

「日本漢文文献画像データベース(仮)」の構築について Constructing the Image Database of *Kanbun* Document

上地 宏一^{†1}, 町 泉寿郎^{†2}
Koichi Kamichi^{†1}, Senjuro Machi^{†2}

^{†1} 大東文化大学 外国語学部, 東京都板橋区高島平 1-9-1

^{†1} Daito Bunka University, 1-9-1, Takashimadaira, Itabashi-Ku, Tokyo

^{†2} 二松学舎大学 文学部, 東京都千代田区三番町 6-16

^{†2} Nishogakusha University, 6-16, Sanbancho, Chiyoda-Ku, Tokyo

あらまし: 二松学舎大学では文部科学省私立大学戦略的研究基盤形成支援事業「近代日本の「知」の形成と漢学」プロジェクトを2015年度より開始し、日本漢学資料の画像データベースを構築中である。本論文ではデータベースに収録を予定する日本漢学資料の特色および画像データベースの概要について報告する。データベースは国際標準規格であるIIIF(International Image Interoperability Framework)にも対応し、世界的な利用が期待されるものである。

Summary: Nishogakusha University started the “Formation of ‘Knowledge’ in Modern Japan and *Kanbun* Studies” Project which is MEXT-Supported Program for the Strategic Research Foundation at Private Universities from 2015, and we constructing an image database of *Kanbun* document. We report on the features of *Kanbun* document contained in the database and the outline of the image database. The database also corresponds to the international standard IIIF, and it is expected to be used worldwide.

キーワード: 日本漢学, 漢文資料, 画像データベース, IIIF

Keywords: *Kanbun* studies, *Kanbun* document, image database, IIIF

1. はじめに

漢学者三島中洲(1831~1919)によって明治10年10月10日に創設された漢学塾をその前身とする二松学舎大学では、近年、「日本漢学」の教育研究によって建学の精神¹⁾の闡明化を図ってきた。2004~2009年度には文部科学省助成による21世紀COEプログラムに採択され、「日本漢文学研究の世界的拠点の構築」をテーマに掲げ、東アジア学術総合研究所が核となって活動を推進した。

「日本漢学」に関する我々の基本的なスタンスは、前近代日本における書記言語としての漢文とそれを通して学ぶ知識(漢学)の重要性に鑑み、漢文を通して日本の学術文化を通時的に捉え直そうとするものである。簡単に言えば、「漢文による日本研究」、「日本学のツールとしての漢文研究」ということになる。

21世紀COEプログラムでは、「①若手研究者の養成」「②日本漢文学資料の公開・データベース化」「③研究者の国際的交流と共同研究」「④漢文教育の振興」を活動の柱に掲げた。データベースに関しては、「日本漢学(日本漢文)文献目録データベース」を構築し¹⁾、日本全国の大学・公共図書館・専門図書館などの各種機関に所蔵する日本漢文文献の書誌情報(書名事項・編著者事項・出版事項等)の公開を行った。二松学舎大学COEプログラムは、同期採択中、「特色あるプログラム」4件の一つに選ばれ、成功裏に終結した。

その後継事業として、2015年度からは文部科学省私立大学戦略的研究基盤形成支援事業に採択され、「近代日本の「知」の形成と漢学」を推進している。研究組織としては「学術研究班」「教学研究班」「近代文学研究班」「東ア

ジア研究班)の4研究班に学内外の研究者を配置し、大学院文学研究科と連携しつつ、国際的な「日本漢学」の教育と研究に取り組んでいる。

立ち遅れが指摘されて久しい日本語文献データベースのなかで、日本古典籍中、過半を占めるとされる漢文文献のデータベースの構築は重要な課題である。我々のプロジェクトとして、情報発信と研究環境整備を図るべく、現在、以下の各種データベースを準備している。

