

データ蓄積とドメイン知識の問題：万葉集を例に

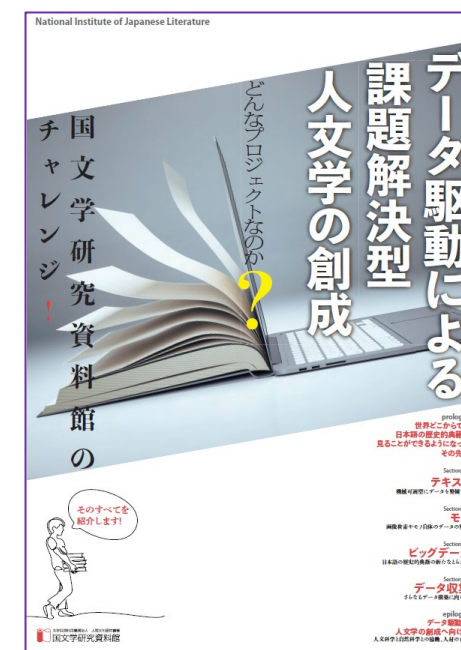
海野 圭介

国文学研究資料館

第7回日本語の歴史的典籍国際研究集会

2021年11月11日（木）18:00-21:30

オンライン：<https://www.nijl.ac.jp/pages/cijproject/sympo2021.html>



この報告の論点

- 国文研の考える新しい時代の大型プロジェクト
- データ駆動型の人文学の構想と視界
- 人文学の成果と知識をデータとして継承するには？

この報告の論点

- 国文研の考える新しい時代の大型プロジェクト
- データ駆動型の人文学の構想と視界
- 人文学の成果と知識をデータとして継承するには？

大規模蓄積からデータ駆動へ

現在推進している大規模学術フロンティア事業
**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**
平成26(2014)年度～令和5(2023)年度



後継プロジェクトとして立案
**データ駆動による
課題解決型人文学の創成**
10ヶ年計画（予算化に向けて取り組みを進めている）



『学術研究の大型プロジェクトの推進に関する基本構想 ロードマップの策定ーロードマップ2020ー』
文部科学省より資料公開（令和2年9月24日）

https://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/toushin/1388523_00001.htm

学術研究の大型プロジェクト

2018年～：（国文研で学術研究の大型プロジェクト立案のためのタスクフォース設置）

2019年2月：日本学術会議が、第24期学術の大型計画に関するマスタープラン公募開始

3月：国文研応募 課題テーマ「データ駆動による課題解決型人文学の創成」

8月：書類審査を経てヒアリング

2020年1月：「マスタープラン2020」公表 重点大型計画として国文研の計画が選定

2月：マスタープラン2020を踏まえ、文部科学省が「ロードマップ2020」公募開始

4月：（国文研は、計画立案タスクフォースを整理し「学術大型計画設置準備室」を設置）

8月：文部科学省学術審議会作業部会によるヒアリング実施

9月：文部科学省が「ロードマップ2020」を公表 全15件を策定 当該計画が人文社会系で唯一残る

2021年4月：（国文研は、人文学データ駆動研究センター設置準備室を併設し計画整備）

150件が
提案される

15件に
絞り込む

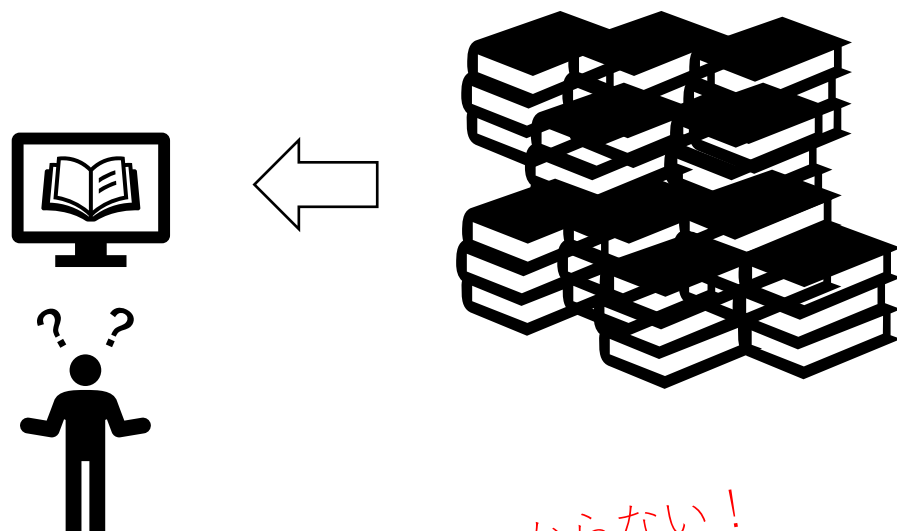
マスタープラン：科学者コミュニティの代表としての日本学術会議が、学術的意義の高い大型研究計画を広く網羅し体系化することにより、我が国の大型研究のあり方について、一定の指針を与えることを目的として策定するもの

(<http://www.scj.go.jp/ja/info/kohyo/kohyo-24-t286-1.html>)

ロードマップ：「マスタープラン」を踏まえつつ、文部科学省 科学技術・学術審議会 学術分科会研究環境基盤部会学術分科会 学術研究の大型プロジェクトに関する作業部会において策定 予算措置を保証するものではないが、関連施策を推進する上で十分考慮すべき資料として、大型プロジェクトの優先度を明らかにするもの

(https://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/021/1412963_00001.htm)

データ駆動ってなに？



必要な情報が見つからない！
こんなデータがあったらなあ！
同じタイトルが多すぎる！
自宅で見ると調査はエンドレス…

「データがそこにある」「いつでも見ること出来る」と「使える！」ことは別！

prologue

世界どこからでも
日本語の歴史的典籍を
見ることができるようになったその先に

データ駆動って？

「データ駆動」型の研究とはどのような研究でしょうか？ この方法は主に自然科学分野の研究領域で行われてきました。研究者が仮説を提示し、実験によって検証するのではなく、大規模データを収集し、そのデータに基づき解析するというアプローチをとります。現在では、人文学や社会科学の分野においても研究のスタイルを変化させ、新たな可能性を切り拓いています。人文学の分野においても、近年、デジタル人文学(digital humanities)が提唱され、人文学分野が生み出す大規模データの蓄積に基づく研究手法の構築と、その環境整備が進められています。人文学分野に存在するデータの駆動性を高め、その活用を促進するために、領域横断的に多くの研究分野の利用に耐えるデータ整備を進めることがいま求められているのです。

課題解決型って？

「課題解決型」とはどのような研究でしょうか？ 現代社会がかかえる大きな課題を見据え、その解決を総合的観点から推進する研究を「課題解決型」研究と呼んでいます。たとえば、環境問題のように、複雑に絡み合った問題には、多くの分野の研究者が参画し、結束して対処してゆかねばなりません。人文学分野が生み出す大規模データを、自然科学や社会科学の領域にも活用しうるように整備することで、さまざまな現代社会の課題に人文学分野の研究からも積極的に参画することができるようになるはずです。

データ駆動ってなに？

データ駆動

- ▶ 大規模データの収集と分析により新たな知を創発
- ▶ この方法は情報学・統計学の領域が推進
- ▶ 情報学分野と協働して人文学資料を分析
(再) 利用可能なデータを蓄積

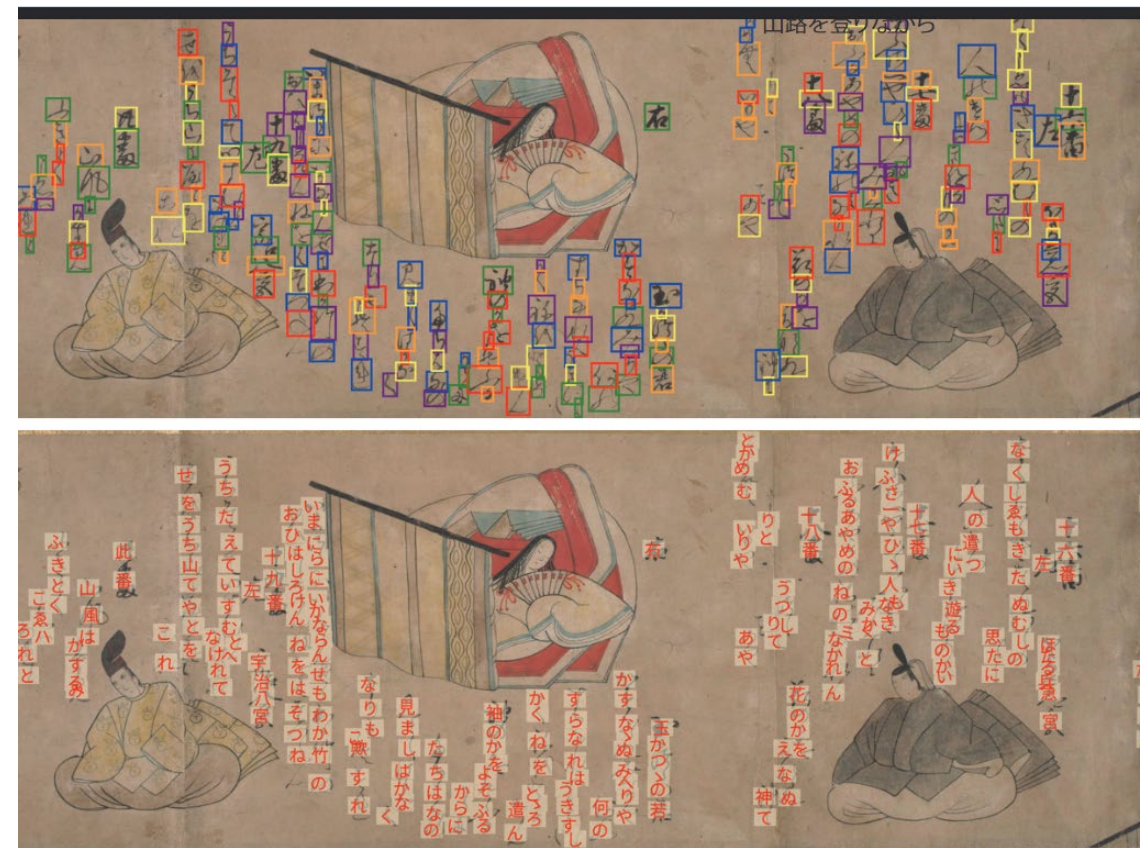
そのメリット

- ▶ 書物 vs データのように考える必要はない
- ▶ 細部へ向かう研究もあればマスで考える研究も
- ▶ 大規模データの存在を前提とした発想
新たなアイデアが研究を切り開く可能性も
- ▶ 大規模データ化は従来も進められている

従来のデータ化を振り返れば…

- ▶ 冊子体の『国書総目録』から古典籍総合目録DBに
- ▶ 国文研マイクロから新日本古典籍総合DBに
- ▶ ジャパンナレッジ、古典ライブラリーなど

→いずれもデータの駆動性が格段に向上



③ AI を用いた認識技術 (人文学オープンデータ共同利用センター (CODH) <http://codh.rois.ac.jp/> 提供)

