
【講 義】

日本語の歴史的典籍の データベースについて

講師 松原恵

(国文学研究資料館)

古典籍共同研究事業センター事務室古典籍データベース係長)

平成29年度 日本古典籍講習会
H30(2018).1.18

日本語の歴史的典籍のデータベースについて

－「日本語の歴史的典籍の国際共同研究ネットワーク構築計画」
(略称：歴史的典籍NW事業)の概要－

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
松原 恵

0.はじめに

- 古典籍資料は、
 - ・ 近現代の刊行図書とは異なり、同じ作品でも一点ごとに内容、形態等が異なっている。そのため、書誌データは記述対象資料ごとに作成されるのが通例。
 - ・ ただし、書誌データに記載できる事項には限りがある。
→ 1丁の行数、挿絵の有無、挿絵の彩色の状態、印記の種類 等々・・・全て記述できる???
 - ・ 書誌データを見比べても、それらが同一の資料かどうかは、**現物を閲覧しない限り確認できない** (ことが多い)。



● 古典籍資料を画像化し、複数の所蔵機関の古典籍画像を一つのデータベースに統合する意義は大きい！



- 検索の手間／検索結果の同定の手間／資料閲覧申請等にかかる手間 などを削減。
- 画像情報を活用した学術研究の可能性の広がり。
- 資料劣化に対する一つの対応策。(現物資料の利用を妨げるものではなく。)
- 資料破損・消失等の危機に対する一つの対応策。(画像データでの長期保存。)

1. 事業の概要

● 名称

日本語の歴史的典籍の国際共同研究ネットワーク構築計画(略：歴史的典籍NW事業)

● 事業実施期間

平成26(2014)年度～平成35(2023)年度の10年間（予定）※H25年度は準備期間

● 本事業の目標と3つの柱

【目標】

- ・開かれた学術研究基盤の構築・共同利用の促進
- ・新たな人文系の共同研究モデルの創出 ー個人間から組織間へ/国文学から異分野へー

【3つの柱】

- (1) 日本語の歴史的典籍画像データベースの構築
→ 大学・専門的機関と協働し、30万点の歴史的典籍画像データの作成・公開
- (2) 古典籍を活用した国際共同研究の推進
- (3) 国際共同研究ネットワークの構築

1. 事業の概要

世界の学術フロンティアを先導する大規模プロジェクトの推進		
<p>目的</p> <ul style="list-style-type: none"> ○ 最先端の大規模研究装置等により人類未踏の研究課題に挑み、世界の学術研究を先導 ○ 国内外の優れた研究者を結集し国際的な研究拠点を形成するとともに、研究活動の共通基盤を提供 		
<p>推進方策</p> <ul style="list-style-type: none"> ○ 日本学術会議において科学的観点から策定したマスタープランを踏まえつつ、文部科学省において戦略性・緊急性等を加味し、ロードマップを策定。その中から実施プロジェクトを選定。 ○ 原則10年間の年次計画を策定し、専門家等で構成される委員会で評価・進捗管理 ○ 大規模学術フロンティア促進事業として、国立大学運営費交付金等の基盤的経費により戦略的・計画的に支援 		
<p>主な成果</p> <ul style="list-style-type: none"> ○ ノーベル賞受賞につながる画期的研究成果(受賞歴:H14小柴先生、H20小林先生、益) ○ 年間約1万人の共同研究者(その約半数が外国人)が集結し、国際共同研究を推進(共) ○ 産業界と連携した最先端の研究装置開発により、イノベーションの創出にも貢献(すばる望遠鏡の超高感度カメラ⇒医療用カメラ) 	<p>人文社会科学分野として初めての 大規模学術フロンティア事業</p>	
<p>「究極の科学技術イノベーション」核融合の実現に向けた学理の追求</p> <p>超高性能プラズマの定常運転の実証 〔自然科学研究機構核融合科学研究所〕 我が国独自のアイデアによる「大型ヘリカル装置(LHD)」により、高温高密度プラズマの実現と定常運転の実証を目指す。また、将来の核融合炉の実現に必要な学理の探求と体系化を目指す。</p> 	<p>大規模学術フロンティア促進事業</p> <p>アインシュタインが予言した重力波(時空の歪み)観測による重力波天文学の創成 大型低温重力波望遠鏡(KAGRA)計画 〔東京大学宇宙線研究所〕 一辺3kmのL字型のレーザー干渉計により重力波を観測し、ブラックホールや未知の天体等の解明を目指すとともに、日本欧による国際ネットワークを構築し、重力波天文学の精進を目指す。</p> 	<p>歴史的典籍を活用した異分野融合研究の醸成と日本文化の国際的発信</p> <p>日本語の歴史的典籍の国際共同研究ネットワーク構築計画 〔人文文化研究機構国文学研究資料館〕 日本語の歴史的典籍30万点を画像データベース化し、新たな異分野融合研究や国際共同研究の発展を目指す。古典籍に基づく過去のオーロラの研究、江戸時代の食文化の研究など他機関や産業界と連携した新たな取組を開始。</p> 

