

タイトルページ

タイトル：TEI ガイドラインとは何か

著者：永崎研宣

著者所属機関：一般財団法人人文情報学研究所

責任著者：永崎研宣 [nagasaki@dhii.jp](mailto:nagasaki@dhii.jp)

キーワード：

人文情報学

デジタル・ヒューマニティーズ

デジタル人文学

電子テキスト

テキストデータベース

テキスト構造化

抄録

人文学のための電子テキスト構造化の国際的なガイドラインである TEI ガイドラインについての歴史と現状について延べ、ガイドラインの主な章についての解説を行った。そして、これらを踏まえて人文情報学における TEI ガイドラインの意義を論じた。

# TEI ガイドラインとは何か

永崎研宣（一般財団法人人文情報学研究所）

## 抄録

人文学のための電子テキスト構造化の国際的なガイドラインである TEI ガイドラインについての歴史と現状について延べ、ガイドラインの主な章についての解説を行った。そして、これらを踏まえて人文情報学における TEI ガイドラインの意義を論じた。

## はじめに

TEI (Text Encoding Initiative) ガイドラインは、人文学研究者が中心となってデジタル資料の構造的記述を行うことを目指して始まったものであり、開始当初の思想と状況については当事者達の講演録が日本語訳されている<sup>1</sup>のでそちらをご覧いただきたい。ここでは、TEI がどのようなものであり、どのように有用であるかについて、いくつかの事例とともに概観する<sup>2</sup>。

## TEI 登場の文脈

電子テキストに対してタグ付けによる注釈を行うことには少なからぬ有用性がある。テキストへのタグ付けは、1980年代後半にはすでにそれなりにできるようになっており、2018年現在ではかなり自由かつ便利な形で利用可能となっている。しかしながら、この種のコトは、技術的にできるだけでは十分ではない。各自が異なるルールでこのような記述をしてしまうと、共通のツールで利便性を高めたり、それぞれの成果を共有したりすることがきわめて難しくなってしまう。研究としては、誰も試みたことがない新しい記述手法に取り組むことには一定の意義がある。しかし、そのような記述手法は他の誰も使ったことがないので、そのように記述されたテキストデータの活用のためには新たに活用ツールも開

---

<sup>1</sup> Nancy Ide, C. Michael Sperberg-McQueen, Lou Burnard, TEI: それはどこからきたのか。そして、なぜ、今もなおここにあるのか?, デジタル・ヒューマニティーズ, 2018, 1 巻, p. 3-28, 公開日 2019/01/18, Online ISSN 2189-7867, [https://doi.org/10.24576/jadh.1.0\\_3](https://doi.org/10.24576/jadh.1.0_3), [https://www.jstage.jst.go.jp/article/jadh/1/0/1\\_2/\\_article/-char/ja](https://www.jstage.jst.go.jp/article/jadh/1/0/1_2/_article/-char/ja).

<sup>2</sup> 本稿は永崎研宣「歴史データのさまざまな応用」『歴史情報学の教科書』文学通信, 2019 を踏まえ加筆したものである。

発しなければならない。新しい記述手法を誰かが開発するたびにそれに合わせた活用ツールも開発しなければならないというのでは、いつまで経っても効率化を図ることができない。

かなり深刻な問題にもつながり得る話であり、それを回避するためには、それほど目新しくなくても、むしろ皆が共通で使える記述手法を定めた方がよいということになる。欧米でデジタルテキストの活用に関わる研究者たちはこれに気がついて対処を始め、それが一つの大きな流れになったのは1987年のことであった。

1987年の冬、ニューヨーク州ポキプシーに集まった研究者らは、長い議論の末に、一つの原則を共有するに至った。これは、会議の地の名を冠し、ポキプシー原則と名付けられた<sup>3</sup>。以下に引用してみよう。

1987年11月13日、ニューヨーク、ポキプシー

1. ガイドラインは、人文学研究におけるデータ交換のための標準的な形式を提供することを目指す。
2. ガイドラインは、同じ形式でテキストのデジタル化をするための原理を提案することも目指す。
3. ガイドラインは、以下のことをすべきである。
  - 形式に関して推奨される構文を定義する。
  - テキストデジタル化のスキーマの記述に関するメタ言語を定義する。散文とメタ言語の双方において新しい形式と既存の代表的なスキーマを表現する。
4. ガイドラインは、様々なアプリケーションに適したコーディングの規則を提案するべきである。
5. ガイドラインには、そのフォーマットにおいて新しいテキストを電子化するための最小限の規則が入っているべきである。
6. ガイドラインは、以下の小委員会によって起草され、主要なスポンサー組織の代表による運営委員会によってまとめられる。
  - テキスト記述
  - テキスト表現
  - テキスト解釈と分析
  - メタ言語定義と、既存・新規のスキーマの記述
7. 既存の標準規格との互換性は可能な限り維持されるだろう。
8. 多くのテキスト・アーカイブズは、原則として、交換形式としてのそれらの機能に関

