

展 望:デジタル人文学の可能性

或るイスラーム研究者とデジタル ヒューマニティーズの歩み

石田 友梨

(岡山大学)

はじめに

新しい研究分野である「デジタルヒューマニティーズ Digital Humanities」は、さまざまな分野を取り込みながら今まさに発展しているその途上にあり、その不定形な全体像を捕捉することは困難である。本稿では、筆者のこれまでの経歴と重ねつつ、デジタルヒューマニティーズの歴史を概観していくが、筆者の観点はイスラーム研究とテキスト研究に偏ったものである。筆者とデジタルヒューマニティーズの出会い、2013年に遡る。当時は出会った相手がデジタルヒューマニティーズという学問であるという自覚はなく、実際のところ出会ったと思っている相手は幻に過ぎないのかもしれない。デジタルヒューマニティーズの基礎概念を示す「方法論の共有地 Methodological Commons」¹⁾が、雲のような各専門分野に支えられていることが象徴的であるように筆者には思える。デジタルヒューマニティーズという自覚をもつ各研究分野の専門家が集まることで成立する学問なのであろう。

第1節では、「デジタルヒューマニティーズに係る個人史」として、2013年から2024年現在までの筆者の学術活動と、デジタルヒューマニティーズの関わりについて述べる。第2節「名称からみたデジタルヒューマニティーズ略史」では、デジタルヒューマニティーズという名称に着目しながら、この研究分野の成立と現在までの発展を概観する。第3節の「イスラーム研究におけるデジタルヒューマニティーズ」では、イスラーム研究分野でデジタルヒューマニティーズと関連するプロジェクトや書籍の一部を紹介する。第4節「アラビア語テキストデータを利用したデジタルヒューマニティーズ」では、デジタルヒューマニティーズの研究手法を取り入れようとした際に、アラビア語テキストの特徴によりもたらされる問題点を指摘する。その後、アラビア語テキストデータを活用した研究の可能性について、具体的な事例を示しながら提案したい。

1. デジタルヒューマニティーズに係る個人史

1-1. 2013年、博士論文執筆時に直面した課題

2013年、筆者は博士論文の執筆をしていた。博士論文の対象は、18世紀インドの思想家シャー・ワリーウッラー Shāh Walī Allāh (1703-1762年) という人物である。彼の思想の形成に影響を及ぼした要因のひとつとして、師弟関係について考察するつもりであった。シャー・ワリーウッラーはデリーに生まれ、父親から学問の手ほどきを受けた。その父を17歳で亡くした後も学問に励み、30歳手前となった1731年7月頃、シャー・ワリーウッラーはイスラーム教徒の義務のひとつであるマッカ巡礼へ旅立つ。続いて、マッカの北に位置する第2の聖地マディーナでアブー・ターヒル Abū Tāhir (1670-1732年) などに学んだ。再びマッカに巡礼し、1732年6月頃に帰国の途に就いた²⁾。このマディーナでの師弟関係というものが複雑である。先行研究やシャー・ワリーウッラーの自伝によれば、シャー・ワリーウッラーの師匠であるアブー・ターヒルの師匠は、父親でもあるイブラーヒーム・クーラーニー Ibrāhīm Kūrānī (1616-1690年) である。しかし、兄弟子のアブドゥッラズール・バルザンジー ‘Abd al-Rasūl al-Barzanjī (1630-1692年)、ハサン・ウジャイミー Ḥasan al-‘Ujaymī (1639-1701年)、アブドゥッラー・バスリー ‘Abd Allāh al-Baṣrī (1638-1722年) らも師匠である。このように師匠と弟子が一对一の関係でない上に、マッカとマディーナのあるアラビア半島へ行ったりエジプトへ行ったりと、各人が動き回りながら師弟関係を結んでいる。学者たちの師弟関係と移動経路を図にまとめようとしたものの、手書きはもとより、普段使用している文章作成ソフト「ワード Word」や発表資料作成ソフト「パワーポイント PowerPoint」ではうまく表現できず、途方に暮れていた。

そんなある日、京都大学東南アジア地域研究研究所にいらっしゃった小島敬裕先生(現津田塾大学)の発表をお聞きする機会があった。ミャンマーの仏教僧たちが国境を跨いで寺を移っていく様子を地図上に示されるのをみて、同じ手法でイスラーム教徒たちの移動経路も描けるのではないかと希望をもった。しかし、提出の締め切りが迫っていた博士論文では、師弟関係と移動経路を表計算ソフト「エクセル Excel」で一覧表にして掲載するのが精一杯であった。この表は公聴会において「何をしたいのか全く意味が分からない」と評されたが、現時点で理解されないことにこそ将来性があるに違いないと、この道をさらに進めてみ