1. 日本漢学画像データベース: 二松學舎大学所蔵等の「日本漢籍」(日本漢文・準漢籍)その他関連資料
2. 日本漢学本文テキストデータベース: 北京大学「儒蔵」編纂事業と協力した「日本漢籍」精華編。三島中洲・山田方谷・川田甕江の著作
3. 日本漢学人物情報データベース: 医家人名録『日本医譜』や各種儒者・漢学者人名録類のテキストデータベース

本報告では、現在構築中であり2016年度内に第1期分を公開予定である1.の日本漢学画像データベースについて紹介する。報告者の上地は、2005年から二松學舎大学21世紀COEプログラムに学外協力者として参画し、「日本漢学(日本漢文)文献目録データベース」の構築を担当した。また2007年から1年半にわたり21世紀COEプログラム研究員として二松學舎大学に所属し、それ以降も学外協力者としてデータベース構築等に携わっている。報告者の町は、前述の支援事業「近代日本の「知」の形成と漢学」の研究代表者である。

2. データベースの概要

日本漢学画像データベースは、一般的な画像データベースと同じく、日本漢学に関するさまざまな資料を写真撮影し、それをインターネット上で公開するものである。計画では、2016年度公開の第1期と2017年度以降公開の第2期にわけて段階的に構築する予定である。

2-1. 収録資料

日本漢学画像データベースに登録準備中の資料から、特徴的な資料について以下3群に分けて紹介する。

資料群1. 三島中洲法律関係文書(学校法人二松學舎所蔵): 三島中洲は明治5~10年に司法省に出仕し民事の裁判官として勤務した。漢学者である三島がフランス民法を学んだ時の資には、中洲自身による大量の書入れが残されており、近代的な法律用語・法概念の定着過程を示す貴重な資料である。その一部を表1.に示す。なお項目中のIDについてはデータベース内でそのまま利用しているものである。資料群1.はすべて第1期分として公開を予定している。

表1. 収録資料群1.の一部

ID	資料名	冊数
7003	仏蘭西法律書 民法	16
7005	国法汎論	11
7008	国立銀行条例附成規	1
7014	仏蘭西法律書 訴訟法	8
7018	仏蘭西法律書 治罪法	5
7019	民法會議筆記	1
7020	仏蘭西民法講義聴書	1
7021	民法會議筆記	24
7022	仏蘭西法律書 民法講義	1
7023	民刑法律聞見随録	1
7024	民法	1
7028	民法草案	1
7030	(民法)第6章売買の契約を廃棄及び解除する他	4
7036	前加条民法仮法則	1
7039	仏国ブスケー氏商法講義聞書	1
7046	聴訟一課心得	1
7048	物権人權の別	1
7063	仏蘭西民法講義 売買の部	3

資料群2. 芳野金陵旧蔵資料(東アジア学術総合研究所所蔵): 芳野金陵(1802~1878)は幕末明治初期の儒者として知られる。はじめ亀田鵬斎・綾瀬に入門し、江戸市中で開塾。駿河田中藩儒を経て、昌平坂学問所の教授に抜擢され、明治維新以降も殆ど唯一の遺老として大学にも出講した。近年、遺族から二松學舎に寄贈された資料には、500点余の漢籍を中心とした書籍のほか、家塾の門人録、昌平坂学問所関係文書、清国公使館員との交流の記録など、江戸後期・明治初期の漢学を知るうえで、極めて貴重な資料群である。その一部を表2.に示す。資料群2.は第2期分として公開を予定している。

