



資料1

科学技術・学術審議会学術分科会研究環境基盤部会  
学術研究の大型プロジェクトに関する作業部会  
(第82回) R1.6.18

科学技術・学術審議会 学術分科会 研究環境基盤部会  
学術研究の大型プロジェクトに関する作業部会

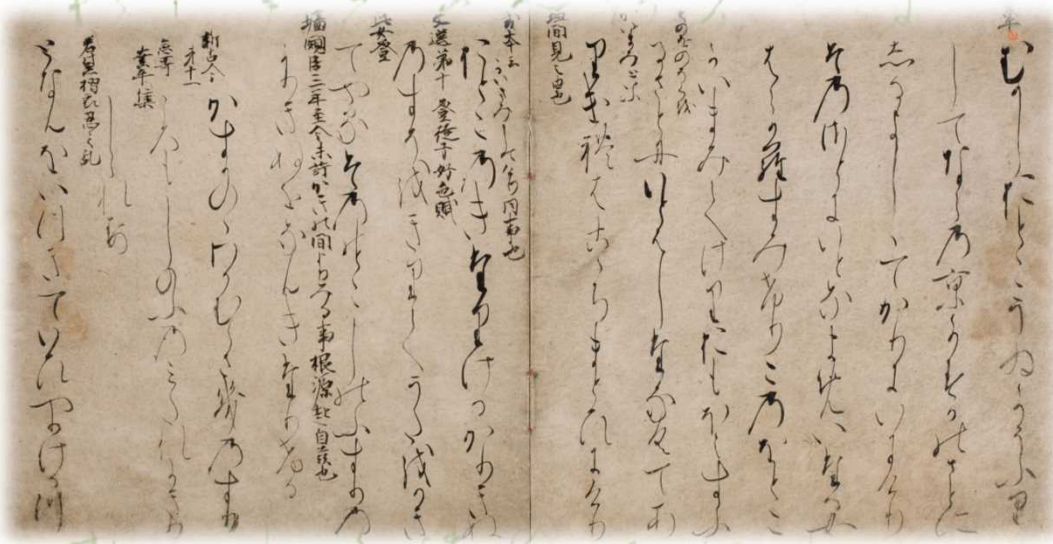
# 日本語の歴史的典籍の 国際共同研究ネットワーク構築計画

—概要と進捗状況—

令和元年6月18日



『浦しま』

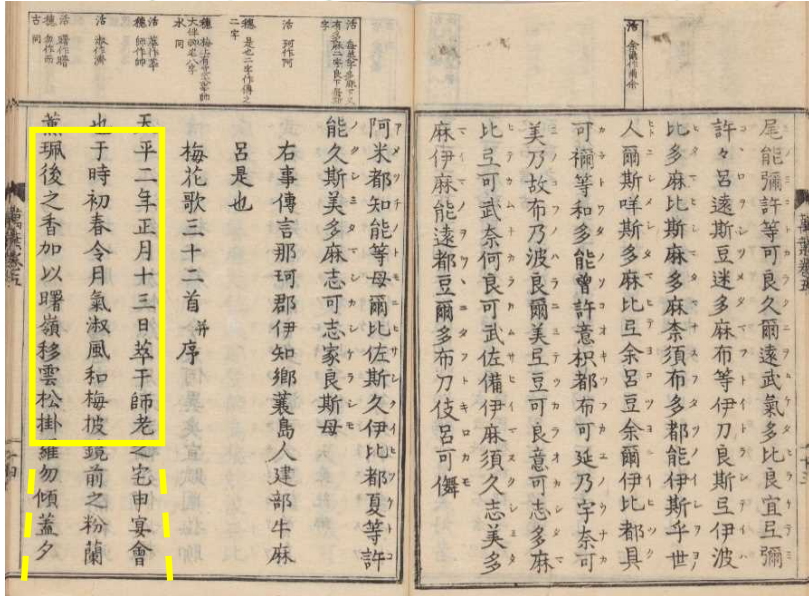


『伊勢物語』

はじめに



## &lt;古典籍とは?&gt;

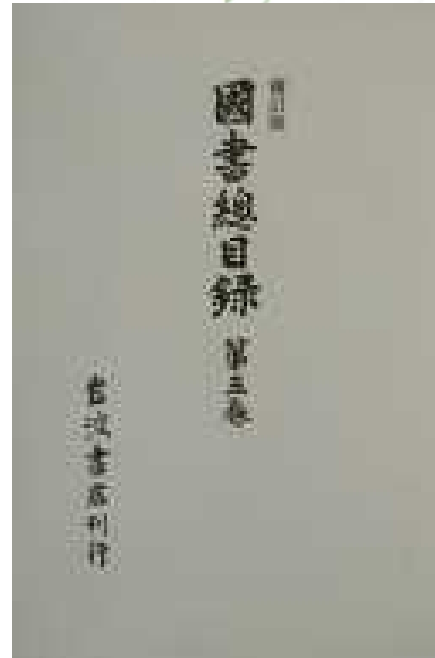


『万葉集』巻五「梅花歌三十二首并序」  
初春令月、気淑風和部分

十二年正月十三日萃于師老之宅申宴會  
丁時初春令月氣淑風和梅披  
珮後之香加以曙嶺移雲松掛

## 日本語の歴史的典籍

近代以前(江戸時代末まで)に、日本で出版・書写された書物のこと。「**古典籍**」あるいは「**国書**」ともいう。



国書総目録(こくしょそうもくろく)  
古代から慶応3年(1867年)までの間に日本人により著述・編纂・翻訳された書籍約50万点が掲載されており、その分野は多岐にわたる。

分野	割合	分野	割合
文学	19%	医学	4%
仏教	15%	武術	3%
歴史	13%	理学	3%
政治・法制	8%	思想	3%
音楽・演劇	5%	その他	23%
神祇	4%	(地理、諸芸、教育、産業等)	



## &lt;古典籍とは?&gt;



『鴨長明方丈記抄』



『安政風聞集』

平安、鎌倉、江戸時代の古典籍には、**例えば**、オーロラ・彗星などの天文情報、地震・旱魃・洪水など自然災害に関する情報が数多く残されている。

しかしながら、我が国の文字は時代が進むにつれ、いわゆる変体仮名や漢字のくずし字が淘汰された結果、近代150年の達成と引き換えに、江戸時代までの書籍に用いられていた文字が全く読めなくなった。

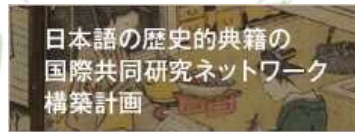
日本のこと、日本で起こったことは古典籍や石碑などに記録として残っているが、文字が読めなくなったため、一般的な日本人は記憶に残るほんの一部しか知ることができない。

