平成26（2014）年度活動記録と研究会の概要

金田房子（平成26年度運営担当）

※各研究会の内容は、井黒・片岡の作成した議事録をもとに作成した。

【研究組織】

研究代表者：山本 和明　（国文学研究資料館古典籍共同研究事業センター・特任教授）
研究分担者：高木 元　（千葉大学文学部・教授）
相田 満　（国文学研究資料館研究部・准教授）
北村 啓子　（国文学研究資料館古典籍共同研究事業センター・准教授）
金田 房子　（同センター・特任准教授）
岩橋 清美　（同センター・特任准教授　※10月より）
増井 ゆう子（同センター事務室・副室長）
中村 美里　（同センター事務室・データベース係長）
研究協力者：神作 研一　（国文学研究資料館研究部・教授）
井黒 佳穂子（国文学研究資料館古典籍共同研究事業センター・特任助教）
片岡 耕平　（同センター・プロジェクト研究員　※6月より）

招請講演者：上岡 英史　（芝浦工業大学工学部・教授）
武田 英明　（国立情報学研究所・教授、JaLC運営委員会・委員長）

【活動記録】

1　第1回「古典籍コード（略称）」研究会

日時　平成26年6月25日　15：00～17：30
場所　国文学研究資料館二階　第2会議室
出席者　神作・相田・北村・金田・増井・中村・井黒・片岡

【次第】
①顔合わせ
②研究の概要及び成果目標
③本年度スケジュールについて
④コード原案作成に向けての基本方針などの討議
※討議中に参考資料としての報告二件を含む。

【討議の流れ】
日本古典籍コードの国際標準化へ向けて
〇論文等に引用しやすく、海外の研究者から見ても使いやすいコードの開発。
〇利便性を考える上で目指すべき二つの方向性。
a　引用しやすさ（短い、簡略である）
b　視認性の高さ（コードからある程度の情報が引き出せる）
　　所蔵・分類・成立年代・本の形態・など
・両立させるのはかなり困難なので、どちらかに絞ることになる。
・所蔵は移動する可能性があり、分野も人によって区別が曖昧になりやすい。
これらを永続的なコードとして用いるのは難しいのではないか。
→サーバや所蔵の移動に左右されない永続的なコードを定義することが求められる。
・重要性が高く、区別が明確で、動く可能性がない項目の割り出しが必要。
〇大学及び研究機関、青空文庫などは独自のコードを付している。
→作者／書名で分類しているものが多い。それぞれ独自に行っているので、体系化は困難。新しく付け直すにはコストが掛かりすぎるため、国文研／所蔵／所蔵機関固有の番号で検討すべきか？
・国文研の画像データのファイル構造は文庫コードや請求記号を利用している。

【結論】コードに意味付けする場合、複雑なものになると予想される。
積極的に引用してもらいたいなら、特定できることに絞るべきである。

〇〈報告１・井黒〉書物に関連するコード付与のサンプル例
「既存の書籍コード（ISBN・ASIN・JAN・アマゾン・青空文庫など）について」
・Amazon、GoogleはISBNを利用（kindle版も同一）
→国（言語圏）／出版社／書物／チェック記号。
〇〈報告２・相田〉「マンガ情報の組織化」研究について
・原正一郎氏「TOPIC MAPSを利用したマンガ情報の組織化」
《FRBR「書誌レコードの機能要件」モデル》
→Work（著作）／Expression（表現形）／Manifestation（体現形）／Item（個別資料）。
《識別子について》
「『神尾葉子』には同名異人が存在しうるように、氏名に限らず、名詞は識別子として不適切である。そこでTopic MapsではIKI(Internationalized Resource Identifier)というユニークな識別子を利用する。この識別子を主題別識別子（Subject Identifier）といい、公開された主題識別子をPSI(Published Subject Identifier)という。」

２．第2回「古典籍コード研究会」

日時　平成26年9月26日　15：00～17：30
場所　国文学研究資料館二階　第2会議室
出席者　上岡・高木・神作・相田・山本・北村・金田・井黒・片岡・江草宣友（国文学研究資料館学術情報係）

【次第】①第1回研究会の確認。
②講演「デジタルアーカイブのコンテンツ及び有効利用法」（上岡英史氏）
③報告「国会図書館永続的識別子について」（金田）
④質疑応答

