

日本における漢籍デジタルアーカイブの現況

木村麻衣子[†]

概要：多くの機関がさまざまな文化資源を対象としたデジタルアーカイブを構築している中、特定の資料種別、例えば漢籍に限った横断検索ができれば便利である。しかし、国内に漢籍を対象とするデジタルアーカイブがどの程度あり、またそれぞれのデジタルアーカイブがどのようなデータを作成しているのか等、詳細は明らかではない。発表者は2017年9月から11月にかけて、コンテンツに漢籍画像を含むデジタルアーカイブを運営していると見られる日本国内の計113機関に対し、漢籍デジタルアーカイブの運営状況に関する質問紙調査および訪問調査を実施した。68機関から、77件の漢籍デジタルアーカイブに関する回答を得たので、その結果を報告する。

Current status of digital libraries for Chinese classic books in Japan

MAIKO KIMURA[†]

1. はじめに

これまでに行われてきたデジタルアーカイブ(以下、DA)の実態調査では、資料種別の設定が大雑把なものが多く[a]、古典籍や漢籍を対象とするDAの数は明らかではない。一方で利用者側には、対象とする資料種別を絞り込んで検索したいニーズがあると考えられる。発表者は2017年5月～7月にかけて、全国3,980機関の図書館を中心とする文化学術機関を対象に郵送法による質問紙調査(以下、一次調査)を実施し、全国に古典籍を対象とするDAの数と、その対象資料種別を明らかにした[1]。その後、古典籍を対象とするDAの中でも、特に漢籍を対象に含むDA(以下、漢籍DA)について詳細な運営状況を把握し、最終的に漢籍のみを対象とした横断検索を実現するための問題点を整理することを目的に、二次調査を実施した。本研究において「漢籍DA」とは、文化・学術機関等において漢籍の画像データを蓄積し、検索および/または閲覧可能としたシステム(漢籍画像とそれ以外の資料の画像・書誌データが混在している場合も含む)を指す。例えば、図書館のオンライン目録データベースや、機関リポジトリ等に、一般書の書誌データや画像データと漢籍画像データが混在している場合、それらのオンライン目録データベースや機関リポジトリも漢籍DAに含まれる。検索機能がなくリスト形式で画像を公開しているものや、Webサイトで公開されていないものも含む。

2. 調査方法と対象

一次調査において漢籍を古典籍DAの収録対象資料に含

むと回答した機関が81機関存在した。これらの機関と、一次調査において漢籍を所蔵し、かつ古典籍DA[b]を複数構築していると回答した機関の古典籍DAのうち、一般に公開されているものを閲覧し、明らかに漢籍を収録対象資料としていないものを除外した結果、72機関が実際に漢籍DAを運営中であると確認された。さらに、①発表者が2016年に実施した公共図書館のウェブサイト調査[2]により、漢籍DAを運営していることが判明しており、上記72機関と重複しない6機関、②「全国古典籍等画像公開データベース一覧」[3]に掲載されており、公開画像の中に漢籍が含まれ、かつ上記①と重複しない30機関、③その他、上記①②の調査の過程で漢籍DAを運営中であることが判明した5機関を合わせた計113機関を、二次調査の対象とした。なお、一次調査において今後の調査協力不可との回答があった機関は調査対象に含めていない。

訪問調査は、11機関を対象とし、2017年9月4日から2017年11月30日の間に実施した。質問紙調査は、103機関を対象とした。質問紙を郵送し、封書、メールのいずれかで回答いただく方法とした。調査期間は、2017年10月30日から2018年1月31日としたが、調査結果には、2018年2月23日までの回答を含む。訪問調査は当初、質問紙調査の調査項目の妥当性を確認するための予備調査として開始したが、質問紙調査開始後も、質問紙では聞き取れない各機関特有の事情等について把握する目的で、一部の機関に対し個別に依頼して実施した。1機関には、訪問調査と質問紙調査の両方にご協力いただいた。