ることにした。

1-2. 2014年、デジタルヒューマニティーズの「発見」

2014年3月に博士号を取得した後、筆者は着任した早稲田大学の中央図書館において、楊曉捷他編『デジタル人文学のすすめ』（勉誠出版、2013年）という本を偶然手に取った。そして自分の目指している研究が「デジタル人文学」、英語では「Digital Humanities」と呼ばれる分野であることを知り、この分野で多用される技術の勉強を始めた。非ラテン文字用「光学的文字認識 Optical Character Recognition」(OCR)である「クラーケン Kraken」³⁾を動かすためにプログラミング言語「パイソン Python」と⁴⁾、師弟関係を分析するためにネットワーク描写ソフト「ゲフィ Gephi」⁵⁾を学ぶ際には、萩原淳氏の協力があつた。氏は京都大学大学院アジア・アフリカ地域研究研究科の先輩であり、サウジアラビア研究者であるが、ペンシルヴァニア大学でコンピュータサイエンスの修士号を取得した経歴をもつ。筆者のその後の研究の広がりをもたせてくれた氏には、感謝の言葉もない。

1-3. 2020年、教育する立場へ

現在デジタルヒューマニティーズ研究者の多くは、①「人文学で博士号を取得した後に、研究上の必要から情報学を取り入れるようになった研究者」と、②「情報学を専門とし、人文学研究に役立つツールなどを開発するようになった研究者」のどちらかに当てはまる。筆者は前者①の典型であるが、2020年以降は新しい世代となる③「デジタルヒューマニティーズを専門とし、博士号を取得した研究者」の育成に携わっている。

2005年にロンドン大学キングスカレッジが「デジタルヒューマニティーズ」の博士号を取得するためのプログラムを開始すると、欧米を中心に「デジタルヒューマニティーズ」の学位が取得できる大学が増えていった⁶⁾。現在日本には「デジタルヒューマニティーズ」という学位を授与する教育機関はないものの、プログラム修了証を発行する東京大学大学院横断型教育プログラム「デジタル・ヒューマニティーズ」⁷⁾ (2012年度開設)、千葉大学を拠点校とする卓越大学院プログラム「アジアユーラシア・グローバルリーダー養成のための臨床人文学教育プログラム」⁸⁾ (2020年度開設)がある⁹⁾。後者はデジタルヒューマニティーズとアジアユーラシア研究のふたつを柱とする5大学連携の大学院教育で、筆者は連携校のひとつである岡山大学にお

いて当該プログラムの担当をしている。

また、人間文化研究機構は2022年度よりデジタルヒューマニティーズの推進を掲げている¹⁰⁾。同機構の呼びかけにより、デジタルヒューマニティーズの研究や教育を行う国内の大学・研究機関から成る「デジタル・ヒューマニティーズ (DH) 組織ネットワーク協議会」¹¹⁾が設立された。2024年7月の第1回協議会には、岡山大学を含む23組織が参加した。これまで研究者や研究機関が個々に進めてきたデジタルヒューマニティーズであるが、国内に一元化された窓口ができたことにより、データベースなどの研究基盤の提供や共有が容易になっていくだろう¹²⁾。同機構が主催する「DH若手の会」¹³⁾も盛況であり、デジタルヒューマニティーズは今後さらに発展していくことが期待される研究分野である。

2. 名称からみたデジタルヒューマニティーズ略史

2-1. 名称の誕生

ここまで定義をせずに使ってきた「デジタルヒューマニティーズ」という言葉であるが、この名称の歴史を振り返ることで、それが指す意味について確認していきたい。デジタルヒューマニティーズの起源とされるのは、人文学研究にコンピュータを導入したロベルト・ブサ神父の研究である。彼は、1946年にトマス・アクィナスの著作の索引をデジタル化するプロジェクトを構想し1949年にIBM社の協力を得ることとなる¹⁴⁾。

