

<https://wakatext.org/>

2026.01

第1号

和歌文学
テキスト
データベース
構築計画

Contents

【巻頭言】ニューズレター刊行にあたって 加藤弓枝	……1
【次世代のテキストデータベース構築に向けて】	
和歌文学研究 次の百年の基盤づくり——TEIによる「構造化と共有化」—— 近藤泰弘	……2
テキストの立体化と深化、そして共有化へ——来たりくる和歌文学研究の未来のために—— 田淵句美子	……4
【TEI入門講座参加記】はじめの一步 荻原大地	……6
【和歌文学テキストデータベース構築計画 NEWS】プロジェクト報告・テキストデータ公開情報・TEI入門講座のお知らせ	……8

【巻頭言】

ニューズレター刊行にあたって

かとうゆみえ
加藤弓枝

本文研究は、あらゆる古典文学研究の基盤をなすものです。近年、国内外の古典籍がデジタル化され、多数の資料に容易にアクセスできるようになりました。その中には、これまで知られてこなかった伝本も含まれています。言い換えれば、和歌に限らず、現在広く用いられている古典文学作品の本文そのものを、あらためて検討し直すべき時期を迎えていると言えるでしょう。また、デジタル時代を迎え、従来の紙媒体やPDF公開を前提とした本文提示と、今後のデジタル公開を想定した本文提示とは、方法が異なってきたべきです。

こうした問題意識のもと、このたび「和歌文学テキストデータベース構築計画」という共同研究を開始しました。本プロジェクトでは、TEI (Text Encoding Initiative) ガイドラインに則した和歌文学のテキストデータを構築し、公開していきます。この共同研究は、基盤研究(C)「次世代の翻刻校訂モデルを搭載した中世歌合データベースの構築と本文分析の実践的研究」(P22K0303)を発展的に継承した、基盤研究(B)「次世代の日本文学テキストデータベース構造化と歌論歌学データベース構築に関する実践的研究」(P25K0439)の採択を受けて遂行しています。研究組織は、研究代表者・加藤弓枝(名古屋大学)、研究分担者・海野圭介(早稲田大学)・青山英正(東京大学)・菊池信彦(国文学研究資料館)・舟見一哉(実践女子大学)・田口暢之(鶴見大学)・甲斐温子(静岡大学)、研究協力者・幾浦裕之(文部科学省)・館野文昭(埼玉大学)の計九名で構成されています。現在は、中世歌合・和歌集のほか、歌論書・歌学書などのテキストデータベース構築を進めています。

良質な論文や注釈に恵まれるかどうかは、古典の存続そのものにも影響を与えます。そのため、本文だけでなく、それに付随する研究成果も広く公開され、長く継承されることが求められます。歴史と未来を包摂し、利用者がそのあいだを自在に行き来できるテキストをいかに構築し公開していくか——この課題に、本計画を通して取り組んでいきたいと考えています。

このニューズレターでは、本研究プロジェクトにおいて現在取り組んでいることや、今後考えていきたいことを、多くの方々のご協力を得ながら発信してまいります。創刊号では、共同研究のメンバーも多数登壇した和歌文学会特別例会「和歌文学研究、この次の百年」をお聞きくださった近藤泰弘氏・田淵句美子氏より、次世代のテキストデータベースに関するご寄稿を、また荻原大地氏より、TEI入門講座に関するご寄稿をいただきました。年三回の刊行を予定しておりますので、今後ともぜひご注目ください。

和歌文学研究 次の百年の基盤づくり

— T E I による「構造化と共有化」 —

近藤泰弘
こんどうやすひろ

■和歌研究の成果を次世代に継承するために

二〇二四年に開催された特別例会は、『千五百番歌合』へのT E I適用を中心に、和歌文学研究とデジタル・ヒューマニティーズ(DH)の交点が具体的な形で示されたという点でまさに画期的な催しであった。T E I (Text Encoding Initiative) は、一九八〇年代末に欧米の文献学者たちが立ち上げたテキスト記述の国際標準である。最初はS G M L、現在ではX M Lという汎用性の高いタグ付け方法を、文学作品などで使いやすいように標準化したものである。日本でも、永崎研宣氏らの献身的な努力によって「文学資料を世界的な枠組みの中で共有するための共通語」として定着しつつある。これまでの和歌研究は、本文校訂や注釈、伝本比較を、「紙の上の精密な作業」で行うことによって支えられてきたが、いま、その成果を次世代に継承するための形式としてT E Iが浮上している。私自身は、T E I 3 (3版) が、一九九四年に日本に紹介された時から、ずっとT E I には関心を持って来たので、感慨深いものがあった。