表2. 収録資料群2.の一部

資料名	状態
[昌平坂学問所日記] 残本([文化10年(自3月至12月)])	42丁
[昌平坂学問所日記] 残本([文政8年(自3月至12月)])	34丁
[昌平坂学問所日記] 残本([天保10年(自4月至12月)])	43丁
[学問所御儒者]人数分限高書付(慶応3年2月15日 金陵筆)	仮綴 2丁
寄宿南楼頭取以下取締之者申上書付并臨時申達書付類 附稽古人より差出候書付類(自安政5年至文久2年)	1冊 169丁
出役稽古人向之儀諸覚	仮綴 22丁
[学問所関係芳野金陵覚書]([幕末])	横本1冊 罫紙40丁
官事 譽事([自文久2年至慶応3年])	仮綴 31丁
起請文 1枚 起請文前書 5枚([幕末])	
御番方部屋住内試一件(自天保11年至12年)	仮綴 23丁
[学問所吟味の出題短冊 易上経泰卦]([幕末])	1枚
御番方部屋住学問考試一件 下(自天保14年至嘉永5年)	仮綴 56丁
学問所書生寮書生共学問試業之儀ニ付申上 草稿([幕末])	仮綴 2丁
戊午春秋試品等 寄宿・南楼・北楼(安政5年)	仮綴 1冊 罫紙10丁
壬戌長崎考試品等簿・壬戌佐渡考試品等簿(文久2年)	仮綴 4丁
[学問所諸褒賞記録](自文政6年至天保4年)	仮綴 85丁
書生寮規定([天保14年10月])	仮綴 6丁
書生寮規定(天保14年)	仮綴 7丁
書生寮式例([幕末])	仮綴 21丁
昌平齋書生寮学規(文久慶応中)	仮綴 13丁
[御書籍拝借規則]([幕末])	仮綴 5丁
書生会業勤怠簿(慶應3年)	仮綴 12丁 末に1丁貼付
学生御引立方之儀ニ付奉申上候書付([幕末])	仮綴 4丁
[学問所課業出精記録](自天保14至嘉永6年 前畝)	残本 墨付 56丁
[通稽古人氏名]([天保13年])	1枚
服忌届福王源十郎([幕末])	1枚
[書生寮入寮願書の草稿]	2丁
奥村季五郎宛 芳野金陵書状 控([安政末~万延]9月20日付)	
[学政更張建白書](慶応4年正月)	仮綴 26丁
[学政更張建白書草稿]([慶応4年正月])	仮綴 19丁

[学政更張建白書草稿]([慶応4年頃])	残本 16丁
学政振興建白書のうち諸会業学科概略の草稿	前畝、存自 第4丁至第19丁
[学政更張建白書草稿]([慶応4年頃])	仮綴 4丁

資料群3. 加藤復齋旧蔵書(東アジア学術総合研究所蔵):二松学舎で三島中洲に明治20年代に学んだ塾生加藤信太郎(号復齋、仙台出身)の旧蔵資料。三島中洲の学統や交友を示す資料が多く含まれ、特に中洲自身による漢籍の講義を記録した『文章軌範』『唐宋八家文』などは、明治期の漢学教育の内容を伝える貴重な資料である。その一部を表3.に示す。資料群3.は第2期分として公開を予定している。

表3. 収録資料群3.の一部

資料名	状態
(三島中洲)史記管晏列伝(段解)	写 1冊
(三島中洲)老子講義	写 1冊
(三島中洲)朱王両学異同・大学説	写 1冊
(三島中洲)尚書講義筆記	写 2冊
川田剛 甕江文鈔	写 3冊
齋藤拙堂 梅溪游記	写 1冊
阪谷素 評釈孟子荀卿列伝	写 1冊
佐藤一斎 論語欄外書	写 1冊 存上
佐藤一斎 大学摘説	写 1冊
佐藤一斎 孟子欄外書	写 2冊
佐藤一斎 伝習録欄外書	写 3冊
三島中洲 日本政記論文(段解)	写 1冊
三島中洲 瓊浦筆談	写 1冊
三島中洲 尚書古今文九家系表	写 1冊
三島中洲 (中洲文稿)	写 1冊
森田節齋 竹窓夏課	写 1冊
安井衡注 大学	写 1冊
山田方谷 孟子養気章或問図解	写 1冊
山田方谷 古本大学講義	写 1冊
(三島中洲講義) 点註唐宋八家文読本 明治刊	16冊中、存13冊 欠巻1・13・14・19・20
(三島中洲講義) 正文章軌範 寛政刊	6冊