文部科学省ウェブサイト「学術研究の大規模プロジェクトの推進」より
http://www.mext.go.jp/a_menu/kyoten/1383666.htm (2018/1/5 閲覧)

1.事業の概要

●参考：共同研究について



<http://www.nijl.ac.jp/pages/cijproject/>



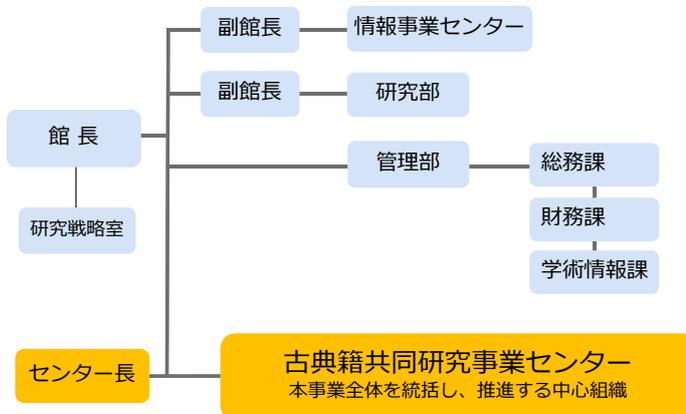
極地研等との異分野融合研究や
NIIとの絵検索に関する共同研究など
多岐に渡る。詳細は

<http://www.nijl.ac.jp/pages/cijproject/research.html>

1.事業の概要

●事業の推進体制（館内）

平成25(2013)年4月1日：古典籍データベース事業センター 発足
 平成26(2014)年4月1日：古典籍共同研究事業センター へ改組
 本事業は、センターを中心として、館内部署と連携しながら様々な取組を実施



1.事業の概要

- 事業の推進体制（国内・海外の連携機関等）



プロジェクト実施期間中は、事業の進捗に応じて国内外の大学・研究所等と連携を進める予定。

1.事業の概要

- 国内の拠点大学

- ・古典籍を多く保有する国内20大学。
- ・各大学2名（教員1名＋図書館員1名）の委員により「拠点連携委員会」を構成。
- ・H27年度は5大学、H28年度は10大学で古典籍資料の撮影を実施。



2.新日本古典籍総合データベース

●新たなデータベースの構築・公開

- ・画像30万点の収集を見据えて、新たな画像データベースを公開。
- ・従来の「日本古典籍総合目録データベース」の豊富な書誌情報を引き継ぎつつ、画像を見やすく活用しやすい「新日本古典籍総合データベース」

【公開】平成29(2017)年10月

【URL】 <https://kotenseki.nijl.ac.jp/>



- ・「日本古典籍総合目録データベース」の豊富な書誌情報と、当プロジェクトで作成・収集した高精細画像をあわせて見ることができる
- ・誰でも無料でアクセスできる
- ・スマホ、タブレットにも対応