上岡氏の講演

第２回研究会風景

【講演の要点】
・デジタルコンテンツの有用性としてセレンディピティの補助。
　無名・人気のないコンテンツも簡単に入手・閲覧ができることから、予期せぬ発見の可能性が生まれる。
・2つのビジネスモデル（フリーミアム・ロングテール）の紹介。
　ほぼ無限の書庫　→　新研究課題発見の可能性
・ロングテール成立の条件として、
　多数のコンテンツの確保
　　多数のアクセスの確保←ニッチ市場への対応による収益確保
　　関連情報へのリンクシステムの構築
　　適切な推薦システムの構築
　　利便性の高いユーザインタフェースの構築
　　国際化
　　→古典籍データベースとの適合性。
・多数のコンテンツ確保のために、現物所有機関との合意形成の必要

＝Win-Winビジネスモデル構築の必要→フリーミアム導入の可能性。
・リンク情報を管理する方法としてDOI（Digital Object Identifier）の紹介。
特徴として、国際的な識別子
書誌情報と所在情報を一括
ドメインの変化に対応し、永続性を担保
日本では、4機関共同運営のJaLC（Japan Link Center）が管理
JaLCは、研究者業績の包括的な管理を志向

【報告の要点】
・国立国会図書館のDOI付与方針の紹介。
（例）　『伊勢物語』（請求記号　に－４４）　享保14年刊
書誌ＩＤ　000007276231
DOI　10.11501/2533076
↑　↑
国会図書館　国会図書館デジタルコレクションの永続的識別子からプレフィックス（固定）　info:ndljp/pid/　を除いたもの
URL　http://dl.ndl.go.jp/info:ndljp/pid/2533076/1
（末尾の /1 は、画像の何枚目かを示す）

【提出された論点】
DOIに関して
・JaLCとの関与は、国文研という機関単位の問題になる。
・古典籍のどのレベルに付与するか。書誌単位orファイル単位。
・完全に置き換える必要の有無。国文研の書誌コード・請求記号活用の可能性。
・QRコードの導入など、未来の展開を視野に入れる必要。
・所蔵先を組み込むと、名称変更への対応が必要になることを想定する必要。

タイトルについて
・同名状態違いの本の扱い。
→オントロジーによる関連付けの可能性。
・入力の充実・機械学習の必要⇔正確性の限界。
→修正を閲覧者に委ねる方法⇔検索システムを機関として公開する責任。
・個別タイトルの扱い。
→日本古典籍データベース「統一書名」の有効性。
⇔仏教書は、各宗派によって異なった書名を使用。統一が困難。

３．第3回「古典籍コード研究会」

日時　平成26年12月16日　15：00～17：00
場所　国文学研究資料館二階　第2会議室
出席者　武田・古瀬蔵（国文学研究資料館研究部・教授）・相田・北村・野本忠司（国文学研究資料館研究部・准教授）・金田・中村・片岡・手塚裕美（国文学研究資料館古典資料目録係）・阿部由貴子（国文学研究資料館古典籍共同研究事業センター古典籍データベース係）・布施紀子（同）・丹治麻里子（同）・神岡晶彦（同）・秋元薫（同）

【次第】①講演「研究活動におけるIDの利用の今」（武田英明氏）。
②質疑応答。
③討議。

武田氏の講演

第３回研究会風景

【講演の要点】
・DOIの概要紹介。
・DOI運営の概要紹介。

【提出された論点】
・DOIの利便性について。
→付与に要するコストと有用性とを天秤にかける必要がある。
《メリット》
・持続性　識別子が有効性を維持
・一貫性　識別子が指すものの“保証”
・RA（Registration Agency）への登録について。
→各々が定めるポリシーに適合すれば、どこにでも登録可能。
＝コンテンツの種類・地域などに特徴があるので、目的に応じて選択する。
・DOIの画像への付与について。
→付与対象となる機関リポジトリ・コンテンツの分類で言えば「研究データ」に該当する。
＝登録先としては、DataCite・JaLCが考えられる。ともに現在登録料不要。
メタデータの作成は必須であり、URIは利用者が引用する単位で付与するのが望

ましい。
→実際に付与・登録する実験を実施中⇔提供データが整備済であることが参加の条件。
・DOIと国文研独自のコードの併用について→両者の切り分けが必要。
DOIを構造的に利用することは推奨しない（各ページ単位で付与した場合、それぞれが別のオブジェクトという認識になる）。