---

<sup>3</sup> Text Encoding Initiative, Design Principles for Text Encoding Guidelines, 14 December 1988, <https://tei-c.org/Vault/ED/edp01.htm>

して、そのガイドラインを支持することに賛成した。私たちは、この交換を効率化するためのツールの開発を援助するよう、支援組織に働きかける。

9. 既存の機械可読なテキストを新しい形式に変換することとは、それらの規則を新しい形式の構文に翻訳するということを意味しており、まだデジタル化されていない情報の追加に関して何か要求されることはない。

人文学者や情報工学者、図書館司書たちによって支えられた TEI (Text Encoding Initiative) と呼ばれるこの動向は、その後、TEI ガイドラインを策定するとともに、TEI 協会 (Consortium) を設置し、参加者による自律的で民主的な運営体制の下、ガイドラインの改良を続けていくことになる。この動きがやがて XML の策定に影響を与え、さらにその後、TEI ガイドライン自体も XML をベースとするものに移行することになる。

## TEI ガイドラインの内容

TEI 協会は、一般的な意味での標準規格というものは目指さずに、あくまでもガイドラインを提示するというを当初より決めていたようである。この理由には、人文学が研究成果刊行の手段として著書の出版にこだわるのが深く関わっているように思える。人文学においては、しばしば、議論を正確に展開するために、用語とその定義、そしてそれらの関係を、一般的な用法とは必ずしも一致しない形で厳密に定義することがある。言うなれば、術語体系が、著書等のひとまとまりの研究業績ごとに異なっているという状況があり得るのである。もちろん、研究資料となる資料においても同様の状況があり得る。厳密に定められた術語体系を強要するのではなく、十分に議論した結果をガイドラインとして提示して実際の用法は利用者・利用者コミュニティに委ねるという TEI の手法は、このような人文学のあり方に寄り添ったものとして捉えることができる。

とはいえ、まったく標準規格と無縁というわけではない。部分的には国際標準化機構 (ISO) の標準規格になっている。ガイドライン第 18 章 素性の構造 (Feature Structures) は ISO 24610-1:2006 において参照されており、ガイドライン第 9 章 辞書 (Dictionaries) ISO 24613-4:2021 において TEI serialization として取り込まれている。また、参照する規格としては、言語コードを定める ISO 639 や ISO 8601 等、よく用いられる国際標準規格はガイドラインからも参照されている。

本稿執筆時点での TEI ガイドラインは、P5 のバージョン 4.3.0 となっており、非常に多くの XML タグ・属性等を定義している<sup>4</sup>。ガイドラインの目次を見ることでその全体像を

---

<sup>4</sup> TEI Consortium, P5: Guidelines for Electronic Text Encoding and Interchange, 31st August 2021, <https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html>

る程度把握することができるので、以下にそれを概観してみよう。

- 1 The TEI Infrastructure
- 2 The TEI Header
- 3 Elements Available in All TEI Documents
- 4 Default Text Structure
- 5 Characters, Glyphs, and Writing Modes
- 6 Verse
- 7 Performance Texts
- 8 Transcriptions of Speech
- 9 Dictionaries
- 10 Manuscript Description
- 11 Representation of Primary Sources
- 12 Critical Apparatus
- 13 Names, Dates, People, and Places
- 14 Tables, Formulæ, Graphics and Notated Music
- 15 Language Corpora
- 16 Linking, Segmentation, and Alignment
- 17 Simple Analytic Mechanisms
- 18 Feature Structures
- 19 Graphs, Networks, and Trees
- 20 Non-hierarchical Structures
- 21 Certainty, Precision, and Responsibility
- 22 Documentation Elements
- 23 Using the TEI