2.新日本古典籍総合データベース

●公開点数（2018.1現在）

- ・書誌：約600,000作品
 - ・画像：約74,000作品
- ➡ 30万作品の画像公開を目指す

●公開中の所蔵者

- ・約80機関
- ・公開中所蔵者一覧：<https://kotenseki.nijl.ac.jp/page/usage.html>

所在地	文庫番号	所蔵者	公開点数	撮影方法	一覧ページへ	画像利用について	備考
北海道	0001	北海道大学附属図書館	154	マイクロフィルム	画像一覧	案内ページ	拠点大学
北海道	HOKUJ	北海道大学附属図書館	477	デジタル	画像一覧	案内ページ	拠点大学
北海道	0056	道庁中央図書館	528	マイクロフィルム	画像一覧	案内ページ	
青森	0096	八戸市立図書館	911	マイクロフィルム	画像一覧	案内ページ	
青森	0272	弘前市立図書館	2278	マイクロフィルム	画像一覧	案内ページ	
青森	HRSJ	弘前市立図書館	184	デジタル	画像一覧	案内ページ	
岩手	0281	もりおか歴史文化館	2669	マイクロフィルム	画像一覧	案内ページ	
山形	0026	酒田市立歴史文庫	1623	マイクロフィルム	画像一覧	案内ページ	
茨城	0006	筑波大学附属図書館	902	マイクロフィルム	画像一覧	案内ページ	拠点大学
群馬	T8	矢口丹波記念文庫	664	マイクロフィルム	画像一覧		
東京		国文学研究資料館			画像一覧	案内ページ	
東京	0020	宮内庁書陵部	3914	マイクロフィルム	画像一覧	案内ページ	当該所蔵フィルム
東京	KSRM	宮内庁書陵部	3197	マイクロフィルム	画像一覧	案内ページ	書陵部所蔵フィルム
東京	0218	国立国会図書館 貴重書	105		画像一覧	案内ページ	NDL画像へのリンク
東京	0002	東京大学総合図書館	177	マイクロフィルム	画像一覧	案内ページ	拠点大学
東京	TOKY	東京大学総合図書館	169	デジタル	画像一覧	案内ページ	
東京	0003	東京大学総合図書館 秋葉文庫	374	マイクロフィルム	画像一覧	案内ページ	拠点大学
東京	0004	東京大学文学部国文学研究室	336	マイクロフィルム	画像一覧	案内ページ	拠点大学
東京	0357	東京大学文学部国文学研究室	833	マイクロフィルム	画像一覧	案内ページ	拠点大学
東京	S6	東京理科大学近代科学資料館 下流文庫	831	マイクロフィルム	画像一覧	案内ページ	
東京	KEIO	慶應義塾大学図書館メディアセンター 富士川文庫	833	デジタル	画像一覧	案内ページ	拠点大学
東京	KNJK	研京会図書館	108	デジタル	画像一覧	案内ページ	
東京	OCHA	お茶の水女子大学図書館	551	デジタル	画像一覧	案内ページ	拠点大学
東京	AJNH	味の素の文化センター	282	デジタル	画像一覧	案内ページ	
新潟	0324	新潟大学附属図書館 佐野文庫	1168	マイクロフィルム	画像一覧	案内ページ	