情報学分野との協働のもっとも象徴的な事例は、
AIの導入によるくずし字認識技術の発展だろう
画像から文字データを検索して確認できるようになった

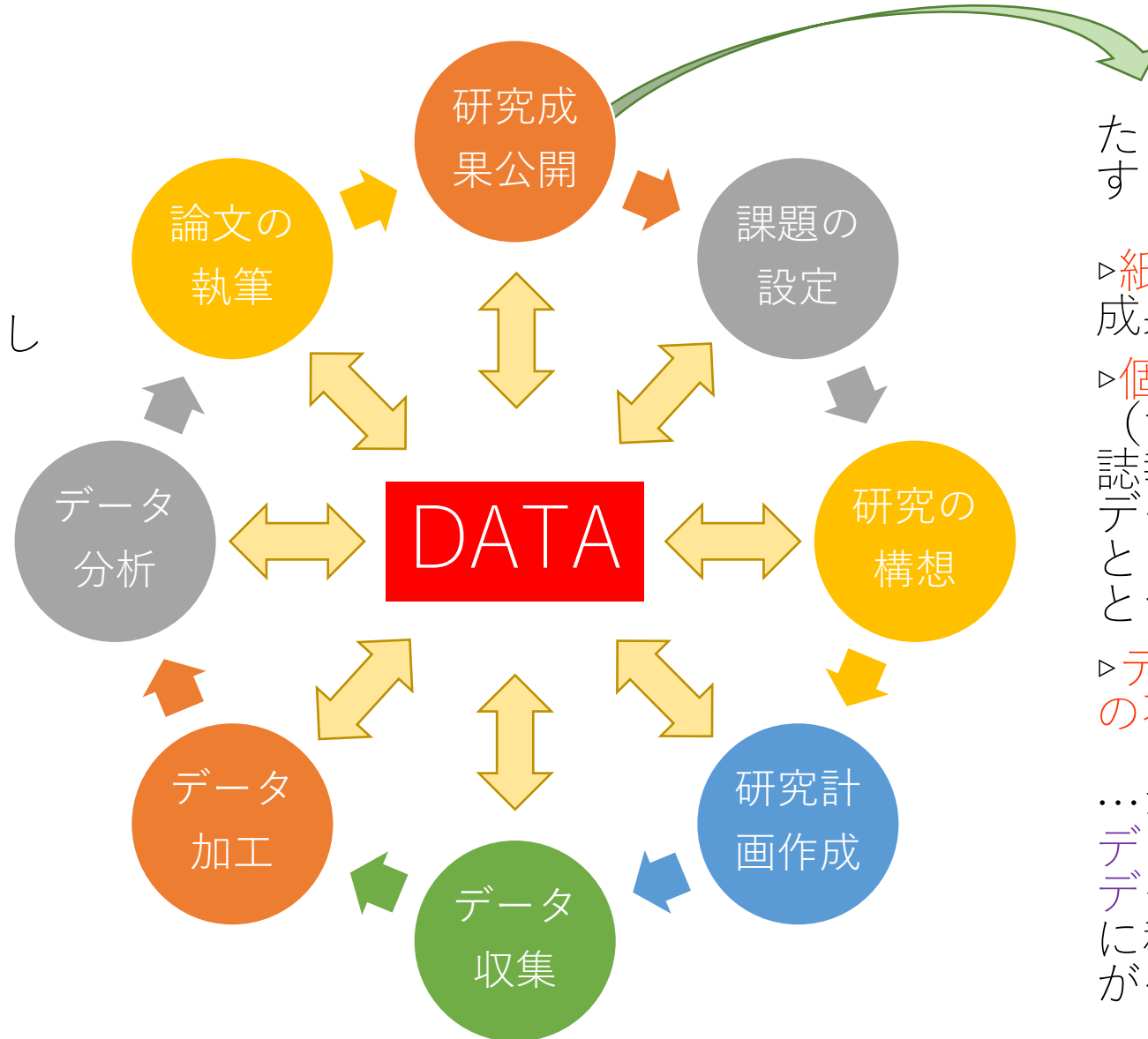
大規模蓄積からデータ駆動へ

たとえば
書物を閲覧し
研究を構想し
書誌カードや翻刻を作成し
内容を分析し
論文を書く

いずれの段階でも
データが出来る

…しかし

中間生成物としての
生データは表に出ない
→研究の過程からの
あらたな展開の派生
は望みにくい



たとえば「論文」ですら…

▷紙媒体による研究成果の蓄積

▷個々人の研究蓄積
(たとえば翻刻や書誌報告などの基盤的データも) はデータとして継承されることが現状では難しい

▷データ継承ルールの不在

…など、
デジタル環境での
データ蓄積
に移行しにくい要素
が少なくない

この報告の論点

- 国文研の考える新しい時代の大型プロジェクト
- データ駆動型の人文学の構想と視界
- 人文学の成果と知識をデータとして継承するには？

大規模蓄積からデータ駆動へ

日本語の歴史的典 国際共同研究ネットワーク構築計画

- ▶ 30万点の日本語の歴史的典籍の全文画像のオープンデータ化



- ▶ 国文研マイクロの自宅書庫化



- ▶ あらゆる分野の書物を公開
だからこそ…→文学研究の課題の複雑化・方法論の多様化や複合化へも対応可能

…歴史学・思想史・美術史・建築史・宗教史などとの連携・乗り入れはすでに通常の状態に
その先には？

データ駆動による 課題解決型人文学の創成

- ▶ 蓄積されたデータは人文学の重要なインフラ & データ

- ▶ 個別の研究蓄積を将来にわたって共有できるデータへ

- ▶ 人文学に求められる課題の複雑化への対応
→分野を横断し協働する視点や協働のための新たな方法が求められる

…そのためには、
共有できるデータ、当該分野に精通しない研究者にも参照可能なデータ構造が不可欠



データ駆動型人文学の視界：体感的には

こんな試みも…



ロックダウン下ではデジタル化が進まないと研究が進まない…という経験

細分化の進展、当該領域に着実に後継者を据えることが難しい現状
蓄積された成果は誰が使う？



進む実体社会のデジタル化

こんな試みも…



一方で…

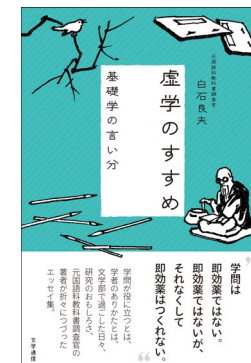


大学院生の減少…
学会員の減少…
それは…
古典文学に関心を持つ国民の減少??

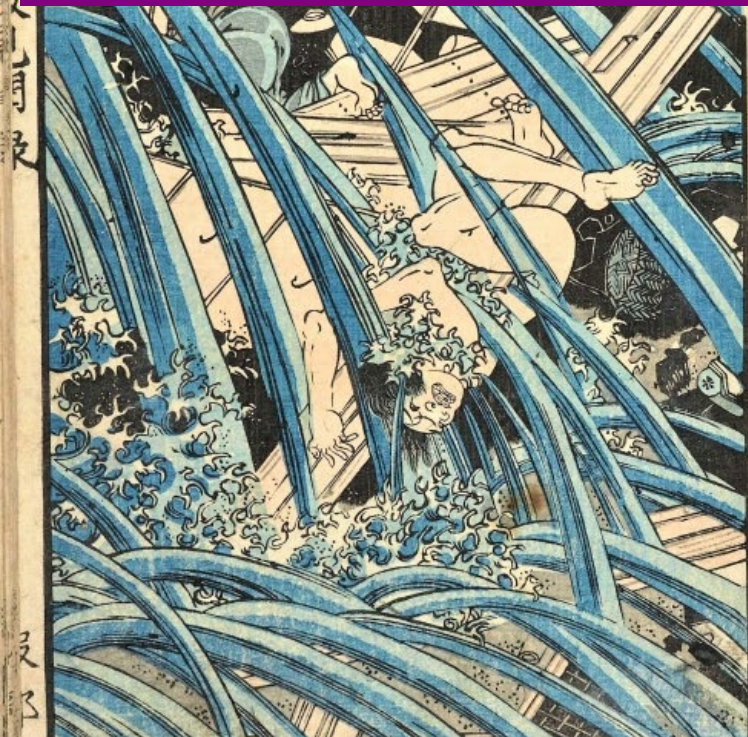
文系vs理系？



一方で…

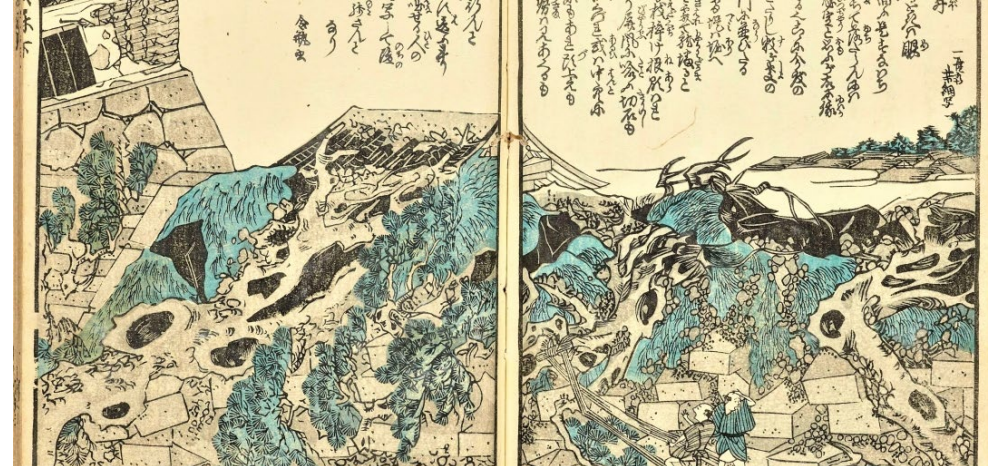


未曾有のパンデミックに 日本人は如何に対処してきたか？



中正長昆金

月喜



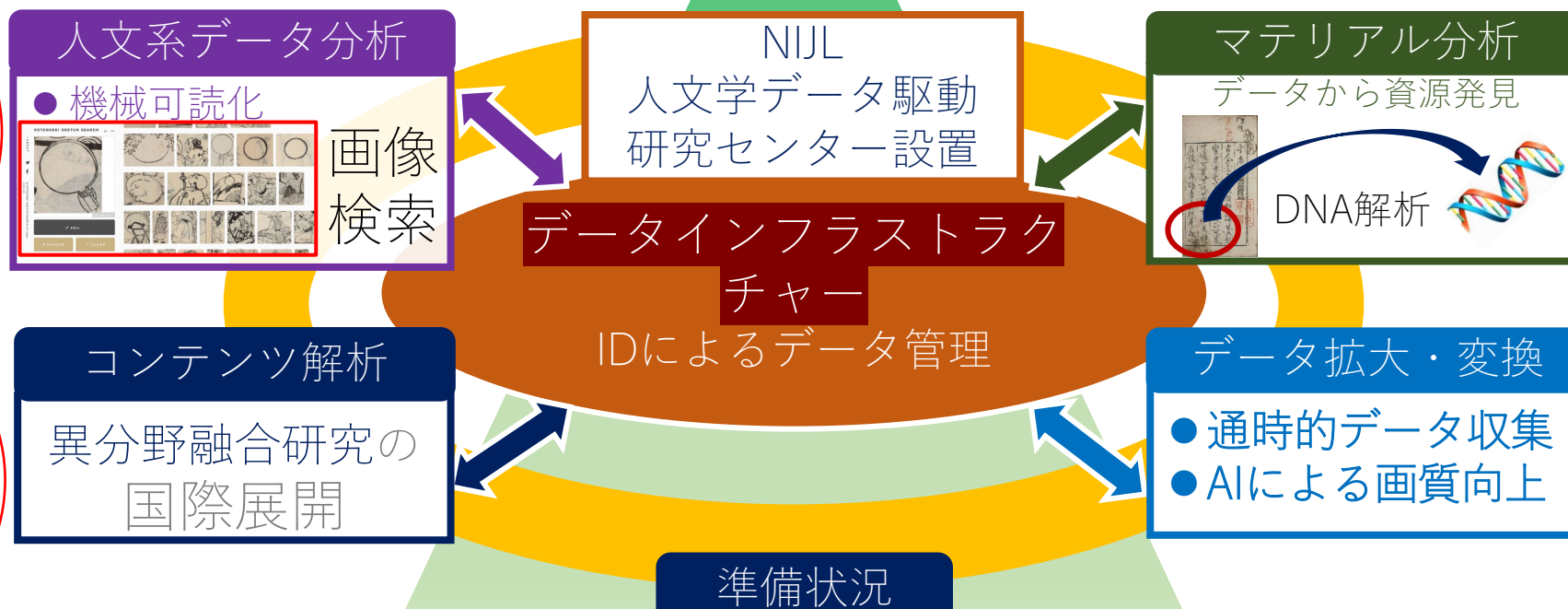
日本語の歴史的典籍の記録と記憶は、
同時に経済的・産業的・国際的価値を持つデータ

データ駆動型人文学の視界：共有・協働

典籍に基づく課題解決型人文学で持続可能な社会の実現へ

4つのパースペクティブから▲現代社会の課題解決へ

データ駆動型人文学の創成



例えば…
データ継承の
ルールは？

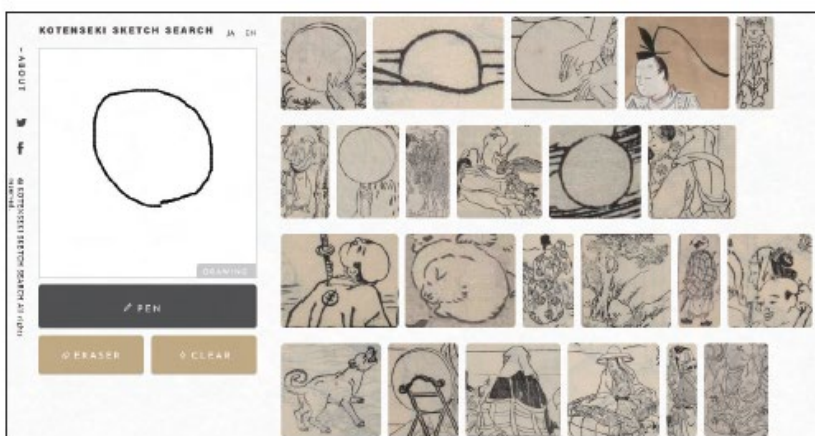
データ処理
の技術革新
ガイドライ
ンの策定
etc.

課題解決型
融合研究の
展開
etc.