欧米諸国と違って、日本人は自らの歴史風土を自在に行き来する能力を失ったとも言える。



### <古典籍とは?>

古典籍は、大学や図書館等公的機関のほか、蔵書家の個人財産として各所に残っている。



ステップアップ



活動実績

### <支持>

近世文学会

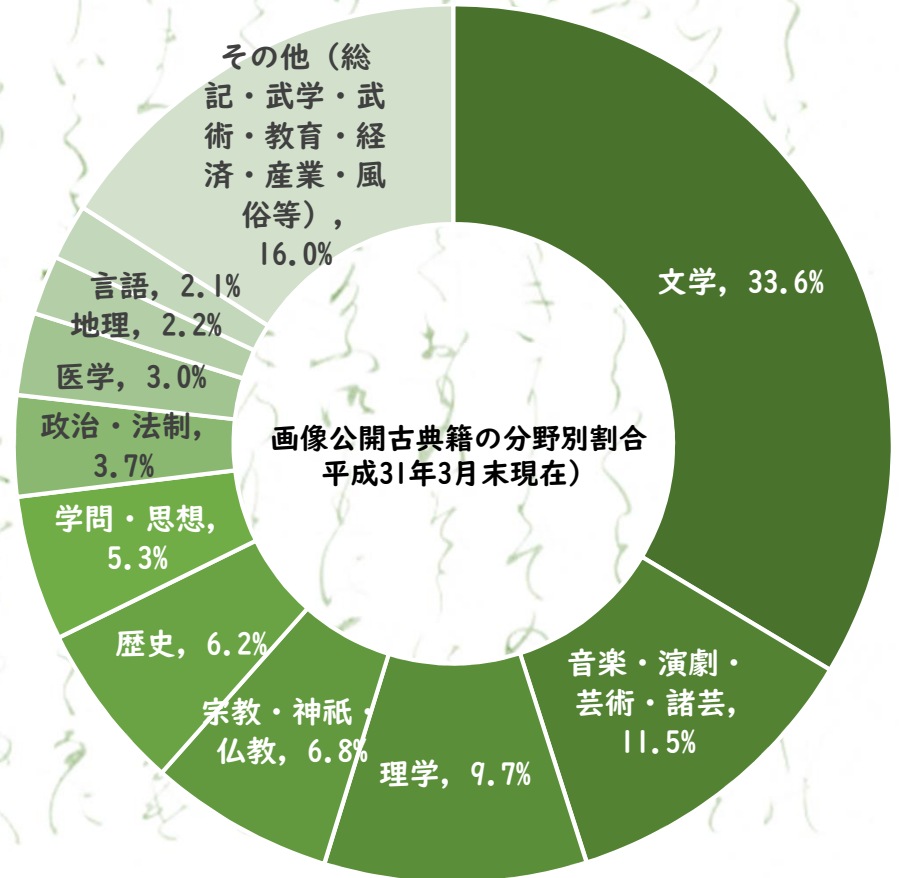
中古文学会

中世文学会

Etc.

研究者コミュニティ

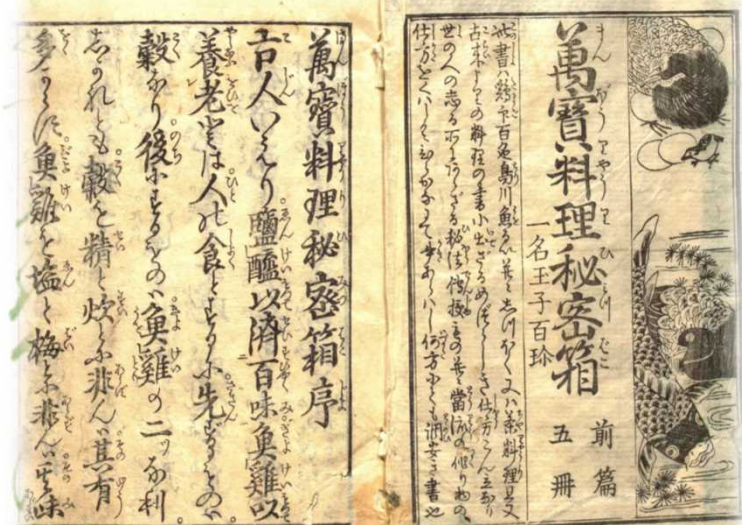
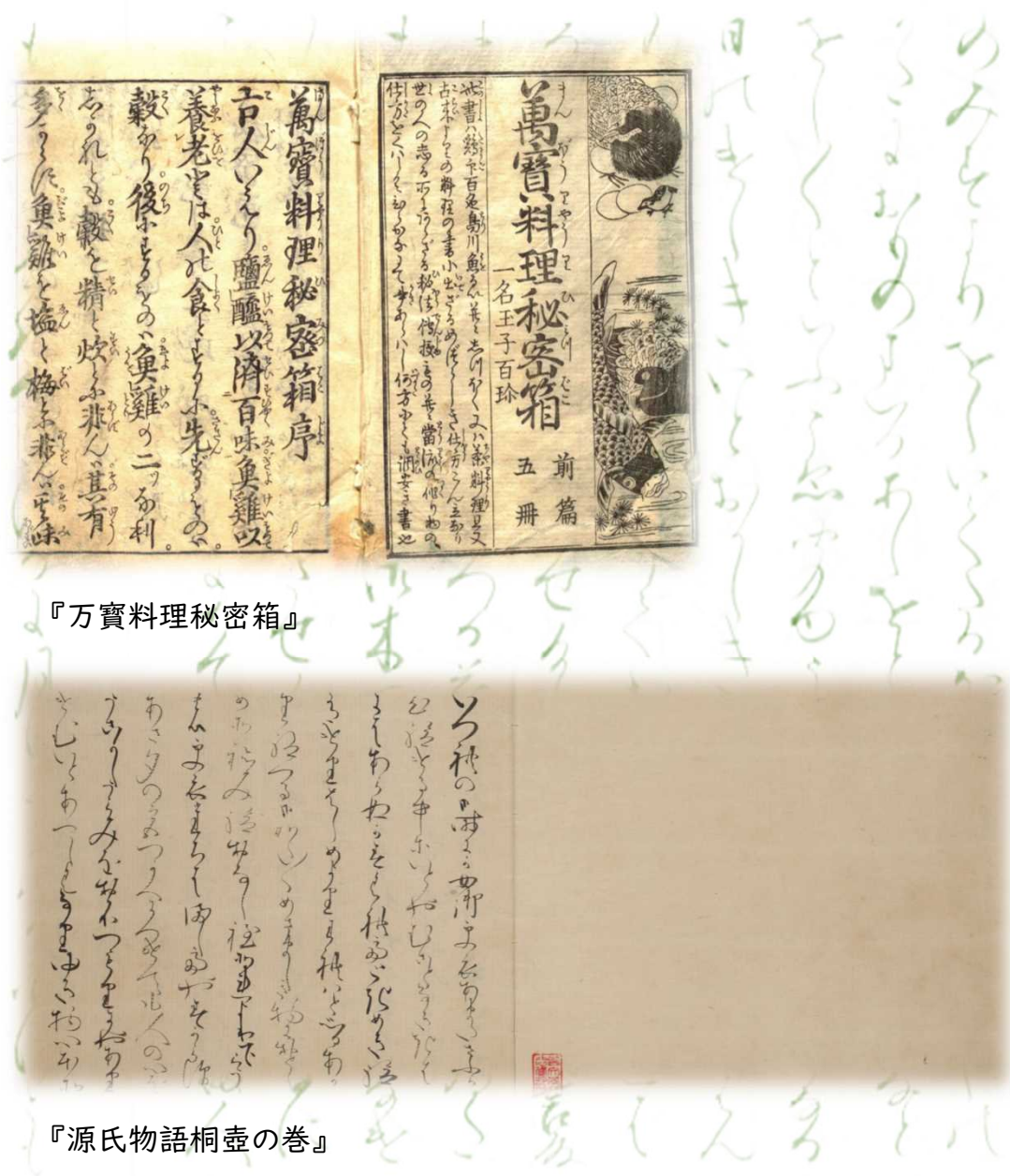
国文学研究資料館では、こうした「古典籍に埋もれた知」を世界に発信するため、国文学以外の古典籍も画像化している。