2.新日本古典籍総合データベース

● 特長その1：探しやすい

- ・タイトルやキーワードでの検索はもちろん、**画像に付けられたタグ**、本文テキスト、おすすめキーワード、ピックアップコンテンツなど、多彩な検索ができる。

● 特長その2：引用しやすい

- ・国際的識別子**DOI**を採用し、永続的なアクセスを保証。論文などへの引用の際に便利。

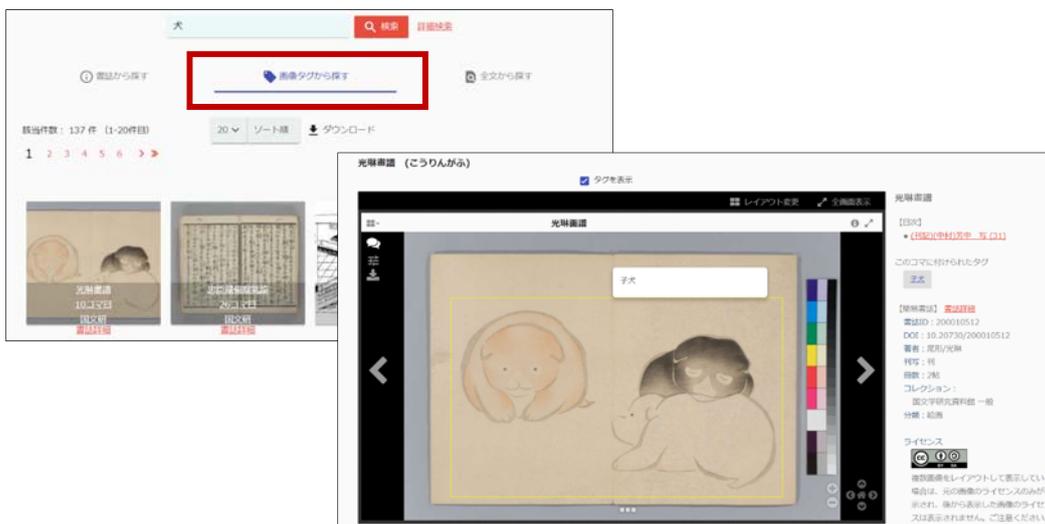
● 特長その3：活用しやすい

- ・デジタルアーカイブの新しい規格である**IIIF** (International Image Interoperability Framework) を採用。異なるデジタルアーカイブ間での画像共有や再利用が容易。
- ・国文研所蔵の古典籍は**CC BY-SA**で公開。（他機関所蔵資料は各機関の利用条件による。）

2.新日本古典籍総合データベース

● 画像タグからの検索

- ・画像タグとは、古典籍画像の見出し情報や翻刻テキスト、挿絵などに付与したキーワード。
- ・書名などがわからなくても、様々な古典籍と出会うことができる。



2.新日本古典籍総合データベース

●DOIによる永続的アクセス

- ・DOIは、デジタルオブジェクトに対する国際的な識別子。
- ・データベースのURLが変更されても、永続的にアクセス可能。
- ・論文等で古典籍画像を参照・引用した際にDOIを明記することで、**読者は元の古典籍画像をいつでも確認することができる。**

伊勢物語	
書籍ID	200024143
DOI	10.20730/200024143
URL	http://kotenseki.nijl.ac.jp/biblio/200024143
記載書名	1. 伊勢物語 (いせものがたり) 帙
刊行	写
出版事項	横覧 函込4
形態	90丁,24.0×18.5cm,半
冊数	1冊
書誌注記	(形) 茶色地花唐草文金襴表紙,金泥地秋草結見返し。 (巻) 桐蔭本第二系結唐草。(巻背)「此本非原本之録石本根 源殊可/庶幾の書写了秘製物也録日可異/書者也/寛元四年丙 午三月廿八日御注了/明教」・「相傳本為人被遺失之屋二種 彼岡本所/廣写也于天文永九年十二月十二日記之/定円」・ 「應永四年十一月一日書写了是則(…)/西方上人願寛」。 (伝) (印記)「鉄心斎文庫」「不忍文庫」「阿波国文庫」 「月明荘」。
コレクション	国文学研究資料館 鉄心斎文庫
請求記号	98-9

<https://doi.org/10.20730/200024143>

2.新日本古典籍総合データベース

●IIIFに対応したビューア

- ・International Image Interoperability Framework (IIIF) はデジタルアーカイブの新たな国際規格。国内機関でも徐々に採用され始めている。
- ・従来のデジタルアーカイブは、公開元のアーカイブでしか見ることができなかったがIIIFに対応したアーカイブでは、異なるアーカイブ間での画像共有や再利用が可能になる。

国文学研究資料館 伊勢物語 (いせものがたり)

レイアウト変更 全画面表示

伊勢物語 伊勢物語

6コマ目 7コマ目 8コマ目 6コマ目 7コマ目 8コマ目

URI <http://kotenseki.nijl.ac.jp/biblio/200024143/viewer/1>

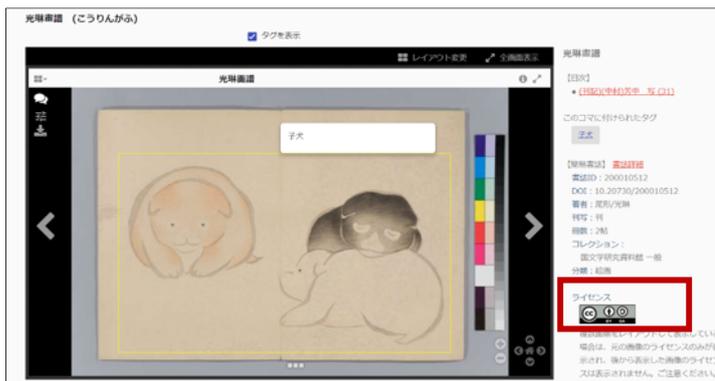
Manifest URI <http://kotenseki.nijl.ac.jp/biblio/200024143/manifest>