第1章では TEI ガイドラインが提示する仕組みの全体像を示しており、技術面に興味がある人は一読してみると面白いだろう。

第2章はヘッダーについての解説である。ヘッダーは、TEI が登場した際のきわめて重要な要素であった。テキストファイルにはしばしば、「このデータがどういうものであるか」「どのようなルールでテキストデータ化されたのか」といったことについての説明が欠けており、どのように扱えばよいか判断が難しいことがあった。TEI では、ヘッダーを必須化し、その種の情報をテキストファイルの中に詳細に記述できるようにしたのである。ヘッダーの役割はそれにとどまらず、様々な資料に関する情報を記載できるようになっており、とりわけ、貴重資料に関しては、詳細な書誌事項を記載でき、欧米の研究図書館で貴重資料の詳細なメタデータ記述に広く用いられている。

第3章は、すべての TEI 準拠文書で使えるエレメントの説明である。この章は大変長く、通常の文書で利用するようなエレメント・属性、そしてその使い方の例が豊富に提示されている。2021年3月に導入されたルビのルールもこの章に導入された。

第4章は、テキストの構造の捉え方について、様々なパターンを提示し、タグによってその構造を記述する手法について解説している。テキストのまとまりをどのように位置づけるか、という例示から、一つの本に複数の資料が含まれる事例やアラビアンナイトのような枠物語、詩集の構造など、様々な事例が挙げられている。

第5章は、書字体系や外字等が扱われており、日本語資料を扱う上で生じてくる外字もこのルールに従うことである程度うまく情報が共有できるようになっている。欧米の資料だとアルファベットだけで済むから楽だという話が聞かれることがあるが、中世の資料では字種が多様に存在し、Unicode では表現できない外字もまだ残されていることから、Medieval Unicode Font Initiative が Unicode への外字登録を目指した活動を続けている模様である。Unicode への文字の登録に関しては、近年、コンピュータの処理性能の大幅な向上に伴い、古典籍・古文書等に登場する学術用途でしか使われないような文字・文字体系も積極的に登録されるようになってきている。手続きとしては、まず国際標準規格である ISO/IEC 10646 への追加が承認されてから Unicode 規格もそれに追従することになっており、新しい文字の追加は、ISO/IEC の規格への登録という形をとることになる。カリフォルニア大学バークレー校を拠点とする Script Encoding Initiative という団体がこの動きを幅広くサポートしている。漢字の登録に関しては、IRG という漢字検討の専門グループがいったん検討した上で ISO のワーキンググループに提案するという手順を踏むことになっている。したがって、漢字を登録する場合には、まずは IRG に提案しなければならないのが現状である。ただし、IRG も近年は学術用途の漢字登録に寛容になっており、文字同定や証拠資料に関する所定のルールを踏まえた上で要登録文字であると判断されれば基本的には登録されるようになってきている。時間はかかるものの、Unicode に登録することによるメリットは大きく、その必要がある文字はなるべく登録しておきたいところである。

第6章以降は、韻文詩、戯曲、演説の文字起し、辞書、手稿の書誌情報、一次資料の記述、校訂情報、と、資料の性質に合わせた詳細な記述の仕方が提示されている。とりわけ、手稿の記述の仕方には非常に力が入っており、欧米有力大学図書館の研究司書が中世写本の日録情報をデジタル化したりデジタル画像に書誌情報をつけたりする際に広く用いられている。また、校訂テキストの異文情報の記述の仕方（第12章）も充実している。

第13章は、人名・地名等の固有表現に関する記述の仕方であり、これはどの種類の資料にも適用可能な有用性の高いルールである。

第 17 章では言語コーパスを作成するための単語やフレーズ、文章等の様々な単位に対して付与すべきタグ・属性について解説されている。単語レベルで文法情報を付与していきたい場合にはぜひ参照されたい。

第 20 章では、本来階層構造をとるべき XML のデータを TEI の形式でうまく表現するための様々な工夫が紹介されている。XML はタグの構造としては階層構造でなければならない、人文学向けのデータとしては必ずしも相性がよいとは言えない面がある。それを乗り越えるための手法はすでに長く研究されてきており、手法も確立されている。それが解説されているのがこの章である。

第 21 章は、人文学によるルールであることを象徴する興味深いものである。テキストデータへのタグ付けにおいては、タグ付けされた文書内の様々な要素（固有名詞やその解説等）の信頼度は必ずしも十分に高い場合ばかりではなく、また、専門家によってその信頼度についての判断が異なる場合もあり、さらには、見解を異にする場合もある。そういった場合に、一つの最適解のみを記述するのではなく、それぞれがどれくらい信頼できるのか、そして、誰に責任があるのか、ということを示すのである。そうすることで、信頼度や見解が異なる情報も含めて一つのテキストデータにまとめられるようになる。この章で解説されているのは、それを実現するための XML タグ・属性等の記述の仕方である。