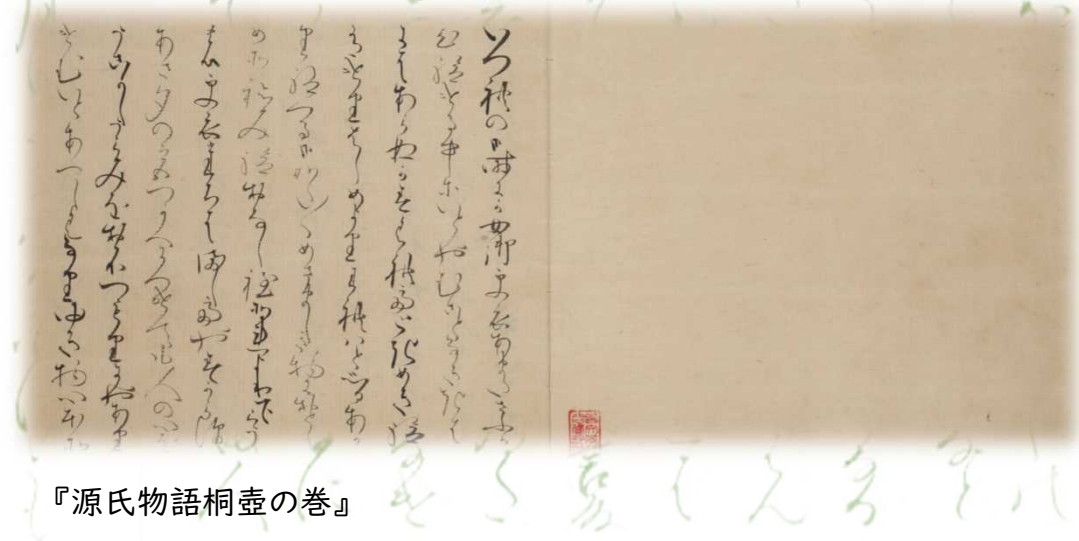




# 概要

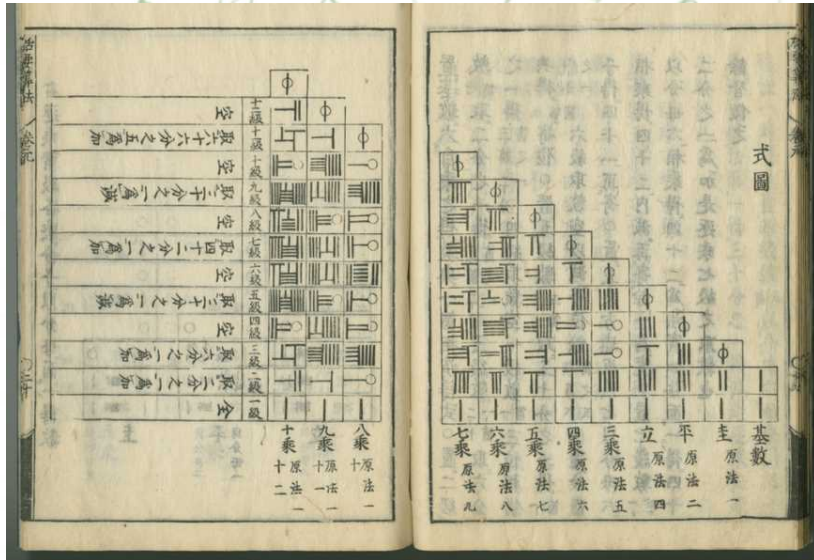


『万寶料理秘密箱』



『源氏物語桐壺の巻』

## &lt;目的・目標&gt;



- 古典籍に集積されてきた膨大な「日本の知」を、大規模画像データベースとして整備
- データベースを自然科学系をも包摂した異分野研究者コミュニティならびに海外研究者コミュニティとの共同・連携により活用
- 日本の文化に関する異分野融合の国際共同研究ネットワークを構築

『括要算法』和算書。関孝和の業績（不定方程式、級数の和、ベルヌーイ数、円および球の求積など）を紹介  
 ★ヤコブ・ベルヌーイより1年早くベルヌーイ数を発見

- 日本文化発信として世界の注目を集めるマンガ、アニメ、和食等の「クールジャパン」が、突然変異的、表層的な現象ではなく、前近代の「知」の伝統の上に花開いた21世紀の「温故知新」であることを発信
- 古典を活きたものとして新たな発想や形に繋げる





## &lt;必要性・緊急性&gt;

europeana  
libraries中国・韓国の国際的地位の向上に伴う欧米諸  
国における**日本文化の位置付けの低下**東日本大震災、西日本豪雨などの災害がもたらす日本  
語で記された文字記録=「日本の知」の喪失(**日本文化  
の損失**)国民の幸福が経済的価値だけに限定されとは限らない現代において  
は、「意味」や「価値」を探求する学問である国文学の振興が不可欠  
「Society 5.0」においても、科学技術の社会実装に向け人文学・社会  
科学に期待

- 極めて大きい大規模画像データベースの構築並びに日本文化の基礎データの公開  
の必要性・緊急性

緊急を要する

- 「オープンデータ」・「オープンサイエンス」化による新しい研究環境、文化環境の整備
- 文学研究における細分化と異分野にまたがる共同研究体制の未成熟の克服

## 大きく後れをとる書籍等の大規模デジタル化

米国デジタル公共図書館 (DPLA)	1,600万点以上
Europeana	1,200万点以上
中国国家図書館 (CADAL)	276万点
韓国国立中央図書館	63万点





## &lt;三つの柱&gt;

## 「日本語の歴史的典籍データベース」の構築

- 30万点<sup>(※)</sup>の画像データの作成 ⇒ 約11万点(平成30年度末現在)
- 大規模情報提供システムの運用 ⇒ 平成29年10月正式公開
- 検索機能の向上・多言語対応 ⇒ 国立情報学研究所等と共同研究