訪問調査での質問項目はDAの構築状況などによって機

[†] 慶應義塾大学文学部
School of Library and Information Science, Keio University

a) 例えば、国立国会図書館は2009年に、全国の文化・学術機関4,302機関を対象とした「文化・学術機関におけるDA等の運営に関する調査研究」による質問紙調査を実施した(国立国会図書館「文化・学術機関におけるデジタルアーカイブ等の運営に関する調査研究」『カレントアウェアネスポータル』2011.3.31. http://current.ndl.go.jp/FY2009_research, (参照2018-07-02)。この調査においては、553機関が計815件のDAを運

営しており、このうち30.4%のDAが「古文書・和漢本・稀本」を収録対象としていることが分かっているが、「古文書・和漢本・稀本」の内訳は不明である。

b) 一次調査における「古典籍」には、和古書、漢籍、古地図、古文書、絵画、日記など、日本ならばおおむね江戸期以前、中国ならばおおむね清朝以前に書写・刊行されたあらゆる資料を含んでいる。典籍でない資料をも対象に含めたのは、一次調査の段階で対象資料種別を限定しすぎず、網羅的に回答を得たいと考えたためである。

関ごとに異なるものとしたが、ほぼすべての機関に、古典籍（または漢籍）所蔵点数、DA 構築の経緯、古典籍（または漢籍）DA への入力項目、メタデータの出典、メタデータを作成したのは誰か、メタデータ作成の際の基準類、外部のデータベースとの連携状況、デジタル画像の作成方法などを尋ねた。これらの項目と、質問紙調査の調査項目の設定に当たっては、デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会が 2017 年 4 月に発行した『デジタルアーカイブの構築・共有・活用ガイドライン』[4]（以下、『ガイドライン』）を参考とし、さらに漢籍 DA の横断検索を実現する上で問題となりそうな事項を発表者の判断で複数含めた。

3. 結果(1):概要

3.1 機関種別有効回答率

質問紙調査に対し、謝絶を含め何らかのお返事があった割合は 72.8%であった。この中には、一次調査への回答の修正（DA に漢籍は含まれていなかった、等）も含まれる。漢籍 DA を運営中または準備中[c]であって、白紙以外の有効な回答をお寄せくださったのは 58 機関であり、質問紙調査の有効回答率は 56.3%であった。訪問調査対象の 11 機関を含めると、二次調査全体で 68 機関から 77 件の漢籍 DA に関する有効なご回答をいただいたことになり、有効回答率は 60.2%であった。質問紙調査と訪問調査を合わせた、機関種別有効回答率を示す(表 1)。表 1 の「その他」には、大学研究室、研究所、国立国会図書館等が含まれる。

3.2 所蔵/収録漢籍のタイトル数/冊数

所蔵古典籍および/または漢籍のタイトル数および/または冊数を尋ねたところ、いずれも不明、または無回答であったのは 68 機関中 17 機関であった。所蔵古典籍のタイトル数および冊数がいずれも不明または無回答であったのは 24 機関、所蔵漢籍のタイトル数および冊数がいずれも不明または無回答であったのは 29 機関であった。所蔵漢籍の冊数のみ判明しているのが 23 機関、タイトル数のみ判明しているのが 29 機関であ

表 1 機関種別有効回答率

	公共図書館	大学図書館	博物館・美術館	文書館	その他	合計
対象機関数	22	52	27	3	9	113
有効回答機関数	20	28	11	3	6	68
有効回答率 (%)	90.9%	53.8%	40.7%	100.0%	66.7%	60.2%

c) 構築準備中なのは訪問調査を実施した 1 機関 1 件の漢籍 DA のみである。二次調査は原則として、漢籍 DA を運営中の機関を対象としたが、構築準備に当たり直面している困難等を聞き取るため、あえて 1 機関のみ準

た。1 タイトル当たりの冊数は、タイトルによって異なり、本来一律に換算できるものではないが、所蔵漢籍のタイトル数と冊数をともに回答した 13 機関の所蔵タイトル数と冊数から、1 タイトル当たりの平均冊数を算出すると 3.4 冊であった。そこで、所蔵漢籍のタイトル数と冊数のうちいずれかが判明する機関 39 機関については、不明のタイトル数または冊数は仮に 1 タイトル当たり 3.4 冊として所蔵漢籍のタイトル数および冊数を算出し階層別に示した(図 1)。