国文学研究資料館の『伊勢物語』の2点を並べて表示した様子。IIIF対応ビューアで公開している異なるアーカイブの画像をこのように表示させることもできる。(Manifest URIが公開されている必要がある)

2.新日本古典籍総合データベース

●クリエイティブ・コモンズ・ライセンスの採用

- ・デジタルアーカイブで公開された画像の利用条件は機関ごとに異なるため、画像を再利用したいとき、そのデジタルアーカイブの利用条件を調べる必要がある。
- ・国文研では、**CCライセンスの採用により、利用条件を明確化し、わかりやすくしている。**(データベースで公開中の他機関の画像については各機関の希望による。)
- ・**国文研所蔵の古典籍画像はCC BY-SA。**国文研所蔵であること等を明記することで再利用可能。(年賀状、ポスター、商用/非商用デザイン、出版物、放送etc…)

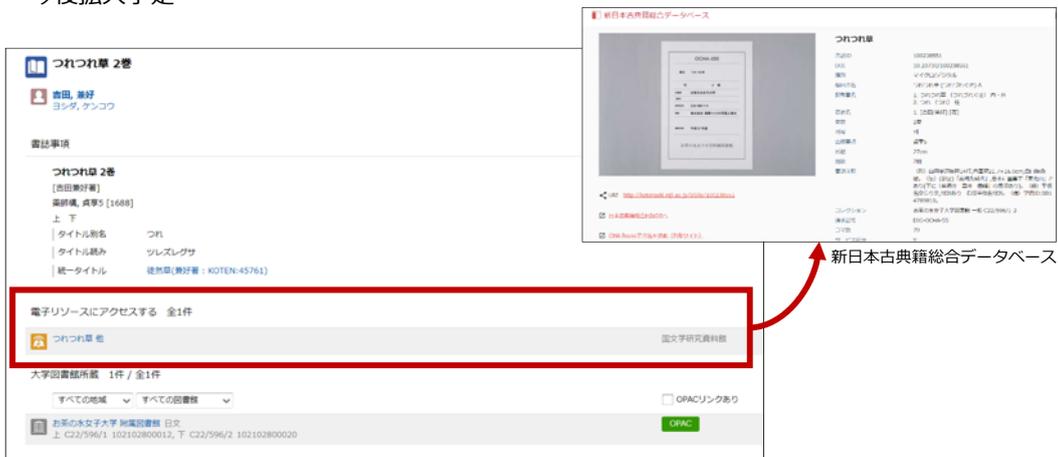


クレジット表示例：
『光琳画譜』
国文学研究資料館所蔵
クリエイティブ・コモンズ 表示
4.0 ライセンス CC BY-SA

2.新日本古典籍総合データベース

●外部サービスとの連携

- ・CiNii Booksとの連携 (2017年10月から)
- ・今後拡大予定



CiNii Books

<https://ci.nii.ac.jp/ncid/BB14789810>

3.オープンデータ

●日本古典籍データセット

- 公開時期：平成27(2015)年～
- 公開点数：1767作品（2018.1現在）
- 公開条件：クリエイティブコモンズライセンスの「CC BY-SA」（表示-継承）
- 公開場所：「人文学オープンデータ共同利用センター（CODH）」
- 内容：
 - ・古典籍画像と書誌データをセットで公開。
 - ・一部の作品には翻刻データや、タグデータを付加。



<http://codh.rois.ac.jp/pmjt/>

3.オープンデータ

●公開その1 日本古典籍データセット

- ・古典籍1767作品分の画像・書誌データ等をセットにして提供。
- ・各作品を丸ごとダウンロード可能。

●公開その2 日本古典籍字形データセット

- ・日本古典籍データセットの画像をもとに、翻刻テキストを制作する過程で生まれるくずし字の座標情報などを含めた、1文字ずつの字形データを提供。
- ・約4,000文字種、約400,000文字（2017.12時点）

●公開その3 江戸料理レシピデータセット

- ・公開日：平成28(2016)年11月24日
- ・『万宝料理秘密箱』に記された、100種類以上の卵料理に関する記述を翻刻。
- ・そのうち20品を現代語訳し「クックパッド」に掲載。うち5品は実際の調理例も掲載。これらのデータをセットにして提供。 ※平成27(2015)年度のアイデアソンがきっかけ