このように、TEI ガイドラインの目次を見ることで TEI のおおまかな内容が見えてくる。タグをなるべく統一的にすることで人文学としての汎用性を持たせようとするものの、一方で個別の資料や個別の研究手法の固有性を無視することはできない。したがって、TEI ガイドラインとしては、全体に共通するものと個別の分野に関わるものとを分けてそれぞれをケアする形になっているのである。

## TEI ガイドラインのアップデートと「国際化」

TEI ガイドラインは現在、半年に一度、アップデートを行っている。なぜアップデートが必要なのかと言えば、主にコミュニティの拡大と要求の高度化のためである。TEI ガイドラインは、人文学のためのガイドラインとして策定されているものの、人文学のすべてに対応できているわけではない。コミュニティを基盤とする標準としてガイドラインを策定しているため、コミュニティに参加する研究分野が増えると、これまでになかったタグや構造が必要になることがある。一方で、特に技術が進歩していくと、それにあわせて新たなタグや構造が必要になることもある。そこで、TEI ガイドラインは着々とアップデート

を継続しているのである。

これを支える TEI 協会では、現在は、会員からの選挙に基づき理事会と技術委員会のメンバーを選出し、協会の運営は理事会、ガイドラインの策定は技術委員会が担当している。また、メンバーの要求に応じた分科会を設置の手続きが用意されており、個別の分野・手法におけるガイドライン拡張については、分科会で検討されることが多い。その検討の成果が会員総会で発表されるとともに、ガイドラインに導入されることになる。書簡に関する分科会や写本に関する分科会、オントロジーに関する分科会、図書館に関する分科会などは、大きな貢献を行ってきた。

近年の興味深い貢献としては、書簡に関する分科会の活動の成果として 2015 年のバージョン 2.8.0<sup>5</sup>に追加された<correspDesc>である。これは書簡に特徴的な構造としての宛先・差出人・差出地・送受信日時等を記述するタグをまとめるタグである。それ以前もそういった情報を個別に書くための構造は TEI ガイドラインで提供されていたものの、文書の中にタグが分散してしまい、それらを探して処理するのは必ずしも容易なことではなかった。バージョン 2.8.0 では、そのような情報を<correspDesc>以下にまとめるというルールが設定され、それにより、書簡の TEI/XML 準拠データを処理したい場合、宛先・差出人等の書簡に特徴的な情報は<correspDesc>以下を探索すれば見つけられることになったのである。この構造を決めるにあたっては、世界中の書簡デジタル化プロジェクトに取り組む人々がこの分科会に集い議論が重ねられ、それが最終的に TEI ガイドラインに反映されたのである。

コミュニティの参加者により民主的に策定される標準としての TEI ガイドラインは、コミュニティへの参加が少なかった非欧米圏の資料や手法への対応は、当然のことながらあまり手厚くはない。欧米の文献研究に取り組む欧米の研究者が主体であったため、必然的に欧米の資料と研究手法が暗黙的にも明示的にも前提となっている面がかつては強かった。一方で、ガイドラインをより汎用的なものにしていくためには世界中の研究者を巻き込んでいく必要があるという認識もあったようであり、TEI ガイドラインにおけるタグの個別説明の部分だけはガイドライン P5 が公開された 2007 年の後に複数言語に翻訳され、このときは鶴見大学の矢一志氏が取り組んだ。

実際のところ、非欧米圏からの参加は少なかったとは言え、それでも、近代日本の資料であれば多くの状況に対応可能であり、少し無理をすれば振り仮名や漢文の返り点にも対応できないことはなかった。しかしながら、古典籍・古文書になると、くずし字の連綿体やヲコト点など、ガイドラインに沿うだけでは構造化が難しい資料が増えてくる。そういっ

---

<sup>5</sup> <https://tei-c.org/Vault/P5/current/doc/tei-p5-doc/readme-2.8.0.html>



た事情と対応の必要性が TEI 協会においても共有されてきた結果、東アジア／日本語分科会が 2016 年に TEI 協会に設置された。この分科会では、TEI ガイドラインの翻訳・日本語による日本語のためのテキスト構造化ガイドライン策定・日本語資料を適正に構造化するための TEI ガイドラインの改訂案提出を目指して活動しており、遠隔ビデオ会議システムを活用して世界各地の有志により作業が進められているところである。その成果として、2021 年春の改訂時には日本語のルビのルールを TEI ガイドラインに導入することに成功し、その後も追加・改良すべき要素について検討と提案を継続している。