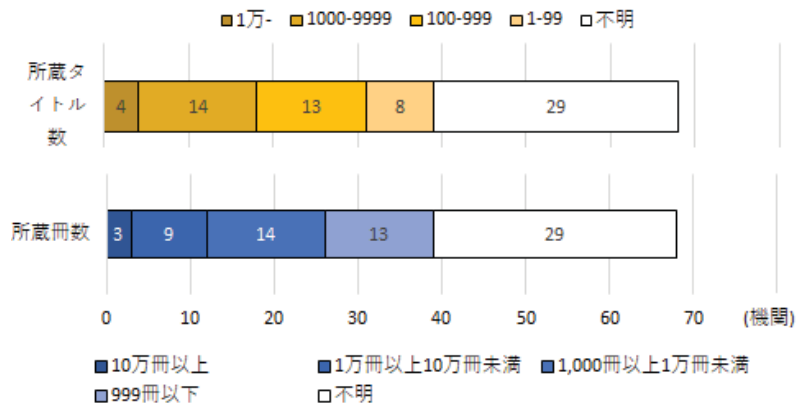


図 1 所蔵漢籍冊数/タイトル数別機関数

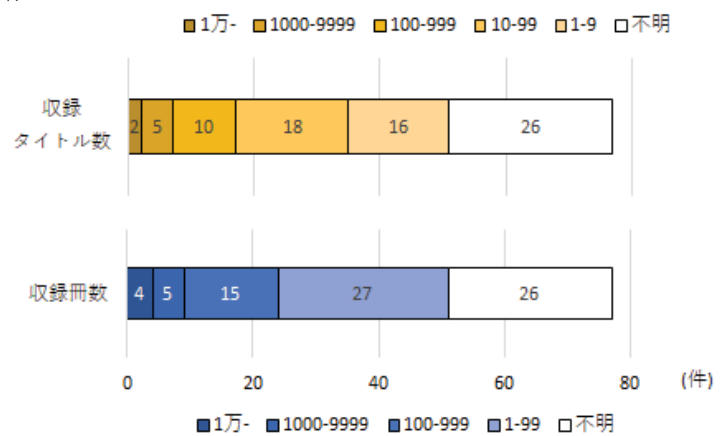


図 2 収録漢籍冊数/タイトル数別 DA 件数

さらに、77 件の漢籍 DA について、収録漢籍のタイトル数および/または冊数を尋ねたところ、いずれも判明しない DA が 26 件、タイトル数のみ判明する DA が 25 件、冊数のみ判明する DA が 3 件であった。タイトル数と冊数のうちいずれかが判明する 51 件の DA について、不明のタイトル数または冊数は仮に 1 タイトル当たり 3.4 冊として収録漢籍

備中の機関を訪問した。

タイトル数および冊数を算出し階層別に示した(図 2)。

3.3 課題と感じていること

68 機関に対し、自機関の漢籍 DA に関して課題と感じていることを複数回答で尋ねた結果を示す(表 2)。「漢籍を扱うための知識・経験が不十分」と回答したのは 44 機関(64.7%)であった。図書館のみ(大学図書館と公共図書館を合わせた 48 機関中)で見ると、25 機関(52.1%)が「漢籍を扱うための知識・経験が不十分」と回答している。

表 2 課題と感じていること(複数回答)

課題と感じていること	機関数 (n=68)
漢籍を扱うための知識・経験が不十分	44
予算不足	33
DA構築や運営のための知識・経験が不十分	31
その他	12
課題と感じていることはない	4
無回答	6

4. 結果(2):画像データ関連

4.1 画像撮影の範囲

本調査では、漢籍の全文でなく、一部のみをデジタル化し収録している場合にも、漢籍 DA として扱った。このため、77 件の漢籍 DA について、DA に収録している漢籍画像の撮影範囲を複数回答で尋ねた結果を示す(表 3)。複数回答としたのは、資料によって撮影範囲が異なる可能性があるためである。複数回答したのは表 3 のうち 8 件であり、いずれも「全ページ」と「一部のページ」の両方を回答していた。

表 3 漢籍画像撮影の範囲(複数回答)

