

デジタル・ヒストリー入門

2018 Spring

Tokyo Digital History Symposium

開催報告

01 情報の入手



02 情報の分析



04 情報の公開



03 情報の表現

Tokyo Digital History 編

編著者一覧

小風尚樹 / 中村覚 / 櫻田宗紀 / 山王綾乃 / 小林拓実 / 清原和之

金甫榮 / 福田真人 / 山崎翔平 / 槇野翔 / 小川潤 / 田中聡

橋本雄太 / 宮本隆史 / 菊池信彦 / 後藤真 / 崎山直樹 / 元ナミ / 加藤諭



Supported by

後援

Historians' Workshop



人文情報学拠点
Digital Humanities Initiative

東京大学大学院人文社会系研究科 次世代人文学開発センター 創成部門
Center for Evolving Humanities, Graduate School of Humanities and Sociology, The University of Tokyo



国立歴史民俗博物館
総合資料学の創成
INTEGRATED STUDIES OF CULTURAL AND RESEARCH RESOURCES



図書出版

文学通信

Bungaku-Report.com

Table of Contents

目次



主旨説明

研究データの海を渡る

小風尚樹

… p. 6



発表予稿集

デジタル時代に史料とどう向き合うか

金甫榮

… p. 9

巨大な史料群のデータを一括入手する

福田真人

… p. 10

データの再現性を担保する

山崎翔平

… p. 11

テキスト群から語の使用傾向を分析する

小川潤

… p. 12

データの表現方法を探索する

山王綾乃

… p. 13

歴史地図にデータを可視化する

小林拓実

… p. 14

デジタル技術で分野を越境する

小風尚樹

… p. 15

データの活用から公開までを展望する

纓田宗紀

… p. 16





歴史研究者のための技術マップ

… p. 17



参考文献一覧

… p. 29



開催実績

告知用ポスター

… p. 37

参加事前申込者の専門分野内訳

… p. 38

シンポジウム当日の動画共有実績

… p. 39

参加者が興味を持った技術のグループ分け

… p. 40

匿名アンケートに見る今後の課題

… p. 41





参加記

アーカイブズ学の立場から	清原和之	・・・ p. 46
人文情報学の教育的側面から	橋本雄太	・・・ p. 48
デジタル・ヒストリーの戦略的側面	宮本隆史	・・・ p. 49
「情報の公開」セクションのチェアを担当して	中村寛	・・・ p. 50
パネルディスカッションに登壇して	菊池信彦	・・・ p. 51
パネルディスカッションに登壇して	後藤真	・・・ p. 52
パネルディスカッションに登壇して	崎山直樹	・・・ p. 53
デジタル・アーカイブとデジタル・ヒストリー	元ナミ	・・・ p. 54
シンポジウム参加記	加藤諭	・・・ p. 56

Preface

主旨説明

研究データの海を渡る

「史料の海に溺れる」という言葉がある
膨大な量の史料に圧倒され、歴史を編む糸口がつかめないことを指す格言だ

では、このデジタル時代、
オンライン上で研究に必要なデータの多くを手に入れられるようになった今、
歴史研究者は果たして「研究データの海を渡る」航海術を心得ているだろうか

Tokyo Digital Historyは、この航海術を示すために
歴史研究が生み出されるまでのプロセスを4つに分類する必要があると考えた



この図が表すのは
史料を入手し、内容を理解し、読者への伝え方を工夫し、論文などの形で世に出す
という4つのプロセスから構成される研究のサイクルである

このサイクル自体は、これまでもこれからも変わらないだろう
これまでの歴史研究の成果は、このサイクルを経ていることが当たり前であった

では、デジタル・ヒストリーを論じる場合に
なぜこのサイクルをプロセスに細分化する必要があるのだろうか



前述のように、歴史研究における4つのプロセスを論じる理由は
まず、人文情報学分野においてはこのようなプロセス自体が研究対象であること
そして何より、歴史研究に関連するデジタル技術や知識に習熟するには
プロセス毎に整理する必要があるためである

いま、デジタル・ヒストリーという語を
「デジタル技術や関連知識を手法として採用する歴史研究」だと定義しよう

しかし、例えばKH Coder, IIIF, Python, JavaScript, TEI/XMLなど
デジタル・ヒストリーの関連技術・知識は多いため
「興味はあっても、何から始めれば良いのかわからない」
という方が多いのではないだろうか

そこで本シンポジウムでは
デジタル・ヒストリーに関心を抱く研究者が次の一步を踏み出せるように
関連する技術の見取り図を提示する
その際、前出の4つのプロセスの間をつなぐ、更なる4つのプロセスを重視したい



この図は

情報を入手してから分析する前にデータを処理しやすい形に加工し
情報を表現するための適切なモデルを探し
情報を公開するにあたってライセンスなどを考え
第三者が入手しやすいような形でデータ形式を整える
プロセスが必要であることを表している

本シンポジウムにおける8人の発表者は
歴史研究における4つのプロセスに軸足を置きつつ
デジタル・ヒストリーにおける8つのプロセスを架橋していく

このシンポジウムの参加者が
それぞれの立場でデジタル・ヒストリーの舵取りとなることを願う

文責 | 小風尚樹

Proceedings

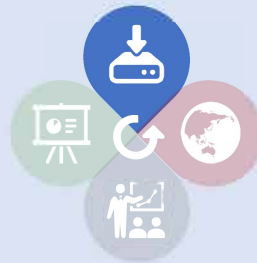
発表予稿集



次のページ以降では、各発表の要旨を掲載する
その際、シンポジウム全体における位置づけを示すために

情報の入手 / 情報の分析 / 情報の表現 / 情報の公開

4プロセスのうち、どこに力点を置くのかを下記のように図示している
この場合は、「情報の入手」プロセスに力点を置いた発表であることを示す



要旨集のページ下部にある **Digital Notes** 欄では
各発表が、どのような技術をどのような順番で用いたのかを示している
各技術の解説については、後述の技術マップおよびその解説に譲る

例えば下記の例では

ISAD (G) の知見に基づき ⇒ Python (のBeautiful Soup 4) でWeb上のデータを取得 ⇒
表計算ソフトにデータを格納 ⇒ 正規表現でデータを整形 ⇒
Python (のJSONモジュール) でデータを構造化 ⇒
最終的にD3.jsでデータの可視化をした

という作業フローを示していることになる





01 デジタル時代に史料とどう向き合うか

金甫榮（渋沢栄一記念財団デジタル・キュレーター | アーカイブズ学）

近年史料のデジタル化が進み、またデジタルで作成される記録の増加により、膨大な情報をウェブサイトを通じて提供する多様なデジタル・アーカイブが登場し、いつでも、どこでも史料情報を利用することができるようになった。このような環境の変化により、歴史研究においても直接史料に接しなくとも、インターネットが使える環境であれば、パソコン上で検索ツールを用いて目録の確認はもちろん、詳細な内容まで閲覧することが可能になりつつある。

本発表では、このような新しい研究環境と向き合う研究者が、情報を入手する際に注意しなければならない点をアーカイブズ学の観点から示す。特に、史料のコンテキスト情報を理解するために、データの出所や、階層構造、記述方法などを把握することが重要であることをアーカイブズ理論に基づき紹介する。そして、コンテキスト情報を把握する一つの手段としてメタデータを見極めることを提案する。

Keywords

デジタル・アーカイブの多義性 / アーカイブズ理論 / メタデータ

Digital Notes



ISAD (G)



02 巨大な史料群のデータを一括入手する

福田真人（東京大学日本史学D3 | 近代日本貨幣史）

日本研究については、人文情報学を活用するために必要なテキストデータが相対的に不足している。そのため、情報の入手セクションを担当する本報告では**Webスクレイピング**技術を用いて分析可能なテキストデータを入手することに力点を置く。対象とする史料は国立公文書館デジタルアーカイブに掲載されている、典型的な明治前期における公文書群の「公文録」である。「公文録」は明治前期の政治・行政・政策についての基本的史料であるが、総件数は10万件を超えており、その全体像を既存の研究では十分に把握できていなかった。他方、オンライン上のメタデータはフォンドの階層構造や原史料に即して記述されている。そこで報告者はこうした大量の情報を一括入手し、分析を試みた。

本報告では、入手のために用いた**Webスクレイピング**技術について論じるとともに、入手したデータの利用可能性について展望し、史料階層の可視化や、計量文献学的分析について言及する。

Keywords

公文録 / 史料の階層 / Webスクレイピング

Digital Notes





03 データ加工の再現性を担保する

山崎翔平（東京大学経済史学D3 | 近代日本経済史）

本報告では、明治期県境変更の補正を事例にデータの整備・加工作業におけるプログラミング言語を用いたバッチ処理の有用性を示す。

データの整備・加工作業は、何らかの方法でデータを入手した後、分析の前準備として不可欠な過程であり、手順が複雑になれば、研究自体の再現性にも支障をきたす。作業手順をプログラミング言語のコードとして記述・公開し、それを用いてバッチ処理することにより、作業の省力化だけでなく、第三者にも加工の過程が検証しやすくなり再現性が担保される。

報告者は、長期府県パネルデータを用いる分析の予備作業として、合併・分割が相次いだ明治前期の県域を現在の47都道府県のそれと一致させるため、**Fukao et al. (2015, A2.6.1)** の手法を用い、「郡」の所属変化を追跡し、郡ごとの人口情報を用いて補正を行うプログラムを開発した。これにより、県域変更により接続が困難であった統計データを用いた長期的な視角による分析が容易に可能となった。

Keywords

府県パネルデータ / バッチ処理 / データ整備・加工
再現性 / 省力化

Digital Notes





04 テキスト群から語の使用傾向を分析する

小川潤（東京大学西洋史学M2 | 古代ローマ属州史）

歴史家にとって最も重要な能力が、史料の精読を通して問題を発見し、これを批判的に考察する能力であることは疑いない。しかし、膨大な一次史料が存在する分野では、自らの研究に関連する記述を一から、網羅的に検索することは困難であろう。そこで、データベースをうまく活用し、史料の検索から計量的な分析に至るまでを、デジタル技術を用いて行うことは、研究の効率化に大きく貢献するはずだ。

本報告では、数多のラテン語著作を収めるデータベースから、特定の語を含む箇所を一括取得し、それを分析ツールにかけるという手順で分析を行った。この際、ラテン語の複雑な格変化や動詞の活用が問題となったが、今回紹介するCLTK (Classical Language Toolkit)を用いた古典言語語処理によって解決した。

デジタル技術を用いることで、頻度表や共起ネットワークという形でテキスト情報を可視化し、文字情報のみでは得られない視点からテキスト解釈を行うことができ、こうした分析と、史料の精読の間を行き来することで、これまでよりも深く、また多角的に史料と対話することが可能になるだろう。

Keywords

Perseus Digital Library / 古典語テキスト解析 / 『ガリア戦記』

Digital Notes





05 データの表現方法を探索する

山王綾乃（お茶の水女子大学西洋史学D1 | 近世フランスアカデミー史）

歴史研究者で、地図やグラフの作成に苦心し、参考資料として先行研究から引用する人は多いだろう。しかし引用した図は、自身の研究結果を提示する上で最適なものだといえるだろうか。

本発表では、18世紀フランスのパリ王立科学アカデミー会員名簿 *Index biographique de l'Académie des sciences* を分析し、高速データ可視化ツール Tableau を使って、出身地と会員区分を地図上に可視化した例を紹介する。フロソポグラフィ研究では、まず人物に関する情報を紙面から抜き取り、手作業で表を作成した後、考えられるすべての視角から情報を分析することが求められる。作業の大部分を占める単調な入力作業をデジタル技術で短縮し、多角的な分析を可視化モデルを使って行えば、マクロな視点からの分析が容易になるはずである。

以上を通じて、フロソポグラフィ研究にプログラミング技術やデジタル・ツールが貢献しうる可能性について、従来の方法論との比較という点から検討したい。

Keywords

会員名簿 / データ可視化 / Tableau

Digital Notes





06 歴史地図にデータを可視化する

小林拓実（東京大学西洋史学M2 | 近代フランス移民史）

従来のフランス移民史では、人間の地理的移動を扱うこと、また歴史人口学との関係が深いことから、地図を用いる研究がしばしば行われている。しかし、管見の限りではそれらの地図は紙媒体に制限されており、またベースとなるデータも公開されていないために、これらの再利用・再検討は困難な状況となっている。一方で、歴史地理学の分野においてはGISを用いた研究はすでに行われており、また近年の技術的発展によって、このようなソフトが容易に扱えるようになりつつある。

本発表では、このような状況を踏まえて、発表者の関心のある第三共和政前半のマルセイユを例として、無料のGISツールであるQGISを用いて歴史地図を作成する。また、この地図に先行研究を含めた種々のデータを実際に重ねて可視化してみると、GISを用いた研究の可能性についても検討し、最後にこれらのデータをどう管理すべきかについて、展望を示す。