## 人文情報学における TEI ガイドラインの意義

TEI ガイドラインは、欧米の人文情報学 (Digital Humanities) においては、長らく中心的な役割を果たしてきた。人文情報学とは、人文学においてデジタル技術を応用する研究を総称するものであり、このことは、国際デジタル・ヒューマニティーズ学会組織連合 (Alliance of Digital Humanities Organizations) がオックスフォード大学出版局から刊行する論文誌のタイトルは Digital Scholarship in the Humanities であることからみとれる。

人文情報学は、方法論の共有地 (Methodological Commons) を体現するコミュニティとして形成されているものである。ここで言う方法論の共有地とは、人間文化の研究に関わるあらゆる分野が自らの方法論を持ち寄り、そこにデジタル技術を適用することで新たな展開を共創する場であり、人文情報学の重要な側面の一つは、それを実現するためのコミュニティ活動である。そして、そのようにして方法論を持ち寄った際、各分野同士で必ずしもかみ合うとは限らない様々な術語や慣習を共通言語とするための役割を果たすのが TEI ガイドラインなのである。

人文情報学における TEI ガイドラインの役割について具体的にみるには、人文情報学についての解像度をもう少しあげてみる必要があるだろう。分類する側の立場や観点によって様々な分類の仕方があり得るが<sup>6</sup>、人文情報学におけるアプローチの仕方には、大別して、(1) 構築系、(2) 共有系、(3) 解析系の 3 つがあると筆者としてはとらえている。

(1) 構築系は、デジタルデータの構築に関わる研究である。データを作るだけでは研究とは言えない、ということはこの種の取り組みにおいて聞かれることがある。確かに、決まったルールに沿ってデータを入力するだけであれば専門知や創造性が発揮される場面は

---

<sup>6</sup>たとえば、後藤真「人文情報学と歴史学」『歴史情報学の教科書』では人文情報学の研究傾向として「発見系・解析系・可視化系」を挙げている。

少なく、結果として、研究である主張できる要素はかなり少なくなってしまうだろう。しかし、データの構築においては様々な専門知が前提となる。構築しようとするデータがどのようなものであり、それはどう分類され注釈されるべきか、という検討を行うにあたっては、データの元になる資料についての専門知の有無が大きく影響するだろう。さらに、構築するデータは、効果的に活用されねばならない。そのためには、効果的な活用とはどのようなものであり、それを実現するためにはどのようにしてデータ構築をすべきなのか、ということについての理解を持っていなければ適切な構築は難しい。TEI ガイドラインが取り組んできたことの中心的な課題はこの点であり、また、欧米の人文情報学においてもこのことは一つの研究テーマとなってきた。

「どのようにデータ構築するか」というテーマは、TEI ガイドラインが扱うようなテキストの構造化に限られない。よりミクロに、文字や、さらに、資料を構成する物質の情報をどう記述するか、という問題も含まれる。文字情報は、記号論を含む深い検討が必要とされるものであり、すでに Unicode という包括的な解決策が共有されているものの、学術的なデジタルデータにおいては様々な問題が残され、議論も盛んに行われている。資料を構成する物質についても、近年、資料の化学分析や高精細デジタルマイクロスコープによる観察・分析が行われるなど、テキスト資料の物体としての研究も盛んになりつつあるが、その分析結果が内容の分析や資料の歴史的文脈においてどのように位置づけられるのか、といったことや、逆に、そこから逆算した分析手法や、それを適切に記述する手法の確立といったことも今後の大きな課題になっていくだろう。さらに、図書館情報学や博物館学、アーカイブズ学におけるメタデータ、文字情報、デジタル画像や 3D、動画などのマルチメディアコンテンツなど、どのタイプのデータにとっても、専門知をどう活かすかという観点から、人文情報学としての研究になり得る。そのような文脈において、TEI ガイドラインは、テキストを含む様々な資料におけるテキストや、物体としての資料に対する注釈等で用いられるテキストをめぐる、人文情報学における構築系の研究課題が集約される場の一つとなっており、世界各地の人文情報学関連の学会研究会などで TEI ガイドラインを巡る議論が行われている。そして、TEI 協会としても査読付きジャーナルを刊行するに至っている。