画像撮影範囲	DA件数 (n=77)
全ページ	61
一部のページ	22
1ページのうちの一部分	1
その他	1

4.2 漢籍画像の公開範囲

本調査では、来館者のみ公開、あるいは機関内職員のみが利用するためのシステムであっても、漢籍が収録されていれば漢籍 DA として扱った。そのため、運営中の漢籍 DA76 件について、漢籍画像の公開範囲を複数回答で尋ね

た結果を示す(表 4)。

少なくとも漢籍の一部をウェブ一般公開している漢籍 DA は 67 件(88.2%)であり、このうち 57 件は全ての漢籍をウェブ一般公開していた。条件付きウェブ公開は、公開

表 4 漢籍画像の公開範囲(複数回答)

公開範囲	DA件数 (n=76)
ウェブ一般公開	67
(うち、単回答)	(57)
条件付きウェブ公開	4
機関内一般公開	9
機関内限定公開	7

にあたってパスワードをつけ、特定の機関や個人のみ閲覧できるようにするケースが見られた。機関内限定公開のうち 5 件は、機関内職員のみ閲覧可としていた。

4.3 画像のフォーマット

77 件の漢籍 DA が採用している公開用および保存用の画像フォーマットを複数回答で尋ねた結果を表 5、表 6 にそれぞれ示す。表 5 の「その他」8 件は特定のビューワに依存するフォーマットであり、このうち 3 件は iPallet/Lime というビューワに対応するものであった。表 6 の「その他」4 件のうち 3 件は、RAW ファイルとの回答であった。

『ガイドライン』において、保存用画像は“非圧縮か可逆圧縮のフォーマットで作成することが望ましい”とされているが、25 件の DA は保存画像として JPG のみを単回答し、JPG は非可逆圧縮方式であることから、望ましいレベルを満たしていない。

表 5 公開用画像フォーマット(複数回答)

公開用画像フォーマット	DA件数 n=77
JPG	61
PDF	15
TIFF	4
JPG2000	3
Flash	2
IIIF	2
その他	8
無回答/不明	2

表 6 保存用画像フォーマット（複数回答）

保存用画像フォーマット	圧縮方式	DA件数 n=77	うち、単 回答数
JPG	非可逆圧縮	45	25
TIFF	非圧縮	43	23
JPEG2000	可逆/非可逆（選択可）	4	0
その他	—	4	1
無回答/不明	—	8	8

4.4 保存用画像の解像度

『ガイドライン』では、解像度について、“原資料・作品の大きさに対して 400dpi~600dpi でデジタル化”することが望ましいとしている。77 件の漢籍 DA の保存画像の解像度は表 7 の通りである。ピクセル数を回答していた場合は、A4 サイズに対する解像度を仮に算出した。KB、MB など画像ファイルの大きさを示す回答は無回答として扱った。複数回答があった場合は、高いほうの解像度を集計した。400dpi 以上の漢籍 DA は 30 件 (39.0%) であった。

『ガイドライン』では画像の質に関し、解像度の望ましい基準を示すほかにも、保存用画像が“カラーチャートで色校され、メジャーを参照できるようにする”ことを必要としている。漢籍 DA の中にはマイクロフィルムをデジタル化した画像が含まれる場合があることや、発表者が過去に指導を受けたことのある複数の漢籍利用者にとって、画像の中で色やサイズが示されることは必ずしも重要でなかった（なぜなら、本文が読めればよいから、または画像がどんなに良くても現物を見る必要があるため）ことから、今回は調査対象としなかった。

表 7 保存用画像の解像度

保存画像解像度	DA件数
1000dpi 以上	2
400dpi 以上999dpi未満	28
300dpi 以上400dpi未満	16
200dpi 以下	7
無回答/不明	24
計	77

4.5 画像の撮影者種別

77 件の漢籍 DA について、画像の撮影者種別を複数回答で尋ねた結果を表 8 に示す。画像撮影を業者に外注してい

る DA が最も多かった (72.7%)。業者への外注のみを単回答した DA は 40 件 (51.9%) であった。

表 8 画像の撮影者種別（複数回答）

画像の撮影者	DA件数 (n=77)	(うち、 単回答)
機関内専任職員	14	(6)
機関内その他の職員	15	(9)
業者に外注	56	(40)
マイクロフィルムをデジタル化	12	(2)
その他	3	(2)