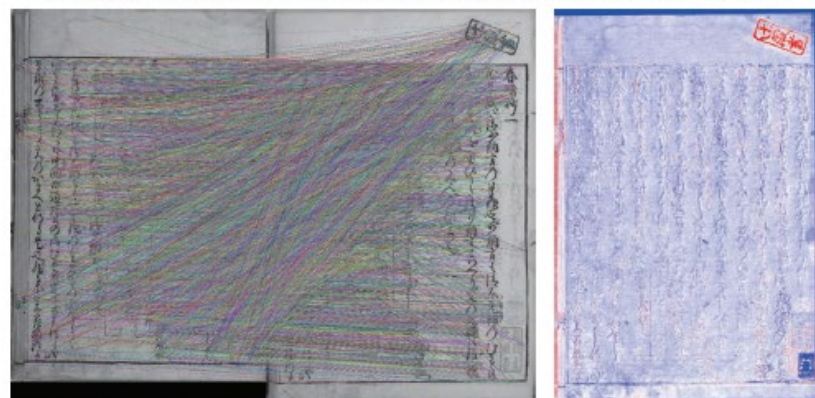
例えば…
デジタル化が難
しい質感は？

物質として
の書物に含
まれるデー
タの解析
etc.

データ拡充
変換技術の
革新
データイン
フラの開発



① 歴史的典籍の画像検索 KOTENSEKI SKETCH SEARCH http://lab.nijl.ac.jp/sketch_search/



② 冊の画像を位置合わせて比較することにより違いを浮かび上がらせる 左図はその処理の流れ 右図はその結果 違いが可視化される (CODH 提供)



③ 『自遣往来』(DOI: 10.20730/200022004)
見返しに毛髪が濃く染み込んでいる



④ 『自遣往来』(DOI: 10.20730/200019682)
左下の黒ずんだ部分は手垢が残った箇所

Section. 02 モノ

画像検索やモノ自体のデータの整備

スケッチで、画像で、検索

日本語の歴史的典籍の中には、文字だけではなく挿絵や模様が添えられているものがあります。こうした非文字情報は、通常考えられている文字列検索では見つけ出すことができません。またモノとしての書物のかたちや表紙の色などについても同様です。頭に浮かんだ図形をスケッチして似たかたちを探す①。国文学研究資料館は、国立情報学研究所と共同で非テキスト情報を探し出す技術の開発を行っています②。

見えない文字を読む、書物の光学解析

文字の上から墨で塗りつぶされた原稿や水害で汚損されて読み取ることが難しい書物に出会うこともあります。こうした、読み取りが困難なものは、そのまま忘れ去られてしまうことも少なくありません。しかし、こうしたものには、創作過程の秘密が隠されていたり、地域の人々の生きた営みの記憶が込められていることもあります。読めない文字を読むために、消えゆく記憶を取り戻すために、国文学研究資料館は、奈良先端科学技術大学院大学光メディアインタフェース研究室、実践女子大学文芸資料研究所と共同で光学的解析技術に関する研究を進めています。

和紙に含まれた毛髪を探る、DNA解析

日本語の歴史的典籍には、その用紙の中に毛髪が含まれていることがあります③。長い期間にわたって大量に作成された実物が保存されている書物は、人間の営みのタイムカプセルです。多くの人々によって読まれた書物の中には、ページの端が手垢で黒く変色したものもあります④。こうした部分からは人間の脂質を採取することができます。その分析から人々の生活環境や日々の営みを知ることができることがわかってきました。歴史的典籍に付着した人に由来する成分分析を通して、歴史的典籍に蓄積された書物と人間に関する歴史的データを抽出する技術開発を行い、歴史的典籍の分析に基づく典籍人類学の確立を目指します。

データから必要な資源を見つけるには、モノ自体の解析も欠かせません。
ようやく先行する諸国に追いついた状況にある日本のデジタルリソース活用を
自然科学・社会科学などの多くの分野に拓いてゆきます。



星石4Dプロジェクト

②星石4Dプロジェクト <https://hoshi-lshi4d.jp/>
国立極地研究所との異分野融合研究 日本語の歴史的典籍の情報から日本最大級の隕石雨の全貌解明を目指す



③歴史的典籍 NW 事業で行われた「津軽デジタル風土記」プロジェクト
<https://tsugaru-fudoki.jp/> 日本語の歴史的典籍の観光資源化を推進



①緊急性を要する地域資料のレスキュー活動



④「津軽デジタル風土記」プロジェクトによるシンポジウム



⑤公益財団法人 廣瀬資料館とその調査

Section.

03

ビッグデータ

日本語の歴史的典籍の新たなとらえ方

NIJL
国文学研究資料館

1000年以上に及ぶ連続性を有する

単一言語で記された書物が残っているということは、世界的に見ても稀な現象であるといえます。

日本語の歴史的典籍は、地球環境史や人間と社会の営みを解明し、現代社会の直面するさまざまな課題を解決するための世界規模のエビデンスデータ=ビッグデータとなるのです。

たとえば典籍防災学として

国文学研究資料館は、茨城大学地球・地域環境共創機構(GLEC)と共同で、日本語の歴史的典籍を防災研究のエビデンスデータとして活用する研究を行っています。歴史的典籍に蓄積された災害関係のデータを整備し、地形データ、GPSデータ等の防災学関連データと照合することを通して、典籍防災学の手法の確立を目指します(①)。

また、国文学研究資料館と国立極地研究所との間では、日本語の歴史的典籍に記載されたデータを、宇宙災害などの地球規模の災害のエビデンスデータとして活用する共同研究を行っています。歴史的典籍に記録された天体・地球環境情報の活用を通して、地球規模の環境問題に関わる研究を進めます(②)。

たとえば文献観光資源学として

かんぎえん 咸宜園という塾が大分県の日田市にありました。儒学者で教育者の廣瀬淡窓が1817年に創設したもので、明治30年代まで活動を続け、高野長英や大村益次郎ら4000人を超える人材が輩出しています。国文学研究資料館では、廣瀬家の資料のデータ収集とともに、その文化的活動についての調査と研究を行っています。和歌や俳諧などの文芸、茶道や華道、雅楽などの諸芸に関わる資料の分析により、廣瀬家がまさに地域の文化拠点だったことがわかってきています。廣瀬家の資料を受け継ぐ公益財団法人 廣瀬資料館と連携し、こうしたデータの蓄積と分析にもとづく観光資源開発のモデルケースを確立します(③~⑤)。

Section. 04 データ収集

さらなるデータ構築に向けて

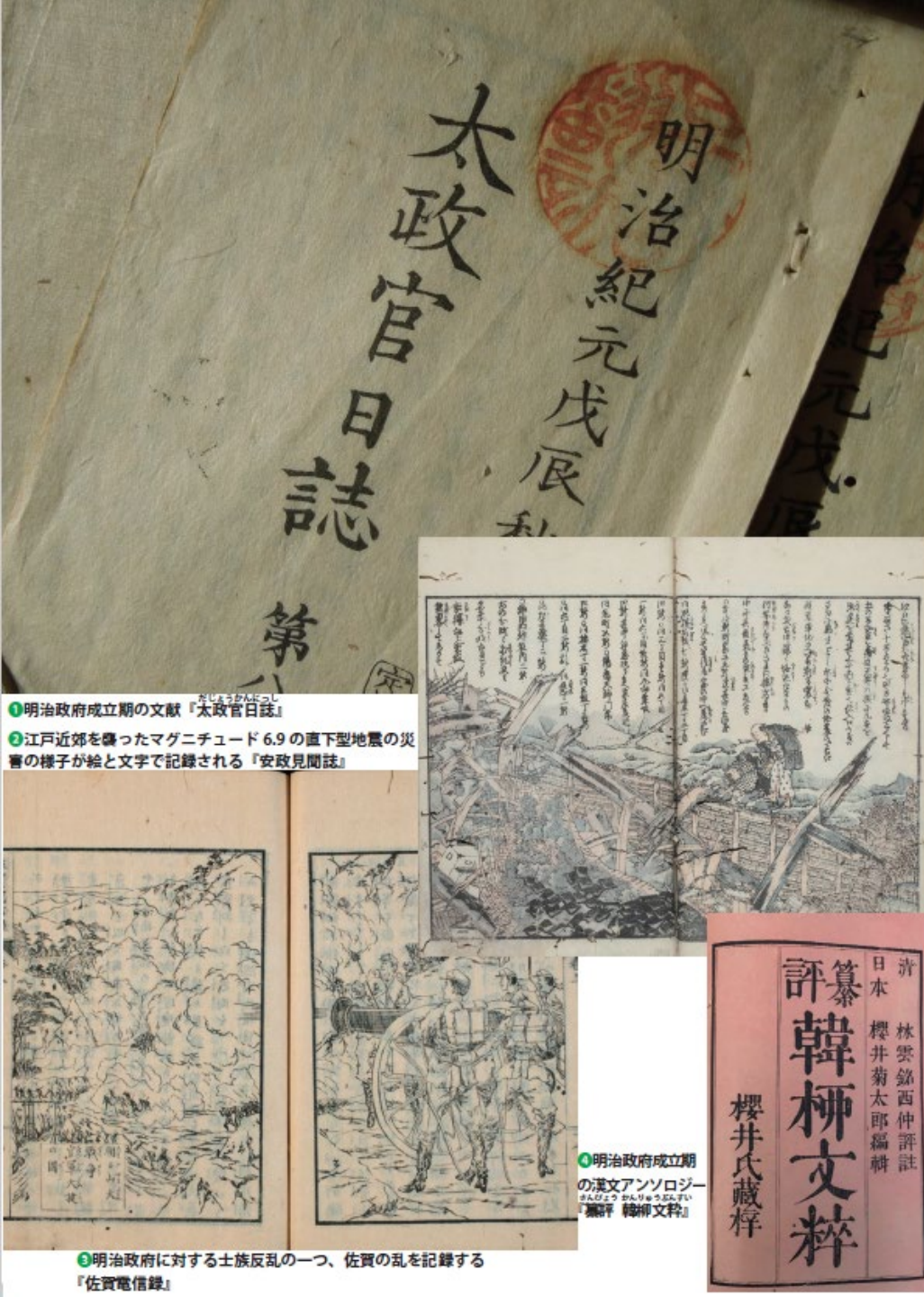
「1000年以上に及ぶ連続性を有する書物が残って」いますが、実はデジタル化という観点で見ると空白期間もあります。また原本が災害等によって失われてしまい、マイクロフィルムや写真としてしか伝わらない書物も存在します。それらを網羅してゆこうとするのがこのプロジェクトです。

大規模データとの連続性

30万点の日本語の歴史的典籍全冊画像という、「日本語の歴史的典籍の国際共同ネットワーク構築計画」のビッグデータの構築を基盤として、さらにデータ集積の範囲を明治時代初頭にまで拡張することを計画しています。書籍館(明治5年[1872]設立)、帝国図書館(明治30年[1897]設置)を経て国立国会図書館へと受け継がれてきた書物群への連続性を確保することで、現在、国立国会図書館から発信されている国立国会図書館デジタルコレクションとの隙間を埋め、1000年に及ぶ通時的なデータをweb上で活用できるようにすることを意図しています。また、個人や寺社などに所蔵され、公開の難しかった書物についても交渉を重ね、一層の典籍画像のデータ蓄積を目指します(①～④)。

デジタル化以前に作成された書物情報

書物の画像収集はマイクロフィルムや写真に撮影することで行われてきました。これらの媒体は時を経るごとに劣化してゆきます。また原本が災害等によって失われてしまい、マイクロフィルムや写真としてしか伝わらない書物も存在しています。そうした書物を保存してゆくためには、マイクロフィルムや古写真の保全処理が必要です。また、すでに劣化してしまった紙質資料から書物の書影を復原する技術やモノクロで残された写真やマイクロフィルムの情報にAIを活用して色調を復原する技術など、蓄積されたメディアの品質劣化による書物情報の消滅を防ぎ、情報を保全する技術開発を進めています。



①明治政府成立期の文献『太政官日誌』
②江戸近郊を襲ったマグニチュード6.9の直下型地震の災害の様子が絵と文字で記録される『安政見聞誌』

③明治政府成立期の漢文アンソロジー
『評纂 韓柳文粹』

④明治政府に対する士族反乱の一つ、佐賀の乱を記録する『佐賀電信録』



① 日本古典籍セミナー（ハワイ大学マノア校）



② 国際研究集会シンポジウム（国文学研究資料館）



③ 北米日本研究資料調整協議会（NCC）
ワークショップ（トロント大学）



④ ヨーロッパ日本研究協会（EJAS）
ワークショップ（リスボン新大学）



⑤ 日本語の歴史的典籍と現代のアーティストとのコラボレーションによる新たな価値の創造を目指した
ないじえる芸術共創ラボによる展示

epilogue

データ駆動型 人文学の創成へ向けて

人文科学と自然科学との協働、人材の育成

データの駆動性向上

日本語の歴史的典籍の30万点に及ぶ画像データを基盤として、データの機械可読化と構造化を進め、データの駆動性の向上を目指します。研究の過程で蓄積された、AI技術を用いた機械可読化技術やデータ分析の手法と、典籍防災学、文献観光資源学、典籍人類学などをモデルケースとして、データサイエンスとしてのデータ駆動型の課題解決型人文学の手法と領域を構築します。