## 国際共同研究ネットワークの構築

- 連携機関との体制構築 ⇒ 共同研究体制(国内にあっては画像作成を含む)の構築
- ⇒ 人文学分野の先導的な共同研究モデルの構築

## 国際共同研究の推進

- 異分野融合研究の醸成 ⇒ 国立極地研究所等と共同研究
- 「総合書物学」の推進・構築 ⇒ 人間文化研究機構内連携共同研究

※ 我が国の古典籍約50万点のうち、国文学研究資料館及び20の拠点大学等が保有する古典籍で、画像化が可能と見込まれるもの(約30万点)を対象



## <ロードマップ>

年次計画

H26

H27

H28

H29

H30

(H31)  
R01

R02

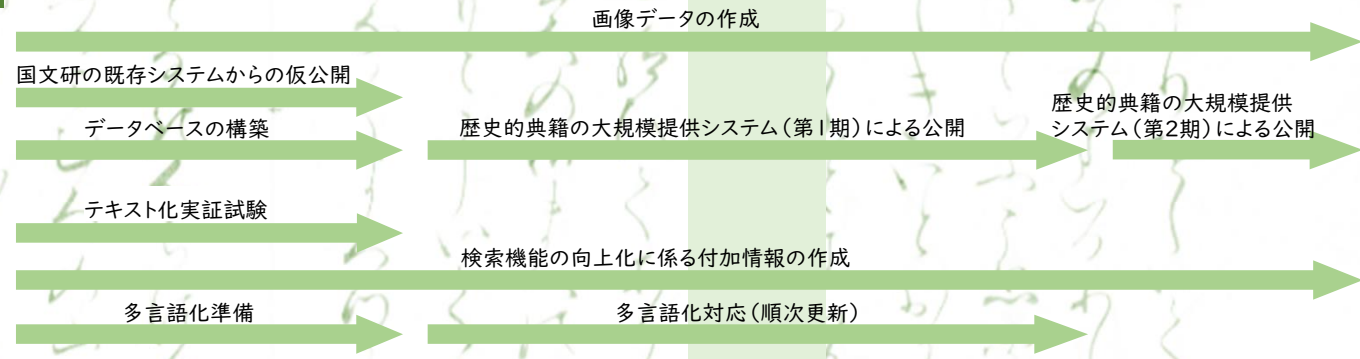
R03

R04

R05

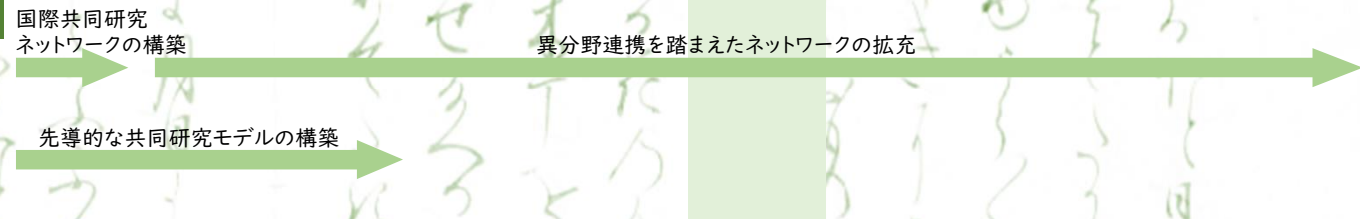
### 1 「日本語の歴史的典籍データベース」の構築

- ◆ 画像収集・データ化による**30万点の画像データの作成**(当面約24万点)
- ◆ 大規模提供システムの運用(第1期・第2期)  
※第1期公開におけるユーザの要望等を踏まえたシステム改修を進め、第2期公開に移行
- ◆ テキスト化実証実験
- ◆ 検索機能の向上化に係る付加情報(タグ)の作成。多言語化対応



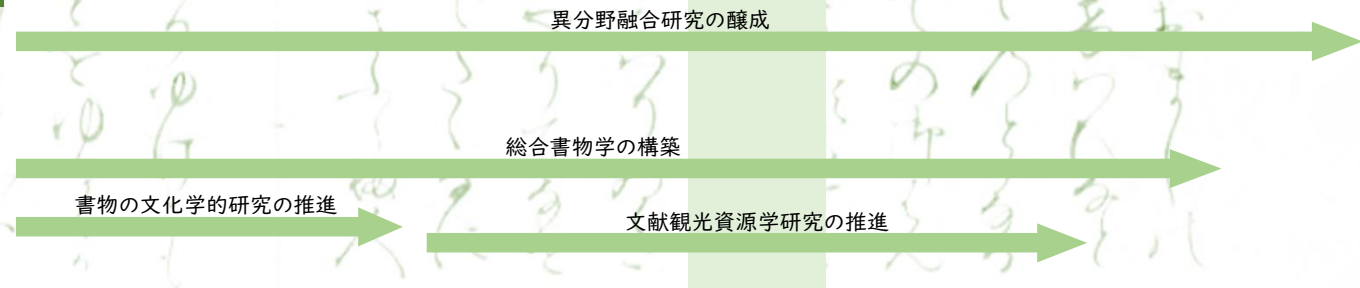
### 2 国際共同研究ネットワークの構築

- ◆ 連携機関との共同研究体制(国内については画像作成を含む体制)の構築
- ◆ 人文学分野における先導的な共同研究モデルの構築(体制構築、**異分野融合研究の開拓・推進**)