■作品の文脈を保った「階層データ」としてデジタル化する

とりわけ今回の『千五百番歌合』のT E I適用(5ページ【図1】【図2】参照)は示唆に富む。歌合という多数のレイヤーを持つ構造——左右の方、判詞、歌番号、作者、判者、出典——をすべてX M Lの階層として表現できる点は、まさにT E Iの真価を発揮する部分である。つまり、和歌本文を単なる「文字列」としてではなく、作品の社会的・文化的文脈を階層的に示したデータとしてデジタル化できる。これにより、研究者は作品を再編集・再配列しながら、過去には見えなかった関係性を動的に探索できるようになるわけである。

■研究そのものの方法を再設計する契機として

しかし今回の催しで印象的だったのは、技術的な側面よりも、むしろ「研究者自身がどのようにデジタル化と向き合うか」という姿勢の問題であった。発表者が、T E Iを単なるアーカイブ化ではなく、「研究そのものの方法を再設計する契機」として語っていたことは注目される。たとえば、同一作者の歌群を自動抽出し、語彙や修辞の特徴を横断的に分析すること、あるいは新たに開発されたT E I和歌ビューワによって、複数の写本の異文を同一フレームで比較し、テキストの変遷を可視化すること。こうした作業は、これまで研究者の記憶やノートに埋もれていた「暗黙知」を形式化する試みともいえる。私自身、A Iを用いた和歌の分析を行っている立場から見ると、T E Iは単なる「表示形式」ではなく、A Iにとつての「理解可能な文脈情報」を与える装置である。A Iは、文の前後関係や語の出現頻度だけでなく、タグによって明示された構造——たとえばauthor、date、ageのような要素——を使って、より厳密な解釈を行う。もちろんA Iには自らこの構造を見いだす能力もあるが、得られた結果の同一性を確保するためには、研究者自身によるタグの定義は重要である。

■ガイドラインに沿った共通形式への整備

今後の課題は、このようなT E I適用の成果を、個々の研究者の手元に閉じ込めず、共同で発展させていく体制をどう作るかである。日本では、大学やプロジェクトごとに、プレーンなテキストファイル、あるいは独自仕様のマークアップが存在し、相互運用性が確保されていない傾向があった。これらをT E Iガイドラインにできるだけ沿った共通形式へ整備することが、次の世代の共有基盤づくりに不可欠である。欧米で

近藤泰弘 [青山学院大学名誉教授]

専門は、日本語史・日本語文法・コーパス言語学。1979年東京大学大学院人文科学研究科国語国文学専門課程修了。2003年に博士(文学)取得。日本女子大学文学部助教授等を経て、2023年まで青山学院大学文学部教授。主著に『日本語記述文法の理論』(ひつじ書房、2000年)がある。第28回金田一京助博士記念賞受賞。

は、GitHubやZenodo[▼]でデータを公開し、誰でも改訂や派生を提案できるエコシステムが整いつつある。和歌研究の世界でも、校訂本文・注釈・書誌情報を、TEI準拠XMLデータ（以下TEIXML）として統合した「共有本文」の構築が求められるだろう。

その際、研究者が自らタグを手入力する必要は必ずしもない。むしろ重要なのは、「どの部分がどのような情報単位であるか」を意識して整理する「構造化の思考」である。実際のマークアップ作業は、AIによってかなりの部分を自動化することが可能になってきている。AIは文脈を解析し、「詞書」「題」「作者」「判詞」などを高精度に判別できる段階にあり、研究者はその結果を確認し、調整するだけでよい。すなわち、人間が必要な構造を決定し、AIが記述を担うという役割分担が現実的な方向であろう。プログラミングを覚えるより先に、「自分の研究対象をどのような構造として提示すべきか」を考えることが肝要である。