【共通事項】

- ・人文学オープンデータ共同利用センター（CODH）からの公開
- ・クリエイティブ・コモンズ・ライセンス「CC BY-SA」の下に提供

3.オープンデータ

●公開データの活用のために・・・アイデアソンの実施

アイデアソン：アイデア（Idea）とマラソン（Marathon）を掛け合わせた造語。

あるテーマについて様々な人が意見を出し合い、アイデア創出などを短時間で実行イベントのこと。国文研では「**古典籍を使って何が出来る？**」というアイデアソンを毎年実施。

【平成27(2015)年度】

「歴史的典籍オープンデータワークショップ～古典をつかって何が出来るか！
じんもんそん2015～」

・平成27(2015)年12月18日(金)@メルパルク京都

・レポート：https://www.nijl.ac.jp/pages/cijproject/report_20151218.html

【平成28(2016)年度】

「歴史的典籍オープンデータワークショップ～使いたおそう！古典籍データ～」

・平成28(2016)年12月9日(金)@国文学研究資料館

・レポート：https://www.nijl.ac.jp/pages/cijproject/report_20161209.html

【平成29(2017)年度】

「歴史的典籍オープンデータワークショップ～切ったり貼ったり、
古典籍からなにを取り出そう？～」

・平成29(2017)年12月8日(金)@大阪市立大学文化交流センター



4.古典籍のデジタル化

●資料の収集範囲

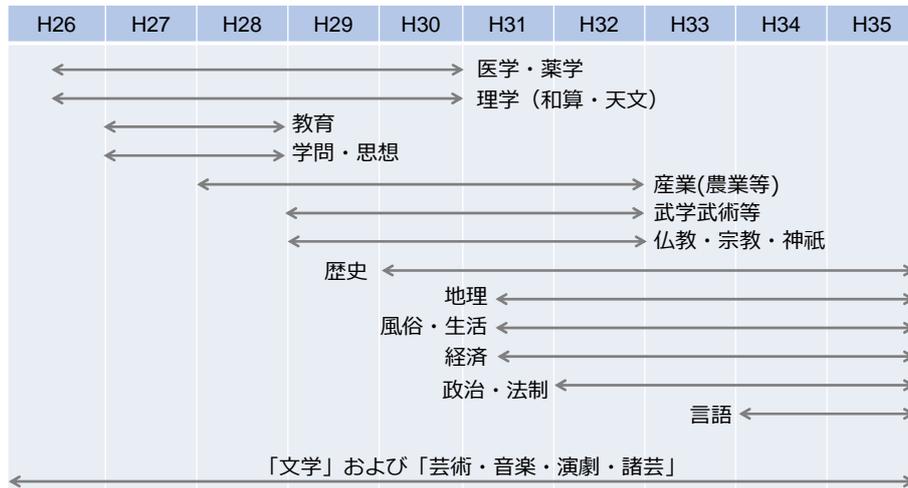
- ・原則として『国書総目録』の収録範囲に準じた収集
- ・古典籍という性質上、1作品1画像に限定してはいない

	区分	対象となる条件
対象範囲	時代区分	・国初から慶応4年（1868年）の期間に成立。
	著者等	・日本人により著編撰訳されたもの。
	形態	・書籍。 ・絵巻物、書画帖を含む。 ・浄瑠璃、長唄の類の歌謡の一篇。
	内容等	・日本人の著作（和文、漢文、欧文を問わずに含む）。 ・日本に帰化したとみなすべき外国人の、日本における著述。 ・外国人の著述を、日本人が改修編纂したもの。 ・日本人が外国書を翻訳したもの、あるいは注釈を施したもの。 （但し、施された注が書き込み程度のものは除く。） ・外国人の著述したものについて、日本人による頭書・首書があるもの。
補足	<ul style="list-style-type: none"> ・対象外となるのは、上記の範囲外にあるもの全て。 ex：漢籍（和刻本漢籍も含む）、 明治時代以降に刊行された資料、 一枚の書画、絵図、地図など、 ・大部な刊本資料（『群書類従』など）は事前調整を行う。 	