Keywords

INDICATEUR MARSEILLAIS / GIS / CCライセンス

Digital Notes





07 デジタル技術で分野を越境する

小風尚樹（東京大学西洋史学D3・歴博研究協力者 | 近代イギリス外交史）

モノを対象とした計量分析を扱う歴史研究では、第三者が元史料の該当箇所すべての文脈・記述にさかのぼり、分析を再現・検証することはきわめて困難である。この困難は、10世紀前後の律令制下の日本における行政マニュアル『延喜式』の研究にも見られる。すなわち、『延喜式』は、貢納品の詳細や財源規定など、当時の社会を知るきっかけを与えてくれる膨大な情報の宝庫であるがゆえに、モノを扱った計量分析事例が多く、分析の妥当性を判断することが難しいのである。

このような困難を解決することをめざすのが、国立歴史民俗博物館「古代の百科全書『延喜式』の多分野協働研究」における、延喜式テキストのTEI化プロジェクトだ。報告者は、TEIのモデル設計・実装を担う研究協力者として携わっている。史料テキストがTEI化されれば、利用者は容易に元史料の文脈・記述を確認したり、関心あるデータを機械的に即座に取得したりできるようになる。

本報告では、今後の歴史研究におけるオープンサイエンスの可能性を開く画期的な事例となるであろう同プロジェクトを紹介する。また、他分野を専門とする報告者が、技術協力者として携わることで開けた世界についても言及したい。

Keywords

延喜式 / データベース構築 / TEI

Digital Notes





08 データの活用から公開までを展望する

纓田宗紀（東京大学西洋史学D3 | 中世教皇史）

本報告では、オンライン・データベース上に公開されたデータを入手・活用し、その成果をみずから公開するまでのプロセスを示す。

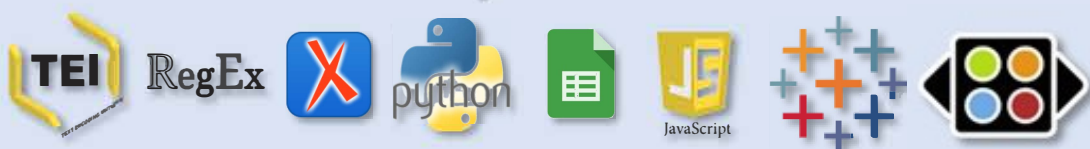
西洋中世史料のデータベース**Regesta Imperii Online**には、中世のドイツ王・皇帝、諸侯、教皇、教皇特使などが発給した文書の要約を**TEI**に準拠して記述した約**13万**件の**XML**ファイルが公開されている。報告者は、この**TEI/XML**ファイル群を用いて、**13世紀**に活動したある教皇特使の移動を可視化した作業を例示する。教皇特使に関連する史料は、滞在先各地に点在しているため、その移動の全体像を把握することには困難がともなう。しかし、テキストに付与された場所や経緯度のメタデータを利用することで、教皇特使の移動ルートを分析し、表現することが可能となる。移動を示す地図と文書内容の表現には、ウェブ上での公開を利用の前提とする年表作成ツール**TimelineJS**を用いた。

以上のように本報告は、本シンポジウムがたどってきた**4つ**のプロセスの循環を展望するものである。

Keywords

Regesta Imperii / 年表・地図 / TAPAS PROJECT

Digital Notes



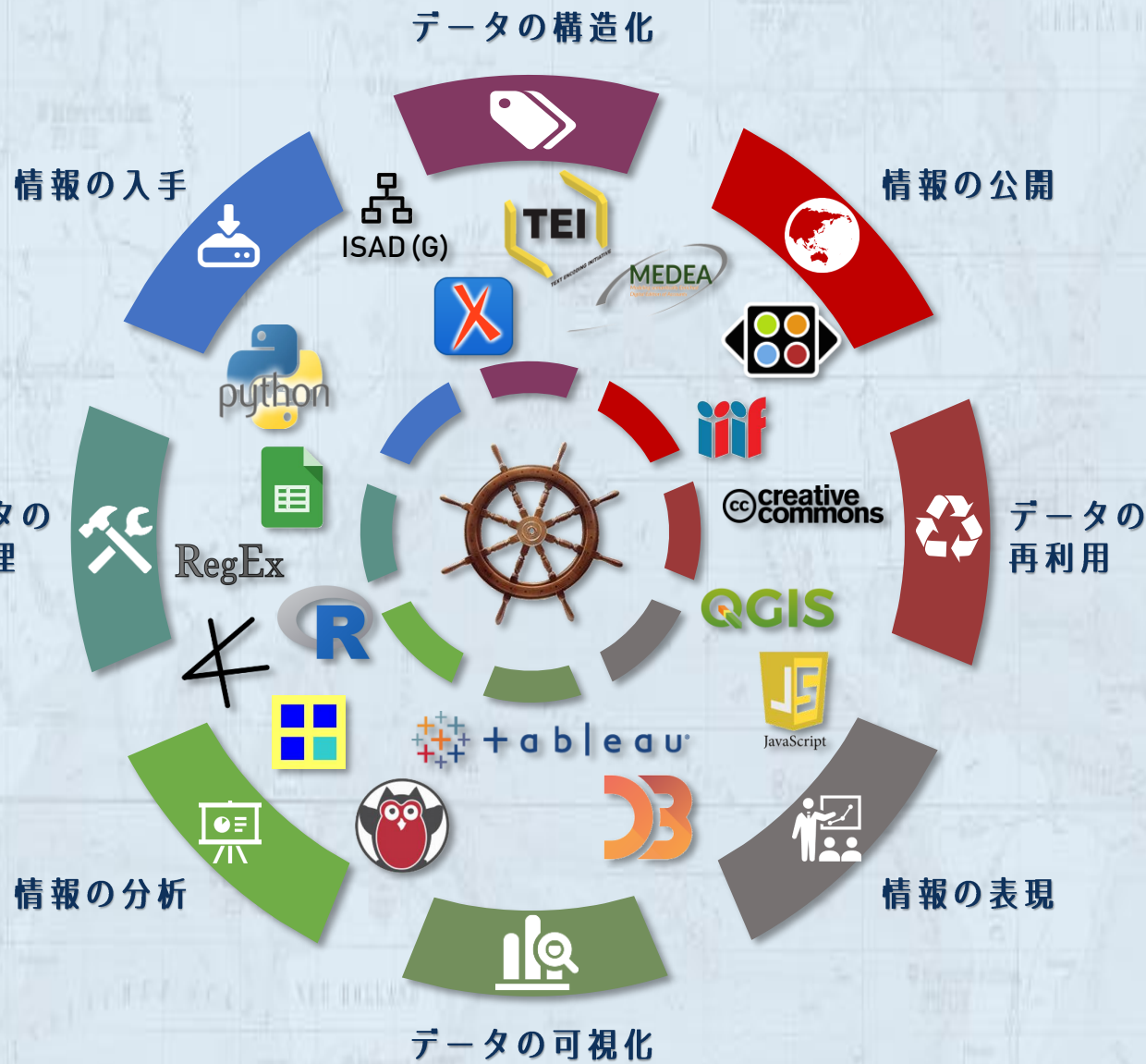
Digital Sketch



















技術マップ

Digital Sketch for Historians by ToDH

Tokyo Digital Historyが提示する歴史研究者のための技術マップ
研究データの海を渡るための「舵」をモチーフにした

歴史研究に関わるデジタルツール・プログラミング言語・関連知識・関連コミュニティについて
あくまでシンポジウムで扱った範囲・用途に限定して関係性を示した
それぞれの項目の説明については、ページ下部の目次を参照されたい
(※マップ中のアイコンについては、公式にロゴが設定されていないものも含む)



	Text Encoding Initiative	...	p. 18		R	...	p. 21		QGIS	...	p. 24
	Oxygen XML Editor	...	p. 18		CLTK	...	p. 21		JavaScript	...	p. 24
	ISAD (G)	...	p. 19		KH Coder	...	p. 22		creative commons	...	p. 25
	Python	...	p. 19		Voyant-tools	...	p. 22		IIIF	...	p. 25
	Google Spreadsheet	...	p. 20		+ a b e a u	...	p. 23		MEDEA	...	p. 26
	RegEx 正規表現	...	p. 20		D3.js	...	p. 23		TAPAS	...	p. 26

legend

凡例



次のページ以降で説明する技術は、それぞれ様々な目的・用途で汎用的に活用できるものがほとんどである

ここでは、歴史研究における諸プロセスにどう関連するかについて中央の技術アイコンを囲む8つのイメージを選択的にハイライトすることによってシンポジウムで扱った範囲における適用可能性を示すこととした

例えば下記のTEIについては
データの構造化・情報の入手
データの再利用・情報の公開
といったプロセスと関係が深いことを示している



各技術を解説するにあたって

正式名称 / 公式サイト / 概要
関連が深いプロセス / 本シンポジウムで扱ったプロジェクト
関連情報 / 学習難易度（5段階） / 参考文献等

について、あくまで目安として同時に示すこととした



名称

Text Encoding Initiative

公式サイト

<http://www.tei-c.org/index.xml>

概要

人文学史資料のテキストを機械可読形式で構造化し、データベースなどの形で共有しやすくするための国際的な枠組み。自分の研究に役立つTEI形式のデータを手に入れるだけでなく、自らもTEI形式でデータを作成することで、研究基盤の構築に貢献できる。デジタル版の刊行史資料を作成する基礎技術と言えよう。

関連が深いプロセス

データの構造化 / 情報の入手
データの再利用 / 情報の公開

扱ったプロジェクト

小川潤 / 山王綾乃 / 小風尚樹 / 櫻田宗紀

関連情報

XML (eXtensible Markup Language)
Oxygen XML Editor
正規表現
MEDEA / TAPAS

学習難易度

★★★★☆

参考文献等

Burnard, Lou & Syd Bauman, 2007, *TEI: P5 Guidelines*. <http://www.tei-c.org/Guidelines/P5/>
Pierazzo, Elena, 2016, 'Textual Scholarship and Text Encoding', in Susan Schreibman, et al., eds., *A New Companion to Digital Humanities (NCDH)*, Chichester, pp. 307—321.
永崎研宣、2015年「仏教文献のための構造的なデジタルテキストの記述と活用」『印度学仏教学研究』63 (2)、1088—1094頁。



名称

Oxygen XML Editor

公式サイト

<https://www.oxygenxml.com/>

概要

TEIに準拠したテキストを作成するための最も便利なテキストエディタのひとつ。正規表現やXPath検索など多機能で、TEI以外にも、XMLやJSONなど様々な形式のデータの構文を検証し、エラーの詳細を表示してくれる。ライセンスの購入が必要だが、30日間の無料トライアル版も提供されている。

関連が深いプロセス

データの構造化 / データの前処理 / 情報の分析

扱ったプロジェクト

山王綾乃 / 小風尚樹 / 櫻田宗紀

関連情報

TEI / XML
正規表現

学習難易度

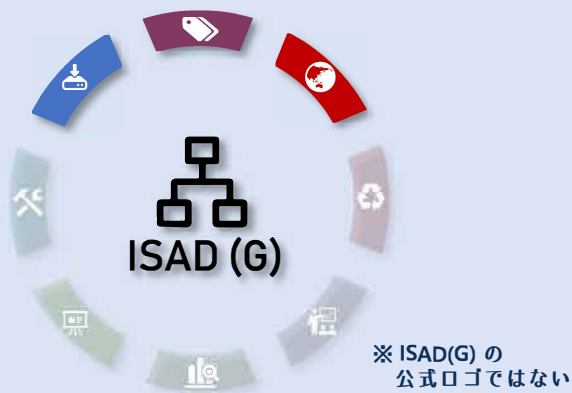
★★☆☆☆

参考文献等

Cummings, James, 2017, 'An introduction to Markup, XML, TEI, and the oXygen XML Editor'. <https://prezi.com/jiwc-yg9wmlq/an-introduction-to-markup-xml-tei-and-the-oxygen-xml-editor/>

永崎研宣、2017年「デジタルアーカイブにおけるテキスト資料の扱い：oXygenで日本語解説付きTEIガイドラインを利用する」。digitalnagasakiのブログ:

<http://digitalnagasaki.hatenablog.com/entry/2017/08/26/095642>



名称

ISAD(G)

公式サイト

<https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>

概要

アーカイブズ資料の編成を記述するための標準。記録史料のコンテキスト・出所を重視し、フォンドというかたまりで資料をとらえ、記録の階層性を反映したマルチレベル記述を採用している。コンテキストや内容、構造、利用制限などの7領域において合計26要素のメタデータを提供している。