■次の百年は「構造化と共有化」の時代に

以上のように、TEIXMLは、和歌文学の「次の百年」を形づくる基盤となると思う。AIが生成する解析モデルは、TEIで整備されたデータを入力すると、文体・修辞・語彙選択の特徴を定量的に「示すことができる。つまり、和歌の美的世界が「数値的に読める」時代が来つつある。もちろん、それは人間の直感や感性を置き換えるものではなく、研究の裏付けを強化するツールである。したがって、AIと人文学の協働を成立させるためには、資料の正確な構造化が欠かせない。今回の『千五百番歌合』のTEI適用は、その第一歩である。これを契機に、勅撰集や歌論書、注釈書などの構造化とも連携していけば、和歌文学の全体像を再構成することが可能になる。

さらに、これは研究者だけに必要なものではない。本文と、わかりやすい現代語訳を対として構造化したTEIテキストは、文学研究の裾野を広げるためにも重要なものとなる。また、XMLは、電子出版のための基礎データとしても重要である。百年前、国文学が「本文校訂」という方法を確立したように、次の百年は「構造化と共有化」の時代となる

だろう。TEIデータは、そのための共通言語であり、AI時代における新たな「本文批評の場」となることが期待されるのである。

*1 今も多くの出版はDTP化されていて、ソフトウェア上で組版が行われる。このデータも、辞書のようにフォーマットが定まっている場合には、XMLで記述すると再利用性が高まる。国文学の研究書、特に注釈書や本文校訂なども、今後XML化が課題となる。

④デジタル人文学関連用語の紹介

▼デジタルヒューマニティーズ(DH)：人文学にデジタル技術を適用する学際的、国際的な研究。各領域の対象や方法を「方法論の共有地」の理念のもとに相互に参照し合い研究の深化と方法の共有を目指す。

▼TEI(Text Encoding Initiative)：人文学資料を構造的にデジタル化するための標準策定を目的としたガイドライン(タグ付けのルールを提案する一覧)と、その策定を行うコンソーシアム。

▼SGML:Standard Generalized Markup Language(標準汎用性マーク付け言語)。データの長期的利用のために開発されたマークアップ言語で、現在ウェブで使用されるHTMLのもとになった。

▼XML:Extensible Markup Language(拡張可能なマーク付け言語)。TEIで使用するマークアップ言語。Microsoft Office(ワード、エクセル、パワーポイント)の裏側でもこれが使用されており、データの変換が容易である。

▼タグ付け:PersName(人名)など半角の▽記号で対象となるテキストを挟み、入れ子状の階層構造を付与し、機械に対象の構造や意味を認識させる。マークアップともいう。

▼TEI-PS(3版)：SGMLを使用していた頃のガイドラインの第三版。書籍としても刊行された。現在の最新ガイドラインは二〇〇七年に公開されたPSでXMLを使用する。

▼GitHub：データを共有してバージョン管理を行うための、プラットフォームとなるリポジトリ。誰でも無料でアカウントを作成でき、TEIXMLを解説とともに公開できる。

▼Zenodo：オープンサイエンスを推進するリポジトリ。研究データを登録するとDOI(デジタルオブジェクト識別子)URLが付与され、研究参照されやすくなる。(幾浦)

column

日本において
TEIはどのように
導入され普及したの？

TEIは欧米を中心とした人文学研究者、図書館情報学者、研究司書が中心となって設立、運営されるプロジェクトです。欧米圏外の資料への適用は不十分でしたが、鶴見大学の矢一志氏によるタグの個別説明の日本語訳が行われました。2016年にはTEIコンソーシアムに東アジア/日本語分会が設置され、永崎研宣氏を中心としてガイドラインの日本語訳や普及のための講習会、シンポジウムの開催、ビューワの開発が行われています。2024年にはTEI古典籍ビューワが永崎氏とフェリックス・スタイルの本間淳氏によって開発され、日本語資料のTEI/XMLがデータ作成中にも確認しやすくなりました。(幾浦)

テキストの立体化と深化、そして共有化へ

来たりくる和歌文学研究の未来のために――

田渕句美子

■デジタル時代の節目の時に

和歌文学研究において、『新編国歌大観』のデータ化は大きな節目・転換であった。一九九六年に発売された最初のCD-ROMを作る際、データ化することの有用性を最初に指摘され、データベースを作ること、を積極的に提言されたのは、故久保木哲夫氏であったという。久保木氏の柔軟な思考と先見性には驚くばかりである。その時と同じように、現在も研究方法が一変するような大きな節目の時と言えよう。このデジタル時代の大きな波の中で、「和歌文学テキストデータ構築計画」の計画力と行動力、先進性に驚き、これを推進する加藤弓枝氏をはじめとするメンバーの方々に、期待と信頼を抱いている。このニューズレターを議論の場としたいということも、秀逸なアイデアである。