4.6 画像利用条件の明示

運営中の 76 件の漢籍 DA のうち、公開範囲を機関内限定としている 5 件を除く 71 件の DA について、画像の利用条件の明示をどのように行っているかを複数回答で尋ねた。

『ガイドライン』では、“著作権保護期間が満了しているなど著作権による制限がないデータは、パブリック・ドメイン・マーク (PDM) などを利用してそのことを明示することがよい”としている。漢籍の著作権は、成立年代から既に保護期間が満了しているため、米国 Creative Commons が制定した PDM[d]を表示することが可能であると考えられる。しかし、PDM または著作権・データの所有権の保有者がこれらの権利を放棄していることを示す CC0[e]を表示していると回答した漢籍 DA は 0 件であった。クリエイティブ・コモンズ・ライセンス(以下、CC ライセンス)を付与していると回答した漢籍 DA は 4 件 (5.6%) であった。そのほか、独自の方法で表示していると回答した DA が 44 件 (62.0%)、利用条件の表示を行っていないと回答した DA が 26 件 (36.6%) であった。

5. 結果(3):書誌データ関連

5.1 書誌データ作成者種別

77 件の漢籍 DA について、書誌データ作成者種別を複数回答で尋ねた結果を表 9 に示す。「その他/不明/無回答」の中には、機関外の研究者やボランティアが作成しているケース 3 件が含まれる。本調査項目は、もともと冊子体目録やカード目録が作成されていた場合でも、それらの紙媒体の目録を電子的な書誌データに変換した際のデータ作成者を尋ねる意図であったが、冊子体目録等の作成者に関する回答も見られた。機関内の大学院生等をアルバイトとして雇用し書誌作成に従事させた場合は、「機関内その他の職員」

d) “パブリック・ドメイン・マーク 1.0”,
<https://creativecommons.org/publicdomain/mark/1.0/deed.ja>, (参照 2018-07-20).
e) 『ガイドライン』では、PDM と CC0 を合わせて“パブリック・ドメイン・ツール”と称しており、質問紙調査においても、“パブリック・ドメイン・ツール”という用語を用いた。なお CC0 は、“CC0 について”
<https://creativecommons.jp/sciencecommons/aboutcc0/>, (参照 2018-07-20) によ

れば、“関連する著作権・データの所有権をすべて保有している自分の作品に対してか、または必要な権利をすべて保有している第三者の作品に対してのみ、適用することが可能”なものであって、漢籍に対し適用することは本来、適当ではないと考えられる。

として扱った。

表 9 漢籍書誌データ作成者種別 (複数回答)

書誌データ作成者	DA件数 (n=77)	(うち、単 回答数)
機関内専任職員	36	(26)
機関内その他の職員	35	(26)
業者に外注	7	(4)
その他/不明/無回答	7	(5)
書誌データなし	4	(4)

5.2 準拠目録規則等

77件の漢籍DAについて、漢籍書誌データを作成する際に準拠している目録規則等を複数回答で尋ねた結果を表10に示す。表10のうち、「NACSIS-CAT マニュアル」とは、NACSIS-CAT『和漢古書に関する取扱い及び解説』[f]、『コーディングマニュアル(和漢古書に関する抜粋集)』[g]のうちどちらか又は両方を指す。「その他」12件のうち5件が、京都大学人文科学研究所附属漢字情報研究センター編『漢籍目録:カードのとりかた』[h]を挙げた。別の1件は、日本目録規則1987年版改訂2版を挙げた。「日本目録規則1987年版改訂3版」と回答した17件のうち、単回答したのは5件のみであり、5件は独自基準との併用、7件はNACSIS-CAT マニュアルとの併用であった。質問紙調査の際、独自マニュアルを使用している機関に対して、さらなる調査のための独自マニュアル提供可否を伺った。その設問への回答内容から、明確にマニュアルが存在しないことが確認されたのは、不明/無回答/特になしと回答した17件のDAのうち、4件であった。

表 10 準拠目録規則 (複数回答)