名称

Python

公式サイト

<https://www.python.org/>

概要

画像認識をはじめ人工知能分野でも活用されている汎用プログラミング言語で、様々な目的に応用される。本シンポジウムでも、Webデータの一括取得やデータの前処理、自然言語処理の技術として、各方面で大車輪の働きを見せてくれた。構文がシンプルで読み書きがしやすく、歴史研究者がこれからプログラミングを始めるならPythonがオススメである。

関連が深いプロセス

データの構造化 / 情報の入手 / 情報の公開

扱ったプロジェクト

金甫榮 / 福田真人

関連情報

ICA | International Council on Archives
ISSAR(CPF) / ISDF / ISDIAH
EAD

学習難易度

★★★☆☆

参考文献等

アーカイブズ・インフォメーション研究会、2001年『記録史料記述の国際標準』北海道大学図書刊行会。
安藤正人、1998年『記録史料学と現代：アーカイブズの科学をめざして』吉川弘文館。
五島敏芳、2003年「日本の記録史料記述EAD/XML化と記録史料管理：記録史料管理過程におけるEAD利用の位置をめぐる」『情報知識学会誌』12(4)、3—21頁。
森本祥子、2003年「アーカイブズの編成と記述標準化」国文学研究資料館史料館編『アーカイブズの科学（下巻）』柏書房、236—260頁。

関連が深いプロセス

情報の入手 / データの前処理 / 情報の分析
データの可視化 / 情報の表現 / データの再利用

扱ったプロジェクト

福田真人 / 山崎翔平 / 小川潤 / 山王綾乃
小林拓実 / 小風尚樹 / 櫻田宗紀

関連情報

Beautiful Soup 4 | Webスクレイピング
正規表現
TEI / XML / JSON
CLTK / NLTK | 自然言語処理

学習難易度

★★★☆☆

参考文献等

Pythonチュートリアル
<https://docs.python.jp/3/tutorial/>
Lubanovic, Bill（斎藤康毅監訳、長尾高弘訳）、2015年『入門Python3』オライリー・ジャパン。
Sweigart, Al（相川愛三訳）、2017年『退屈なことはPythonにやらせよう：ノンプログラマーにもできる自動化処理プログラミング』オライリー・ジャパン。



名称

Google Spreadsheet

公式サイト

https://www.google.com/intl/ja_jp/sheets/about/

概要

Googleが開発した無料の表計算ソフトで、Microsoft Excelと同レベルの関数計算やグラフ作成ができる。サーバー上で即時保存しながら動くため、処理速度ではExcelに劣るが、Tableauなど他サービスとの連携に優れ、複数端末・ユーザー間でのファイル共有が容易である。

| 関連が深いプロセス
すべて

| 扱ったプロジェクト
山王綾乃 / 小風尚樹 / 櫻田宗紀

| 関連情報
Microsoft Excel
Tableau
TimelineJS
SQL

| 学習難易度
★☆☆☆☆

| 参考文献等
武井一巳、2011年『ポケット百科Gmail & Googleドキュメント：知りたいことがズバツとわかる本』翔泳社。
武田雅人ほか、2017年『はじめてのスプレッドシートの教科書（Kindle版）』Tekuru Inc。
立山秀利、2017年『仕事のExcelが1日でざっくりわかる本：ネコの手を借りるより、3つの「自動化」で』SBクリエイティス。



名称

正規表現 | Regular Expression

公式サイト

特になし

概要

高度な検索・置換を可能にする文字列の表現方法。通常の文字と、プログラムで特別な意味を持たせた「メタ文字」を組み合わせることで、文字列のパターンを表現する。各種プログラミング言語やテキストエディタで使用できるため、低い学習コストで高い汎用性が期待できる。特にOCRをかけたテキストを整形する作業（データクリーニング）や、半自動で大量のTEIタグを付ける作業などに便利。

| 関連が深いプロセス
データの構造化 / データの前処理

| 扱ったプロジェクト
福田真人 / 山王綾乃 / 小林拓実
小風尚樹 / 櫻田宗紀

| 関連情報
ワイルドカード
TEI / Oxygen XML Editor
Python / JavaScript

| 学習難易度
★☆☆☆☆

| 参考文献等
O'Hara, Laura Turber, 2013, 'Cleaning OCR'd text with Regular Expressions'. The Programming Historian:
<https://programminghistorian.org/lessons/cleaning-ocrd-text-with-regular-expressions>
岩谷宏、2008年『入門正規表現：検索・置換・テキスト処理に強くなる!』技術評論社。
Friedl, Jeffrey E. F. (株式会社ロングテール・長尾高弘訳)、2008年『詳説正規表現（第3版）』オライリー・ジャパン。



| 名称
R

| 公式サイト
<https://www.r-project.org/>

| 概要
統計解析向けのオープンソース・フリーソフトウェアであり、データ可視化や高度な分析手法にも対応し、初心者から上級者まであらゆるデータ・サイエンティストの需要に応える。Pythonほど汎用性は高くないが、統計分析の手法についてはPythonよりも解説が多く、情報が手に入りやすい。



| 名称
CLTK | Classical Language Toolkit

| 公式サイト
<http://cltk.org/>

| 概要
その名の通り、ラテン語・ギリシア語はもちろん、古エジプト語、ヘブライ語、サンスクリット語など、地域を問わず数多くの古典語に対応している。韻律分析や活用形の表示、語幹や見出し語への変換など、多様な自然言語処理を行うことができる。公式サイトにモデルコードが掲載されているため、環境設定さえ乗り越えれば容易に活用することができる。

| 関連が深いプロセス
情報の入手 / データの前処理 / 情報の分析
データの可視化 / 情報の表現 / データの再利用

| 扱ったプロジェクト
山崎翔平

| 関連情報
Python / Pandas
GIS
ネットワーク分析
KH Coder

| 学習難易度
★★★★☆

| 参考文献等
Munzert, Simon et al. (石田基広ほか訳)、2017年『Rによる自動データ収集：Webスクレイピングとテキストマイニングの実践ガイド』共立出版。
本橋智光、2018年『前処理大全：データ分析のためのSQL/R/Python実践テクニック』技術評論社。

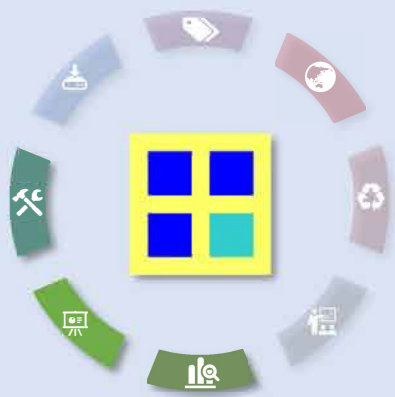
| 関連が深いプロセス
データの前処理 / 情報の分析 / データの再利用

| 扱ったプロジェクト
小川潤

| 関連情報
Python
自然言語処理
NLTK | Natural Language Toolkit
Voyant-tools

| 学習難易度
★★★★☆

| 参考文献等
Bird, Steven et al. (萩原正人ほか訳)、2010『入門自然言語処理』オライリー・ジャパン。
Johnson, Kyle P., 2016, CLTK Docs
<http://docs.cltk.org/en/latest/>
浅尾仁彦・李在鎭、2013年『言語研究のためのプログラミング入門：Pythonを活用したテキスト処理』開拓社。



| 名称
KH Coder

| 公式サイト
<http://khc.sourceforge.net/>

| 概要
樋口耕一氏が開発し2001年に公開された計量テキスト分析用ソフトウェア。多変量解析などを用いる客観的分析から、コーディングルールを用いて分析者の主観を反映した分析にまで幅広く対応している。フリーソフトとして公開されており、2018年4月現在、これを利用した研究事例は2000件を超えた。

| 関連が深いプロセス
データの preprocessing / 情報の分析 / データの可視化

| 扱ったプロジェクト
福田真人

| 関連情報
Voyant-tools
自然言語処理

| 学習難易度
★★☆☆☆

| 参考文献等
樋口耕一、2014年『社会調査のための計量テキスト分析：内容分析の継承と発展を目指して』ナカニシヤ出版。
——、2017年「計量テキスト分析およびKH Coderの利用状況と展望（特集号：テキストマイニングをめぐる方法論とメタ方法論）」『社会学評論』68 (3)、334—50頁。



| 名称
Voyant-tools

| 公式サイト
<https://voyant-tools.org/>

| 概要
デジタルテキストを読解・分析するためのWebサービス。テキストをアップロードするだけで頻度表、ワードクラウド、共起ネットワークなど、様々なテキスト解析の結果を可視化してくれる。プログラミングの高度なスキルがなくても共起範囲の指定、ストップワードの作成などを容易に行えるため、初心者にもオススメのテキスト解析ツールである。

| 関連が深いプロセス
データの preprocessing / 情報の分析
データの可視化 / 情報の表現

| 扱ったプロジェクト
小川潤

| 関連情報
KH Coder

| 学習難易度
★☆☆☆☆

| 参考文献等
Sinclair, Stefan & Geoffrey Rockwell, 2016, 'Text Analysis and Visualization: Making Meaning Count', in *NCDH*, pp. 274—90.
小風尚樹、2017年「歴史研究者のためのデジタル道工具箱」『クリオ』31、2—9頁。
<http://hdl.handle.net/2261/72555>



名称
Tableau Software

公式サイト
<https://www.tableau.com/ja-jp>

概要
データ分析とチャート作成に特化した分析ツールで、他の表計算・データベースサービスとの連携を必要とする。年間ライセンスの購入が必要だが、無料のアカデミック版がある。マウス操作で動かせるためプログラミングの知識は必須ではない。スピーディーに、多彩なデータ表現方法を試せる。

関連が深いプロセス
情報の分析 / データの可視化 / 情報の表現
データの再利用 / 情報の公開

扱ったプロジェクト
山王綾乃 / 小風尚樹 / 櫻田宗紀

関連情報
SQL | データベース言語
Google Spreadsheet / Microsoft Excel
R

学習難易度
★★☆☆☆

参考文献等
小野泰輔ほか、2017年『Tableauデータ分析：入門から実践まで』秀和システム。
木田和廣、2016年『TableauビジュアルWeb分析：データを収益に変えるマーケティングの武器』インプレス。
Steele, Julie, Noah Iliinsky & 増井俊之、2011年『ビューティフルビジュアライゼーション』オライリー・ジャパン。



名称
D3.js | Data Driven Documents

公式サイト
<https://d3js.org/>

概要
単純なチャート（棒・円・折れ線など）だけでなく、ネットワークや階層図など複雑なグラフをWebブラウザ上に描画することができる。JavaScriptを基に設計されており、各種の前提知識が必要。ユーザの操作に対応する動的なデータの表現も可能なので、データ可視化の技術としての人気は非常に高い。

関連が深いプロセス
情報の分析 / データの可視化 / 情報の表現
データの再利用 / 情報の公開

扱ったプロジェクト
福田真人 / 小風尚樹

関連情報
HTML / CSS / JavaScript
SVG | Scalable Vector Graphics
Tableau

学習難易度
★★★★★

参考文献等
Dale, Kyran（嶋田健志監修・木下哲也訳）、2017年『PythonとJavaScriptではじめるデータビジュアライゼーション』オライリー・ジャパン。
Murray, Scott（長尾高弘訳）、2014年『インタラクティブ・データビジュアライゼーション：D3.jsによるデータの可視化』オライリー・ジャパン。
古籟一浩、2014年『データビジュアライゼーションのためのD3.js徹底入門』SBクリエイティブ。



| 名称
QGIS

| 公式サイト
<https://www.qgis.org/ja/site/>

| 概要
地理的・空間的情報を統合的に管理・操作・可視化するGIS (Geographical Information System) アスリケーションの一つ。OSGeo (Open Source Geospatial foundation) のプロジェクトで、オープンソースかつ無料で使用可能なことが最大の特徴。他の高機能なGISソフトと比べて遜色がなく、また多くの追加機能がプラグインとして追加できる。



| 名称
JavaScript

| 公式サイト
<http://www.tei-c.org/index.xml>

| 概要
主にWebブラウザ上の動きを制御するプログラミング言語。Javaと混同されやすいが別物である。Webデザインに必要な技術なので、HTMLやCSS, jQueryなどと合わせて学習したい。
本シンポジウムでは、D3.jsやTimeline.jsを活用し、複雑なデータを効果的に可視化するための基礎技術として用いている。

| 関連が深いプロセス
情報の分析 / データの可視化 / 情報の表現
データの再利用 / 情報の公開

| 扱ったプロジェクト
小林拓実

| 関連情報
ISO19139
CSDGM | Content Standard for Digital
Geospatial Metadata
INSPIRE Metadata Directive
JSIG2.0

| 学習難易度
★★★★☆

| 参考文献等
QGISオフィシャル・ドキュメント(version 2.18)
<https://www.qgis.org/ja/docs/index.html>
QGISユーザーマニュアル(version 2.18)
https://docs.qgis.org/2.18/ja/docs/user_manual/
「QGISの基本的な操作」
<https://qiita.com/adamski/items/0b7c65e1447973ed720a>

| 関連が深いプロセス
データの前処理 / 情報の分析 / データの可視化
情報の表現 / データの再利用 / 情報の公開

| 扱ったプロジェクト
福田真人 / 小風尚樹 / 櫻田宗紀

| 関連情報
D3.js / jQuery
HTML, CSS
Webデザイン

| 学習難易度
★★★★☆☆

| 参考文献等
狩野祐東、2015年『確かな力が身につくJavaScript「超」入門』SBクリエイティブ。
山田祥寛、2016年『(改訂新版) JavaScript本格入門：モダンスタイルによる基礎から現場での応用まで』技術評論社。
吉田真麻、2015年『HTML5/CSS3モダンコーディング：フロントエンドエンジニアが教える3つの本格レイアウト：スタンダード・グリッド・シングルページレイアウトの作り方』翔泳社。