私自身はTEIについて、いくつかの論文や和歌文学会特別例会、シンポジウムなどでごく粗々知っているだけの全くの素人であり、このエッセイの依頼があった時には正直驚いた。が、TEIに疎い側からの見方も必要という広いお考えだろうと勝手に解釈して、誤解・失考も覚悟しつつ、たどたどしい感想を書かせていただきたい。

■TEIで古典文学研究が進展する未来

「和歌文学テキストデータ構築計画」などで既に作成されているTEI標準テキストデータ【図1】の例とそれを表示したTEI和歌ビューワ【図2】を見ると、実に立体的な注釈・研究テキストである。本文、異同、原本画像、注、参考歌などが見やすく構成されていて、縦書きなのもありがたい。直感的な操作もしやすい配置である。これまで一冊の本のあちこちのページ、数冊の本、いくつかの画面など複数ひらいて見ていたものが、一つの画面に集められているので、集中して見たり考え

たりすることができるよう。検索のツールも備わる。しかも付加するデータの分量が自由という点は書籍では不可能であり、これは大きい。底本や対校伝本の本文がクリック一つで切り替えられるのも便利である。勿論こうした異文・画像情報だけに留まらない。盛田帝子氏編『十番虫合絵巻』のように、TEI標準で日本語訳と英語訳が対照されているテキストも増えてほしい。また今後、著作権の問題はあるけれども、あるテキスト本文を掲げて、たとえば古注の集成、古筆切の集成、近現代の注釈書の集成、あるいはさまざまな言語の翻訳の集成も一度に見られるようになっていくのだろうか。実現すれば、テキストを基盤にして研究していく私達にとって、テキストの革命的な立体化・深化であろう。

私達が作品としてテキストを読む時には、将来も書籍などの紙のテキストや注釈を手放すことはないとと思う。紙媒体の研究書も、影印・翻刻も必要であり、残さなければならない。全体をじっくり読むことができる紙の書籍が、駆逐されるという方向には行かないのではないか。むしろTEIの利便性によって古典文学研究の可能性が押し広げられて、紙媒体とデジタルとの往還の中で、テキストデータ・研究データが共有化される傾向が強まり、必要に応じて知的財産として守りながら、研究が進展していく未来が思い描かれる。というより、デジタル情報が急拡大している現在は、既にその途上にあると言えよう。

■サポート機能と確認体制の必要性

とは言え、新たな試みには常に懸念も生まれる。これまで私達の殆どはデータを使う側だったが、データを作る側に立つことはさほど容易ではないような感も抱いてしまう。「和歌文学テキストデータ構築計画」ではそのための配慮や動画の公開を予定しているようだし、TEIの入

田渕句美子 [早稲田大学 教育・総合科学学術院教授]

国文学研究資料館教授などを経て現職。著書に『中世初期歌人の研究』（笠間書院、2001年）、『阿仏尼』（吉川弘文館、2009年）、『新古今集 後鳥羽院と定家の時代』（角川選書、2010年）、『異端の皇女と女房歌人——式子内親王たちの新古今集——』（角川選書、2014年）、『女房文学史論——王朝から中世へ——』（岩波書店、2019年）、『百人一首——編纂がひらく小宇宙』（岩波新書、2024年）ほか。

門書等もあるが、やや壁もありそうだ（スキルの拙い私などは特に）。理系・実験系の研究室ならば日々顔を合わせ、方法や進展を共有できるが、それとは異なる人文系研究者には、今後のデジタル時代には、職場にも何らかのサポート機能のある共同体制が必要かもしれない。また、TEIのテキスト同士を容易に統合できるかどうかにも気になる。TEI準拠テキストの精度、統合、存続などのためには、プラットフォームを作るだけではなく、TEI準拠テキストの確認体制を作り、『新編国歌大観』の時と同じように、専門知識を持つ編集者が関わってまとめていくのが望ましいような気がしている。