現代社会の課題解決

現代の科学が結束して対応する、持続可能で多様性と包摂性のある社会の構築という課題に、先人の英知と記憶・経験を対象に研究を進めてきた人文学分野が参画し、他分野との協働による課題解決型研究のコンセプトを提案し、現代社会の課題解決へと向かい得るデータ駆動型、課題解決型の人文学への展開を促進します。こうした取り組みは、人文学、自然科学、社会科学といった領域の垣根を越えた学問の融合へとつながってゆきます。

若手研究者の育成

データ駆動による課題解決型人文学研究を通して、人文学と自然科学、とりわけ情報学分野の双方に通じ、その手法に精通した若手研究者の育成を行います(①・②)。あわせてそれぞれの研究領域と実体社会とを結ぶ高度な専門知識を持ったインタープリタ、コミュニケーター等を育成し、研究成果の社会還元と普及に努めます(③～⑤)。

「科学」としての「人文科学」への視線

「第6期科学技術・イノベーション基本計画」（内閣府、令和3年3月26日）

本文：<https://www8.cao.go.jp/cstp/kihonkeikaku/6honbun.pdf>

概要：<https://www8.cao.go.jp/cstp/kihonkeikaku/6gaiyo.pdf>

「我が国は、これまでも少子高齢化や過疎化の進展といった課題を抱えてきたが、更に近年、深刻化する自然災害、科学技術の国際競争力低下など新たな社会的課題に直面している。また、若者世代の自己肯定感の低さなど次代を担う人材に関する課題も浮き彫りになっている。それらを解決するためには、自然科学のみならず人文・社会科学も含めた多様な「知」の創造と、「総合知」による現存の社会全体の再設計、さらには、これらを担う人材育成が避けては通れない。」（p.4「はじめに」）

「2020年の第201回国会において、25年ぶりとなる科学技術基本法の本格的な改正が行われた。この法改正では、法律の名称を「科学技術・イノベーション基本法」とし、これまで科学技術の規定から除外されていた「人文・社会科学（法では「人文科学」と記載）のみ」に係るものを、同法の対象である「科学技術」の範囲に位置づけるとともに、「イノベーションの創出」を柱の一つに据えた。」（p.10「2.「科学技術・イノベーション政策」としての第6期基本計画」）

1000年におよぶ日本の記憶と経験を 世界にひらく



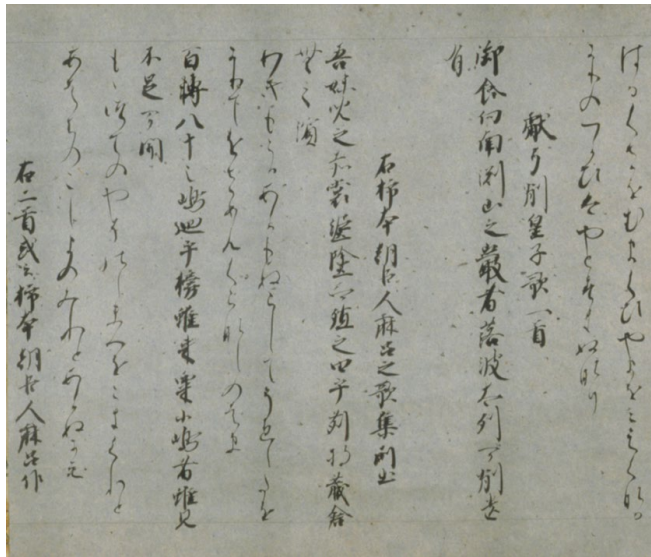
- データの蓄積からデータ駆動へ → 画像から機械可読データへ
* 蓄積されたデータは人文学の重要なインフラでありツール
- 人の手と目からコンピュータへ → 分散型から集約型へ
* 個別の研究者に紐付けられた蓄積を将来へわたって共有できるデータに
- 自然科学・社会科学と共有できるデータ構造 → 多分野融合
* 人間社会をめぐる課題の複雑化、広範囲化に対応するには分野横断で対応
- 人材の育成は急務 → 育成体制も分野融合で
* 人文学を推進し、発展させる方法・手段としてデータとその処理に通じた人材の育成

この報告の論点

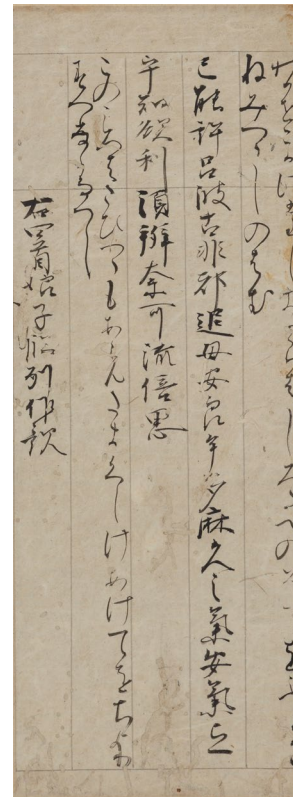
- 国文研の考える新しい時代の大型プロジェクト
- データ駆動型の人文学の構想と視界
- 人文学の成果と知識をデータとして継承するには？

万葉集

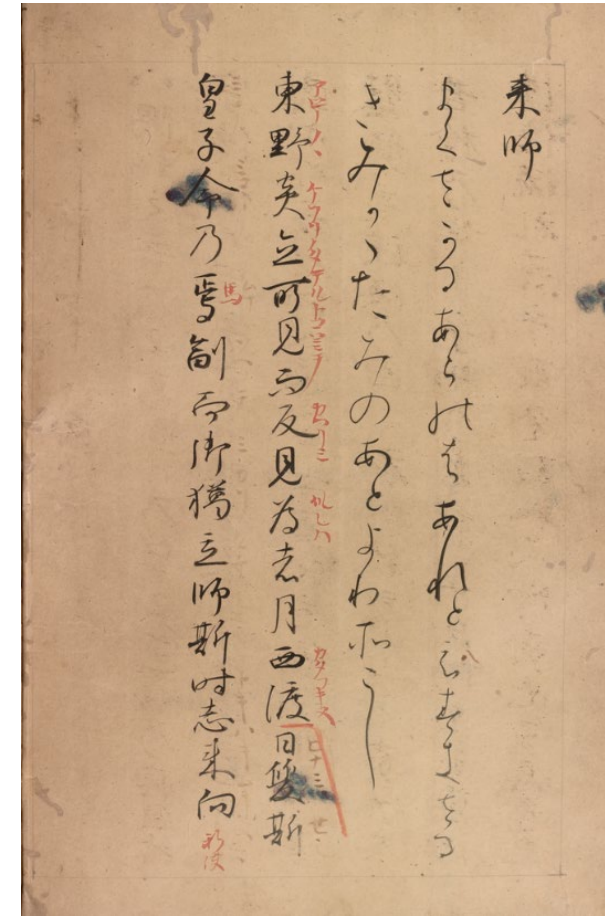
- 7世紀後半から8世紀後半にかけて編纂
- 世間に出たのは、11世紀初め頃。橘俊綱が法成寺宝蔵本を願い出て写したことに始まるという（袋草紙）



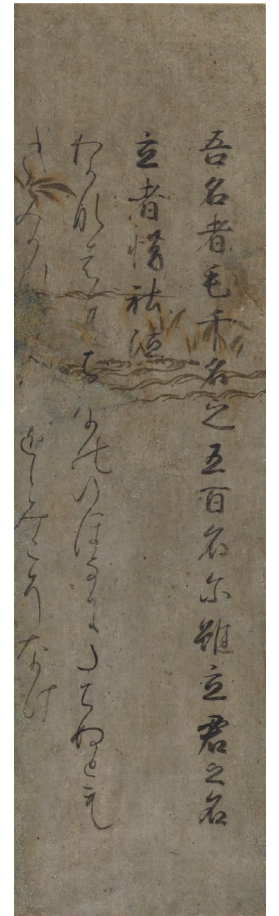
藍紙本
平安時代11世紀
東京国立博物館



天治本
平安時代12世紀
東京国立博物館



元暦校本
平安時代11世紀
東京国立博物館



桂本
平安時代11世紀
東京国立博物館

複数の万葉集

- 奈良時代に成立した『万葉集』 原本や奈良時代に書写された写本は伝わらない
- 平安時代に書写された写本（一部のみ伝わる例を含む）は複数伝わり、互いに記される本文や訓が異なる

→ 本文や訓の異なりが何を意味するのか？

その差異は意味のある差異なのか？

単なるミスなのか？

人文学データの意味は人文学研究者が最もよく理解している

→ 協働で研究を進める意義



→ ドメイン知識（ドメインちしき、英: Domain knowledge）または領域知識は、はっきり限定された、ある専門分野に特化した分野の知識…（Wikipedia）

万葉集の1300年

700 800 900 1000 1200 1300 1400 1500 1600 1700 1800 1900 2000

新日本古典文学大系に新しい訓

斎藤茂吉が万葉集を現代語訳

賀茂真淵が万葉集を研究

僧・契沖が万葉集を研究

この間に、様々な万葉集の写本、版本が作成される

現在の我々が読む本文

関東の僧・仙覚、万葉集を校訂↓仙覚本
藤原定家による万葉集の書写↓廣瀬本
新古今集が成立

この間に、桂本、藍紙本などの万葉集が書写される

梨壺の五人が万葉集を解説

古今集が成立

すでに本文が読めなくなっている

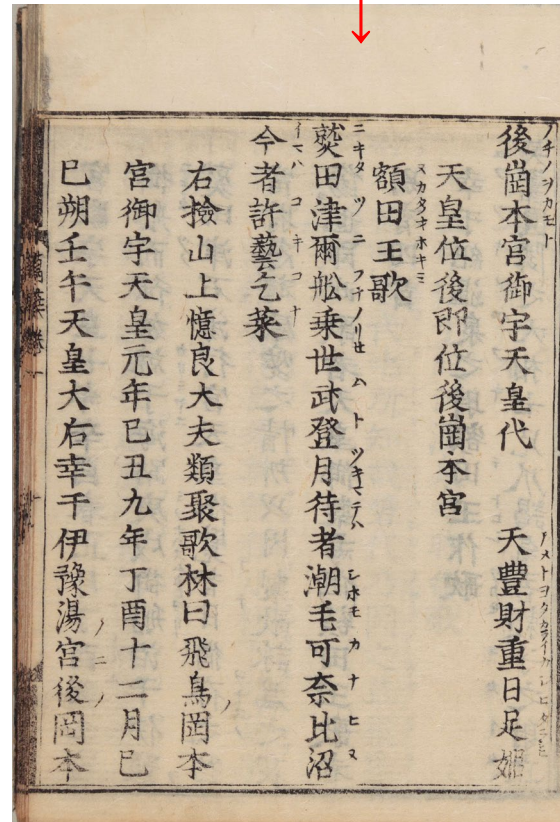
万葉集の編纂が現在見る形に及ぶか？

万葉集の編纂がはじまるか？

万葉集の訓

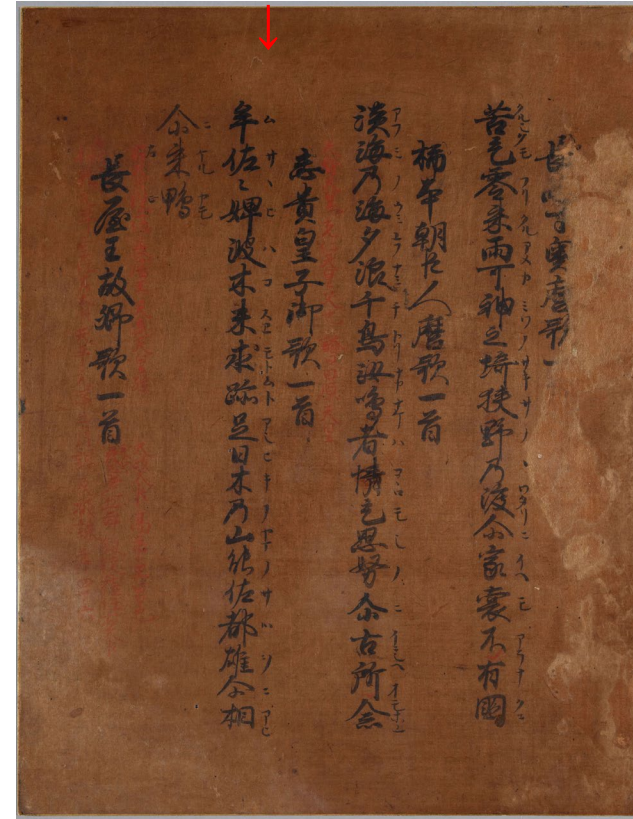
- 非仙覚本（平仮名訓、片仮名訓）
→古い時代の読み方を伝えているが、委細は未詳なものが多い
- 仙覚（1203-?）による訓
→現代へと繋がる読み方が提示される