名称

Creative Commons License

公式サイト

<https://creativecommons.org>
<https://creativecommons.jp/>

概要

著作物の適正な再利用を目的とした新たな著作権ルールで、著作者が自らの著作物をどのような条件で使えるか意思表示するツール。著作権を保持したまま、クレジット表示や改変禁止、同じ利用条件を保持すること、営利利用の禁止の4つの条件・計6通りのライセンスを付与することができる。

関連が深いプロセス

情報の入手 / データの再利用 / 情報の公開

扱ったプロジェクト

小林拓実

関連情報

パブリック・ドメイン

著作権法

RightsStatements.org

<http://rightsstatements.org/en/>

学習難易度

★☆☆☆☆

参考文献等

永崎研宣、2013年「人文学分野とサイバーインフラストラクチャ: デジタル・ヒューマニティーズにおける現状と課題」『情報の科学と技術』63 (9)、369—76頁。

福島幸宏、2015年「歴史資料のデジタル化とオープンデータ化の実践と理念」『情報の科学と技術』65 (12)、515—18頁。



名称

IIIF | International Image Interoperability Framework

公式サイト

iiif.io/

概要

画像を公開するリポジトリ間の相互運用を支援する主要な研究図書館による国際的な枠組み。IIIF対応のビューワを利用することで、ユーザーは1つのインターフェースで画像を閲覧・比較し、註釈をつけられる。絵巻物の全貌を一覧したり、挿絵が切り抜かれて散逸した写本の原型をブラウザ上で再現するなど、IIIFで画像閲覧の常識が覆るだろう。

関連が深いプロセス

データの構造化 / データの前処理 / 情報の公開

扱ったプロジェクト

小風尚樹

関連情報

API | Application Programming Interface

Web Annotation (W3C)

Omeka

TEI / XML

学習難易度

★★★★☆

参考文献等

北本朝展、2017年「IIIF Curation Viewerの開発と利用」IIIF Japanシンポジウム: デジタルアーカイブにおける画像公開の新潮流。

鈴木親彦・高岸輝・北本朝展、2017年「IIIF Curation Viewerが美術史にもたらす「細部」と「再現性」: 絵入本・絵巻の作品比較を事例に」『じんもんこん2017論文集』157—164頁。

永崎研宣、2017年「デジタル文化資料の国際化に向けて: IIIFとTEI」『情報の科学と技術』67 (2)、61—66頁。



名称

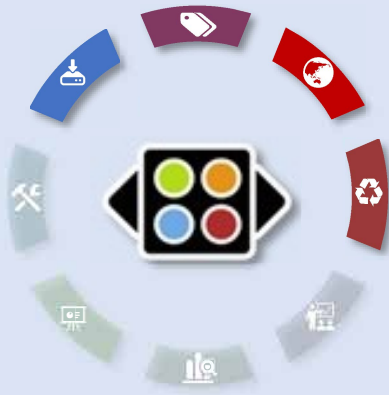
MEDEA Workshop |
Modeling semantically Enriched
Digital Edition of Accounts

公式サイト

<https://medea.hypotheses.org/>

概要

複式簿記に代表される財務記録史料を構造化するための理論を構築する国際的ワークショップ。アメリカとドイツの大規模な研究助成金で運用されている。本シンポジウムでは、延喜式テキストのTEI化プロジェクトとの関係が深い。



名称

TAPAS Project

公式サイト

<http://tapasproject.org/>

概要

TEIファイルを保存・公開・共有するためのリポジトリ。個人やグループ単位で作成するプロジェクトページでTEIファイルを持続的に管理し、他のユーザーを招待して議論することができる。様々な種類の人文学資料をマークアップするためのガイドやテンプレートも提供する。

関連が深いプロセス

データの構造化 / 情報の入手
データの再利用 / 情報の公開

扱ったプロジェクト

小風尚樹

関連情報

TEI / XML
TAPAS
Transactionography

学習難易度

★★★★☆☆

参考文献等

Tomasek, Kathryn & Syd Bauman, 2013, 'Encoding Financial Records for Historical Research', *Journal of the Text Encoding Initiative*, Issue 6. DOI: [10.4000/jtei.895](https://doi.org/10.4000/jtei.895).

———, 2016, 'Modeling Semantically Enriched Digital Edition of Accounts', [hcommons.org](https://hcommons.org/deposits/item/hc:12317/). <https://hcommons.org/deposits/item/hc:12317/>

小風尚樹、2016年「歴史研究とデジタル・ヒューマニティーズの相関」東京大学学術機関リポジトリ。 <http://hdl.handle.net/2261/60849>

関連が深いプロセス

データの構造化 / 情報の入手
データの再利用 / 情報の公開

扱ったプロジェクト

小風尚樹 / 櫻田宗紀

関連情報

TEI / XML
Oxygen XML Editor
リポジトリ

学習難易度

★☆☆☆☆

参考文献等

Flanders, Julia & Scott Hamlin, 2013, 'TAPAS: Building a TEI Publishing and Repository Service', *Journal of the Text Encoding Initiative*, Issue 5. DOI: [10.4000/jtei.788](https://doi.org/10.4000/jtei.788).

南山泰之、2016年「研究データ管理における機関リポジトリの可能性」『大学図書館研究』103、17—23頁。

Selected Bibliography

参考文献一覧



次のページから、参考文献一覧を示す
シンポジウムを企画するにあたって示唆を受けたものと
各発表に関する参考文献を
人文学分野と技術分野とに分けて挙げることにした



author-date方式を採用し
人文学・技術分野ごとに欧語文献・邦語文献の順に挙げる
技術分野の文献は、欧語・邦語限らず
アルファベット順→五十音順の順に挙げた
また、DOI (Digital Object Identifier)やパーマリンクがあるものは
積極的に表記することとした
URLのアクセス確認日時はすべて2018年4月12日である



なお、参考文献一覧表を作成するにあたっては
文献管理ソフトのうち、Zoteroを採用した



グループで文献の書誌情報を管理し
それぞれのメンバーに関する文献を各自の名前とともに
タグ付けすることによって
巻末文献目録を作成しやすいようにしている

Tokyo Digital HistoryのZoteroグループライブラリのURLは
以下の通りである

[https://www.zotero.org/groups/
2077256/tokyo_digital_history](https://www.zotero.org/groups/2077256/tokyo_digital_history)

こうしたZoteroの活用方法は、コミュニティやグループで
書籍を出版する際などに有用であろう

シンポジウム全体に関わるもの

- American Historical Association, 2015, 'Guidelines for the Professional Evaluation of Digital Scholarship by Historians'. <https://www.historians.org/teaching-and-learning/digital-history-resources/evaluation-of-digital-scholarship-in-history/guidelines-for-the-professional-evaluation-of-digital-scholarship-by-historians> (菊池信彦・小風尚樹・師茂樹・後藤真・永崎研宣訳、2016年「歴史学におけるデジタル研究を評価するためのガイドライン」東京大学学術機関リポジトリ、<http://hdl.handle.net/2261/59142>)
- Arguing with Digital History Working Group, 2017, *Digital History & Argument White Paper*, Roy Rosenzweig Center for History and New Media. <https://rrchnm.org/argument-white-paper/>
- Cohen, Daniel & Roy Rosenzweig, 2005, *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*, Philadelphia.
- Guldi, Jo & David Armitage, 2014, *The History Manifesto*, Cambridge, UK.
<https://www.cambridge.org/core/what-we-publish/open-access/the-history-manifesto> (平田雅博・細川道久訳、2017年『これが歴史だ！：21世紀の歴史学宣言』刀水書房)
- Schreibman, Susan, Ray Siemens & John Unsworth, eds., 2016, *A New Companion to Digital Humanities* (hereafter cited as *NCDH*), 2nd Edition, Chichester.
- Siemens, Ray, 2016, 'Communities of Practice, the Methodological Commons, and Digital Self-Determination in the Humanities', *Digital Studies/Le Champ Numérique*. DOI: [10.16995/dscn.31](https://doi.org/10.16995/dscn.31)
- The Programming Historian. <https://programminghistorian.org/>
- Tufte, Edward R., 2001, *The Visual Display of Quantitative Information*, Cheshire.
- 青木祐一、2017年「《シリーズ 歴史家とアーキビストの対話》アーカイブズとアーキビストをめぐる3つの誤解」『歴史学研究』960、29—31頁。
- 青山英幸、2002年『記録から記録史料へ：アーカイバル・コントロール論序説』岩田書院。
- 新井由紀夫、1987年「パソコンを利用した歴史学のデータベースについて」『クリオ』2、23—43頁。
<http://hdl.handle.net/2261/72654>
- 安藤正人、1998年『記録史料学と現代：アーカイブズの科学をめざして』吉川弘文館。
- 菊池信彦、2013年「《特集 デジタル時代の歴史学》デジタルヒューマニティーズ/デジタルヒストリーの情報源：デジタル時代の歴史学を考えるために」『現代史研究』59、55—68頁。
- 、2015年「西洋史 DH の動向とレビュー：西洋史学はウェブ情報をどのように位置づけているのか：『研究入門』を題材に」『人文情報学月報』46。 <https://www.dhii.jp/DHM/dhm46-1>
- 清原和之、2015年「アーカイブズ資料情報の共有と継承：集合記憶の管理を担うのは誰か」九州史学会・史学会編『過去を伝える、今を遺す：歴史資料、文化遺産、情報資源は誰のものか』山川出版社。
- 国文学研究資料館編、2014年『アーカイブズの構造認識と編成記述』思文閣出版。

古賀崇、2017 年「「デジタル・アーカイブ」の多様化をめぐる動向：日本と海外の概念を比較して」『アート・ドキュメンテーション研究』24、70—84 頁。

<http://opac.tenri-u.ac.jp/opac/repository/metadata/4389/>

後藤真、2018 年「日本における人文情報学の全体像と総合資料学」国立歴史民俗博物館編『歴史研究と〈総合資料学〉』吉川弘文館、1—16 頁。

———・田中正流・師茂樹、2009 年『情報歴史学入門』金壽堂出版。

小風尚樹、2017 年「歴史研究者のためのデジタル道具箱」『クリオ』31、2—9 頁。

<http://hdl.handle.net/2261/72555>

———、2017 年「歴史研究におけるデジタル利活用の可能性」第 33 回歴史コミュニケーション研究会発表資料。<https://www.slideshare.net/NaokiKokaze/ss-79827098>

近藤和彦、1992 年「ABC に H を加味して：歴史家とパソコン」『クリオ』6、80—84 頁。

<http://hdl.handle.net/2261/72675>

———、2004 年「修正主義をこえて」史学会編『歴史学の最前線』東京大学出版会、101—120 頁。

長野壮一、2015 年「〈特集 デジタル時代の歴史学〉デジタル歴史学の最新動向：フランス語圏におけるアーカイブ構築およびコミュニティ形成の事例紹介」『現代史研究』61、39—47 頁。

中村覚・大和裕幸・稗方和夫・満行泰河・鈴木淳・吉田ますみ、2018 年「Linked Data を用いた歴史研究者の史料管理と活用を支援するシステムの開発」『情報処理学会論文誌』59 (2)、267—277 頁。

橋本雄太、2015 年「〈特集 オープンデータ〉人文学資料オープンデータの可能性と現状」『情報の科学と技術』65 (12)、525—530 頁。DOI: [10.18919/jkg.65.12_525](https://doi.org/10.18919/jkg.65.12_525)

逸見龍生・小関武史、2018 年『百科全書の時空：典拠・生成・転位』法政大学出版局。

三浦崇・伊東幸宏・小西達裕・田村貞雄・赤石美奈・中谷広正・阿部圭一、1997 年「歴史学研究支援システムの構築」『情報処理会研究報告人文科学とコンピュータ』48 (1997-CH-034)、25—30 頁。

南山泰之、2015 年「データジャーナル：研究データ管理の新たな試み」『カレントアウェアネス』325、19—22 頁。DOI: [10.11501/9497651](https://doi.org/10.11501/9497651)

宮本隆史、2011 年『デジタル・ヒストリー：スタートアップガイド』風響社。

吉澤誠一郎・岡本隆司・村上衛・村田雄二郎・西英昭・石川禎浩・斎藤希史、2012 年「座談会：近代中国研究の現状と課題」岡本・吉澤編『近代中国研究入門』東京大学出版会、233—263 頁。

Dale, Kyran (嶋田健志監修・木下哲也訳)、2017 年『Python と JavaScript ではじめるデータビジュアライゼーション』オライリー・ジャパン。

Lima, Manuel (三中信宏訳)、2015 年『THE BOOK OF TREES—系統樹大全：知の世界を可視化するインフォグラフィックス』ビー・エヌ・エヌ新社。

Sweigart, Al (相川愛三訳)、2017 年『退屈なことは Python にやらせよう：ノンプログラマーにもできる自動化処理プログラミング』オライリー・ジャパン。

鈴木雅彦・鈴木嘉右、2015 年『データビジュアライゼーションのデザインパターン 20：混沌から意味を見つける可視化の理論と導入』技術評論社。

永原康史、2016 年『インフォグラフィックスの潮流：情報と図解の近代史』誠文堂新光社。

Section I 情報の入手

01 金甫榮 | デジタル時代に史料とどう向き合うか

Brown, Adrian, 2013, *Practical Digital Preservation: A How-to Guide for Organizations of Any Size*, London.