『源氏物語』の論文情報データベースはできないか

この「和歌文学テキストデータ構築計画」が大きな扉をひらいてくれたが、TEIはさまざまな使い方ができそうなので、その可能性を常に諸学会でも共有することが大切だと思う。和歌関係ではないが、一案としてTEIガイドラインを活用して『源氏物語』の論文情報データベースはできないだろうか。『源氏物語』の本文を掲げ、そこにその箇所を論じている論文（著書）が掲げてあれば、至便であろう。近年『源氏物語』や作中和歌の論文を数編書いたが、取り上げたい箇所や和歌について論じている論文を探すのに、毎度多大な時間と労力が必要となる。キーワードだけではたどり着けないのだ。『源氏物語』の研究者人口は極めて多く、論文数は膨大である。苦勞している研究者は多いはずだ。それに、機会があればこの魅力的な作品の一角にアプローチしたいと思う研究者は、他分野にも多いだろうし、研究の活性化、学際化、共有化に繋がる。これは著者が各々、随時簡略に付けていくようにできないだろうか。勿論これは『源氏物語』に限らない。

■人文学研究から理解と連携を強化する

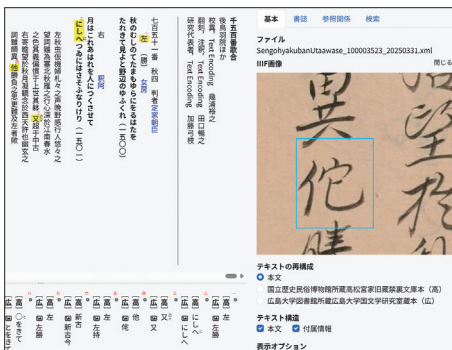
和歌文学研究の分野で、TEIを用いて精密に研究成果を織り込んでいくことは、新たな研究的可能性を帯びており、その成果と拡大は楽しみである。と同時に、それを研究者の業績として評価し、守らなければ

ならない。そして和歌関係以外の分野・領域においても、TEIという方法がどういうふうに使えるか、何が始まっていくのか、私達研究者の側から、古典文学、日本文学、日本語学、及び人文学全体を見渡して、理解と連携を強化して壁を低くしていくことができらばと思う。何と言っても、私達人文学研究者は何らかのテキストを基盤に研究を進めていくという点で共通しており、大きなチームなのだから。

来たりくる和歌文学研究の未来のために、協働・共有を見据えて計画された「和歌文学テキストデータ構築計画」を見守りつつ、心からのエールを送りたい。（ただし皆さんオーバーワークにならないようにお願いします）

*1 久保木哲夫（一九三二―二〇二三年）……都留文科大学名誉教授。専門は中古文学（和歌文学、古事切）。平安時代私家集の注釈を多数刊行。『私家集大成』CD化委員会委員、『和歌文学大辞典』編集委員。

*2 WEBサイト「十番虫合絵巻」(<https://juban-mushi-awase.dhii.jp/>)では、原本画像のほか、校訂本文・現代語訳・英訳などを公開している。なお、書籍版『江戸の王朝文化復興』ホノルル美術館所蔵レイン文庫『十番虫合絵巻』を読む（文学通信、二〇二四年）も刊行されている。



(上)【図1】『千五百番歌合』XMLデータ
(下)【図2】『千五百番歌合』TEI和歌ビュー
※ IIF 画像は広島大学図書館所蔵本
公開作品の詳細は8ページを参照

column

和歌文学本文は
どのように参照されて
きたの？

前近代の歌人は先行作品を参照する必要があったため、早くから多様な歌書が編まれてきました。平安時代の主題別歌集『古今和歌六帖』、院政期の語別分類『類聚古集』、鎌倉時代の検索書『撰句抄』『撰集佳句部類』、江戸時代の『古今類句』などがその例です。明治期には歌番号を付すなど画期的な『国歌大観』が登場し、現在は全文検索も可能な『新編国歌大観』や『新編私家集大成』が研究の基盤となっています。典型的な古典和歌の研究には膨大な暗記が前提でしたが、『国歌大観』の登場により、研究の裾野が大きく広がりました。（加藤）

はじめの一步

■TEIは新しい翻刻のカタチになるのではない

TEI (Text Encoding Initiative) という言葉を、ここ数年、耳にする機会が増えた。私がTEIを意識するきっかけになったのは、幾浦裕之氏の「百首歌・題詠・画中歌・絵入本のTEIマークアップの試み——天和三年刊・菱川師宣画『絵入藤川百首』を例として——」(日本近世文学会春季大会、二〇二三年六月)という口頭発表^{※1}だった。文学領域と情報科学領域の交錯という点で興味を惹かれたし、単なる文字列以上の情報をテキストに盛り込める点も画期的だと感じた。何より「これは新しい翻刻のカタチになるのではないか」とワクワクしたのである。