2-2. 収録データ形態

データベースに収録される画像データは、資料をカメラ撮影し、デジタル化したものである。もともとの画像データは400dpiで5760ドット×3840ドットフルカラーの無圧縮

TIFF形式であるが、それをそのまま公開するにはWebサーバ資源(特にストレージ容量とネットワーク転送量)に限界があるため、品質を20%に落としたJPEG形式に変換し、同時に二松学舎大学のロゴをウォーターマークとして埋め込んでいる。書き込みなどの文字が判別できる限界まで品質を落とし、ファイルサイズを小さくしたため、紙の色合いや色シミ等は劣化してやや再現できていない部分がある。

資料によっては、注書きの付箋などを貼り付けているものがあり、付箋にさらに小型の付箋を貼り付けているケースも見られた(図1)。これらの付属資料の扱いについて検討したところ、大事な情報であり収録対象としたいが(実際には全て写真撮影を行っている)、データベースの構造が複雑になるため、現時点では取り外した状態の写真のみをデータとして収録することと判断し、将来的な検討課題として残すこととした。

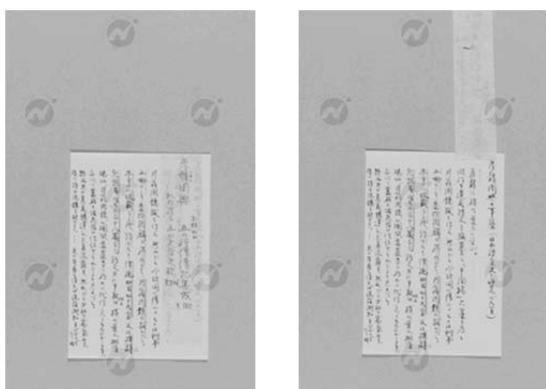


図1 (上)付箋がついた状態、(下左)付箋を取り外したものの、(下右)付箋についた小付箋をめくったところ

2-3. 画像データのライセンス

データベースを公開するにあたっては、データのライセンス決定が非常に重要であり、利用者側の立場にとってみればオープンライセンスであることが重要である。その検討にあたっては、他の学術機関が公開している同種データベースのそれも参考に検討を行い、クリエイティブコモンズ^[2]のCC BY-SA 4.0ライセンスまたはCC BY-NC 4.0ライセンスのいずれかとする方向で最終調整を行っている。

2-4. データベース・フロントエンド

データベースはインターネット上で公開するものであるが、そのフロントエンドは基本的な画像閲覧機能をJavaScriptで実装している。ページ・巻の移動、画像表示倍率の変更(拡大・縮小)、回転などをマウスで操作してデータを閲覧する(図2)。



図2 画像データ閲覧画面



図3 収録資料リスト表示画面

データベースはトップページにアクセスすると簡単な紹介と収録資料のリストへのリンクが示される。収録資料リストについては資料名に対する簡易検索機能が用意されていて、リスト表示の対象を絞り込むことができる(図3)。表示されたリストの項目をクリックすると画像閲覧ページに飛ぶ形態となっている。表示される画像は単純な JPEG 形式の画像ファイルであり、再利用をできなくするような仕掛けは特に施していない。

3. 第1期公開

先述のように、2016年度内に資料群1を含めた第1期分の公開を予定している。資料数は333点、写真カット総数は48,812枚であり、データ量は25.4GBである。すでにデータベース・フロントエンドの構築は完了し、データベースへの画像ファイル収録も完了している。データベース紹介の文面準備など細かい作業を残し、公開に向けた準備は順調に進んでいる。

データベースの設置は一般の Web ホスティングサービスを契約し、そこにファイル群を設置する形態としている。自由度やスケーラビリティの面でやや非力ではあるが、Web サーバを管理できる予算と人員がないための判断である。