Duranti, Luciana & Patricia C. Franks, 2015, *Encyclopedia of Archival Science*, Lanham/ Boulder/ New York/ London.

Sabharwal, Arjun, 2015, *Digital Curation in the Digital Humanities: Preserving and Promoting Archival and Special Collections*, Oxford.

小川千代子・高橋実・大西愛編著、2003 年『アーカイブ事典』大阪大学出版会。

記録管理学会・日本アーカイブズ学会、2006 年『入門アーカイブズの世界：記憶と記録を未来に（翻訳論文集）』日外アソシエーツ。

後藤真、2008 年「文化遺産学における「デジタル」序説：保存と共有 活用と表現」『情報処理学会研究報告人文科学とコンピュータ研究会報告』79、57—64 頁。

坂口貴弘、2010 年「≪特集 書誌コントロール再考≫アーカイブズの編成・記述とメタデータ」『情報の科学と技術』60 (9)、384—389 頁。DOI: [10.18919/jkg.60.9.384](https://doi.org/10.18919/jkg.60.9.384)

森本祥子、2011 年「≪特集 日本アーカイブズ学会 2011 年度大会企画研究会報告：広がりゆく「デジタルアーカイブ」とアーカイブズ≫コメント 伝統的なアーカイブズとデジタルアーカイブ：発展的な議論を進めるために」『アーカイブズ学研究』15、55—60 頁。

02 福田真人 | 巨大な史料群のデータを一括入手する

石渡隆之、1977 年「「公文録」の一覧」『北の丸』9、30—80 頁。

国立公文書館デジタルアーカイブ：公文録

https://www.digital.archives.go.jp/DAS/meta/Fonds_F2005032421074303276

中野目徹、2000 年『近代史料学の射程：明治太政官文書研究序説』弘文堂。

———・熊本史雄、2009 年『近代日本公文書管理制度史料集』岩田書院。

西川誠、1992 年「左院における公文書処理：左院の機能に関する一考察」『日本歴史』528、66—80 頁。

———、1996 年「参事院の創設：明治 14 年政変後の太政官における公文書処理」『書陵部紀要』48、42—63 頁。

Mitchell, Ryan（黒田利明訳）、2016 年『Python による Web スクレイピング』オライリー・ジャパン。

Wang, Weixin et al., 2006, 'Visualization of Large Hierarchical Data by Circle Packing', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 517—520.

加藤耕太、2016 年『Python クローリング&スクレイピング：データ収集・解析のための実践開発ガイド』技術評論社。

Section 2 情報の分析

03 山崎翔平 | データ加工の再現性を担保する

Fukao, Kyoji, et al., 2015, *Regional Inequality and Industrial Structure in Japan: 1874—2008*, Maruzen.
石田諭司、2013 年「府県の変遷（改訂版）」。<http://www.tt.rim.or.jp/~ishato/tiri/huken/huken.htm>
袁堂軍・攝津斉彦・Jean-Pascal Bassino・深尾京司、2009 年「戦前期日本の県内総生産と産業構造」、
Global COE Hi-Stat Discussion Paper Series、No. 71、一橋大学経済研究所。

福代和宏、2010 年「Painting Maps in terms of Statistical Data: R で地図の塗り分け」。

<http://ds0.cc.yamaguchi-u.ac.jp/~fukuyo/r-map.html>

古谷知之、2011 年『R による空間データの統計分析』朝倉書店。

松浦寿幸、2015 年『Stata によるデータ分析入門 第 2 版 経済分析の基礎からパネル・データ分析まで』
東京図書。

本橋智光、2018 年『前処理大全：データ分析のための SQL/R/Python 実践テクニック』技術評論社。

04 小川潤 | テキスト群から語の使用傾向を分析する

Leveau, Philippe, 1993, « Agglomérations secondaires et territoires en Gaule Narbonnaise », *Revue archéologique de Narbonnaise*, 26 (1), pp. 277—99. DOI: [10.3406/ran.1993.1438](https://doi.org/10.3406/ran.1993.1438)

Tarpin, Michel, 2009, « Organisation politique et administrative des cités d'Europe occidentale sous l'Empire », *Pallas. Revue d'études antiques*, 80, pp. 127—45. DOI: [10.4000/pallas.1783](https://doi.org/10.4000/pallas.1783)
———, 2003, *Vici et Pagi dans l'Occident Romain*, Rome.

Bird, Steven, Ewan Klein & Edward Loper（萩原正人・中山敬広・水野貴明訳）、2010 年『入門自然言語処理』オライリー・ジャパン。

Johnson, Kyle P., 2016, The Classical Language Toolkit. <http://cltk.org/>

Sinclair, Stefan & Geoffrey Rockwell, 2016, 'Text Analysis and Visualization: Making Meaning Count', *NCDH*, pp. 274—90.

浅尾仁彦・李在鎬、2013 年『言語研究のためのプログラミング入門：Python を活用したテキスト処理』開拓社。

樋口耕一、2017 年「≪特集 テキストマイニングをめぐる方法論とメタ方法論≫計量テキスト分析および KH Coder の利用状況と展望」『社会学評論』68 (3)、334—350 頁。

Section 3 情報の表現

05 山王綾乃 | 統計データの表現方法を探索する

Institut de France, 1979, *Index biographique de l'Académie des sciences, 1666—1978*, Paris.

McClellan, James E., 1985, *Science Reorganized: Scientific Societies in the Eighteenth Century*, New York.

———, 1981, 'The Académie royale des sciences, 1699—1793: A Statistical Portrait', *Isis*, 72 (4), pp. 541—567.

Roche, Daniel, 1989, *Le siècle des lumières en province: académies et académiciens provinciaux, 1680—1789*, 2 vols., Paris.

大野誠、2013 年「フロソポグラフィ考」『化学史研究』40 (1)、34—41 頁。

Steele, Julie (牧野聡訳)、2011 年『ビューティフルビジュアライゼーション』オライリー・ジャパン。

小野泰輔・前田周輝・清水隆介・三好淳一・山口将央、2017 年『Tableau データ分析：入門から実践まで』秀和システム。

木田和廣・できるシリーズ編集部、2016 年『Tableau (タスロー) ビジュアル Web 分析：データを収益に変えるマーケティングの武器』インプレス。

06 小林拓実 | 歴史地図にデータを可視化する

Liens, Georges, 1967, « Les "Vêpres marseillaises" (juin 1881), ou la crise franco-italienne au lendemain du traité du Bardo », *Revue d'histoire moderne & contemporaine*, 14 (1), pp. 1—30. DOI: [10.3406/rhmc.1967.2933](https://doi.org/10.3406/rhmc.1967.2933)

Lopez, Renée & Emile Témime, 1990, *Histoire des migrations à Marseille. Tome 2: L'Expansion marseillaise et l'invasion italienne: 1830—1918*, Aix-en-Provence.

Regnard, Céline, 2009, *Marseille la violente: criminalité, industrialisation et société, 1851—1914*, Rennes.

國府久郎、2002 年「≪西洋史特集：移動する人々≫マルセイユにおける移住現象（1806—1911 年）：国内移住から「移民」の時代へ」『駿台史学』114、51—82 頁。

Bodenhamer, David J. et al., 2010, *The Spatial Humanities: GIS and the Future of Humanities Scholarship*, Bloomington.

Gregory, Ian N. & Alistair Geddes, 2014, *Toward Spatial Humanities: Historical GIS and Spatial History*, Bloomington.

Onoda, Kazuyuki et al., 2013, 'Historical Geography in Japan since 1980', 『人文地理』, 65, pp. 1—28.

川口洋、2009 年「≪特集 歴史地理情報システムの活用≫『Historical GIS の地平』シンポジウムから」『情報の科学と技術』59 (11)、545—50 頁。

矢野桂司・瀬戸寿一、2013 年「地理情報システムを用いた地理学と歴史学の連携：歴史 GIS の試み」『アリーナ』15、13—19 頁。

Section 4 情報の公開

07 小風尚樹 | デジタル技術で分野を越境する

国立歴史民俗博物館編、2018 年『歴史研究と〈総合資料学〉』吉川弘文館。

小山修三、1970 年「古代アワビ産業の発達」『国史学』81、18—39 頁。

虎尾俊哉・日本歴史学会、1995 年『〈新装版〉延喜式』吉川弘文館。

林和弘・村山泰啓、2015 年「科学技術動向研究：オープンサイエンスをめぐる新しい潮流（その 3）研究データ出版の動向と論文の根拠データの公開促進に向けて」『科学技術動向』148、4—9 頁。

宮原武夫、2014 年『古代東国の調庸と農民』岩田書院。

小風尚樹、2016 年「歴史研究とデジタル・ヒューマニティーズの相関」東京大学学術機関リポジトリ。

<http://hdl.handle.net/2261/60849>

Tomasek, Kathryn & Syd Bauman, 2013, 'Encoding Financial Records for Historical Research', *Journal of the Text Encoding Initiative*, Issue 6. DOI:[10.4000/jtei.895](https://doi.org/10.4000/jtei.895)

———, 2016, 'Modeling semantically Enriched Digital Edition of Accounts (MEDEA)', hcommons.org. <https://hcommons.org/deposits/item/hc:12317/>

08 櫻田宗紀 | データの活用から公開までを展望する

Kaufhold, Martin, 2000, *Deutsches Interregnum und europäische Politik. Konfliktlösungen und Entscheidungsstrukturen 1230—1280*, Hannover.

Sassen, Johannes H. H., 1908, *Hugo von St. Cher, seine Tätigkeit als Kardinal, 1244—1263*, Bonn.

Sisson, Keith D. & Atria A. Larson, eds., 2016, *A Companion to the Medieval Papacy: Growth of an Ideology and Institution*, Leiden.

Flanders, Julia & Scott Hamlin, 2013, 'TAPAS: Building a TEI Publishing and Repository Service', *Journal of the Text Encoding Initiative*, Issue 5. DOI: [10.4000/jtei.788](https://doi.org/10.4000/jtei.788)

Kuczera, Andreas & Dieter Rübsamen, 2006, 'Verborgene, vergessen, verloren? Perspektiven der Quellenerschließung durch die digitalen Regesta Imperii', in *Forschung in der digitalen Welt. Sicherung, Erschließung und Aufbereitung von Wissensbeständen*, hrsg. v. Rainer Hering et al., Hamburg, pp. 109—123.

Schulz, Julian, 2017, 'Regesta Imperii Online', *RIDE: A Review Journal for Digital Editions and Resources*, 6. DOI: [10.18716/ride.a.6.5](https://doi.org/10.18716/ride.a.6.5)

Weller, Tobias, 2014, 'Die Regesta Imperii Online', *Rheinische Vierteljahrsblätter*, 78, pp. 234—241.