だが、「興味を持つ」から「実際にやってみよう」とはならなかった。分からないことだらけだったからである。

TEIのタグ付けに便利な^{オキシジェン}XMLなるソフト(下記コラム参照)をどこで入手すればよいのか、分からない。TEIの記述のルールがどのようなものか、分からない。どのタグをどうやって使えばいいか、分からない。タグ付けしたテキストデータを、どうすればデジタル本文として表示できるか、分からない……。もちろん、これらの疑問は、誰かに聞けば、たちどころに解決する疑問である。だが、TEI研究会にわざわざ問い合わせるような大問題でないことは、素人でも分かる。となると、「多少詳しい人」を探すわけだが、誰が「多少詳しい人」なのか分からない。結局、「なんだかよく分からない。また今度にしよう」と、TEIを遠巻きに眺める日々が続くのであった。

それだけに、今回の「イチからはじめるTEI入門講座」の開催は待ち望んでいたものであった。巷間に「入門講座」と銘打ちながら、まったく入門的な内容ではないと感じる講座も多いなか、今回の講座は紛う方なき「入門講座」であった。

■講義+実践形式の講座

「イチからはじめる」と題されていた通り、講座の内容は「TEIとは何か」「タグ付けとは何か」という概略を学んだ後、①講師の幾浦氏のデモンストレーション②幾浦氏による具体的なタグ付けに関する説明③受講生が実際にタグ付け①②③の繰り返しという形で進行的な実践。講師のデモンストレーションと説明で作業手順をイメージした後、自分で実際に手を動かすという講義+実践形式だったため、「Aというタグを使うと、Bができるようになる」と肌感覚で分かり、TEIのシンプルな操作性とそこから生み出される面白さを実感できた満足度は大きかった。

また、主催者の加藤弓枝氏のご配慮によって、専門領域の近い参加者がひとまとまりになっており、「ここ、分からないんだけど」「ちょっと見せて」といった、ささいな助け合いが自然とできたのもよかった。

TEIは、TEI Text Encoding Initiative 協会が運営する、人文学を中心とするテキスト集約のための国際標準です。

やってみよう!

特別なスキルは必要ありません! どなたでも参加できます!

参加者大募集!

「イチからはじめる」TEI入門講座

【日時】2025年2月15日(土) 13:00~17:00

【会場】名古屋大学文学部講義室128号講義室

【定員】25名 (先着順) 抽選あり

【講師】幾浦裕之 (名古屋大学文学部助教授、TEI研究会代表)

【協賛】名古屋大学文学部、名古屋大学文学部助教授、TEI研究会

【申込】2月8日(土) 18:00まで

【申込先】名古屋大学文学部助教授 加藤弓枝氏

【申込先】中世総合DBプロジェクト shinhoshigaku@gmail.com

スケジュール

13:00~13:15 開会

13:15~14:00 講義①-基本操作をマスターしよう!

14:00~15:00 実習①-タグ付けの基礎知識を学ぶ

15:00~16:00 講義②-タグ付けの応用知識を学ぶ

16:00~17:00 実習②-自分の研究に活かすには?

こんな講座です!

TEIガイドラインに準拠した人文学のテキストデータ構築方法を学ぶ入門講座です。特別なスキルは必要ありません。ノートPCを持参でき、Wordなどの文章作成ソフトで文章を入力できる方ならどなたでもご参加いただけます。テキストデータ構築に興味はあるけれど、これまでその機会がなかった方、やってみたくてもできなかった方など、大歓迎です。この講習会では、タグ付けをすることが内閣府に認められていることを知り、複数の作品でタグ付けすることでその利点を体感することを目的とします。

「イチからはじめるTEI入門講座」
2025年2月15日 名古屋大学にて開催
☞ 次回の入門講座の詳細は8ページをご覧ください

荻原大地 [昭和女子大学専任講師]

早稲田大学大学院教育学研究科教科教育学専攻単位取得後退学。博士(学術)(早稲田大学)。愛知淑徳大学文学部助教、早稲田大学教育・総合科学学術院助教を経て、現職。論文に、「『佐倉惣五郎物』実録の系譜——『佐倉花実物語』の位置づけをめぐる——」(『近世文藝』107号、2018年)、「大塩平八郎物実録の展開とその受容」(『近世文藝』112号、2020年)などがある。

