

国文学研究資料館古典籍データ駆動研究センターにおいては、大規模学術フロンティア促進事業「データ駆動による課題解決型人文学の創成」（以下、「プロジェクト」という。）の趣旨に即して、国文研プロジェクト型共同研究を実施しています。

ついては、令和 8（2026）年度に新規実施する共同研究（重点課題研究）の提案を募集します。

記

1. 募集テーマ

以下に挙げる各テーマに関する研究課題について募集します。

- ①古典籍画像からのテキストデータ自動生成のための手法開発と実践
- ②自動生成テキストデータを活用した検索手法の開発
- ③多様なテキスト資源を活用したテキスト校正方法の検討と実証
- ④自動生成テキスト及び校正テキストの活用事例研究

2. 申請資格

申請資格を有する者は、次の（１）（２）のいずれかに該当する者としてします。

- （１）大学又は研究機関に所属する常勤の研究者
- （２）上記（１）以外の研究者（非常勤講師、日本学術振興会特別研究員、等）

※ただし、府省共通研究開発管理システム（e-Rad）で資格停止の措置を受けている研究者は申請できません。

3. 研究組織

「研究代表者」と「研究分担者」で構成します。

○研究代表者（申請者）：共同研究の遂行に当たってすべての責任を持つ者

○研究分担者：研究代表者とともに、共同研究の遂行に責任を持つ者

※「研究分担者」には、「館内担当者」として当館教員（１名）を含めてください。館内担当者は研究代表者と連絡を取りつつ共同研究の遂行をサポートしますので、「研究代表者」は、申請書を提出する前に館内担当者と十分な打ち合わせをしてください。なお、当館の教員が館内担当者として参画しない共同研究は、審査の対象外とします。

※大学院生（研究開始時点で大学院に在籍することが見込まれる者も含む。ただし、研究開始時点において在籍していない場合は除外する。）も「研究分担者」に含めることができます。若手研究者育成のためにも、大学院生の積極的な参加を期待します。

※研究会や資料調査に「研究協力者」を随時参加させることができます。「研究協力者」は研究組織に含めません。

4. 研究期間

令和8（2026）年度～令和9（2027）年度の範囲

※研究の成果等の状況に基づき、上記を超えて研究期間を延長できる場合があります。

※令和9（2027）年度以降の研究継続の可否については、研究の進展等に基づき改めて審査されます。今年度の審査結果をもって、次年度以降の研究期間が保証されるものではありませんので、ご注意ください。

5. 研究経費

（1）各年度分の上限は、100～300万円を目安とします。

※金額設定の上限は以下を目安としてご検討ください。

- ・研究目的が、技術基盤の開発等のための研究：100万円程度
- ・研究目的が、技術の実用化に向けた実証研究：300万円程度

※必要に応じて上限を超える申請も可能とします。

※実際の配分額は減額調整する場合があります。また、予算の有効利用のため、進捗に応じて分割して配分する場合があります。

※次年度分は申請時点での予定を記入してください。ただし、次年度の要望額を制約するものではありません。

※本共同研究は、大規模学術フロンティア促進事業の趣旨に即して実施することから、同事業の仕組み上、次年度以降の配分額を保証するものではありません。

（2）研究経費の配分方法（送金等）や留意点については、別紙1「国文研プロジェクト型共同研究の研究経費配分・留意点等について」をご参照ください。

6. 採択予定件数

募集テーマ毎に若干件数

7. 申請書類

「研究計画書（重点課題研究）」

8. 留意点

（1）研究目的、研究計画の策定にあたっては、別紙2「重点課題研究の立案にあたっての留意点・重要事項」に明示した事項に沿ったものとなるよう、ご注意ください。

（2）「連絡調整員」（菊池信彦・国文学研究資料館准教授）が、プロジェクトと本共同研究の間の連携を図ります。

9. 提出方法

申請書類を電子データにて【pjsuishin（あっと）nijl.ac.jp】宛てにお送りください。

※スパムメール等の対策のため、e-mail アドレスには、「@」の代わりに「(あっと)」を

入れています。メール送信の際は、「(あっと)」を「@」に換えて送信してください。

10. 申請期限

令和8（2026）年2月2日（月）17時（必着）

11. 採否結果の通知

- （1）採否及び配分額は、古典籍データ駆動研究センターにおける書面審査及び国文学研究資料館諸会議における審議を経て決定し、その結果を申請者にメールで通知します。
- （2）採択された研究課題の構成員には、当館の「共同研究員」としての委嘱を別途行います。