Event Records

開催実績

2018 Spring Tokyo Digital History Symposium

Introduction 小風尚樹

情報化の波は歴史研究の世界にも押し寄せてきている
これまでの歴史学の蓄積と対話しながら、この研究データの海を渡るにはどうしたら良いか
本シンポジウムでは、歴史研究が生み出されるまでのプロセスを4段階に分け
それぞれに関連する情報技術やツールおよび知識を整理し、具体的な研究実践例を提示する
歴史研究におけるDigital Humanitiesの可能性を探る画期的なシンポジウムとなるだろう

Section 1 情報の入手 | Chair 清原和之

歴史研究に必要な史料は大量のデータとして入手できるようになった
アーカイブズ学の知見を活かし情報入手の作法を考察する

Section 2 情報の分析 | Chair 橋本雄太

入手した大量のデータをどう理解すればいいか
データ解析技術を活用して多面的な情報分析の手法を探る

Section 3 情報の表現 | Chair 宮本隆史

分析で得た気づきや成果をどうわかりやすく表現すればいいか
視覚的なイメージも含めて多彩な情報表現の可能性を探る

Section 4 情報の公開 | Chair 中村寛

研究の基礎データは他の研究者による二次利用を促す形式で共有されるのが望ましい
検証可能性やデータ再利用の観点から情報公開のあり方を学ぶ



Date

2018/04/15 (Sun) 13:00~18:00

Venue

東京大学本郷キャンパス
経済学研究科学術交流棟
小島ホール1階 第2セミナー室

Participation form

定員は50名程度になります
下記から事前申込をお願いします

<https://goo.gl/forms/9EDiVmGQqDSwPAxw2>

fee

参加無料

Wifi

eduroam, UTokyo WiFi / Guest

Contact Us

tokyodigitalhistory@gmail.com
Twitter @DHistory_Tokyo

Supported by

Historians' Workshop
東京大学人文情報学拠点
国立歴史民俗博物館
メタ資料科学研究センター
図書出版 文学通信

Outline and Keywords of the Main Section

渋沢栄一記念財団デジタル・キュレーター
金甫榮 | アーカイブズ学

デジタル・アーカイブスの多義性 / アーカイブズ理論
デジタル時代に史料とどう向き合うか

東大日本史D3

福田真人 | 近代日本貨幣史

公文録 / 史料群の階層 / Webスクレイピング
巨大な史料群のデータを一括入手する

東大西洋史D3

櫻田宗紀 | 中世教皇史

Regesta Imperii / 年表・地図
データの活用から公開までを展望する

東大経済史D3

山崎翔平 | 近代日本経済史

府県パネルデータ / Python / Stata
データ加工の再現性を担保する

東大西洋史D3・歴博研究協力者

小風尚樹 | 近代イギリス外交史

延喜式 / データベース構築 / TEI
デジタル技術で分野を越境する

東大西洋史M2

小川潤 | 古代ローマ属州史

Perseus Digital Library / 古典語テキスト解析
テキスト群から語の使用傾向を分析する

お茶大西洋史D1

山王綾乃 | 近世フランスアカデミー史

会員名簿 / データ可視化 / Tableau
統計データの表現方法を探索する

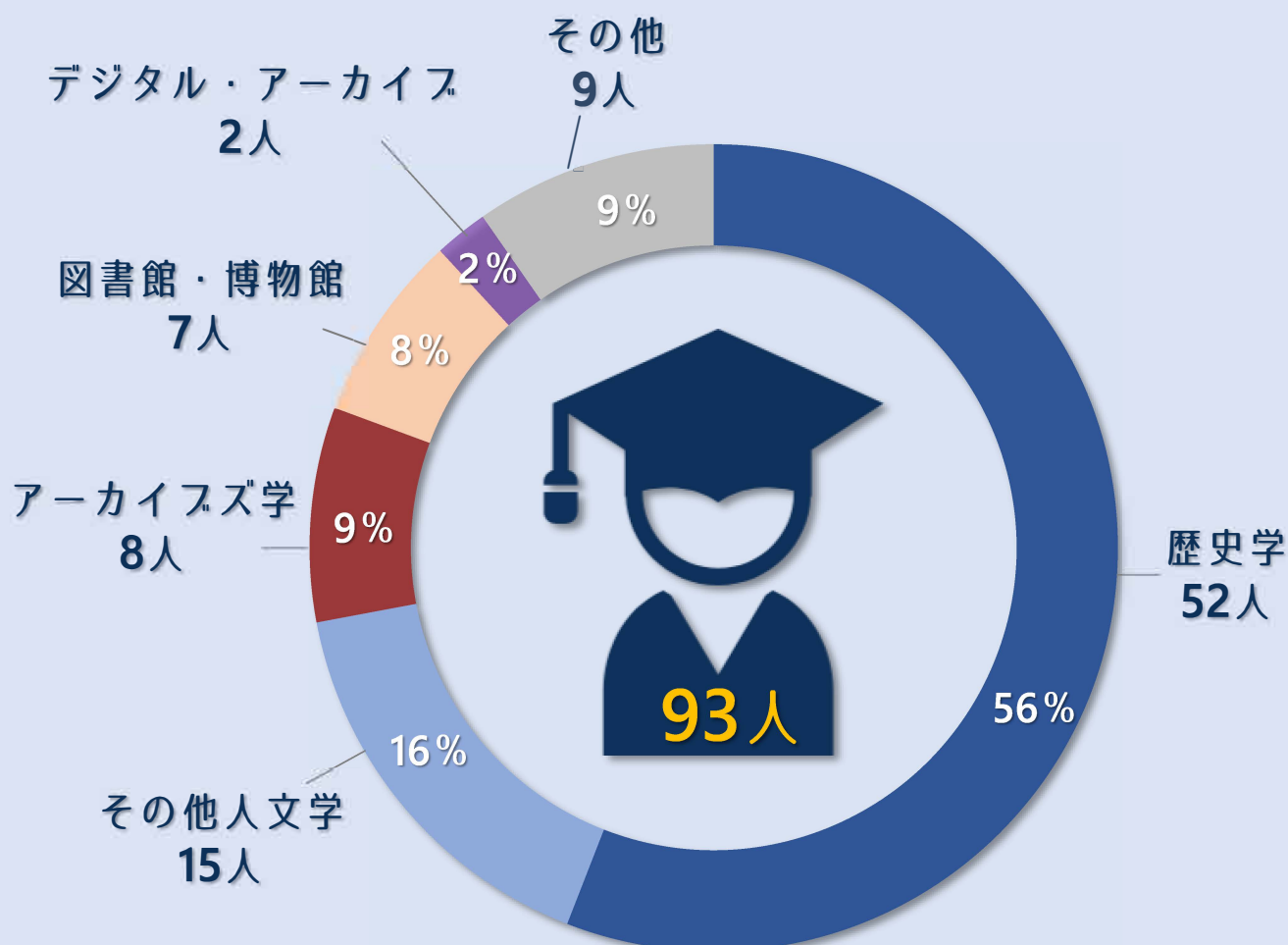
東大西洋史M2

小林拓実 | 近代フランス移民史

Indicateur Marseillais / GIS / CCライセンス
歴史地図にデータを可視化する

Panel デジタル・ヒストリーの可能性 | 菊池信彦 後藤真 崎山直樹

参加事前申込者の専門分野内訳



上図の通り、参加事前申込者93人のうち過半数は、データ利用者側の歴史学・人文学研究者であったがデータ提供者側の学術機関からも一定数の関心を得た

興味深いことに、宗教学や教育学、精神医学を専攻する参加者もありDH系のイベントとしての学際的性質が顕著に出る結果となった

なお、当日の来場者数は、およそ70人であった

シンポジウム当日の動画共有実績



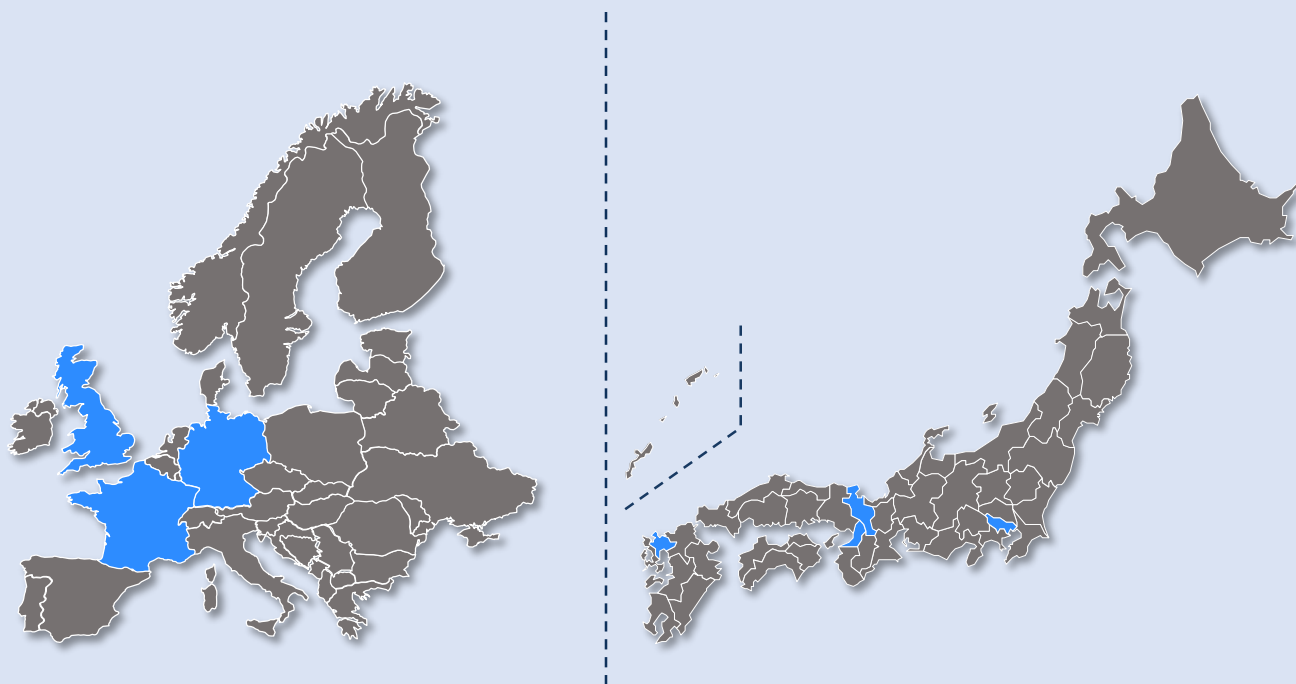
シンポジウム当日は、遠方や国外からの参加希望者に対応すべく
WebミーティングサービスZoomを用いて、個別対応で動画中継を行った

zoom

<https://zoom.us/>

Zoomミーティング参加者の所在地は、下図の通りである

@海外	イギリス・フランス・ドイツ
@国内	東京・大阪・京都・佐賀

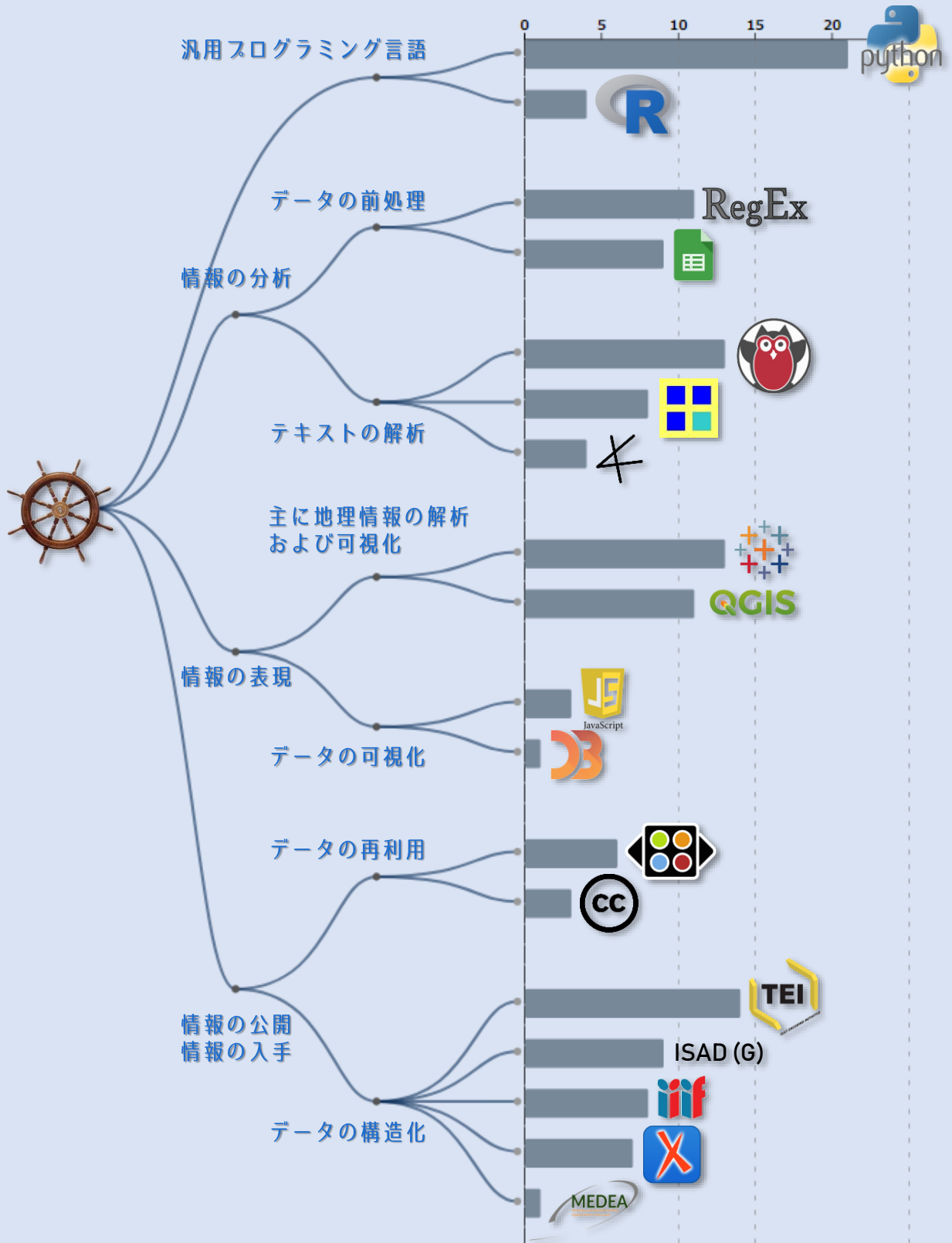


なお、Zoomミーティングのホストを務めてくださったのは
人文情報学研究所の主席研究員永崎研宣氏（当日はロンドン出張中）である
シンポジウム当日、会場のカメラ・マイク調整は、榎野翔が担当した
この場を借りてお礼申し上げたい