おぎはらだいち
荻原大地

質疑応答の時間からも考えさせられることは多かった。「TEIのテキストデータは研究実績として認められるのかどうか」「散文・韻文・和歌・俳諧・漢詩など、分野ごとに異なるビューワの乱立を招かないか」

私見を述べれば、TEIテキストデータを、少なくとも、従来の翻刻と同程度の研究実績と評価することは可能だろう。基本的に文字列のみの提供に留まる従来の翻刻に対して、一面で本文異同や本文画像の提示、原文と校訂本文の切り替えなどを実現できるTEIテキストデータの方が情報量に富み、従来の翻刻▽TEIテキストデータとする根拠がないからである。ビューワについては、小学館の『新編日本古典文学全集』が採用した「注釈・本文・現代語訳の三段組」を再現できるビューワがあれば、もつとも汎用性が高いと思われる。このあたりは、実際にTEIを使える人が増えていくことで、合意形成が可能になるだろう。

今回のTIE入門講座はTIEの使い方だけでなく、TIEを取り巻く環境に至るまで、様々な学びを得られた講座であった。一方で、半日という時間の制限があったため、どうしても作業が駆け足になり、出遅れてしまった参加者へのフォローが弱かった感は否めない。これも講座の充実度ゆえであろうが、丸一日、あるいは複数日間開催でもよかったのかもしれない。

最後に、私の勤務校では、プロジェクト型学習科目が設けられている。その科目内で、学生に今回のT E E入門講座の内容を紹介し、俳諧作品のタグ付け【図3】を試みたところ、たちどころにxmindのタグ付けまでできてしまった。学生いわく、「わりと余裕」とのこと。今回の入門講座の内容がいかに分かりやすいものだったかが、お分かりいただけるだろう。

*1 幾浦裕之「百首歌・題詠・画中歌」
絵入本のTEIマークアップの試み——
天和三年刊・菱川師宣画『絵入藤川百首』
を例として——『近世文藝』一一九号、
二〇二四年一月に掲載。

*2 ただし、作業時間の観点では、従来の翻刻の方が手軽であり、そもそもTEIで様々なタグ付けを行えるほどの研究の積み重ねがない作品も多い（特に近世文学作品）。しばらくは、従来の翻刻で作品紹介を行い、ある程度の研究が積み重なったら、TEIでエンコーディングという流れになるのではないかと考える。



(右) TEI 入門講座当日の様子

テキストデータを作成するソフトは
何が良い？

TEIに沿ったマークアップには Oxygen XML Editor が便利ですが、無料で使える VS Code (Visual Studio Code) などでも基本的な編集は可能です。「Scholarly XML」や「XML (by Red Hat)」などの拡張機能を入れると作業がぐっと楽になります。ただし、Oxygen には便利な機能が多く、プロジェクトや丁寧なチェックには安心です。Oxygen は公式サイトから購入でき、個人や教育向けのライセンスもあります。また、アンテナハウスという日本の企業からも入手可能です。(加藤)

プロジェクト報告

詳細は、和歌文学テキストデータ構築計画 WEB サイト「研究成果」<https://wakatext.org/publications/> をご覧ください。

論文・書籍・研究報告

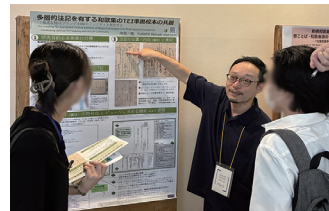
- 「特集・座談会 テキストデータのこれから——歴史と未来をつつみこむ TEI」『武蔵野文学』第七二集（2024 年 12 月、武蔵野書院）座談会：舟見一哉〔司会〕、幾浦裕之、海野圭介、加藤弓枝、田口暢之 ほかに論文掲載
※同誌は武蔵野書院公式サイトのお問い合わせフォームからお申し込みいただくことで、在庫があれば無料で入手できます。
- 和歌文学会 令和 6 年 11 月特別例会「和歌文学研究、この次の百年」『和歌文学研究』第 131 号（2025 年 12 月）
加藤弓枝「和歌文学研究、この次の百年——不易と流行——」／幾浦裕之「文字列一致を超えた和歌研究の可能性」／田口暢之「『千五百番歌合』藤原定家判の注釈と TEI」／海野圭介・舟見一哉・持田玲「次世代の和歌研究データに求められること」
- 幾浦裕之「趣意および総括」中古文学会 2025 年度春季大会 シンポジウム「AI【で】できること・データサイエンス【で】できること——いかに古典を“読み解く”か——」『中古文』第 116 号（2025 年 11 月）
- 舟見一哉「デジタル化に伴うテキスト遺産の変容——本文研究という営為と研究書を例として」『近現代日本を生きるテキスト遺産（アジア遊学 305）』（2025 年、勉誠社）
- 海野圭介「人文学データの大规模蓄積と研究資源としての利活用に向けて」日本学会会議「日本文学の伝統と現代社会」分科会（2025 年 6 月）研究報告