片仮名による訓



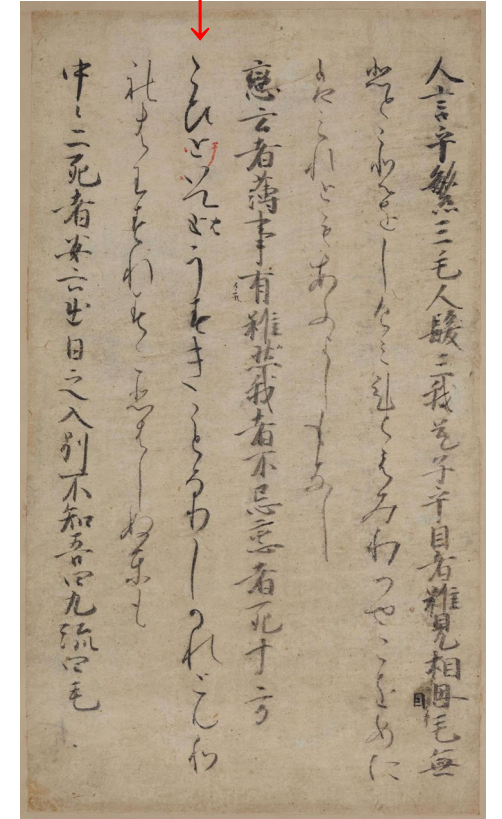
国文学研究資料館蔵『万葉集』
寛永20年刊
DOI: 10.20730/200016534

片仮名による訓



国文学研究資料館蔵柘枝切『万葉集』
鎌倉時代
DOI: 10.20730/200022522

平仮名による訓

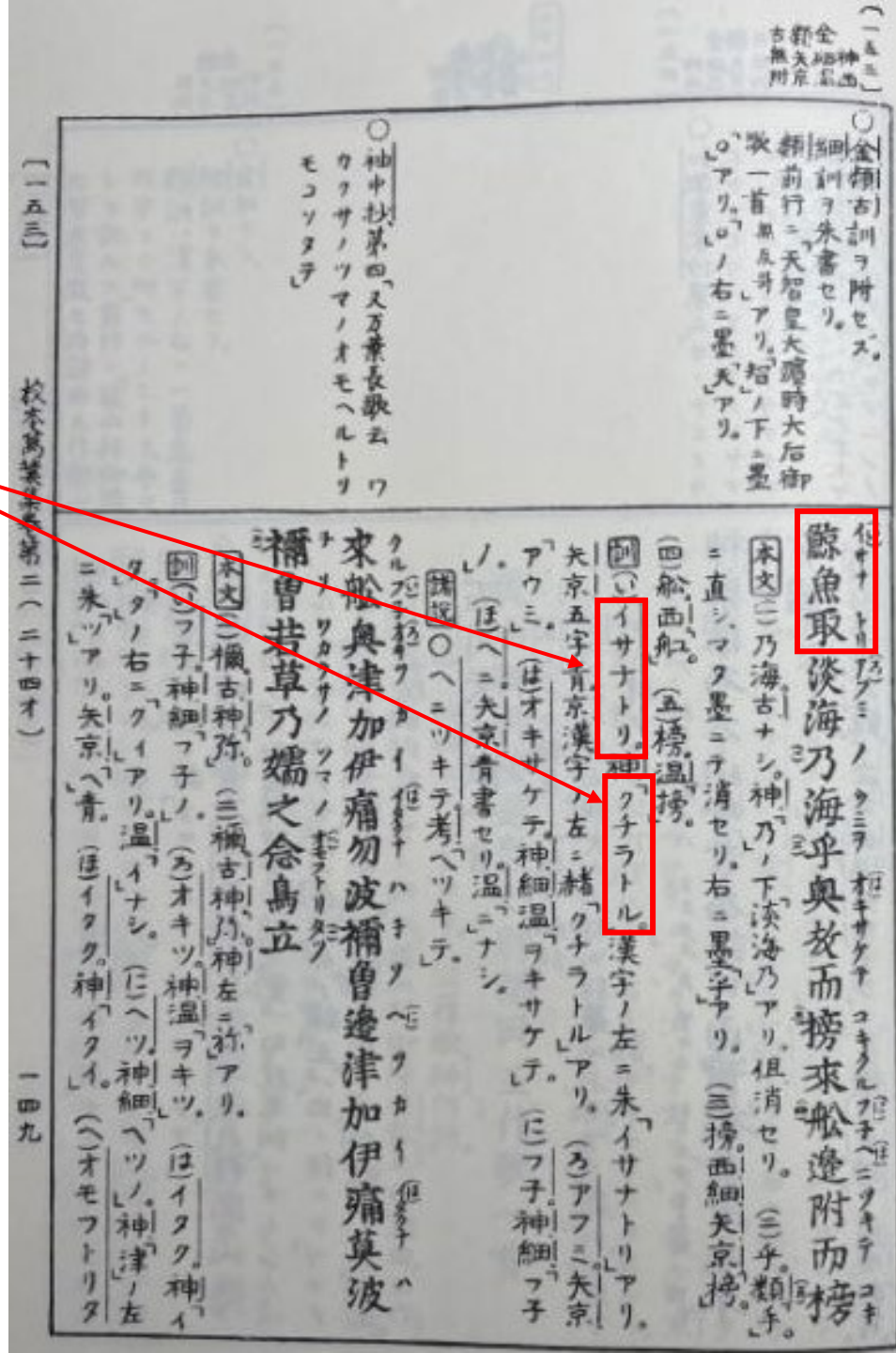


国文学研究資料館蔵尼崎切
『万葉集』平安時代
DOI: 10.20730/200017724

校本万葉集

「イサナトリ」という読み方と「クチラトル」という読み方があったことが示されている

鯨魚取



校本萬葉集刊行会, 1924-1925

電子化された万葉集

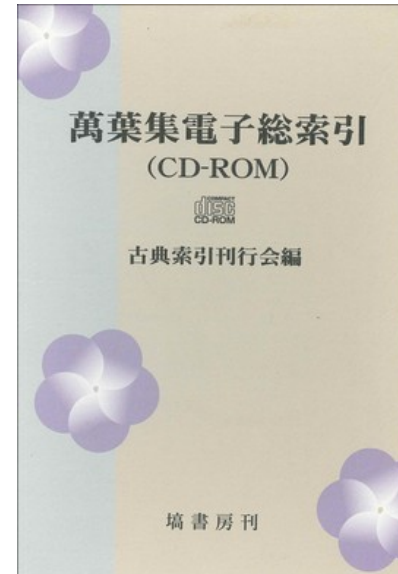


『日本文学web図書館』
古典ライブラリー

『新編国歌大観』（角川書店、1983-1992年）
に収められた前近代の和歌約45万首のデータを
web経由で検索

歌句（5文字列、7文字列）と語彙検索が可能

データの基づくテキスト：西本願寺本



古典索引刊行会編
『万葉集電子総索引CD-ROM版』
塙書房、2010年

塙書房刊『万葉集索引』に基づき、語彙
索引、単漢字索引、歌番号の検索が可能。
総語彙約16万件について、見出語句の
他、訳文漢字、原文、自立語・付属語の
別、および品詞や語性の検索が可能

データの基づくテキスト：西本願寺本に
もとづく校訂本文



『廣瀬本万葉集』の翻刻およびTEI
マークアッププロジェクト
関西大学アジアオープンリサーチセン
ター

データの基づくテキスト：廣瀬本

なぜ複数のテキストが作成されるのか

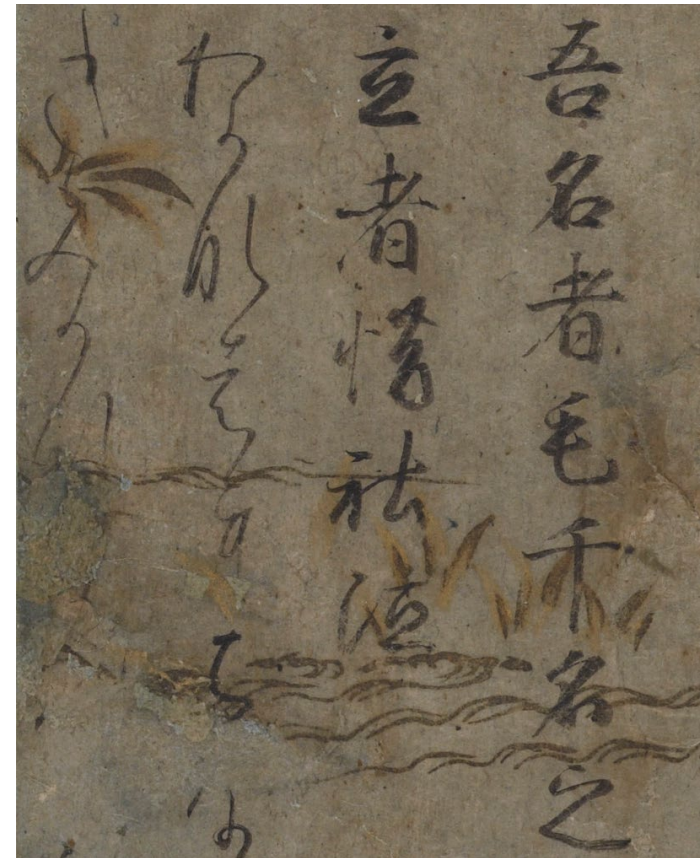
- 万葉仮名で書かれた本文自体の違い
 - 西本願寺本（古典ライブラリー・万葉集電子総索引）
 - 廣瀬本（関西大学アジアオープンリサーチセンター）
- 訓の違い
 - 仙覚以前の古い訓
 - 仙覚による訓
 - （→ 現在読まれている訓）



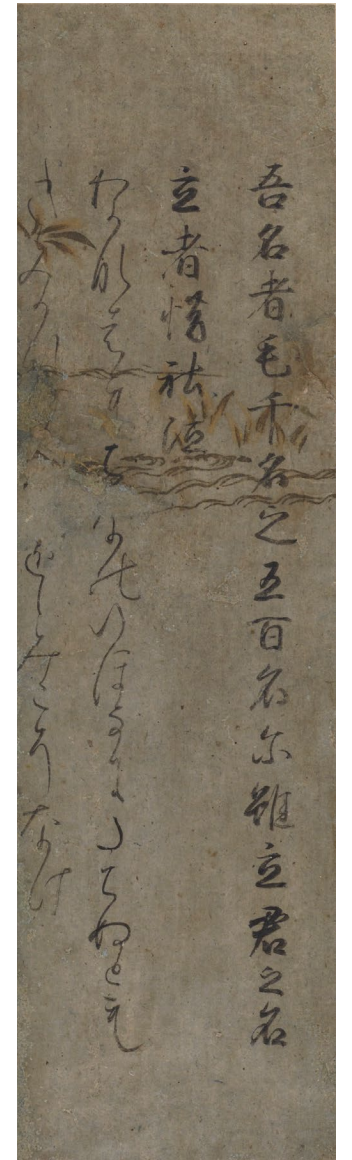
→ テキストが異なる
→ 目的も異なる
→ わかることも異なる

桂本万葉集・杵尾切・杵尾類切

- 桂本『万葉集』：現存最古の『万葉集』写本 卷4のみ伝存（御物、零巻）
- 杵尾切（とがのおぎれ）：桂本『万葉集』から切り出した断簡。巻第四のみ伝存。40葉ほどの断簡が古筆切として伝わり、杵尾切と呼ばれる
- 杵尾類切（とがのおるいぎれ）：『万葉集』巻7、8の断簡。桂本『万葉集』を思わせる下絵があり、書風も桂本に似ている。桂様万葉集切とも



東京国立博物館蔵杵尾切『万葉集』
平安時代（11世紀）
Colbase (<https://colbase.nich.go.jp>) による



杵尾類切

- 室町時代写？
→鎌倉～江戸時代まで諸説あり
- 桂本『万葉集』に似た書風
- 桂本『万葉集』に似た下絵がある
→桂本『万葉集』の模写か？
- テキストは非仙覚本
→田中大士による再調査あり
→テキストは仙覚本らしい
- もとの形は冊子本？
→現存の杵尾類切の寸法がほぼ同じ
→しかし卷子の特徴である巻皺あり

桂様万葉集切 かつらやうまんえふしふぎれ 〔古筆〕
伝源順筆。宗尊親王とも。室町時代写か。『万葉集』巻七、八の断簡。一面八行。名称は、桂本『万葉集』を思わせる下絵があり、書風も桂本に似ていることによる。杵尾類切とも称する。平仮名別提訓で、題詞は歌よりも低い。非仙覚系。もとは冊子本か。現在一〇葉が知られている。縦二一・七×横一三・七cm。

【参考文献】『万葉集事典』佐佐木信綱（平凡社 1956）＊『校本万葉集 新增補』（岩波書店 1980）＊『古筆学大成 12』小松茂美（講談社 1990）（田中大士）（『和歌大辞典』古典ライブラリー、web版）