### 3 国際共同研究の推進

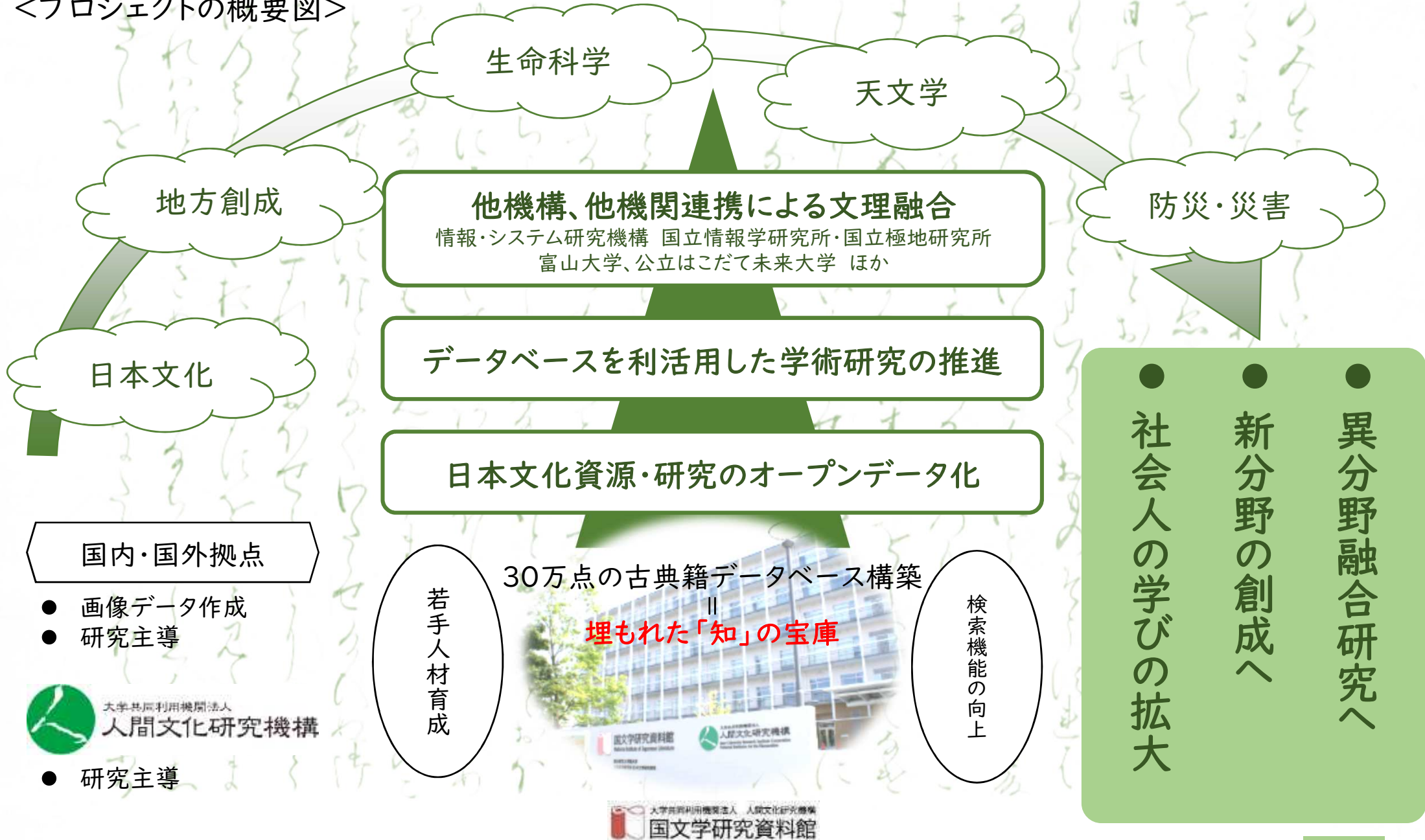
- ◆ 国際共同研究ネットワークを活用した異分野融合研究の醸成(文学と宗教など、幅広い領域(言語学、自然科学や社会科学にアプローチ、貢献しうる相互理解(相互連携))の構築
- ◆ 同時代に著作された古典籍を横断的に集約し、その時代時代の人々の精神思考を解明する総合書物学の推進・構築  
総合書物学の研究の一領域として、書物の文化学的研究や文献観光資源学等の推進







<プロジェクトの概要図>



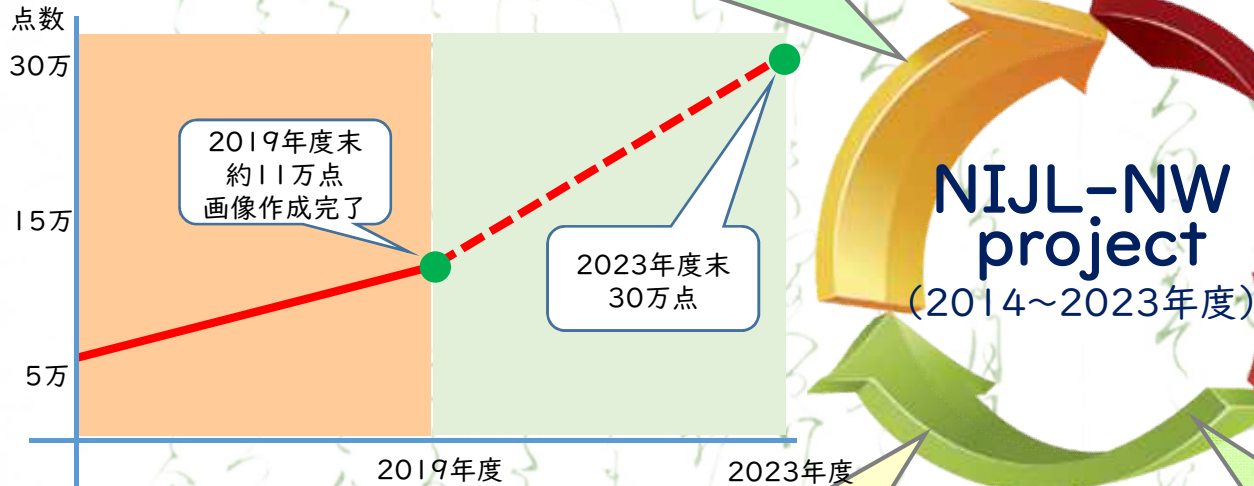
## &lt;平成31（令和元）年度の主な取組&gt;

画像作成の効率化

- ◆ 画像作成の**内製**による**コストを削減**、既存画像の活用を図るなど効率化**37,000点**（うち内製化16,000点）

未来志向の  
研究基盤整備新領域の開拓

- ◆ これまでに実施した**国際共同研究**の成果を踏まえて研究課題を設定した**国際共同研究**を引き続き実施し、日本古典籍の新たな魅力を世界に発信

NIJL-NW  
project  
(2014~2023年度)

宗教

産業

武術

新日本  
古典籍  
総合DB

理学

医学

共同研究による  
現代的課題の解決シチズンサイエンスに  
繋ぐ活動推進成果の発信から往還へ

- ◆ 分野を超えて研究と社会を繋ぐ若手人材の養成
- ◆ シンポジウムやワークショップの開催・参加を拡充
- ◆ ホームページやSNS、**オンライン・ジャーナル**による情報発信等を継続