12. 研究成果の公開及び報告

- （1）「研究代表者」は、合同成果報告会（年度末開催、ハイブリッド（予定））に参加し、活動状況及び研究成果等の報告を行うものとします。（研究代表者が参加できない場合は、研究組織構成員から代理を選出いただきます。）
- （2）「研究代表者」は、各年度末に「共同研究年次報告書」を作成し、それぞれ所定の期日までに提出するものとします。なお、「共同研究年次報告書」に記載された内容は当館ウェブサイト等で公表する場合があります。
- （3）本共同研究で作成されるデータ・プログラム等の研究成果は、共同研究契約に基づき、原資料の所蔵者による制限等がない限り、原則としてオープンデータとして公開することとします。

13. 公的研究費の不正使用防止等に関する取組への協力

人間文化研究機構では、文部科学省が定めた「研究機関における公的研究費の管理・監査のガイドライン（実施基準）」及び「研究活動の不正行為への対応等に関するガイドライン」の趣旨に鑑み、機構における不正防止体制等を整備し公表しております。

<http://www.nihu.jp/ja/opendoor/fusei>

「共同研究員」は、共同研究を行うに当たって、機構及び当館の規程その他の定めを遵守しなければなりません。また、機構及び当館が実施する不正防止に関する取組を遵守いただき、監査等に協力していただきます。

なお、「共同研究員」の委嘱の際に、公的研究費の不正使用防止に関して「誓約書」の提出をお願いしております。

14. 申請書類に含まれる個人情報の取扱い等

本募集に関連して提出された個人情報については、審査の目的に限って利用し、審査終了後、全ての個人情報は責任を持って破棄します。

15. 問い合わせ先

〒190-0014 東京都立川市緑町10-3

大学共同利用機関法人 人間文化研究機構

国文学研究資料館 管理部総務課 データ駆動型研究係

電話 050-5533-2988 (直通)

FAX 042-526-8604

e-mail pjsuishin (あっと) nijl.ac.jp

※スパムメール等の対策のため、e-mail アドレスには、「@」の代わりに「(あっと)」を入れております。メール送信の際は、「(あっと)」を「@」に換えて送信してください。

★当館は、「国文学研究資料館におけるダイバーシティ宣言」に基づき、ダイバーシティに関する取組を推進し、採用を行います。

国文研プロジェクト型共同研究の研究経費配分・留意点等について

1. 共同研究契約及び機関への送金について

研究代表者又は研究分担者が希望する場合は、当館と所属研究機関との間で「共同研究契約」を締結することにより、研究経費を送金いたします。

送金を希望する場合は、送金先研究機関の共同研究契約担当事務と事前に調整の上、事務的な対応の可否等をご確認ください。不明な点がある場合は、以下の当館担当部署へご相談願います。

2. 留意点

(1) 国文研プロジェクト型共同研究においては、研究経費として、直接経費（共同研究の遂行に必要な経費（研究成果の取りまとめに必要な経費を含む。))を配分します。間接経費（共同研究の実施に伴う研究機関の管理等に必要な経費）は配分されません。

※上記 1. に基づき送金を希望する場合、本要件についても、送金先研究機関の共同研究契約担当事務へ事前にお伝えくださいますよう、ご留意願います。

(2) 直接経費は、次の費用として使用できません。

- ①建物等の施設に関する経費（直接経費により購入した物品を導入することにより必要となる据付等のための経費を除く。）
- ②共同研究遂行中に発生した事故・災害の処理のための経費
- ③研究代表者又は研究分担者の人件費・謝金
- ④上記のほか、直接経費を使用することが適切ではない経費