杵尾切 とがのをぎれ 〔古筆〕
伝源順筆。宗尊親王とも。平安時代写。『新撰古筆名葉集』に「源順杵尾切 四半万葉真名力ナ哥二行書金銀下絵花鳥岩水草等アリ」とある。『万葉集』巻四の断簡。料紙は、八色に染めた斐紙に金銀泥で花鳥草木等の装飾下絵をほどこす。題詞は歌よりも高い。平仮名別提訓。非仙覚系。長歌には訓がない。『万葉集』の現存最古の写本である『桂本万葉集』（零巻が現存・宮内庁侍從職蔵）から切り出された断簡。縦二六・九cm。もと卷子装。源兼行の書写とされる。現在四三葉（目録を含む）が確認される。断簡の影印集成に『桂之落葉』（竹柏会 1931）がある。

【参考文献】『万葉集事典』佐佐木信綱（平凡社 1956）＊『校本万葉集 新增補』（岩波書店 1980）＊『古筆学大成 12』小松茂美（講談社 1990）＊『桂本万葉集』（二玄社 日本名筆選 27 1994）（田中大士）（『和歌大辞典』古典ライブラリー、web版）

杵尾類切は最古の『万葉集』の一部か？

- 【仮説1】 杵尾切（巻4）と同じ『万葉集』の巻7・8を他の書写者が写したもの ←「室町時代か」という従來說と抵触
- 【仮説2】 桂本『万葉集』の巻7・8を転写したもの ←「室町時代か」という従來說にも抵触しない
- 【仮説3】 桂本『万葉集』とはまったく別の『万葉集』の断簡←なぜ似ているのか？ [ニセモノ？]
- 【仮説4】 桂本『万葉集』に似せて作成されたもの 例えば桂本『万葉集』の欠巻を補うために補写された？

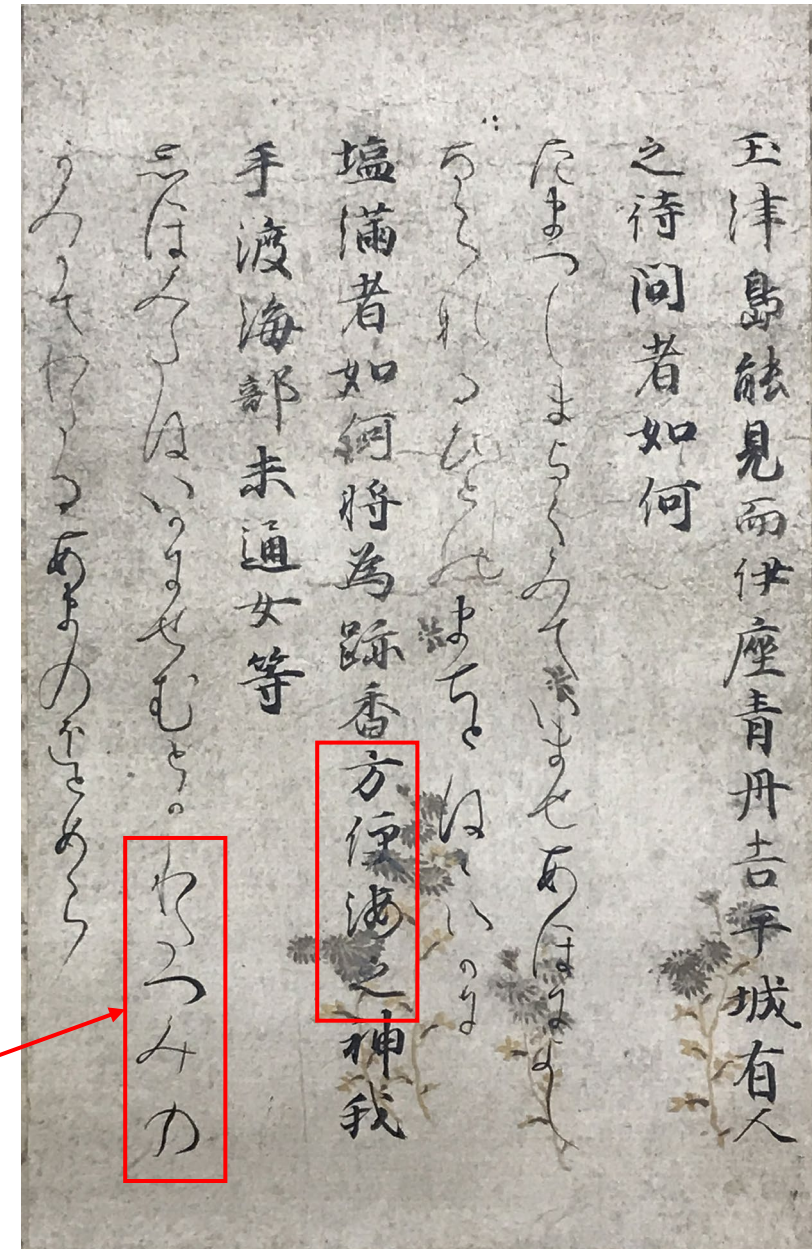
本文と下絵をめぐる研究の進展

- 田中大士による研究（「桂様切（梅尾類切）万葉集の再検討」『万葉集研究』38、2018年）
 - ▷記される本文は実は仙覚本系統
 - ▷しかし、解決しない訓の対立（ウメームメ）も残る
 - ▷「平安風の写本に仕立てるにはかなりの見識が必用」

- 徳植俊之による研究（「「梅尾類切」考」汲古72、2017年12月）
 - ▷下絵は後入れらしい（マイクロスコープによる観察）
 - ▷和歌本文に関連した絵柄が選択されて描かれているらしい
 - ▷一面に見える巻皺の一部は墨で書き入れられているらしい

田中説の指摘

仙覚によって鎌倉時代
（13世紀）に考案された
訓が記される
そのため
仙覚以前の写本ではあり
えないと判断される



さまざまなテキスト

- ながい期間にわたって読み継がれてきた古典のテキストは一様ではない
- より成立期に近いテキスト（証本などと呼ばれてきた）
より広く読まれたテキスト（流布本などと呼ばれてきた）
通例とは異なるテキスト（異本などと呼ばれてきた）
- 原典へと遡及する方向性
享受の側面を重視する方向性

→人文学の研究者は同一作品でも、テキストを使い分けている

→むしろ、同一作品であるのに異なるテキストが残されていることで、
明らかになることが多い

テキストからわかること・わからないこと

- 【仮説1】 杵尾切（巻4）と同じ『万葉集』の巻7・8を他の書写者が写したもの ←「室町時代か」という従來說と抵触
- 【仮説2】 桂本『万葉集』の巻7・8を転写したもの ←「室町時代か」という従來說にも抵触しない
- 【仮説3】 桂本『万葉集』とはまったく別の『万葉集』の断簡←なぜ似ているのか？ [ニセモノ？]
- 【仮説4】 桂本『万葉集』に似せて作成されたもの 例えば桂本『万葉集』の欠巻を補うために補写された？

→ 杵尾類切はどのようにつくられているのか？

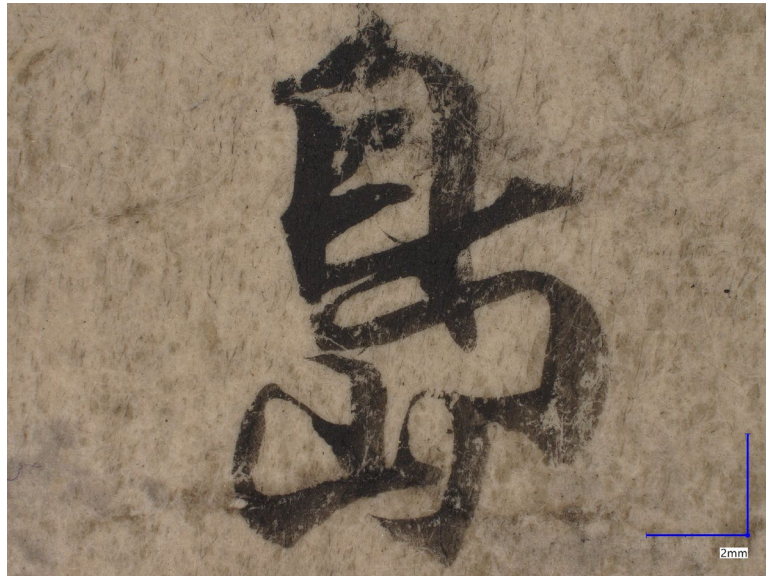
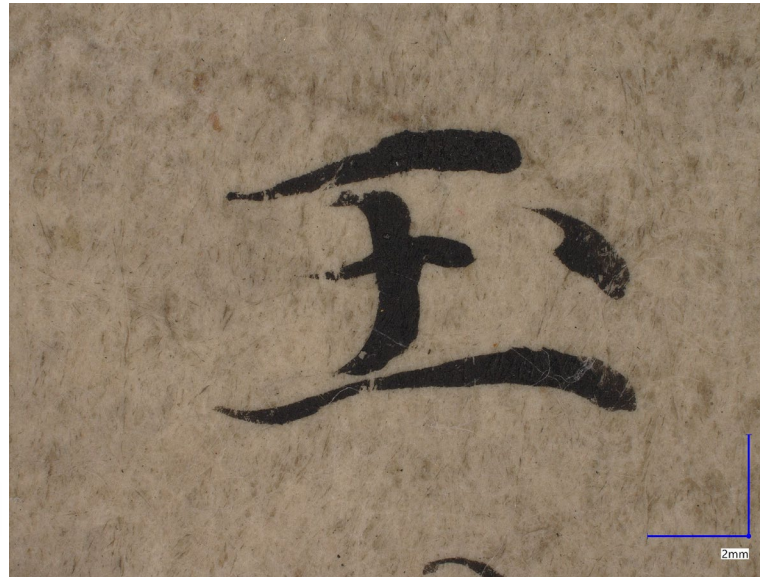
検討にはテキストの領域を越えるデータが必用

本文と下絵をめぐる研究の進展

- 田中大士による研究（「桂様切（杣尾類切）万葉集の再検討」『万葉集研究』38、2018年）
 - ▷記される本文は実は仙覚本系統
 - ▷しかし、解決しない訓の対立（ウメームメ）も残る
 - ▷「平安風の写本に仕立てるにはかなりの見識が必用」
- 徳植俊之による研究（「「杣尾類切」考」汲古72、2017年12月）
 - ▷下絵は後入れらしい（マイクロスコープによる観察）
 - ▷和歌本文に関連した絵柄が選択されて描かれているらしい
 - ▷一面に見える巻皺の一部は墨で書き入れられているらしい



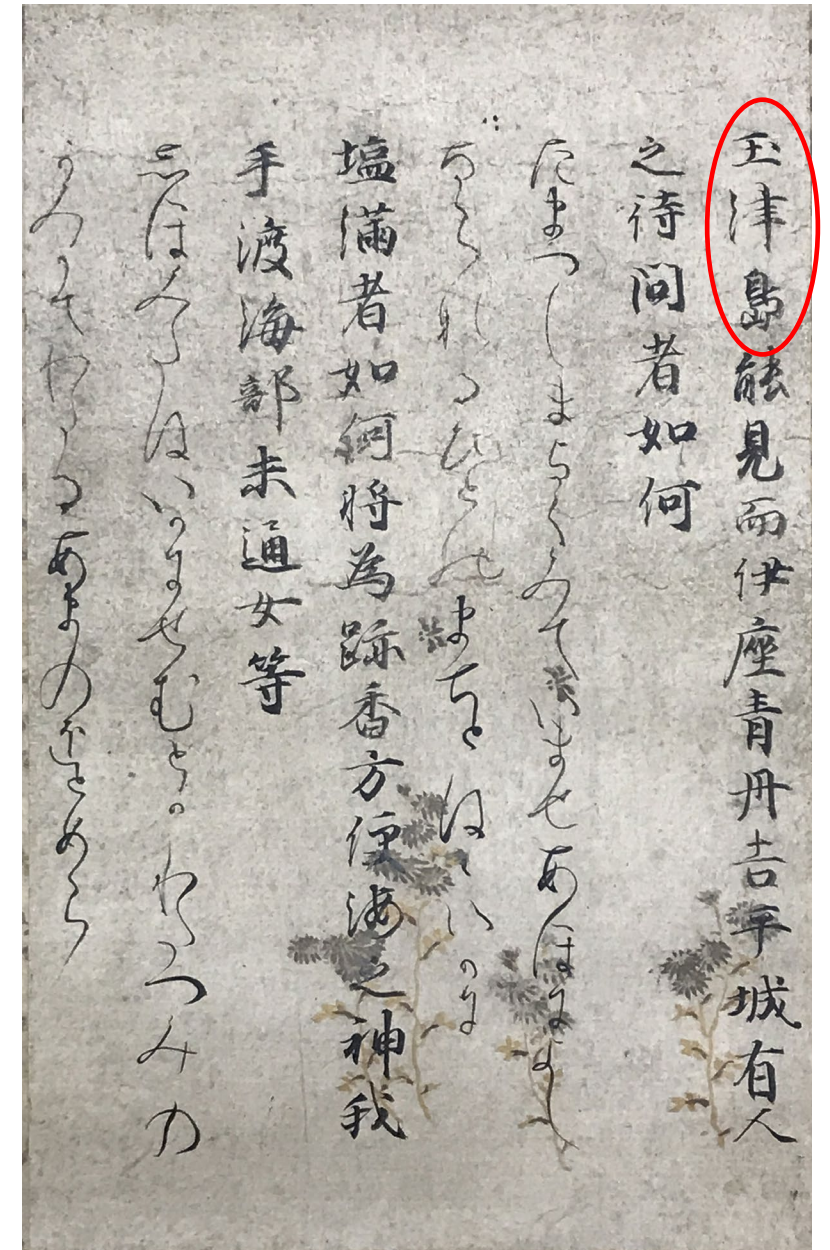
CKEYENCE VHX-8000導入 (2021年度)