準拠目録規則	DA件数 (n=77)
自機関独自の整理基準、マニュアル等	29
日本目録規則1987年版改訂3版	17
NACSIS-CATマニュアル	15
その他	12
不明/無回答/特になし	17
書誌データなし	4

5.3 メタデータスキーマ

77件の漢籍DAが採用している入力用メタデータスキーマ

f) 国立情報学研究所. “和漢古書に関する取扱い及び解説”.
http://catdoc.nii.ac.jp/pdf/wakan_toriatsukai.pdf, (参照2018-07-18).
g) 国立情報学研究所. “コーディングマニュアル(和漢古書に関する抜粋集)”.
http://catdoc.nii.ac.jp/pdf/wakan_cm.pdf, (参照2017-07-

マと、運営中の76件の漢籍DAが採用している出力用メタデータスキーマを、それぞれ複数回答で尋ねた結果を表11、表12に示す。

入力用メタデータスキーマを複数回答したDAは1件(2種類)のみであった。「その他」3件は、MARC21, TRC-MARC, DC-NDLであった。65件(84.4%)の漢籍DAが独自フォーマットを採用していた。出力用メタデータスキーマは、表形式で出力可と回答したDAが50件(65.8%)であった。

表 11 入力用メタデータスキーマ(複数回答)

メタデータスキーマ(入力)	DA件数 n=77
独自	65
Dublin Core	3
MODS	1
その他	3
不明/無回答	2
メタデータ無し	4

表 12 出力用メタデータスキーマ (複数回答)

メタデータスキーマ(出力)	DA件数 n=76
独自フォーマットを表形式で出力	50
DC-NDL	3
Junii2	1
その他	8
出力不可/提供せず/不明/無回答	14
メタデータ無し	4

5.4 画像に対する書誌データの作成単位

複数の漢籍DAの横断検索を想定した場合、タイトル単位の書誌データと物理単位の書誌データが混在することで、横断検索がしづらくなる恐れがある。そのため、運営中の76件の漢籍DAに対して、画像データに対する書誌データの作成単位を尋ねたところ、1タイトルにつき1件の書誌データを与えている漢籍DAが43件(56.6%)、物理的な1冊に対し1件の書誌データを与えている漢籍DAが25件(32.9%)、両方与えていると回答したDAが4件、書誌データなしが4件であった。

5.5 ユニコード対応

漢籍の書誌データは、漢籍目録等では繁体字にて表示さ

18).
h) 京都大学人文科学研究所附属漢字情報研究センター編. 漢籍目録:カードのとりかた:京都大学人文科学研究所漢籍目録カード作成要領. 朋友書店, 2005, 139p.

れることが一般的であり、特に国際的な書誌データの相互運用を視野に入れると、ユニコード対応は必須である。運営中の76件の漢籍DAに対しユニコードに対応しているかどうかを尋ねた結果、対応していると回答したDAが39件(51.3%)、対応なしが29件(38.2%)、対応検討中1件、不明/無回答3件、書誌データなしが4件であった。

5.6 漢籍書誌データのみ抽出可否

本研究が想定している、対象を漢籍のみに限った横断検索を行うためには、さまざまな資料のメタデータが混在する各DAの中から、漢籍書誌データのみを識別可能でなければならない。運営中の76件の漢籍DAに対し、収録されている全資料のメタデータの中から、漢籍の書誌データのみを出力/抽出することが可能かどうかを尋ねた結果、抽出不可と回答したDAが36件(47.4%)であった。抽出可の漢籍DAは合わせて33件(43.4%)であり、内訳は、漢籍のみを対象としたDAのため抽出可が4件、請求記号等を条件として抽出可が13件、収録件数が少ため目視で抽出可が16件であった。その他2件、無回答1件、書誌データなしが4件であった。

6. 考察

6.1 画像データに関する課題

今回の調査の範囲では、少なくとも漢籍の一部をウェブ一般公開している漢籍DAが9割弱に上り、一部のページのみを公開するものよりも、全ページを公開している漢籍DAのほうが多かった。