(3) 送金した研究経費については、送付先研究機関において当該年度に係る『収支報告書（様式自由）』を作成いただき、翌年度上旬（4月末を予定）までに当館へご報告いただくこととなります。本要件についても、送金先研究機関の共同研究契約担当事務へ事前にお伝えくださいますよう、ご留意願います。

(4) 上記のほか、研究経費の執行に当たって不明な点がある場合は、以下の担当部署へご相談願います。

<当館担当部署>

管理部総務課データ駆動型研究係

E-mail pjsuishin (あっと) nijl.ac.jp

※スパムメール等の対策のため、e-mail アドレスには、「@」の代わりに「(あっと)」を入れております。メール送信の際は、「(あっと)」を「@」に換えて送信してください。

重点課題研究の立案にあたっての留意点・重要事項

国文研プロジェクト型共同研究(重点課題研究)は、大規模学術フロンティア促進事業「データ駆動による課題解決型人文学の創成」(以下、「プロジェクト」という。)の趣旨に即し、プロジェクトにおける年次計画に関連する重要事項を研究テーマとするものである。

このことから、研究目的、研究計画の策定にあたっては、以下に明示する留意点・重要事項を踏まえ、これに沿ったものとなるよう十分留意すること。

記

基本的事項

プロジェクトにおいては以下の基本的事項に従って事業を実施することとしており、本共同研究にあたってはこれらを前提とする。

ただし、L2 は今回の重点課題テーマの対象には含まれないことに注意されたい。

- 1) プロジェクトにおいては、歴史的典籍(以下、「資料」という。)のテキストデータ作成を、以下の L0、L1、L2 のレベルに分けて実施する。

- OCR によるテキストデータ自動生成

- L0)OCR/機械処理等によるおおまかな全文テキスト化

- 校正テキストデータの作成・公開

- L1)プレーンな翻刻(人による確認を経たもの)

- L2)校訂本文、釈文、現代語訳、翻訳等(TEI 形式のデータを含む)

- 2) 各レベルにおいて、それぞれ次の資料を、テキスト化の対象とする。

L0)国書データベースから画像公開するもの全点。

L1)以下の①～③に該当するもの。教科書掲載箇所等、作品の一部を含む。

①L0 のうち、教育目的または研究への波及効果が高いと当館が認めるもの

②当館が保有または今後作成する翻刻テキスト等が利用可能なもの

③外部機関・研究者等から入手可能な既存または今後作成予定の翻刻テキスト等が利用可能なもの

L2)L1 のうち、当館が認めるもの

- 3) 各レベルにおいて、それぞれ次の形式でテキストデータを作成する。

L0/L1)資料のページ単位、および記載された行単位でテキストを作成する。

L2) TEI (Text Encoding Initiative) 等、用途目的に沿った形式でテキストを作成する。

テーマ毎の事項

各テーマに関して想定される事項を以下に記す。研究計画の立案にあたっては、必ずしもこれらに厳密に従う必要はないが、趣旨に沿った提案を期待する。

① 古典籍画像からのテキストデータ自動生成のための手法開発と実践

古典籍画像を対象とした既存の OCR を上回る精度の実現や、これらで十分な精度を得ることが難しい種類の古典籍画像に適用可能な新たな手法の開発、複雑なレイアウトのテキストの自動整列手法などの実現を目指す。対象とする古典籍画像データは当館より提供する。

② 自動生成テキストデータを活用した検索手法の開発

古典籍画像から自動的に生成したテキストデータを利用して、通常のキーワード検索や全文検索だけでなく、高度な検索機能や生成 AI を利用した対話インタフェースなどの実現を目指す。