4. IIIF への対応

当初の計画では前章までの内容を基本的なデータベース構築としていたが、構築の過程で国際的なデジタルアーカイブの規格である IIIF (International Image Interoperability Framework)^[3]に対応させたらどうかという提案があった。もともと21世紀 COE プログラム事業として「日本漢学(日本漢文)文献目録データベース」を公開するにあたって、いかに公開したデータを有意義に利用してもらうかは大きな課題であった。今回公開を予定している画像データベースにしても、データを広く活用してもらう手段として IIIF への対応は非常に魅力的である。そこで第1期公開に合わせてこの IIIF への対応についても計画に含めることとなった。

現状で IIIF Image API に対応するためには、当初想定していたデータベースのように Web ホスティングサービスに画像ファイルを置くだけでは実現できない。そこでリモート管理が可能な VPS サービスを利用して仮想サーバ上に IIIF 対応の Web サービスを本体の画像データベースとは別に構築した。構築にあたっては先行して国内のデータベースの IIIF 対応とその普及に取り組んでいる研究者のブログ記事^[4]を大いに参考とした。IIIF Image API の処理については、速度にやや難があると評されているものの、導入が容易な Loris IIIF Image Server^[5]を利用している。IIIF Presentation API の manifest データ作成については Perl スクリプトを書いて自動生成した。

データベース・フロントエンドの資料リスト項目に IIIF のロゴ画像を配置し、この画像を IIIF 対応のビューア (Mirador^[6]など) に対してドラッグ・アンド・ドロップすることで画像データを閲覧可能としている。大学で用意したデータベース・フロントエンドでは、同時に1枚の画像しか表示できないが、Mirador では日本漢学画像データベースに限らず、複数の画像を同時に比較するなど、より高度な閲覧が可能で、例えば今回第1期分として公開する資料の中にも同名で旧所蔵の異なる資料が存在するが、これらを1画面内で比較するといった行為が容易に可能である(図4)。また国際標準に準拠して公開することで、国内に限らずより多くの利用者に対して本データベースをアピールできるものと期待している。

5. 今後の計画

まずはデータベース第1期公開分の完成が第一の目標であるが、その後については公開後の反応を見てシステムの改良を行い、また第2期公開分の準備を行うものである。

それ以外には、2017年度の計画として、すでに公開中の「日本漢学(日本漢文)文献目録データベース」とのデータ項目のリンク付けを行い、目録データベースの検索結果から画像データベースの当該項目にたどれるように2つのデータベースの一体化を計画している。画像データベースは検索性に劣るが、すべての資料に対してテキ

スト(全文)データを用意するほどの予算がないため、少なくとも表題などのメタ情報を画像データに付与できる機構の検討も行う。

最終的には、「近代日本の「知」の形成と漢学」プロジェクトとして計画中のテキスト(全文)データベースとの融合も考えている。公開中の文献目録データベース、本報告の画像データベース、計画中のテキスト(全文)データベースというこれら3者を融合させた「日本漢学総合データベース」の構築がプロジェクト(あるいはその後)の最終ゴールとなる。

付記

本報告は平成28年度～平成32年度(予定)科学研究費補助金(基盤研究(B))、研究代表者:二階堂善弘、課題番号:16H03351の成果の一部である。

注

- 1) 「東洋固有の精神による人格の陶冶」「己を修め人

を治め以て一世に有用の人物を養成する」

文献

- [1] 日本漢文文献目録データベース, <http://www.nishogakusha-kanbun.net/database/>, 2017年1月27日閲覧
- [2] クリエイティブ・コモンズ・ライセンスとは, <https://creativecommons.jp/licenses/>, 2017年1月27日閲覧
- [3] IIIF (International Image Interoperability Framework), <http://iiif.io/>, 2017年1月27日閲覧
- [4] digitalnagasaki のブログ, <http://digitalnagasaki.hatenablog.com/>, 2017年1月27日閲覧
- [5] Loris IIIF Image Server, <https://github.com/loris-imageserver/loris>, 2017年1月27日閲覧
- [6] Mirador, <http://projectmirador.org/>, 2017年1月27日閲覧



図4 Mirador にて『秋苑日涉』12巻の「文化4年江戸北澤伊八等刊本」と「安政四年数庫堂修本」の比較