画像データの公開用フォーマットには、JPGやPDFなどの一般的なフォーマットが多く用いられていた。Flashは、閲覧のためのソフトウェアAdobe Flash Playerの開発および提供が2020年末に終了することが予定されているため[5]、早期の対応が必要である。IIIFは、画像データの相互運用性を保証するための規格として急速に広がりつつあるとされるフォーマットであるが[6]、今回の調査の範囲では採用率はまだ低かった。他方、3割強の漢籍DAが、保存用画像フォーマットとして非可逆圧縮方式のJPGのみを採用しており改善が望まれる。

撮影を業者に外注している漢籍DAは7割以上に上り、全面的に外注しているものも半数以上であった。画像の質に関しては、『ガイドライン』の示す解像度の最低基準である400dpiを満たす漢籍DAは4割弱であった。画像の利用目的によっては、今後再撮影が必要になる可能性がある。単純なことであるが、解像度が高くてもピントによるボケがあっては意味がないので[i]、自前で撮影する場合はもちろん、外注の際にも粗悪な画像の納品がなされないよう留意する必要がある。

i) 『ガイドライン』p.9においても画像作成の際、“ピントによるボケ”や電子的に生まれるノイズ、天地等への注意喚起がなされている。

画像の利用条件表示について、PDMは使用されておらず、CCライセンスを使用している漢籍DAも少数であった。例えば「京都大学貴重資料デジタルアーカイブ」では、“平面複製である電子化画像の作成においては作成者に著作権が発生しないという法解釈に基づき”[7]、CCライセンスを採用していない。東京大学附属図書館も、同総合図書館が公開するデジタル画像の利用条件として、CCライセンスを使用せず、代わりに“CC BY”相当の条件を適用している[8]。漢籍画像データの二次利用にあたって所蔵機関名の表示などを求めたとしても、著作権保護期間が満了している漢籍に対してはCCライセンスを採用できず、結果として各機関が独自に利用条件を表示しているものと思われる。CCライセンスは、こうした所蔵機関側のニーズに充分に対応されていないと言える。

6.2 書誌データに関する課題

書誌データは、画像データと異なり自機関で作成されている場合が多かった。日本目録規則1987年版は、改訂3版より、和古書・漢籍に関する条文が追加された[9]が、日本目録規則1987年版改訂3版を準拠目録規則として挙げた漢籍DAは2割強であった。準拠目録規則が不明、特になし、または無回答であったのも同じく2割強であった。自機関独自の整理基準またはマニュアルに基づいて書誌データを作成している漢籍DAが4割弱で最も多かった。

入力用メタデータスキーマは8割強の漢籍DAが独自のものを採用していた。6割五分程度の漢籍DAが表形式でメタデータを出力可能であることは横断検索を目指す上で好材料ではあるが、ある程度のメタデータ項目の統一あるいは標準化の可能性も検討していきたい。

書誌データを、1タイトル分の画像に与えるか、1冊分の画像に与えるかという書誌データ作成単位は漢籍DAによってばらつきがあった。現状では、これらの漢籍DAを横断検索する際、検索結果画面に作成単位の異なるメタデータが混在し、利用者にとってわかりづらくなる恐れがある。他方、メタデータ作成単位は、漢籍以外の所蔵資料やこれまで作成してきた書誌データに左右されるもので、どちらかが誤りというものではない。横断検索システム実装の際に、タイトル単位の書誌データをばらしたり、物理単位の書誌データに対しタイトル単位の書誌データを付与するといった工夫が必要である。

ユニコードに対応していないシステムでは、漢籍書誌データに現れる文字列を正確に表示できないことがあるので、そのようなシステムに対応するため、書誌データに対して、日本の漢字に直す、ゲタ文字を使用する等、本来であれば不要な処理が行われてしまう可能性がある。未対応の漢籍DAには早急なユニコード対応が望まれる。

漢籍書誌データのみを抽出できる漢籍 DA と、抽出できない漢籍 DA の数は同程度であった。目視で抽出可であっても、機械的に抽出できない場合、検索対象を漢籍に限定した複数 DA の横断検索は不可能である。これを実現するためには、書誌データに漢籍という資料種別を示す識別子等を含める必要がある。