一方、現在の用法としてのデジタルヒューマニティーズという名称が誕生したのは、2001年のことである。その年、2004年にブラックウェル社から『デジタルヒューマニティーズの手引書 A Companion to Digital Humanities』として出版されることになる本¹⁵⁾の編集会議において、3名の編者うちレイ・シーメンス Ray Siemens は、その時点でよく使われていた「人文学コンピューティング Humanities Computing」を書名に使用することを希望し、出版社側は「デジタル化された人文学 Digitized Humanities」を希望していた。そして、これも編者のひとりであったジョン・アンズワース John Unsworth が、「単なるデジタル化」から重点を移すために「デジタルヒューマニティーズ」を提案したという¹⁶⁾。「グーグル社書籍 N グラム閲覧装置 Google Books Ngram Viewer」¹⁷⁾が描き出す図1に明らかのように、『デジタルヒューマニティーズの手引書』

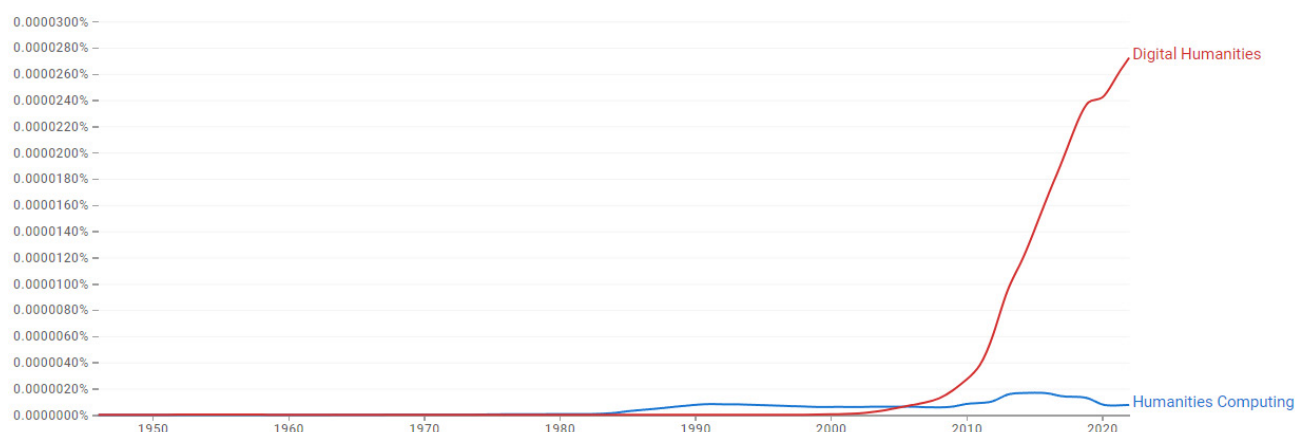


図1 1946年以降の英語コーパスにおける「Humanities Computing」と「Digital Humanities」

が出版された2004年以降、「デジタルヒューマニティーズ」という言葉の使用が増加していく。

2-2. ADHOの役割

現在『デジタルヒューマニティーズの手引書』は、「デジタルヒューマニティーズ組織連合 Alliance of Digital Humanities Organizations」(ADHO)によって無料公開されている¹⁸⁾。ADHOはデジタルヒューマニティーズ分野の世界最高峰の国際会議を開催している主体であり、ヨーロッパ、アメリカ、カナダ、オーストラリア、日本、南アフリカ、台湾、韓国などの各国・各地域のデジタルヒューマニティーズ学会や、フランス語、ドイツ語、スペイン語などの各言語のデジタルヒューマニティーズ学会などから構成されている¹⁹⁾。ヨーロッパの「文学と言語学のコンピューティング学会 Association for Literary and Linguistic Computing」(ALLC)とアメリカの「コンピュータと人文科学学会 Association for Computers and the Humanities」(ACH)が、2002年からADHOの設立に向けて動き出し、2006年に「デジタルヒューマニティーズ」を冠する初めての国際会議をパリ第4大学(現ソルボンヌ大学)で開催した²⁰⁾。ヨーロッパのALLCは2012年に「欧州デジタルヒューマニティーズ学会 European Association for Digital Humanities」(EADH)へと名称を変更したが²¹⁾、アメリカは現在もACHの名称のままADHOに参加している。

ADHOの説明に基づけば、デジタルヒューマニティーズは、英語学、歴史学、フランス語学、現代語学、言語学、哲学、演劇学、音楽学、コンピュータ科学、視覚芸術学などの専門家による、テキスト分析、電子出版、資料の符号化、テキストの研究と理論、新しいメディア研究とマルチ