③ 多様なテキスト資源を活用したテキスト校正方法の検討と実証

翻刻テキスト等の既存のテキストを有効活用して L1 作成を行う際に必要となる技術的課題を分析し、対処するための方法を開発することを目指す。例えば翻刻テキストと OCR 結果との統合（融合）を行う場合の技術的手法の実現や実用化のための課題解決を目指す。対象とする翻刻テキスト等については、館内担当者と相談して実際の利用可能性に配慮して選定すること。

④ 自動生成テキスト及び校正テキストの活用事例研究

L0/L1 テキストデータについて、活用事例等の研究、あるいは研究振興につながる研究活動を行う。なお、対外的な実践活動等を行う場合、使用するデータについては館内担当者と相談して選定すること。

利用可能なデータ

各共同研究において利用できる、当館が提供するデータを以下に記す。提供対象の範囲や提供方法については、採否結果の通知後、採択された研究課題の研究代表者と個別に調整する。なお、本データは本共同研究の遂行の目的以外には使用しないこと。

① 古典籍画像データ

国書データベースに収録されているものの中から共同研究で利用可能なものを当館から提供する。

※本共同研究の範囲を超える場合は国書データベースを利用すること。当館所蔵資料の古典籍資料画像はパブリックドメインのため自由に利用できる。他機関所蔵資料の画像の利用条件は原資料の所蔵者によって異なるため、所蔵者が指定した利用条件を遵守し、資料所蔵者の権利及び利益を侵害することがないように十分留意すること。所蔵者毎の利用条件については、以下のウェブページを参照のこと。

<https://kokusho.nijl.ac.jp/page/usage.html>

② 自動生成テキストデータ (L0)

上記の古典籍画像データを対象として、NDL 古典籍 OCR により試行的に作成した自動生成テキストデータ (L0) を当館から提供する。

※本データは自動生成であり、文字認識誤りを含む。

③ 校正テキストデータ (L1)

当館で作成した古典籍の校正済みテキストデータを提供する。

※令和 8 年度当初より順次提供できるデータは古典籍約 340 点分の予定である。当館における L1 データ作成作業の進捗に応じて、令和 9 年度以降の提供点数は順次増加する見込み。

④ 国書 DB 書誌メタデータ

国書データベース上で公開中の国文研所蔵資料に付与している書誌データ

※書誌データとは、「国文研所蔵の『源氏物語』」・「●●大学所蔵の『源氏物語』」等、個別の（物理的な）資料毎のデータを言います。（同一書名だから同じとは限らないもの）

⑤ 国書 DB 著作メタデータ

国書 DB 上で公開中の全著作データを 1 ファイルにまとめたデータ

※著作データとは、『源氏物語』・『古今和歌集』等、作品レベルでのデータです。国書データベースでは、著作データに個々の資料情報（＝書誌データ）が紐付いています。

<書誌と著作について>

国書データベースでは、古典籍の個別資料の書誌情報を、内容によって著作情報に紐づけています。古典籍は、同じ作品（これを「著作」といいます）であっても、その本により異なる書名をもつことがあります（たとえば、『源氏物語』、『けんし』、『光源氏』、『きりつほ』、『うつせみ』など）。また、同じ書名であっても、異なる内容をもつことがあります（たとえば、紫式部の著した『源氏物語』とその注釈書が同じ『源氏物語』という書名である場合など）。そうした個別資料の情報を著作情報でとりまとめているため、同じ著作の個別資料同士を容易に見比べたり、作品の成立年代などの情報を得たりすることができます。

出典：ジャパンナレッジ (<https://japanknowledge.com/contents/kotensekidb/>)

その他に関する事項

本課題の研究組織構成メンバーは、テキストデータ作成事業に関して当館が組織する「連絡調整会（仮称）」の構成員となること。

- ・当館は、共同研究による研究成果の事業等への反映を模索するため、プロジェクトと各研究課題との連携を行う。これを踏まえ、「連絡調整会（仮称）」の構成員は、以下の方法による情報交換、連絡調整を積極的に行うこと。

<連絡調整会の主な活動（予定）>

- ・当館が指定するコミュニケーションツールを用いての日常的な連絡調整
- ・必要に応じて対面・オンラインによる相談会を実施