4.古典籍のデジタル化

●資料の分野別収集

- ・下記の計画表に基づいて画像を収集し公開する予定。
- ・ただし、共同研究計画に合わせた収集や、文庫単位での撮影を実施。



4.古典籍のデジタル化

●デジタル化の方法

(1)業者による撮影

- ・撮影業者による古典籍の撮影・画像作成（撮影の外注）

(2)マイクロフィルムからの画像作成(マイクロコンバート)

- ・マイクロフィルムからTIFF画像やJPEG画像を生成

<特徴>

- ・迅速かつ大量の画像化が可能
- ・安価
- ・モノクロ（マイクロの大半はモノクロ撮影）

(3)画像作成の内製化

- ・館内に設置したブックスキャナーで撮影・画像作成を実施



4. 古典籍のデジタル化

● 画像作成の内製化

- ・平成26(2014)年度から国文研内外で実証試験を実施
- ・平成30(2018)年度から本格実施予定

<特徴>

- ・撮影ノウハウが蓄積できる
- ・撮影を行う日時(時間帯)や対象資料の変更など、柔軟な撮影実施が可能
- ・ブックスキャナでは撮影不可能な資料は対象外
(卷子本や規格外の大きさの資料など → 業者による撮影)

<撮影方法>

- ・古典籍の撮影に適したブックスキャナの開発に関する共同研究を実施し、今後実運用予定
(参照：<http://current.ndl.go.jp/e1954>)
- ・既存のブックスキャナよりもさらに容易に撮影が可能となる

4. 古典籍のデジタル化

● 拠点大学での撮影

平成27(2015)～28(2016)年度で13大学での撮影を実施し、うち3大学では予定全点を撮影完了。

大学名	H27-28 撮影点数	大学名	H27-28 撮影点数
北海道大学 (全点完了)	927	京都大学	934
東北大学	137	奈良女子大学 (全点完了)	1,114
筑波大学	174	神戸大学	218
東京大学	171	九州大学	100
お茶の水女子大学 (全点完了)	1,085	慶應義塾大学	833
名古屋大学	269	関西大学	33
大阪大学	373		

● 拠点大学以外でのデジタル化 (専門機関所蔵資料の収集)

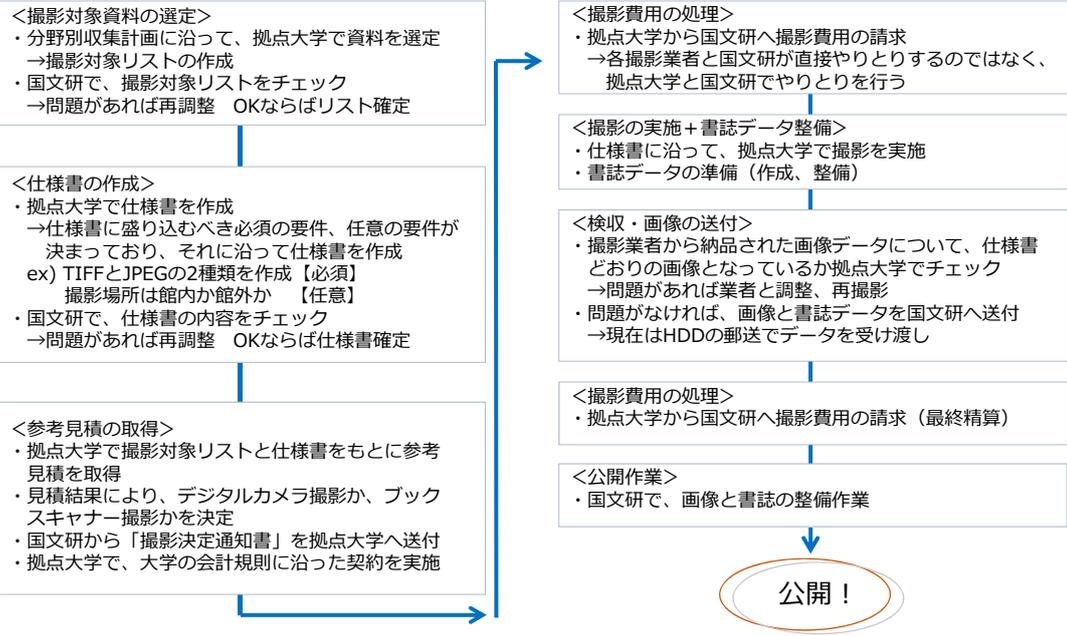
拠点大学以外の機関でも、専門性の高い古典籍を所蔵している学術機関と連携し、分野別収集や共同研究の進捗に応じた撮影を実施。