ロンドンでのシンポジウムの反響については
『人文情報学月報』81号【後編】の編集後記も併せてご参照いただきたい

参加者が興味を持った技術のグループ分け

※29件の参加アンケート（複数回答可）を集計



匿名アンケートに見る今後の課題



デジタル技術を活用して取り組んでみたいと思ったこと

1. テキストの解析・構造化

- ある1人の文学者が使用する語の使用法が年代ごとにどのように変化しているか。TEIやVoyant-toolsを利用して調べてみたい。ただ、私の研究対象は著作権保護期間中なので、大量のテキストデータをどのように入手し、構造化したらよいかという点が困難だと感じた。
- 哲学古典テキストあるいはタームたちの影響関係ないし類似性の空間的表現（あるいは「概念」の地図化）
- 日本古典文学テキストの構造化、デジタルアーカイブとの連携

2. 文化資源の保存

- 歴史的遺物や貴重書の劣化を防ぐために、デジタル情報を活用したい



運営に関するフィードバックなど

1. 構成

- 質問、パネルディスカッションの時間を十分にとる
- 進行の仕方がよかった（2名発表後、コメントを入れ、また2人発表後、コメントを入れ、休憩をとる方法）。オーディエンスに考える時間を与えるから

2. 配信

- 運営や各報告もよく練られたものだったので、次の公開の研究会はweb配信はフルオープンでも良いかもしれない



匿名アンケートに見る今後の課題



今後ToDHに期待するイベント内容

1. 具体的に「次の一步」に踏み込めるような講習会

- ソフトの講習会、歴史学への応用として実際の研究発表
- 実践的な細やかな情報技術についての情報発信提供と自由に参加できる雰囲気
- 以前行われたという「TEI入門セミナー」のような、体験型のイベントがあれば是非参加したい

2. 図書館や博物館などとの横断的内容

- 美術館や博物館の所蔵品解説や展示会カタログのデジタル化、アーカイブ化がどれだけ実現しているかを知りたい
- 図書館・博物館などとの共催イベント

3. デジタル・ヒューマニティーズへの切り込み

- デジタル技術に否定的な意見・消極的な意見というのも広く聞いてみたい。デジタル技術利用における、他の学問領域との違いを知りたい
- いわゆる質的アプローチに対するDHの貢献



匿名アンケートに見る今後の課題



今後のデジタル・ヒストリーに期待することや課題

1. 既存の歴史学をデジタル技術でどう変えるのか

- パネルディスカッションで指摘されていた「デジタル技術によって、研究の質がどのように変化するのか」、という問いへの回答がでると良い。それがおそらく学術的評価の話と連動する
- 「（既存の）歴史学の中での位置づけ」が課題として挙げられていたが、時間が足りなかったこともあり、もう少し深い議論が聞きたかった。今度は、それをテーマの中心に据えて議論しても良いのではないか
- パネルディスカッションについて。あくまでカギカッコ付きだが、「新しい」研究（アスローチ）の方だけで話をするよりも、デジタル・ヒストリーに精通している方と、むしろ違う毛並みというか「昔気質の」実証的研究をされている方が話をする形のほうが、デジタル・ヒストリーの意義（何ができて、何ができないのか）を考えるうえで手がかりになったかもしれないと考えた

2. 技術者との協働を模索する

- 歴史研究者は、プログラミングのできる研究者にもっと声かけてほしい
- 再現性の担保が利点とはいうが、神エクセルならぬ神（というか可読性の低い）コードの問題はあるので、結局、本業（プログラマ）の倫理を援用しなくてはならない（課題というよりは運命とでも言うべきか）

3. アーカイブズ学との協働

- アーカイブズ学との対話。金氏が示されたようなアーカイブ理論に基づく資料の記述と、他の発表者が作り上げたデジタルデータや成果などが、どのように結びつくのかを考えること

4. その他

- 歴史資料のアーカイブ化を十分に進めてほしい。マッピングやデジタル化された資料が一般にアクセス可能な仕組みを考えてほしい。また、インバウンド観光客へのサービスとして、使う可能性を探ってほしい



匿名アンケートに見る今後の課題



その他感想

1. 今後の歴史学の研究活動に与える影響

- 今後、大学での歴史学やゼミのあり方がどう変わっていくのかが気になった。例えば、ワードが一般的でなかった時代、ゼミ風景はどのようなものだったのだろう...とふと思った

2. 「若手」の活動に対する期待

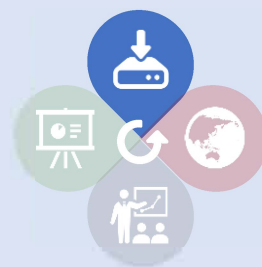
- 院生を中心とした若手研究者の方々の動きにとっても刺激を受けた。今後も積極的な活動を期待したい

3. その他

- 表現化という点での成果はすごいと感じたが、発表されている院生の研究テーマが小さい気がした。高度な技術を導入するのであるから、もっと大きな展開のあるものがないのではと考えた。説明に専門用語が多すぎて理解できない部分が多かった。もう少し具体的な作業手順について触れて欲しかった
- （配布資料をダウンロードする際の）QRコードが機能しなかったのが意外だった。あらためて「文字認識」について考えさせられた
- 非常に熱気があり、最後まで楽しく学ぶことができた
- 勉強になった。海外で働く際にも参考にしたい。また機会があれば参加したい

Event Reports

参加記



アーカイブズ学の立場から

清原和之 | アーカイブズ学

九州大学大学院統合新領域学府ライブラリーサイエンス専攻博士後期課程

4月15日に行われた**Tokyo Digital History Symposium**にチェアの一人として参加した。ここでは、アーカイブズ学の立場から、このシンポジウムに至るまで何度かワークショップにも参加してきた者として、当シンポジウムについて若干のコメントを述べたい。

シンポジウムでは主に歴史研究者の視点から、様々なツールを駆使した情報の入手・分析・表現・公開の仕方が提示された。このシンポジウムの意義は、研究成果として公表される前段階の作業が一研究者を越えて共有されるとき、それはどのようなになされ、いかなる意味をもちうるかが問われたことであろう。データの入手の仕方、分析方法、表現の仕方等の研究成果に至る「経過プロセス」が研究者個人のなかで行われるものではなく、公共的に開かれたものであるということが歴史学のなかでアピールされたことの意義はことのほか大きいように思われる。なぜなら、研究自体が細分化し、学問と社会との乖離が問題視されてきた昨今の状況に対して、分野横断的な研究者との連携や市民に開かれた公共歴史学の可能性を示唆しているからである。

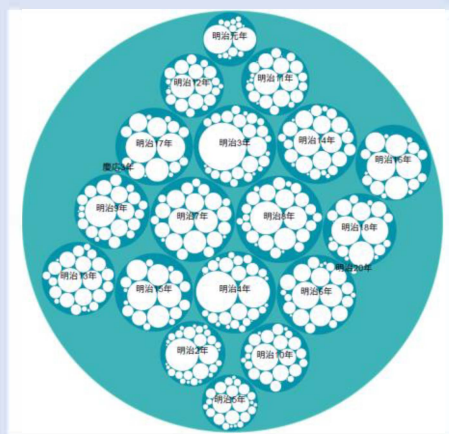
しかしながら、デジタル技術を活用することで可能となるビジュアル化や公開性、再現性の確保といった方向性は支持しつつも、今回は資料情報をデータ化し、いかに分析・活用しうるかという点に力点が置かれていたため致し方なかったかもしれないが、少くないコストを投入してなされる「データ処理のプロセス」自体の妥当性について、もう少し踏み込んだ議論が必要であったように思う。すなわち、データ化の対象となる資料ないし史料とは何か、資料自体の信頼性・真正性の問題、対象となるデータの範囲、データ自体の妥当性、データ処理の手続きの正当性等である。



これらはデータ分析・活用の前提であり、金報告でもふれられたように、アーカイブズ資料自体がどのように提供されうるかという問題と関わるが、この点が十分に議論され、学界内外に理解されていくことがデジタル・ヒューマニティーズ研究やデジタル技術を活用した歴史研究の評価にもつながっていくもの考える。

資料を提供する側としては、資料が「データ」として研究者に利活用されるものであるということ自体をより一層認識する必要があるし、資料および資料情報がウェブ上にどのような処理を経て載せられているのか、**API**等を通じて公開していくことも検討すべきであろう。

また、提供機関が資料情報へのアクセスを提供するユーザ・インターフェースのデザインについても、資料の発見を促す**finding aids**等も含めて再検討されるべき点である。今回、福田報告では「階層型バブルチャート」という資料群全体をデータ化して可視化するデザインが提示された（※[リンク先](#)も参照）。



日本のアーカイブズ機関ではこれまで資料群の階層構造を反映したウェブ・デザインが広く採用されてきたが、研究者や市民といったユーザ側がウェブブラウザをどのように探索し、インターフェイスに何を求めているのか、という点は十分に検討されてこなかったのではないだろうか。

こうした資料（情報）の提供・入手に関わる側面は特に、アーキビスト等資料提供者・研究者等ユーザ・設計者側のエンジニアが協同して課題に取り組んでいく必要がある。その点で、今回のシンポジウムでは、研究者の側から、「資料の提供のあり方はこのようにしてほしい」という意見が聞かれたことはとりわけ有益であった。今後も、歴史学・デジタル人文学・アーカイフズ学間などでの協同の場が継続して設けられ、議論が深められていくことを期待する。

Section 2 情報の分析 Chair



人文情報学の教育的側面から

橋本雄太 | 人文情報学
国立歴史民俗博物館テニュアトラック助教

画期的なシンポジウムであったと思います。史学系大学院生のグループが、自分で手を動かしてデジタル技術を習得・活用し、その成果をシンポジウムとして披露したことはこれまでにないでしょう。シンポジウムの盛況ぶりも含めて、大成功と言えるのではないのでしょうか。

情報技術に触れてまだ数ヶ月という学生さんの発表が多かったため、今後はグループの中からシステム開発や統計処理を突き詰める人が現れてくるとよいな、と個人的には感じています（そういう人の数がとても少ないので）。専門分野の研究との兼ね合いもあり、なかなか難しいかもしれませんが...今後、DHを日本の人文学研究に根付かせていくためには、DHの標準カリキュラムを早期に整備することが最大の課題になると橋本は考えています。標準カリキュラムを編成するには、アルゴリズムやデータ構造といった一般的な情報学の教程から必要なテーマを取捨選択し、デジタルアーカイフやTEI、歴史GISなどDHの各分野のテーマと組み合わせる必要がありますから、大変な作業になるはずです。これは「業界」全体を挙げて取り組まねばならない課題だと思うのですが、DH分野の研究者の数はまだ少なく、まとまった取り組みができていない状況です。

Tokyo Digital Historyは、DHに高い関心を持つ研究者で構成される、国内では非常に稀少なグループです。今後、DH分野のカリキュラム策定においても主導的な役割を果たしていけるのではないかと（大変身勝手ながら）期待しています。

Section 3 情報の表現 Chair



デジタル・ヒストリーの戦略的側面

宮本隆史 | 歴史方法論、パブリック・ヒストリー
東京大学文書館特任助教

このイベントは、歴史研究者がいかにデジタルな道具を使いこなして研究をすることができるかについて、具体的な事例に取り組みつつ考察するというものであった。わたしが大学院に所属していた頃とは時代が変わり、このような若い研究コミュニティが日本でも根付きつつあることは頼もしく、またうらやましくも感じている。そのうえで、こうした場において繰り返し語られるレトリックについて、世紀転換期に歴史家たることを志した者として、ひとつの読みの可能性をここに記しておきたい。