和歌文学会 令和 6 年 11 月特別例会
「和歌文学研究、この次の百年」

パネリスト・ポスター発表

- 海野圭介 2025 年度中世文学会春季大会 シンポジウム「古典とデータ駆動」（2025 年 5 月）パネリスト／軍記・語り物研究会 2025 年度大会 シンポジウム「デジタル時代過渡期の古典文学研究」（2025 年 8 月）パネリスト
- 舟見一哉、持田玲 DH 国際シンポジウム「東アジア／日本における人文学向けテキストデータ構造化のためのガイドライン策定に向けて」（2025 年 9 月）ポスター発表



DH 国際シンポジウム ポスター発表の様子

テキストデータ公開情報

公開中 『千五百番歌合』『石清水社歌合』『古今和歌集』TEI 準拠テキストデータ

TEI 準拠テキストデータを TEI 和歌ビューワ及び XML データで公開中 URL: <https://wakatext.org/viewer/>



予告 『百人一首歌加留多』TEI 準拠テキストデータ

『百人一首歌加留多 烏丸光広卿書』（近世後期写、大本一冊、国文学研究資料館所蔵・田安德川家資料 <https://doi.org/10.20730/200023857>）は、『百人一首』かるた札の書き方の見本帳です。墨付 27 丁の薄様に『百人一首』かるたの読み札・取り札の枠線を設け、百首の上の句と下の句を散らし書きで示します。内題、奥書もなく現在知られる限り孤本です。当プロジェクトでは、和歌の基礎的マークアップの習得も兼ねて 10 人で分担して TEI 準拠でマークアップし、『百人一首』『百人秀歌』との校異を示し、国書データベースの IIIF 画像を参照できるかたちで公開します。

お知らせ

——イチからはじめる—— TEI 入門講座

〔日時〕 2026 年 3 月 7 日（土）13:00～17:00

〔場所〕 早稲田大学国際会議場 第 3 会議室 〔定員〕 25 名 〔参加費〕 無料 〔使用言語〕 日本語

〔講師〕 幾浦裕之（文部科学省）

〔対象〕 人文学を専攻している学部・大学院生・研究者・図書館職員など

〔参加条件〕 ノート PC を持参できる方

〔申込締切日〕 2 月 28 日（土）先着順 〔申込 URL〕 <https://forms.gle/ECUtrGNtArfWwLw9>

特別なスキルは必要ありません。

テキストデータ構築に興味はあるけれど、機会がなかった方、
やってみたけどうまくいかなかった方
など、大歓迎です！

申し込み QR コードは

こちら



和歌文学テキストデータ構築計画ニューズレター 第 1 号（創刊号）【DOI】<https://doi.org/10.18999/wakbtd.1>

発行日 2026（令和 8）年 1 月 31 日 編集発行人 研究代表者 加藤弓枝（名古屋大学） 発行 和歌文学テキストデータ構築計画 <https://wakatext.org/>

刊行サイクル 年 3 回（1 月・5 月・9 月）【謝辞】本研究は JSPS 科研費 JP25K00439 の助成を受けたものです。

画像出典 ○表紙『伊勢集断簡（石山切）』（九州国立博物館所蔵）ColBase https://colbase.nich.go.jp/collection_items/kyuhaku/B14?locale=ja 『光琳画譜』（国文学研究資料館所蔵）CODH <http://codh.rois.ac.jp/iiif/iiif-curation-viewer/index.html?pages=200010512&pos=5&lang=ja> ○2・4・6 頁『伊勢集断簡（石山切）』（東京国立博物館所蔵）ColBase https://colbase.nich.go.jp/collection_items/tnm/B-1301?locale=ja

科研費