思いもつかぬ  
〈発見〉へ検索機能の強化

- ◆ 情報・システム研究機構等、国内の大学・研究機関との共同研究を実施
- ◆ **タグ付けの加速化**
- ◆ **絵検索技術のステップアップ**

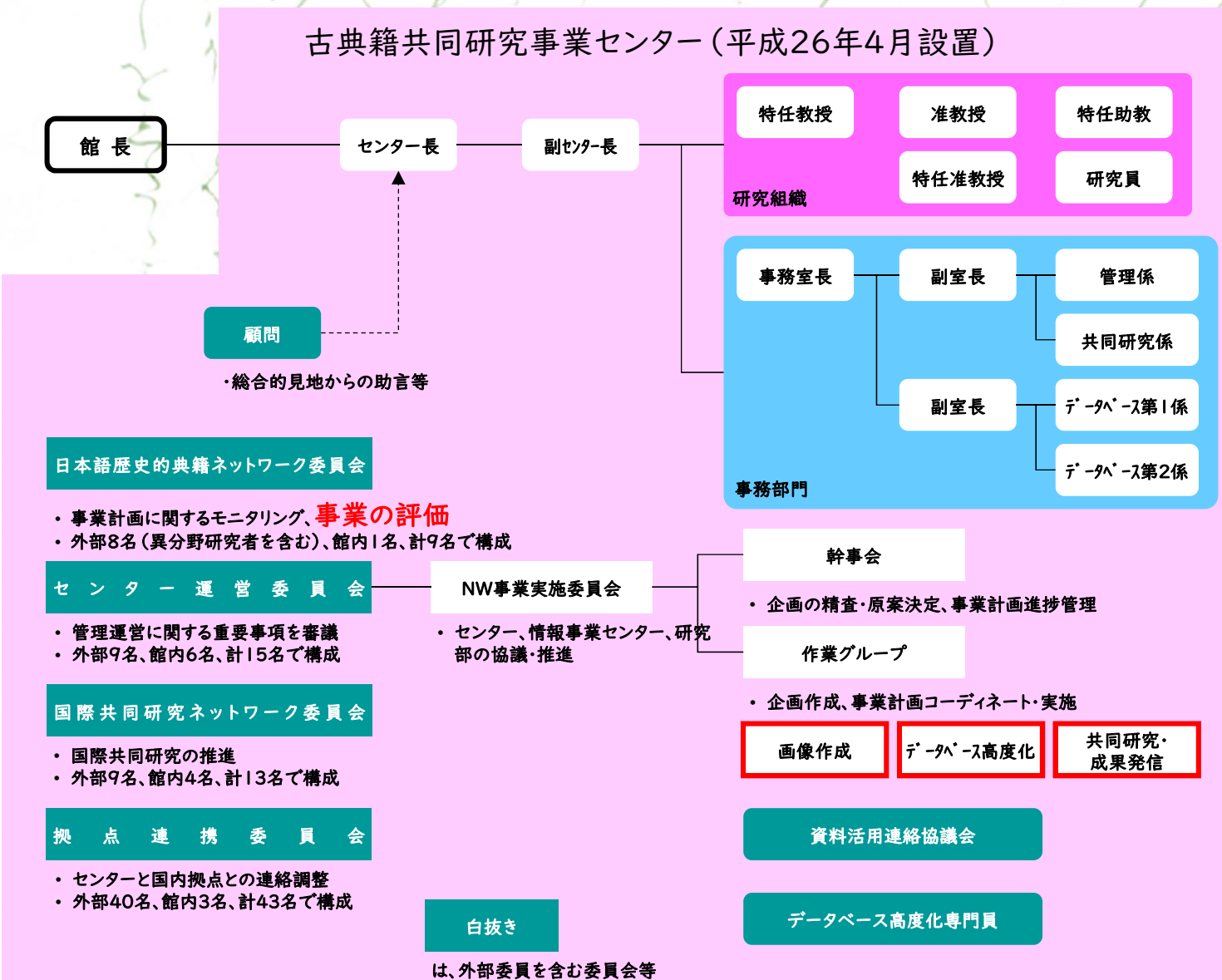


## &lt;大規模データベースの構築により期待される研究の広がり&gt;

関連領域	研究内容例	発展・展開することにより 将来的に得られる成果の例
総合	様々な分野の歴史的典籍の中に記載されている特定の事象や事物に関する情報を総合することによって、その事象や事物に関する <u>失われた知識や技術を復現・再構成</u>	『延喜式』に記載された古代の技術の復現など
防災	気候・天文等に関する歴史的典籍から <u>過去の災害の状況や要因を解明</u>	災害の予知予防に貢献など
医薬	<ul style="list-style-type: none"> <li>● <u>鍼灸や本草学に関連する歴史的典籍から、伝統医療について究明</u></li> <li>● <u>本草学の歴史的典籍から、現在は活用されていない天然資源の有用性等について究明</u></li> </ul>	伝統的な医学的手法の究明による新たな医療法・新薬の開発など
農	<u>農業技術や種苗技術の伝播経路を歴史的典籍から読み解き、その原産環境を特定し、日本ブランドの農作物の品質向上に寄与</u>	農業技術向上の支援など
生活・文化	名称しか伝わっていない料理について、 <u>材料及び調理方法を歴史的典籍から解明</u>	江戸時代の料理の復現など
観光	特定の地域を対象として、様々な分野の歴史的典籍の中に散在する地域情報を、地域の大学・自治体・企業などと連携し、最新のデジタル技術により網羅的に再編成することにより、 <u>新たな地域情報資源を開発し、地方再生に貢献</u>	新たな観光資源の開発など
地理	多様なジャンルの歴史的典籍に記された日本の地名とそれに関する情報を集積・整理し、 <u>歴史の中で育まれた多彩な文化空間としての日本を再構成</u>	総合地誌研究への貢献など



### <組織体制>



### <機構内連携機関>

- 国立歴史民俗博物館
- 国立国語研究所
- 国際日本文化研究センター

### <国内拠点(20機関)>

- 北海道大学
- 筑波大学
- お茶の水女子大学
- 京都大学
- 神戸大学
- 広島大学
- 慶應義塾大学
- 立教大学
- 大谷大学
- 立命館大学
- 東北大学
- 東京大学
- 名古屋大学
- 大阪大学
- 奈良女子大学
- 九州大学
- 國學院大学
- 早稲田大学
- 同志社大学
- 関西大学