(2) 共有系は、デジタルデータの共有をめぐる様々な課題である。実践を中心とした技術的な事柄のみならず、ここからの派生として、法的な課題やメディア論としての議論、あるいはそれらにおける公平性など、どちらかと言えば理論に関わるテーマも含まれる。そもそも人文情報学が新たな装いをまとうきっかけの一つとなった Web2.0 と呼ばれる高度な双方向 Web は、この共有系の議論に親和性が高く、データを共有するだけでなく、データの提供者が利用者にもなるような、集合知が実現される状況を専門知の世界でどう咀嚼していくか、といったことを実証的に研究できる場であり、すでに様々な成果が国内外で発表されている。日本の代表的な例には「みんなで翻刻」がある。

このなかで、TEI ガイドラインが果たす役割は、(1) 構築系研究を踏まえてでてきた成果として、共有のための道具立てを提供することである。この道具立てには、単にデータの構造に関する仕様としてのガイドラインだけでなく、教材や教授法、各分野における具体的な適用の仕方など、様々な要素が含まれる。また、TEI ガイドラインをはじめとする人文学における資料構築とそれをめぐるコミュニティの形成は、それ自体が一つの研究テーマにもなり得るかもしれない。

(3) 解析系は、人文情報学においては花形の研究である。解析し、その結果をグラフ等に整理して可視化し、それによって新たな知見を得たり、それをわかりやすく提示したりする。あるいは、そこまでいかずとも、新たな知見につながるヒントを得られることも多いだろう。ここには、便利なツールやプログラミング技術など、いろいろなものが投入されることになり、ディープラーニング等の人工知能関連技術を活用するものも増えてきている。新たな知見を得るための解析だけでなく、解析した結果を構築に利用する研究、たとえば、日本語のくずし字をコンピュータで読み取って文字起こしをするくずし字 OCR の研究なども行われている。昨今注目を集めつつある Transkribus もその流れに位置するものと言えるだろう。解析系研究における TEI ガイドラインの役割は、解析のための基礎を提供するということになるが、一方で、解析系研究を通じてデータ構造の課題が生じてくることがあるとしたら、それを構築系研究にフィードバックするにあたり、TEI ガイドラインが提供する共通言語という機能は大きな役割を果たすことだろう。

人文情報学は、構築系／共有系／解析系の研究が円環をなして発展していくものであり、その背景として円環運動を支える「方法論の共有地」において、TEI ガイドラインはその典型的なアプリケーションの一つとして機能している。そして、この円環に入り、人文情報学、そしてそれを通じて人文学そのものを発展させていくために、TEI ガイドラインは良質な入り口の一つとなっているのである。

## 終わりに

TEI ガイドラインは、自然にできたものでもなければ、どこかの技術者たちが作ったものでもなく、志を持つ人文学研究者たちが自ら集い、作り上げてきたものであり、そこに込められた様々な意味は、ガイドラインを読み込めば読み込むほど、そして、これに従ってタグを付ければ付けるほど、その試行錯誤の痕跡も含めてよくわかってくる。深く理解した果てで、それでもどうしても改訂すべきだと思ったなら、そうする道もひらかれている。そのような意識で、TEI ガイドラインに取り組んでみていただきたい。

## 参考文献

- [1] Nancy Ide, C. Michael Sperberg-McQueen, Lou Burnard, TEI：それはどこからきたのか。そして、なぜ、今もなおここにあるのか？, デジタル・ヒューマニティーズ, 2018, 1 巻, p. 3-28, 公開日 2019/01/18, Online ISSN 2189-7867, [https://doi.org/10.24576/jadh.1.0\\_3](https://doi.org/10.24576/jadh.1.0_3), [https://www.jstage.jst.go.jp/article/jadh/1/0/1\\_2/\\_article/-char/ja](https://www.jstage.jst.go.jp/article/jadh/1/0/1_2/_article/-char/ja).
- [2] 永崎研宣「歴史データのさまざまな応用」『歴史情報学の教科書』文学通信, 2019
- [3] Text Encoding Initiative, Design Principles for Text Encoding Guidelines, 14 December 1988, <https://tei-c.org/Vault/ED/edp01.htm>
- [4] TEI Consortium, P5: Guidelines for Electronic Text Encoding and Interchange, 31st August 2021, <https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html>
- [5] TEI Consortium, Text Encoding Initiative P5 Guidelines version 2.8.0, 2015, <https://tei-c.org/Vault/P5/current/doc/tei-p5-doc/readme-2.8.0.html>
- [6] 後藤真, 人文情報学と歴史, 歴史情報学の教科書, 文学通信, 2019.

## 利益相反に関する開示

本論文の内容に関し、著者には開示すべき COI はない。

## 著者の貢献

著者が本論文のすべてを執筆した。