名「市村座」での検索結果

DB にはテキストデータを保存する機能もある。また、付属する「NDL OCR」による OCR 機能を使ったテキストデータ化も可能なため、ページ全体を一括でテキストデータ化していくこともできる。必要な記事だけ検索で絞り込んで、さらにテキストデータ化すれば、より効果的な分析が可能になると予想される。

以前のテクニカルサポート通信で紹介した「usermemo」⁶⁾や「Imagenote」⁷⁾も同様に利用できる。これらは、必要な部分だけを抽出して注釈を付与できる機能だが、これを併用して記事の細部を抽出したり、自身のキーワードを付けて蓄積することもできる【図 13】。こうして蓄積したデータを書き出し、独自の DB を構築することも可能である。そこからまた別の抽出をするなど、データを循環的に活用するエコシステムを作り、さらなる研究活用へと広げることができるのである。

図 13 usememo 画面

- 2) デモ用データ URL
https://www.dh-jac.net/db1/mbooks/search_portal.php
- 3) DB によって子・孫書誌「含まない」「含む」というラジオボタンの場合もあるが、「含む」を選択すれば同様に結果が得られる。
- 4) 『歌舞伎(第一次)』データ URL
<https://www.dh-jac.net/db1/mbooks/AN0040756X/portal/>
- 5) 作成データ URL
https://www.dh-jac.net/db1/mbooks/AN0040756X-001_004/portal/
- 6) 戸塚史織「テクニカルサポート通信 ARC オンラインデジタルツール その 1 : UserMemo : マイクロコンテンツ型アノテーション機能による索引型情報蓄積」『アート・リサーチ』vol.23(立命館大学アート・リサーチセンター, 2023)pp. 105-112.
- 7) 戸塚史織、王嘉韻「ARC オンラインデジタルツール その 2 : ImageNote : マイクロコンテンツ型 DB 構築のためのアノテーション機能」『アート・リサーチ』vol.24(立命館大学アート・リサーチセンター, 2024)pp. 39-46.

4. おわりに

ARC が持つ DB の機能として親・子・孫書誌機能の応用的利用について紹介した。大量のデータを蓄積することと、そのデータの中から必要なものを抽出するという過程において、親・子・孫の三段階でグループ化して階層的にまとめられるのは有効な機能となるだろう。

今回は近代書籍 DB を中心に説明を行ったが、文化資源 DB、古典籍 DB など他のリソース DB でも同様にこの機能を利用できる。複数巻あるものを子データ、その記事を孫データとして登録するように、多様な資料形態に対して、資料全体である親データ、分割した中身の子データ、さらに詳細な部分の孫データという構造が適切な資料もあるだろう。この親・子・孫書誌機能を活用して資料それぞれに適した構成で利用していただきたい。

ARC テクニカルサポートでは様々な機能追加と展開を検討している。DB を利用する中で追加したい機能があればご提案いただけると幸いである。

[注]

- 1) 近代書籍データベース トップページ