４．平成27年1月15日（木）～1月18日（日）
韓国国立中央図書館・同デジタル図書館・清州古印刷博物館　訪問
参加者：山本・金田・井黒・片岡・増井・中村

５．第4回「古典籍コード研究会」

日時　平成27年2月13日　13:30～15:00
場所　国文学研究資料館５階　古典籍共同研究事業センター
出席者　古瀬・相田・山本・北村・金田・岩橋・増井・中村・井黒・片岡

【次第】①討議「古典籍コードの作成に向けて」。
②報告「韓国国立中央図書館・デジタル図書館・清州古印刷博物館訪問」

【討議の要点】
・古典籍コードの構造（案）

nijl/○○○○○/△△△△△△△△ ↑A　　↑B　　↑C　書誌ID

A　冒頭に、国文研の画像であることを示す nijl/　を置く。
※この部分は、DOIの場合、Prefix（DOIシステム管理機関が付与する所蔵者記号　例：10.1002 など）に置き換える。　《方針了承》

B　案１，現在国文研で付している、英数大文字四桁の所蔵者記号
（例）広島大学図書館＝HRSM　　愛媛鈴鹿文庫＝EMSK
富加町郷土資料館＝TMKC
→足りなくなる可能性・記号としてはやはり暗号か

案２，独自のコード体系　→数字としては暗号でも、わかりやすい形に通訳可能。
〈例１〉CBG植物分類コード

１桁目：１桁の数値で門　２桁目：１桁の数値で綱　３桁目：１桁の数値で亜綱
４桁目から５桁目：２桁の数値で目　６桁目から８桁目：２桁の数値で科を表す。」
８桁目は予備桁として０を入れる。→数字から翻訳できる。

〈例２〉個別医薬品コード　→英数字がそれぞれに意味を持つ。
都道府県・機関種別などを数字化
参考：ISIL(アイシル)…全世界の図書館をはじめ、博物館・美術館、文書館などの機関に付与し、識別する国際標準ID。日本の国内登録機関は国立国会図書館。

〈例３〉雑誌のJAN２段コード
一段目　F1F2F3　I1I2I3I4I5　Y1Y2　V1V2　C
①フラッグ　②新・雑誌コード　③年度　③号数
二段目　F4F5F6　C1C2C3C4　P1P2P3P4P5　C
①フラッグ　④分類コード　⑤価格
（『将来の出版バーコードのあり方―共通雑誌コード、書籍JANコードのあるべき姿を求めて―』　流通システム開発センター　平成９年５月）

Ｂについての結論：案１とする。既に用いられている記号がある以上、それを用いた方が良い。

Ｃ　◯書誌ＩＤの場合
一位に特定できるが、Ｂの示すものと重複し、Ｂが不要となる。桁数も多い。
◯請求記号の場合
各機関の特性が表れるが、漢字や平仮名が用いられている場合が多い。
◯連番の場合
既に公開済みの画像に用いられている。
作業の手間が非常に大きくなるわけではない。
桁数を最も少なくできる。

Ｃについての結論：所蔵機関ごとに連番とする。

【研究会としての結論】
以上のことから、古典籍コードを次のように策定することを提案する。

nijl/○○○○○/△△△△△△
A　B　C

Ａ：国文研の画像であることを示すnijl。
Ｂ：アルファベット大文字で５桁。
＝所蔵機関名を記号化した４文字+将来の不足に備える1文字。
Ｃ：所蔵機関ごとに6桁の連番を付与。

これで実際の作業が可能か、調査収集係と協議調整を経て、「日本語の歴史的典籍に関する国際共同ネットワーク構築計画」事業実施委員会に諮ることとする。

そのほかの活動

関連ワークショップへの参加（情報収集・出張）

〇第16回　図書館総合展　　金田・増井

「識別子ワークショップ　JaLK　CrossRef　DOI　ORCID そして…」

主催　ジャパンリンクセンター運営委員会

日時　平成26年11月16日　10：00～11：30

会場　パシフィコ横浜　第6会場

〇ジャパンリンクセンター活用の為の対話・共創の場　　井黒

（第2回）～研究データへのDOI登録～

日時　平成27年2月27日（金）15：00～18：00

会場　独立行政法人科学技術振興機構 東京本部 別館2階会議室