当日、デジタル化の効用として繰り返し強調されたのは「検証可能性の向上」であった。たとえば計量文献学的な作業にとっては、全文テキスト化によって計算間違いの可能性がなくなるといった効用があり、もちろんそれを低く評価すべきではない。だが、そうした効用を強調する語りは、ともするとナイーフにも響く「科学としての歴史」のレトリックに回収されるかに見える。

しかし、20世紀後半の歴史学の到達点のひとつは、歴史資料がテキストの外部を指示する保証はないとし、多様で多声的な解釈の可能性を指摘したことだったはずである。だとすれば、現在わたし自身をふくめしばしば用いる「検証可能性」のレトリックは、戦略的科学主義とでも呼ぶべきものなのだろう。そうであれば、わたしたちは何に対してこうした戦略をとっているのか、その史学史上の意味は何かを問わざるをえない。現在、この問いを、わたし自身の課題とかがえ、取り組んでいる。

Section 4 情報の公開 Chair



「情報の公開」セクションのチェアを担当して

中村 覚 | デジタル・アーカイブ
東京大学情報基盤センター学術情報研究部門助教

近年の学術研究において、研究論文だけでなく、研究活動の過程で作成・取得した各種データ（以下、研究データ）についても公開していくことが求められている。研究データを公開・共有することにより、研究の透明性の確保や当該研究領域の深化、学際領域研究への活用等が期待されている。本セクション「情報（研究データ）の公開」は、このような学術研究における先進的な取り組みをテーマとしたセクションであった。

その中で、小風氏は自身が研究協力者として参画している国立歴史民俗博物館「古代の百科全書『延喜式』の多分野協働研究」におけるテキストのTEI化プロジェクトを中心に、研究データの作成に関わる取り組みについて報告した。また、纓田氏は西洋中世史料に関するデータベースで公開されているTEI準拠のデータを活用し、13世紀に活動したある教皇特使の移動ルートを可視化した取り組みについて報告した。

両者の発表の関係はTEIをキーとして、前者が第三者による再利用を促す研究データを作成・公開するプロセス、後者がそのような形で公開されたデータを自身の研究に活用するプロセスであった。今後の発展が期待される研究データの公開と利活用のサイクルについて、両者の発表はその具体的な事例を示した点において大変意義のあるものであったと考える。



パネルディスカッションに登壇して

菊池信彦 | スペイン近現代史、デジタル・ヒューマニティーズ
関西大学アジア・オープン・リサーチセンター特命准教授

菊池のパネル報告では、次の2点の話題を提供した。

1つ目は、西洋史学におけるデジタルヒストリーの位置づけである。ここでは、『〇〇史研究入門』といった各国史別の研究入門書を素材に、そこでデジタルアーカイフ等のデジタルヒストリーに関わるような話題がどのように取り上げられているかを分析・紹介した。その詳細は省くが、結論は西洋史学においてデジタルヒストリーは情報収集の観点からしか論じられておらず、したがって、その意味において、本シンポジウムの意義は「情報入手」段階以降の場面での積極的活用例を示した点にあると考えている。

2つ目は、デジタルヒストリーの「これから」についての問いである。すなわちデジタルヒストリーがどうなっていくのかではなく、歴史研究者としてデジタルヒストリーをどうしていきたいのか。この問いは、**ToDH**メンバーだけでなく、参加者各人に対する問いかけでもあり、そしてもちろん筆者自身もその問いから免れてはいない。

そこで筆者は、デジタルヒストリーの「これから」について2つの考えを述べた。一つは、デジタルヒストリーが、大きなものを大きなままに、すなわち巨視的に歴史をとらえる視点をもたらしたということ、もう一つは、デジタルヒストリーが既存の歴史叙述とは異なる新しい「歴史叙述」の可能性を示しているということである。これらの意味を、歴史学の中で／歴史学として今後議論していく必要があるだろう。



パネルディスカッションに登壇して

後藤真 | 人文情報学、デジタル・アーカイブ
国立歴史民俗博物館准教授

ここでは、当日のコメントの一端を紹介したい。

日本における人文情報学には大きく三つの系統がある（蓄積・解析・可視化）。この中では、可視化と蓄積・解析の場合では問題設計が異なる。可視化は、歴史学のプロセスの結果を可視化するのに対し、蓄積と解析は歴史学のプロセスに大きく関わるためである。これまでの日本のCH/DHにおいては、蓄積系優位であり、この成果に対して歴史学がどのように活用するか考えるものが多かった。

しかし、今回のToDHシンポジウムでは、解析系を意識した研究が多かったように思われる。これは歴史学の側が、自らの研究手法を考え、抽象化していく中の一つの試みであり、新たなトレンドであるという評価をしたい。その点は大変に興味深いことである。さらにこの手法の抽象化は、歴史学のあり方そのものを「見せる」ことにも繋がる。見せることは、社会に学問を開くというアカウンタビリティを果たすとともに、より高度なフラッシュアップを求められるということでもある。

今回、発表いただいた皆様については、より歴史学の「あり方」をフラッシュアップしていただき、新たな歴史学を開くことを強く期待したい。今回のイベントに心から敬意を記す次第である。



パネルディスカッションに登壇して

崎山直樹 | アイルランド近現代史
千葉大学国際教養学部講師

崎山のパネル報告では、自身のテキストマイニングの使用例を紹介し、日本におけるデジタル・ヒストリーの発展のために今後どのような議論が必要なのかについて問題提起を行った。それはデジタル・ヒストリーの導入による方法論の変化および現在の方法論そのものへの批判的な検証の必要性についてである。

デジタル・ヒストリーに取り組んでいる歴史学研究者は、テキストマイニングなどの新技法を導入することで、自身の研究の発展を期待している。現時点では扱える史資料の量的な拡大という点で変化が生じている。では質的な点においてはどうか。個人的にテキストマイニングを試行して考えたことは、共起表現などテキスト全体の傾向を捉えることには有効であるが、前処理を通じて失われる情報や頻度は少ないものの重要な単語の意味づけなどに課題があると感じている。

私が感じた違和感はおそらく歴史学研究者の読解という行為と結びついている。我々が無自覚に行っている読解という行為は、どのような認知行為なのであろうか。デジタル・ヒストリーの発展は、このメカニズムの解明に資するものだと考えている。例えば日本社会学会は、雑誌『社会学評論』(第68巻第3号、2017年12月)にて特集「テキストマイニングをめぐる方法論とメタ方法論」を掲載し、議論をはじめている。歴史学においても同様の取り組みが必要となるであろう。



デジタル・アーカイブとデジタル・ヒストリーの関係

元ナミ | アーカイブズ学
京都大学大学文書館助教

2年ほど前、国際会議などにおいて英語プレゼンテーションをする際の有効なメソッドを学ぶワークショップに参加する機会があった。その際に、発表内容をより効果的に伝えるための視覚化ツールなどについて知ることができた。今回のToDHシンポジウムでもそのようなツールがいろいろと登場した。

私は今も、デジタルヒストリー（以下、**DH**is）については完全に無知に近い程度の知識しかいないため、会場では「デジタル（がわかる者による）歴史学」が進んでいくのだろうか」と、単純に考えていた。**DH**isの定義や歴史学における位置づけも知らなかったため、セクション1（情報の入手）と他のセクション（情報の分析、情報の表現、情報の公開）の間でデジタル・アーカイブ（以下、**DA**）と**DH**isの関係性について考えるヒントが得られず、最後まで**DH**isを「歴史学の研究成果を分析・表現する新しいツールの活用術」のようなものと誤解してしまっていた。

ところが、各報告を振り返って考えてみると、報告者たちが入手した研究データは主に大学や図書館、（公）文書館などが提供するあらゆる「デジタルアーカイブ」が出典となっていた。報告者たちは「信頼できるアーカイブズ機関」が提供する「信頼できる一次資料」をデータとして入手し、自分の研究データとして加工した。**DA**が提供する情報やデータによっては、研究データを活用するための「前処理」に結構な努力（＝時間と苦労）が伴うケースもあったようだが、**API**が簡単に取得できる**DA**や、メタデータの入手が比較的簡単な**DA**の場合は、研究データの処理過程において**DH**is研究に寄与することが多かったように感じた。これからの**DA**は、目録記述の標準化・互換性のあるメタデータスキーマの採択・**API**提供の他、**DA**そのものの長期保存を含め、より具体的に設計しなければならなくなったと実感した。



私自身は、今回のシンポジウムにおける各報告の意図を正しく把握しきれないこともあったかもしれないが、**DHIs**が「デジタル技術が使えるようになった歴史学」ではなく、その成果とデータを広く公開・共有し、共同のデータベース構築や活用術までを視野に入れる歴史学の一分野であることが確認でき、大変勉強になった。さらに言えば、**DHIs**が生み出す新しい情報とデータ、ポーンデジタル記録の保存と活用、そのための**DHIs**と**DA**の連携可能性などについて、さらなる議論が始まっていいのではないかと考える。今後アーカイブズ学からも、これらの課題に真摯に向き合い、**DHIs**のような新しい学問の潮流を意識しながら、「これから求められる**DA**像」について真剣に検討する必要があるだろう。

蛇足であるが、常に「金・人・時間」の不足に苦しんでいるほとんどの日本のアーカイブズ機関では、学問ベースの利用者が満足できるほどの**DA**が提供できないことが多い。とにかく**ICT**（情報通信技術）の専門家とアーキビストが協業できる環境が整っていないし、デジタルアーキビストの不在どころか、**ICT**がわかるアーキビストさえもごく少ないからである。もし可能であれば**DHIs**で経験した**ICT**の技術と研究成果を**DA**側と共有しながら、「**DHIs**に必要な**DA**のあり方」について積極的に意見を打診していけば、新たな学際的研究や実践的プロジェクトが生まれるかもしれないと考える。



シンポジウム参加記

加藤 諭 | 日本近現代史

東北大学学術資源研究公開センター史料館准教授

2018年4月15日、2018 Spring Tokyo Digital History Symposiumが東京大学において開催された。イントロダクションで小風尚樹代表が説明した通り、本シンポジウムはデジタル化された、膨大な研究データを海に例え、デジタル・ヒストリー関連技術を船の舵になぞらえることで、その見取り図を描くことを目的としたものであった。デジタル・ヒストリーの関連技術を、情報の入手から分析、表現、公開に至るプロセスを重視して示し、さらにデータの構造化、前処理、可視化、再利用の観点も含め、その実践例を提示するシンポジウムの構成は、デジタル・ヒストリーが果たす研究基盤構築のプロセスを俯瞰する上で、明快な構成であったといえよう。金甫榮、福田真人、山崎翔平、小川潤、山王綾乃、小林拓実、小風尚樹、櫻田宗紀の新進気鋭の8人の若手研究者の報告、清原和之、橋本雄太、宮本隆史、中村覚の4人の研究者のチェア、パネルディスカッションでは菊地信彦、後藤真、崎山直樹の3人の研究者という、登壇者の世代間バランス、またアーカイブズ学と歴史学双方からの発表というアプローチも、本シンポジウムの理解に厚みを持たせる上で有効であった。

本シンポジウムは大学院生を中心に、デジタル・ヒストリーに関する技術習得が1年未満の報告者も少なくない中、企画段階からおよそ半年かけて、かたちにしていったという。その企画力には率直に敬意を表するとともに、そこにはデジタル技術をどう既存の歴史学研究に還元していくか、若手研究者が抱く既存の歴史学研究の方法論、評価方法へのジレンマがモチベーションの源泉であるようにも感じられた。



Webスクレイピングや、テキスト・マイニングなどは、大規模な記述統計的、計量的分析を可能にする技術であり、GIS等を用いた事象の可視化、データ再現性のためのコードの公開や、TEIファイルのリポジトリでの共有の在り方も、歴史学の新たな表現方法を開拓するものである。一方、報告者が一様に、原典にあたることの重要性もまた指摘していたように、表現方法ではなく、歴史学研究の方法論そのものに、デジタル・ヒストリーが何をもたらすのか、またその評価軸については、パネルディスカッションでも問われたテーマであった。そもそも船の舵が何処を指し示していくべきか、という論点は十分回収されたとはいえなかったが、問題提起こそが本シンポジウムの果たした役割であったともいえよう。

すぐに結論が出る課題ではないが、少なくとも筆者は、デジタル・ヒストリー関連技術はアーカイブズ学のコンテキスト研究と歴史学のコンテンツ研究を架橋する技術になり得ること、歴史学研究においてもモノの移動や、偏差など定量的な分析との親和性が高いことなど、本シンポジウムを通じて、いくつかの可能性を見出すことが出来た。先行する歴史学の成果とデジタル・ヒストリーの技術をどう接合していくのか、今後の取り組みに期待したい。