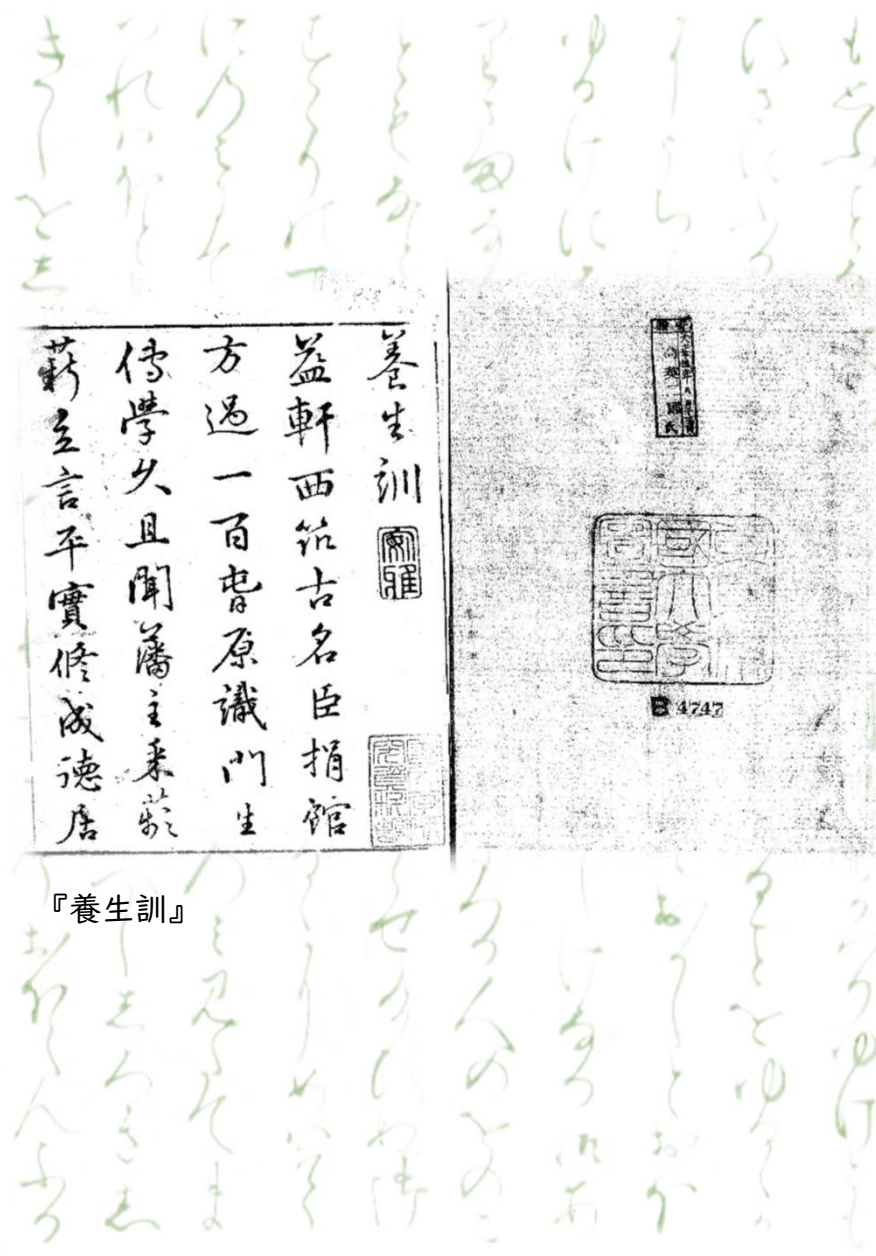
### <国外拠点(18機関)>

- コレージュ・ド・フランス日本学高等研究所
- コロンビア大学
- ライデン大学
- ナポリ大学オリエンターレ
- サピエンツァ・ローマ大学
- ブリティッシュ・コロンビア大学
- ケンブリッジ大学
- カリフォルニア大学バークレー校東アジア図書館
- ハワイ大学マノア校
- ヴェネツィア・カ・フォスカリ大学
- ゲーテ大学フランクフルト・アム・マイン
- 大英図書館
- フィレンツェ大学
- 高麗大学校
- 北京外国語大学
- バチカン市国図書館
- ベルリン国立図書館
- ハイデルベルク大学

### <連携機関>

- 情報・システム研究機構(国立情報学研究所、国立極地研究所、データサイエンス共同利用基盤施設)
- 茨城大学地球変動適応科学研究機関
- 弘前大学
- 国立国会図書館 ほか





『養生訓』

# 進捗状況

## 「日本語の歴史的典籍データベース」の構築

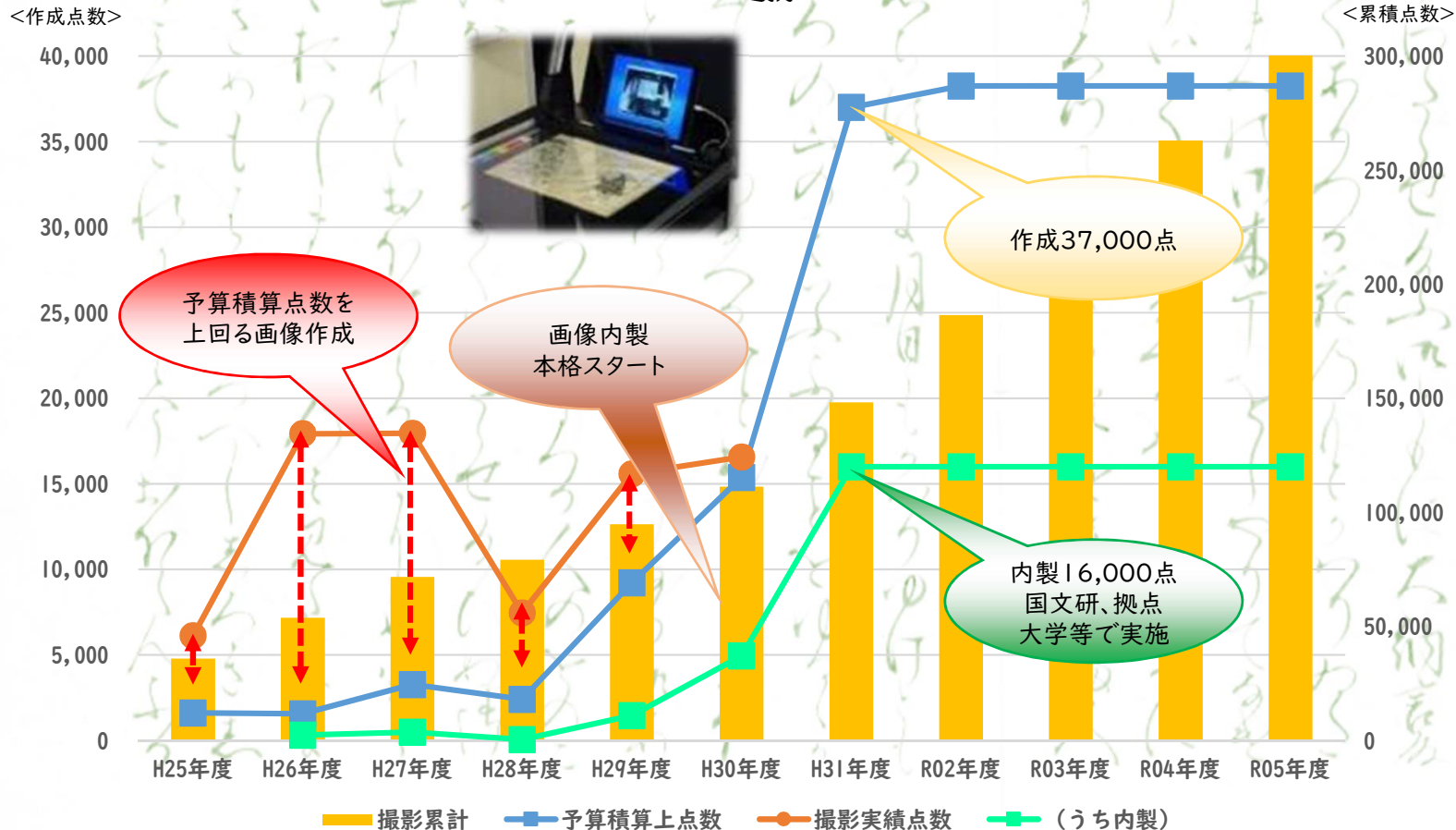


## ● 画像作成における5つの取組

- 業者撮影
- **内製**
- デジタルコンバート
- 既存画像活用
- IIIF連携

## ● 画像公開における4つの取組

- i デジタルオブジェクト識別子 (DOI) の採用
- ii クリエイティブ・コモンズライセンスの表示
- iii トリプルアイエフ (IIIF) の導入
- iv タグや全文テキストによる **多彩な検索方法の導入**