6.3 その他の課題

自機関の課題として、「漢籍を扱うための知識・経験が不十分」と回答した機関が、伝統的に漢籍の所蔵機関であったはずの図書館において5割を超えることは、危機感を持って受け止められるべきであると考えられる。現状、和古書や漢籍を扱うための知識を習得できるような科目は、“司書資格取得のために大学において履修すべき図書館に関する科目”[10]に含まれていない。東京大学東洋文化研究所附属東洋学研究情報センターや、京都大学人文科学研究所附属東アジア人文情報学研究センターでは、毎年図書館職員等を対象に漢籍整理研修を実施しているが、そのような研修を受けられる環境にない機関も存在するものと思われる。自機関の課題を尋ねる選択肢の後に設けた自由記述欄では、“資料が漢籍に当たるか判断できる職員がいない”、“知識のある方から助言や指導を受けたい”といった記述がみられた。本件はDAに限らず、図書館目録中の漢籍の書誌データにも関わることであり、専門家の数が限られる中でどのような対策を講じることができるかは、今後の課題とした。

6.4 終わりに

以上の結果より、残念ながら現時点では、対象を漢籍に限った複数 DA の横断検索を今すぐ実現することは難しいと結論せざるをえない。しかし、本調査によって、漢籍 DA の横断検索を目指す上で解決すべき課題が明らかとなった。

現在、漢籍利用者がどのように漢籍を利用するかについてのインタビュー調査を実施中であり、この結果を踏まえた上で、新しく漢籍 DA を構築する機関に向けて、横断検索を視野に入れた、漢籍 DA のための標準的なメタデータフォーマットの開発等を行っていききたい。

謝辞

質問紙調査および訪問調査にご協力いただいた機関の皆様へ心より御礼申し上げます。また、本研究の遂行をご許可・ご支援くださいました、東京大学東洋文化研究所教授大木康先生ならびに研究所事務室の皆様へ感謝申し上げます。本研究は科研費 17J40023, 18K18329 による研究成果の一部です。

参考文献

- [1] 木村麻衣子. 日本における古典籍デジタルアーカイブの現況: 質問紙調査の結果から. 2017年度三田図書館・情報学会研究大会発表論文集. 2017, p. 1-4.
- [2] 木村麻衣子. 日本の公共図書館における古典籍デジタルアーカイブの現況: ウェブサイト調査を通じて. 2016年度日本図書館情報学会春季研究集会発表論文集. 2016, p. 5-8.
- [3] 千葉真由美, 井黒佳穂子, 金田房子. “資料: 全国古典籍等画像公開データベース一覧”. 共同研究「日本古典籍コードの国際標準化」成果報告書(2014年度). 山本和明編. 国文学研究資料館, 2015, p. 25-46, <http://id.nii.ac.jp/1283/00003164/>, (参照 2018-07-02).
- [4] デジタルアーカイブの構築・共有・活用ガイドライン. デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会, 2017, 43p. https://www.kantei.go.jp/jp/singi/tit eki2/digitalarchive_kyougikai/guideline.pdf, (参照 2018-07-02).
- [5] “Adobe、「Flash」を2020年末に終了へ”. <http://www.itmedia.co.jp/news/articles/1707/26/news044.html>, (参照 2018-07-15).
- [6] 永崎研宣. デジタル文化資料の国際化に向けて: IIIF と TEI. 情報の科学と技術. 2017, 67(2), p. 61-66.
- [7] 赤澤久弥, 大村明美. 京都大学附属図書館における貴重資料画像の二次利用自由化. カレントアウェアネス-E. 2018, no. 343, <http://current.ndl.go.jp/e2004>, (参照 2018-07-20).
- [8] 東京大学附属図書館. “利用条件: 画像データ等の利用について”. <https://www.lib.u-tokyo.ac.jp/ja/library/general/reuse>, (参照 2018-07-20).
- [9] 日本図書館協会目録委員会編. 日本目録規則 1987年版改訂3版. 日本図書館協会, 2006, 445p.
- [10] 文部科学省. “司書資格取得のために大学において履修すべき図書館に関する科目一覧”. http://www.mext.go.jp/comp onent/a_menu/education/detail/_icsFiles/afieldfile/2013/01/07/1320994_6.pdf, (参照 2018-07-22).