×20

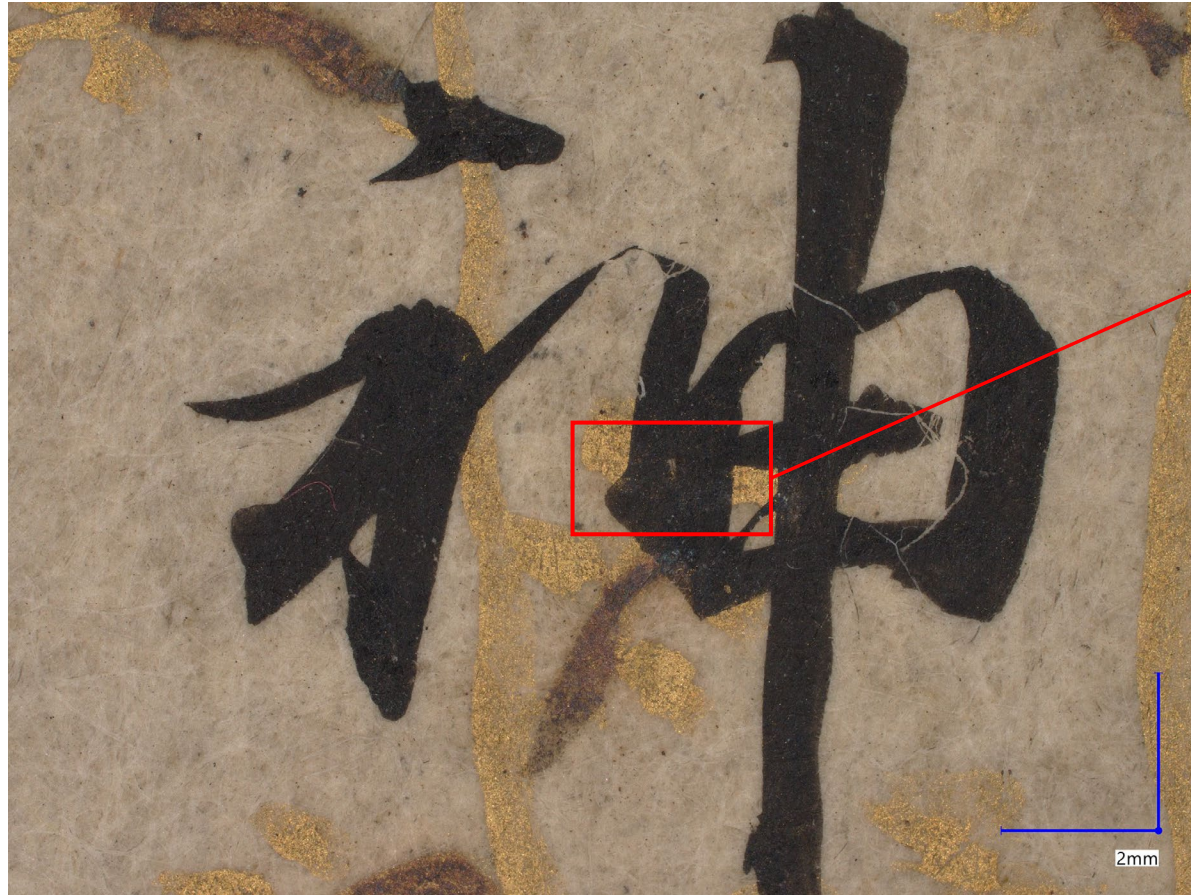


この状態の画像データは通常は流通していない

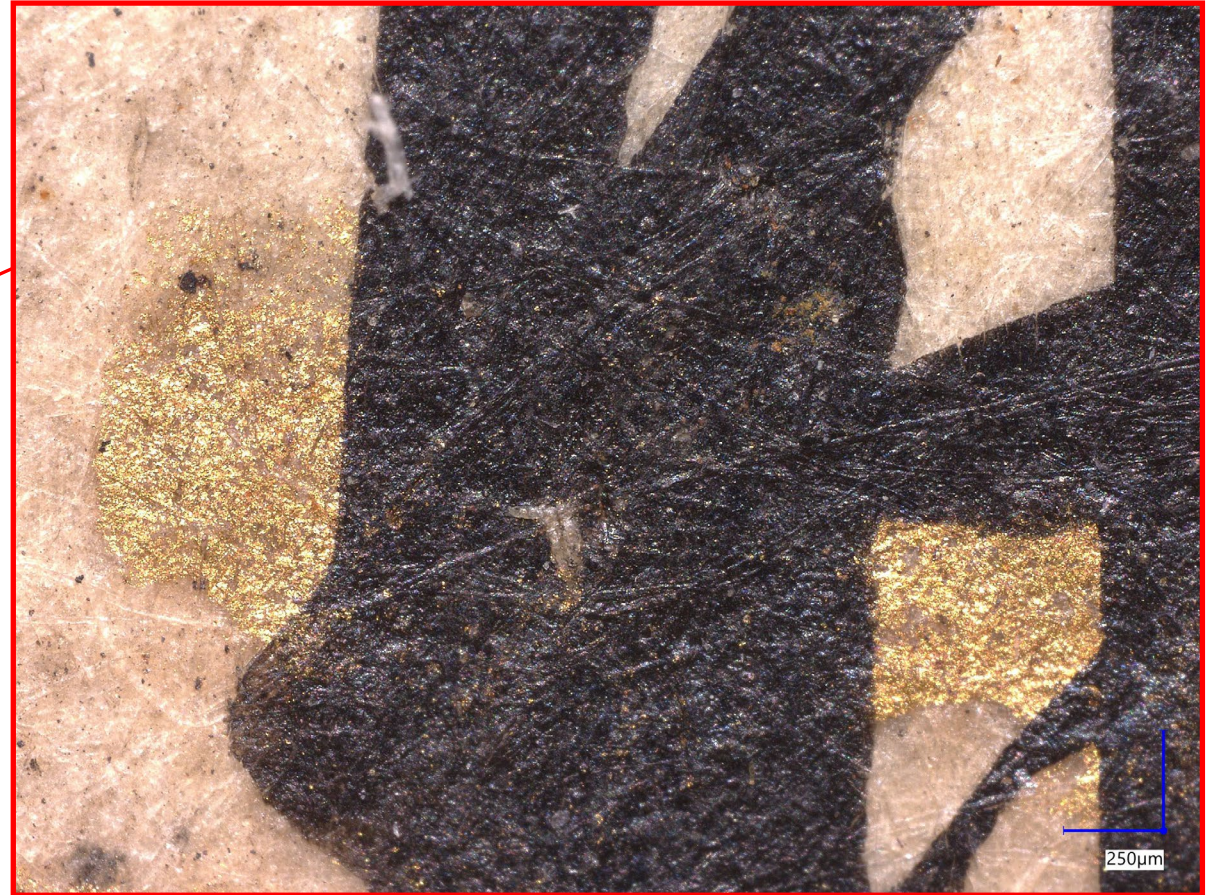


国文学研究資料館蔵梅尾類切『万葉集』

巻7、1215-1216番歌



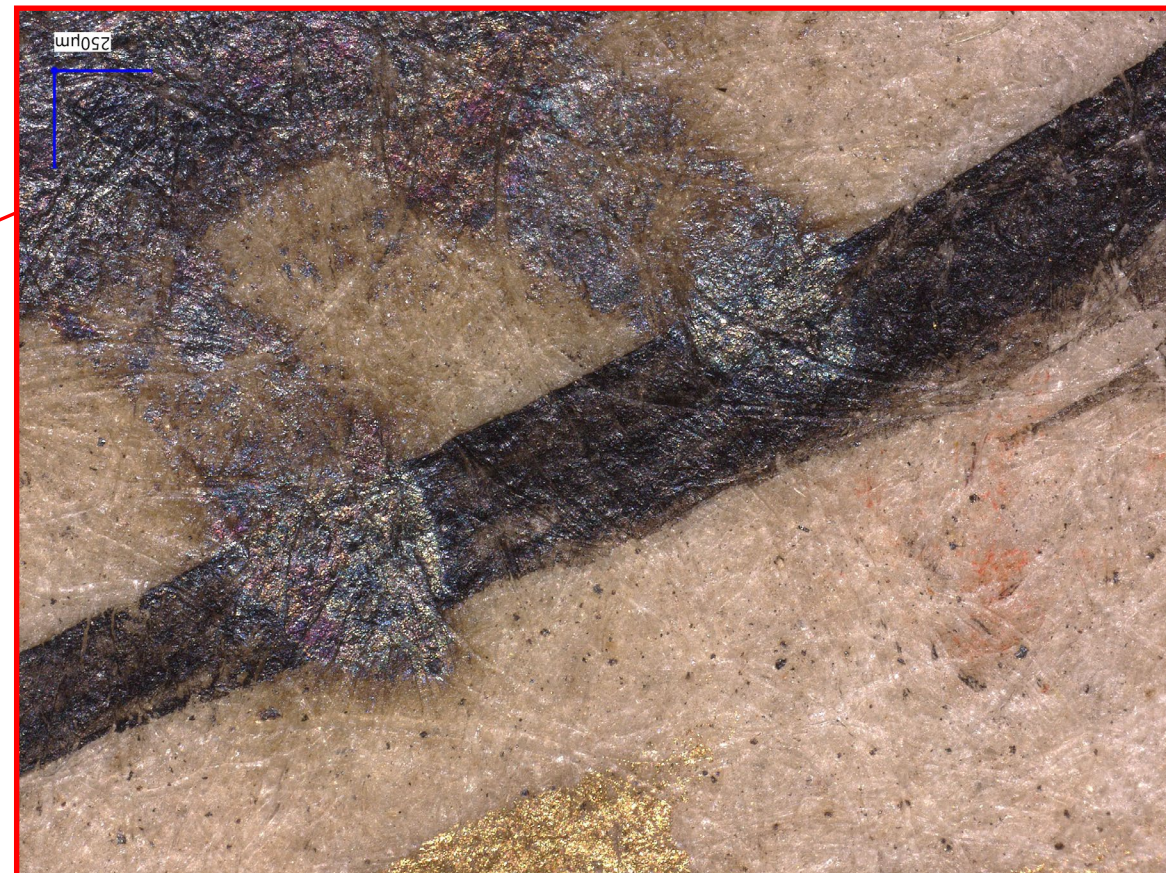
× 20



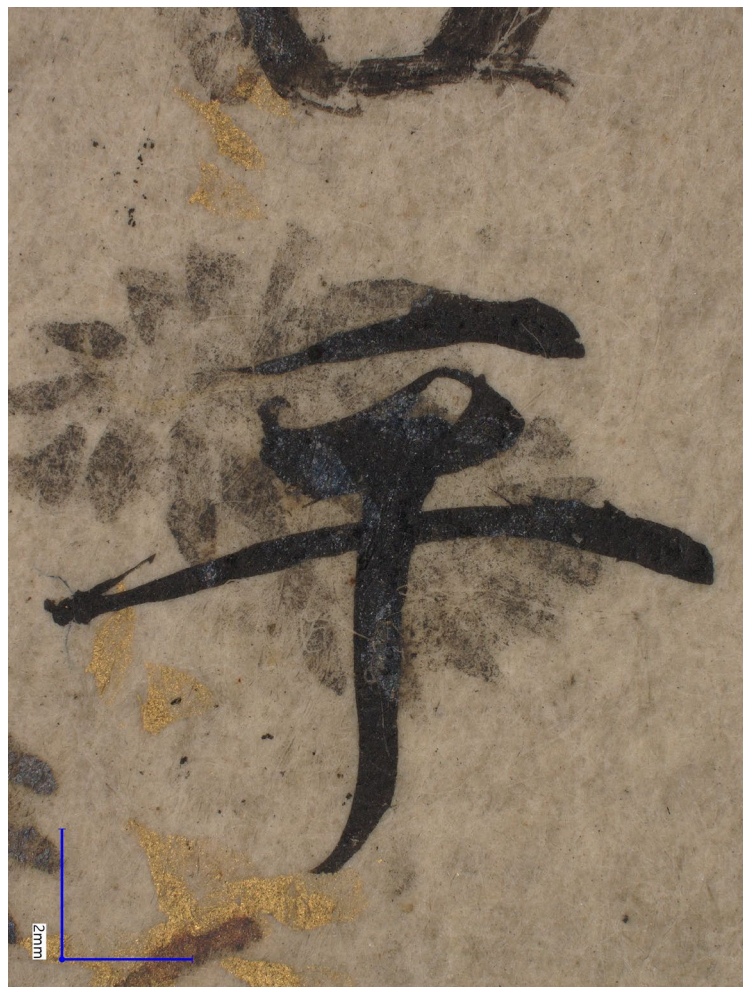
× 100



× 20



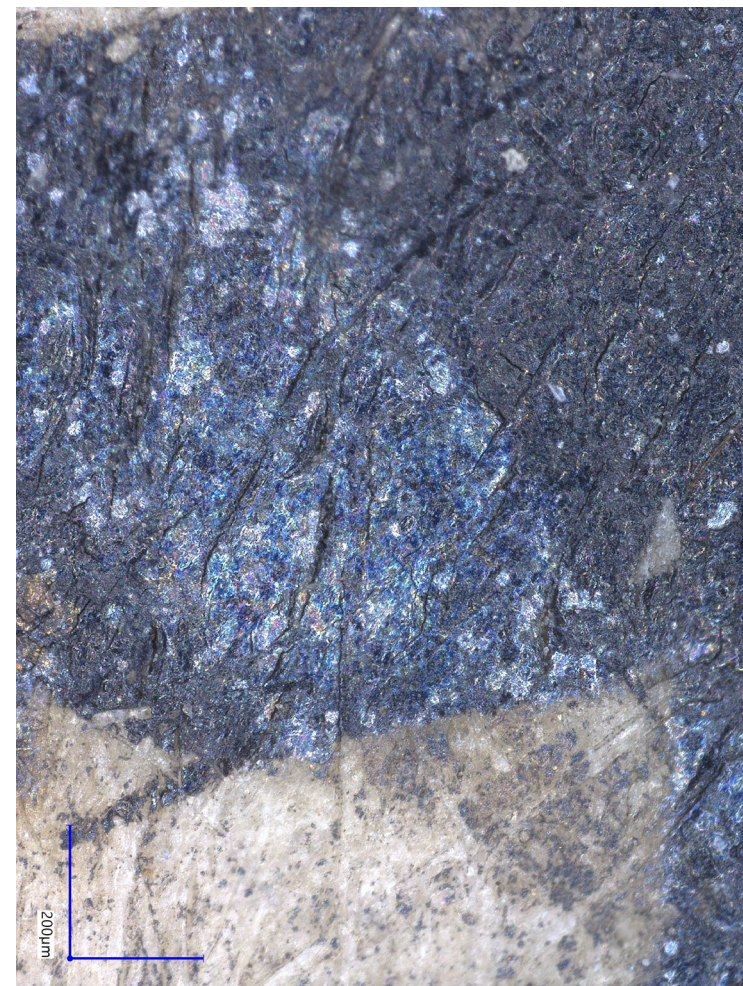
× 100



× 20



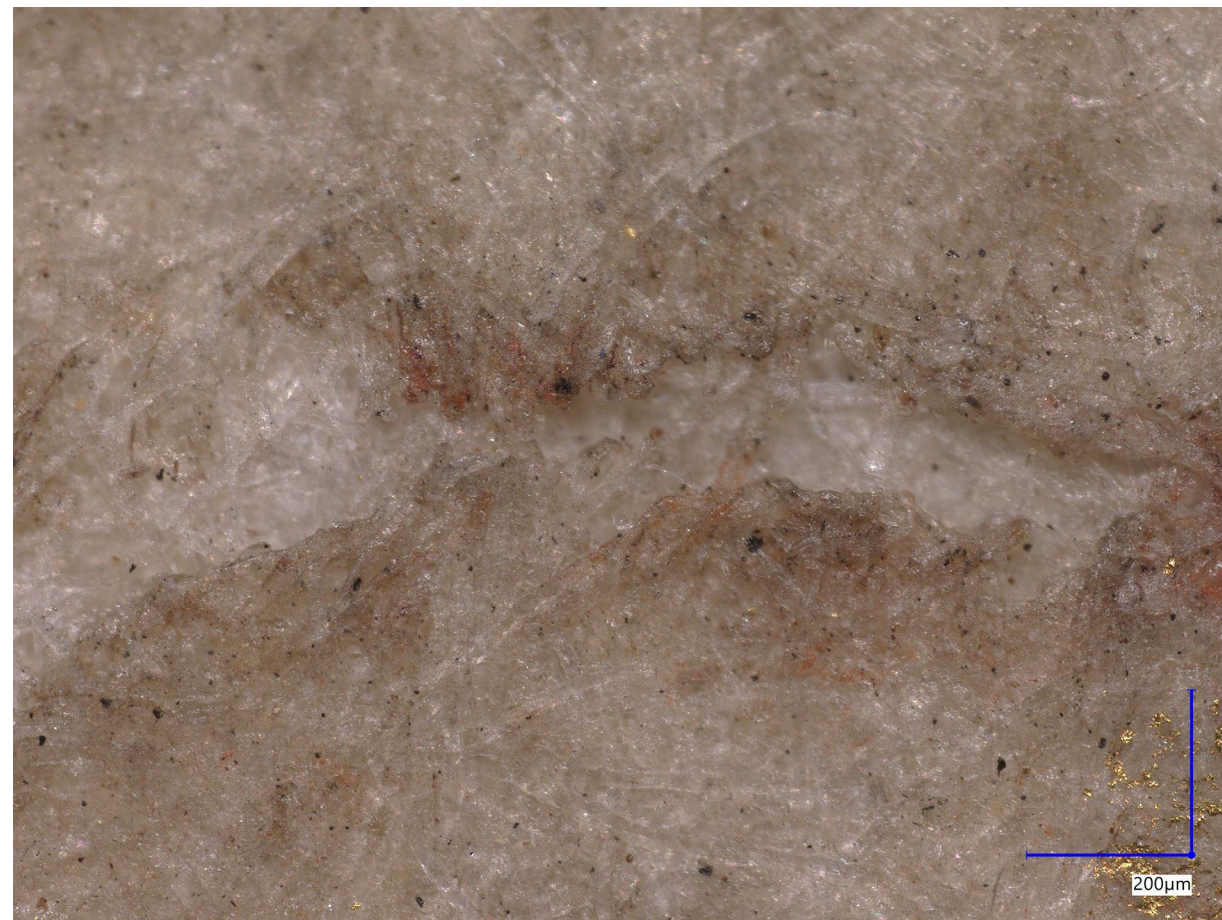
× 40



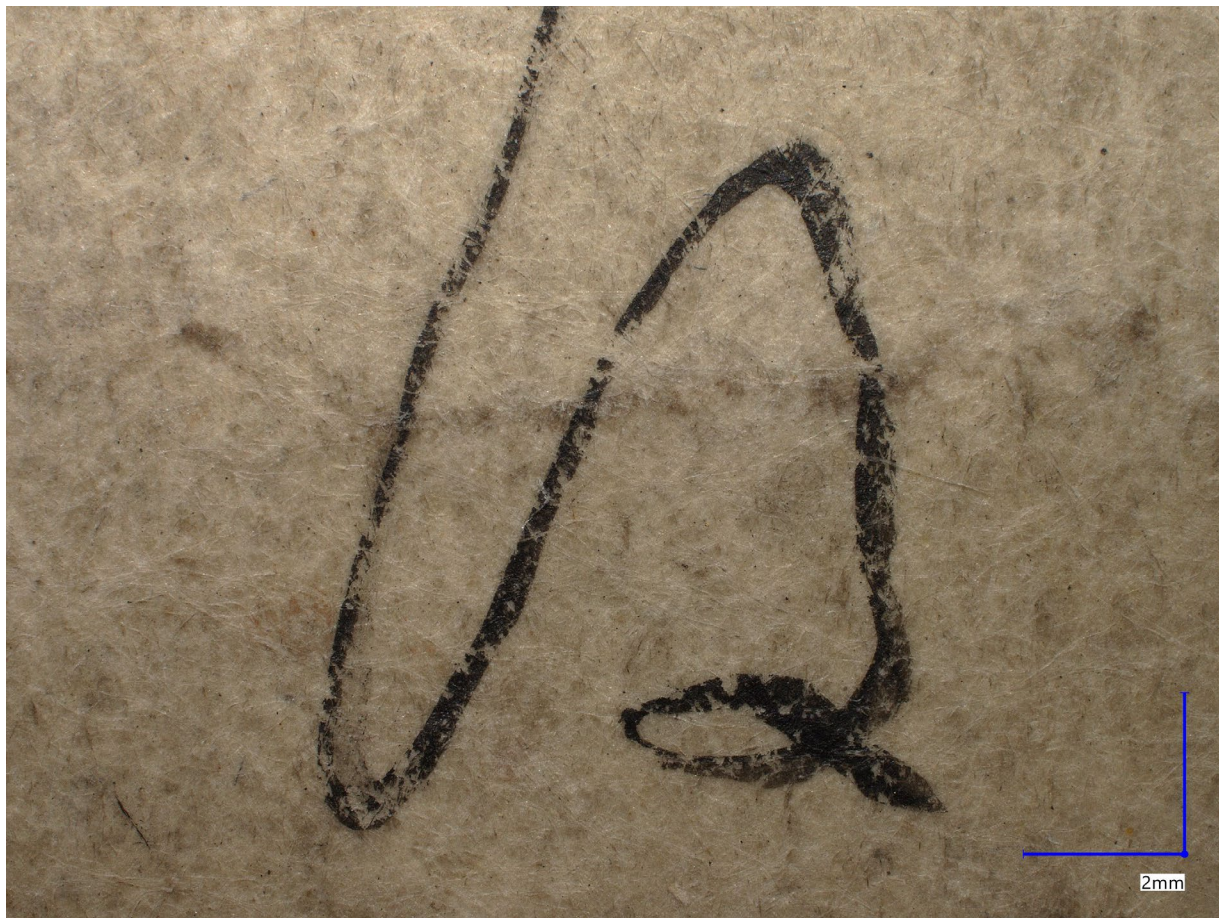
× 200



× 20



× 200



× 20



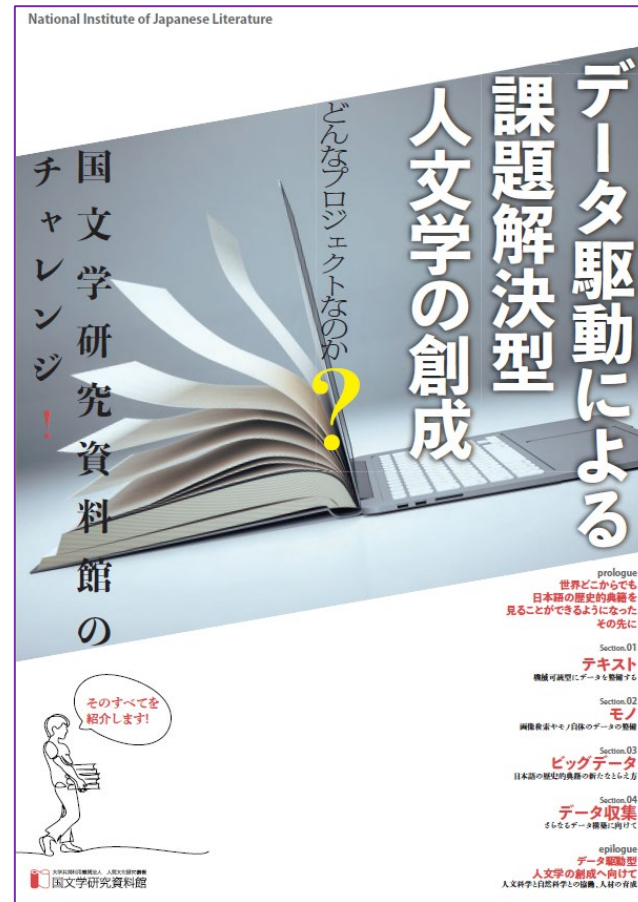
× 40

データ蓄積とドメイン知識

- 人文系の研究者の作成した“科学的”データ
 - 諸般の事情により再現性の確保が難しい
 - 再検証されることのないデータが長く継承されることも
 - さまざまな位相のデータを接続→データの価値向上→データ駆動へ
- テキストの“中身”の専門家である人文学研究者の研究成果
 - そのデータは何を意味しているのか？は自明ではない
 - 画像・テキスト等のデータのメタデータとしてもきわめて重要
 - 人文学研究者が求めるデータとは？→積極的な提言・提案を！
- …しかし、蓄積されたデータは万能ではない
 - 蓄積と分析・積極的な活用を担うことができる人材の育成は急務

ご清聴
ありがとう
ございました

本計画にも
どうかご支援を賜りたく
お願い申し上げます



 [https://www.nijl.ac.jp/
pages/cijproject/imag
es/datakudou.pdf](https://www.nijl.ac.jp/pages/cijproject/images/datakudou.pdf)