## ● 画像公開における二つの連携

JapanKnowledge



日本最大級の辞書・事典サイトである  
**「ジャパンナレッジ」と連携**

CiNii Books  
 全国の大学図書館の所蔵資料を検索する国立情報学研究所「サイニイブックス」との連携

## □ DBへのアクセス数

H29年度	H30年度
566,106	1,233,083

約2.2倍





# 新日本古典籍総合データベース

(平成29年4月試験公開、10月正式公開)



<https://kotenseki.nijl.ac.jp>

## 画像検索機能の向上化

- ・書誌情報(書名著者名)検索
- ・タグ(キーワード)検索
- ・テキスト本文による検索

## くずし字検索を試験的導入(ワードスポッティング)



## <世界中で利用されるデータベース>

	2017.4~2018.3	2018.4~2019.3
アクセス数	566,106	1,233,083
ユーザー数	31,462	48,461
(うち海外)	67カ国・地域 2,858	62カ国・地域 3,643



新日本古典籍総合データベースを活用した講義風景 (於: ヴェネチア大学)

カリフォルニア大学バークレー校では、「新日本古典籍総合データベース」の資料をベースに、世界的利用に供するデータの集積とインストールを構築(国際共同研究「UCバークレー所蔵古典籍資料のインストールとキュレーション」(2018~2020))

## 国際共同研究ネットワークの構築



- 国内外の大学・研究機関等との連携は、国外12機関、国内20機関を拠点としてスタート
- 平成27年度にカリフォルニア大学バークレー校東アジア図書館、平成28年度はベルリン国立図書館、ハワイ大学マノア校、ハイデルベルク大学、平成30年度にゲーテ大学フランクフルト・アム・マイン、さらに令和元年に大英図書館を加え国外18機関と連携しネットワークを拡張

さらに、

- **機構を超えた連携体制の整備**を図るため、国立情報学研究所及び国立極地研究所等との協働関係を構築
- 国立国会図書館とも協定書を取り交わして連携体制を強化
- その他、宮内庁書陵部等の公的機関のほか、日本近世文学会、絵入本学会、日本医史学会、日本武道学会等の研究者コミュニティとの連携体制を構築
- さらに大規模画像データベースの検索機能の向上化等を目指し、凸版印刷株式会社等**産業界との連携**を推進

86 機関と関係深化



## 国際共同研究の推進

## 国際共同研究等実施件数

種別	件数
国際共同研究	5件
公募型共同研究	5件
拠点主導共同研究	4件
国文研主導共同研究	7件
機構内連携共同研究	4件
異分野融合研究	9件
6種	34件
研究開発系共同研究	10件



表紙に人の毛髪が混じっていた江戸時代の書籍

古典籍に混入された毛髪で江戸時代の食生活を分析  
江戸時代の書籍に使われた再生紙に混じる人の毛髪を分析し、当時は現代よりも米、野菜、海水魚に大きく依存していたことを解明。2018年8月、英国オンライン科学誌「Scientific Reports」に掲載



朝日新聞 2017年4月12日

古典籍に埋もれた知の発見(異分野融合研究)  
藤原定家「明月記」の記述から、鎌倉時代における連発巨大磁気嵐の発生パターンを解明、アメリカ地球物理学連合「Space Weather」に掲載、新聞各紙にも取り上げられた(2017年3月)。  
2017年9月、江戸時代のオーロラ絵図と日記の記述から、観測史上最大の磁気嵐キャリントン・イベント(1859年)を上回る史上最大の磁気嵐を推定。再び「Space Weather」に掲載



大規模な磁気嵐発生の問題は巨大磁気嵐に脆弱な現代のIT社会において重視すべき問題で、本研究は1000年ぐらいのスケールで、**どれだけ大きな磁気嵐が起きるかという一つの見通しを提供する可能性があり、社会・経済・安全保障等、広い範囲で大きな意味を持ち社会的意義が高い。**

関係機関数	研究者の分野			
197大学・機関	日本文学	文化人類学	天文学	風俗学
	政治学	人文地理学	香道	農学
	歴史学	中国文学	気候学	食生活学
	民俗学	医学	図書館情報学	植物学
	情報学	数学	建築学	生態学
	経済学	本草学	生物学	社会心理学
	言語学	思想史	薬学	環境学
	芸術学	宗教学	工学	地質学
	哲学	考古学	教育学	土木工学
	博物館学	人間情報学	地理学	情報工学

### <国際共同研究等の成果(一例)>

#### ◆ 深層学習によりスケッチから類似の挿絵を探す技術を開発(国立情報学研究所)

(1) 手描きでスケッチを  
すると

(2) 似ている古典籍を  
データベースから探す

(3) 検索結果をドラッグ&ドロップすると  
「結果で再検索」ができる

#### 江戸時代の料理書・料理本約 300点をオープンデータ公開



The IAPR Best Paper Award of HIP2017  
において Best Paper Award 受賞

「第21回電子情報通信学会PRMUアルゴリズムコンテスト」に、国文研が提供した字形データを利用し「古典籍のくずし字認識」が課題となった。そのコンテストで優勝した研究が、国際会議The 4th International Workshop on Historical Document Imaging and Processing (HIP2017)で高い評価を受けた。

凸版印刷株式会社、株式会社PFUなど  
民間企業と研究開発系共同研究を実施



共同研究のプロセスで得られた市場調査等  
種々の活動が、企業活動に活かされ、さらに  
技術者の人材育成という観点でも貢献



くずし字学習アプリ  
「KuLA」  
50,000超DL



セミナー「江戸の書物から読  
み解く「食」の楽しみかた」  
(於:日本橋三越本店)

クックパッドとのコラボ

三越グルメ!  
MITSUKOSHI  
GOURMET

#### 三越伊勢丹が国文 研とコラボした料 理を販売

江戸時代、天保10年(1839年)と、料理の文化が盛んな江戸時代には、100冊以上の料理書が出版された。当時の人々の生活様式は、今とは異なり、江戸時代特有の生活様式が色濃く残っている。そんな江戸時代の生活様式を、あなただけのオリジナルにアレンジし、あなただけの江戸時代を再現しようとした。『江戸時代のくずし字認識技術』が、あなただけの江戸時代を再現しようとした。『江戸時代のくずし字認識技術』が、あなただけの江戸時代を再現しようとした。『江戸時代のくずし字認識技術』が、あなただけの江戸時代を再現しようとした。