(機関例)

- ・研医会図書館：医学
- ・味の素の文化センター：料理、食文化
- ・信州大学附属図書館繊維学部図書館：産業 (養蚕)
- ・国立天文台：理学
- ・宮内庁書陵部所蔵マイクロフィルムからの画像化：歴史 など

4. 古典籍のデジタル化

● 拠点大学での撮影 手続き



4. 古典籍のデジタル化

● 拠点大学での撮影 仕様書について

- 作成する画像
 - ・ TIFF形式 1セット 国文研での保管用
 - ・ JPEG形式 2セット 拠点大学での活用・保管用+国文研での公開用
- 解像度
 - ・ デジタルカメラ 2,100万画素相当
 - ・ ブックスキャナー 400dpi相当
- 資料以外の撮影
 - ・ ターゲット：各資料の先頭に挿入（資料の基本情報を記したもの）
 - ・ スケール：各資料の1カット目に挿入（メジャー（目盛り））
 - ・ カラーチャート：ターゲットを除く全コマに挿入（色の比較測定用）
- 発注できる範囲
 - ・ 撮影、TIFFとJPEGの画像生成、画像の検収・納品 まで
（例えば「書誌データ作成」などを仕様書に含めることは不可）

<基本方針>

- ・ 各拠点大学で、品質にばらつきのある画像が作成されないようにするため、そして業者発注の範囲が異なるようにするため、一部の項目については仕様書を統一。
- ・ 一方で、各大学の状況に応じた撮影ができるよう任意事項も設けてある。
（ex.撮影場所は館内か館外か／業者に求める資格要件・実績 など）

4.古典籍のデジタル化

●これまでの画像収集



国文研では平成23(2011)年度から古典籍のデジタル化を本格スタート。

(a)所蔵資料の撮影

国文研が所蔵する古典籍（写本・刊本）を撮影し公開。

(b)収集したマイクロフィルムからの画像作成

国文研が所蔵するマイクロフィルムから画像を作成。

→デジタル公開の許諾が得られたものから画像作成、公開。

(c)拠点大学等における画像作成

平成27年から実施



(a)+(b)+(c)

平成29(2017)年12月時点での画像公開数：7万4,000点

4.古典籍のデジタル化

●書誌データの作成

【基本方針】 既存の書誌データがあれば、可能な限り活用する。

※既存の書誌データ例

- (1)各大学でNACSIS-CATあるいは大学独自のデータベース等に登録済みのもの。
- (2)当館の日本古典籍総合目録データベース等に登録されているもの。

※書誌データが作成されていない場合

- (1)画像作成と同時に各拠点大学で作成していただく。
- (2)冊子体目録から入力することで、画像作成と同時に書誌を作成する。
- (3)そのほか、ケースバイケースの対応。

【著作(典拠)と書誌のリンクは当センターで実施】

- 古典資料において著作コントロールは必須。
この点は変更なく運用。
- 拠点大学から提供される書誌データについて、
著作データと書誌データのリンク作業はセンターで行う。
- 当館データベースの入力規則に沿った修正もセンターで実施

5.さいごに

●本事業について（広報）

●ホームページ

国文研HP

<http://www.nijl.ac.jp/>

本事業専用ページ

<http://www.nijl.ac.jp/pages/cijproject/>

国文研 歴史的典籍

検索

●本事業のニュースレター「ふみ」

年2回の発行。本事業の最新動向などをお知らせしています。

また、刊行と同時にウェブ上でも公開しています

https://www.nijl.ac.jp/pages/cijproject/newsletter_fumi_new.html



●国文研公式Twitter @nijlkokubunken

国文研の公式アカウントです。

本事業の最新情報やイベント情報などについても発信しています。



5.さいごに

本事業は、大学や専門的機関と国文研が
共同で行うプロジェクトです。

古典籍を所蔵されている機関の皆さま、多くの情報が
詰まっている古典籍のデジタル化、そして公開への
御協力をよろしく申し上げます。

本件に関する問い合わせ：古典籍データベース係

center_db@nijl.ac.jp