メディア、電子図書館、応用拡張現実(AR)、双方向型ゲームなどの最先端の研究²²⁾となる。個々の研究の詳細については、一般財団法人人文情報学研究所監修『欧米圏デジタル・ヒューマニティーズの基礎知識』(文学通信、2021年)などが参考になるだろう。

2-3. 日本における名称

続いて、国内におけるデジタルヒューマニティーズの歴史を簡単に振り返ってみたい。文部科学省「グローバルCOEプログラム」で立命館大学に「日本文化デジタル・ヒューマニティーズ拠点」²³⁾が設置されたのが2007年、「日本デジタル・ヒューマニティーズ学会 Japanese Association for Digital Humanities」(JADH)が設立されたのが2011年²⁴⁾、JADHがADHOへ加盟したのが2012年²⁵⁾のことである。2011年よりメールマガジン『人文情報学月報』を発行している人文情報学研究所は、「人文情報学」を、Digital Humanitiesを含む言葉として用いる²⁶⁾として、研究所名とメールマガジン名の「人文情報学」を「Digital Humanities」と英訳している。「Digital Humanities」の和訳としては、「デジタルヒューマニティーズ」、単語の区切りに中黒をつけた「デジタル・ヒューマニティーズ」、「人文情報学」、「デジタル人文学」が併存している状態である。

一方、1989年以降「人文科学とコンピュータ研究会」を運営し、JADH発足以前よりデジタルヒューマニティーズに該当する研究分野を担ってきたのは情報処理学会である²⁷⁾。2009年には「デジタル・ヒューマニティーズの可能性」と題するシンポジウムを開催している²⁸⁾。また、デジタルヒューマニティーズの研究手法のうち、最も世間的な認知度が高いテキストマイニングには、1956年創立の

計量国語学会²⁹⁾という先駆者の存在を指摘してよいだろう。デジタルヒューマニティーズという言葉が生まれる以前より、日本においても現在のデジタルヒューマニティーズに該当する研究が行われてきたが、ここでそのすべてを網羅する力量を筆者はもたない³⁰⁾。我が国におけるデジタルヒューマニティーズの歴史の詳細については、稿を改める必要がある。

小括として、第1節と第2節で取り上げたデジタルヒューマニティーズに関連する出来事を、表1の年表にまとめた。1995年にマイクロソフト社がWindows 95を発売して以来、人文学研究者もワードプロセッサ（ワープロ）専用機ではなくコンピュータで原稿を執筆するようになった。この意味では、現代の人文学研究者であれば、誰もがコンピュータを使った研究を行っており、研究はデジタル化されている。しかし、テキストの入力から一歩踏み込み、コンピュータの計算能力を積極的に活用しながら行われる人文学研究の取り組みが、2004年以降デジタルヒューマニティーズという言葉で総称されつつある。

表1 デジタルヒューマニティーズ関連年表

西暦	出来事
1946年	ロベルト・ブサがトマス・アクィナスの著作の索引のデジタル化を構想する
1973年	欧州の「文学と言語学のコンピューティング学会」(ALLC)が設立される
1978年	米国の「コンピュータと人文学学会」(ACH)が設立される
1989年	日本の情報処理学会が「人文科学とコンピュータ研究会」を発足する
1995年	マイクロソフト社がWindows 95を発売する
2001年	ジョン・アンズワースが書名『デジタルヒューマニティーズの手引書』を提案する
2002年	ALLCとACHが「デジタルヒューマニティーズ組織連合」(ADHO)設立に向けて始動する
2004年	ブラックウェル社が『デジタルヒューマニティーズの手引書』を出版する
2005年	ロンドン大学キングスカレッジがデジタルヒューマニティーズの博士学位プログラムを開始する
2006年	ADHOが初めての国際会議 Digital Humanities 2006を開催する
2007年	立命館大学が「日本文化デジタル・ヒューマニティーズ拠点」を形成する
2012年	「日本デジタル・ヒューマニティーズ学会」(JADH)がADHOに加盟する
	ALLCの名称が「欧州デジタルヒューマニティーズ学会」(EADH)に変更される
	東京大学が大学院横断型教育プログラム「デジタル・ヒューマニティーズ」を開設する

2020年	千葉大学が卓越大学院プログラム「アジアユーラシア・グローバルリーダー養成のための臨床人文学教育プログラム」を開設する
2024年	第1回デジタル・ヒューマニティーズ (DH) 組織ネットワークワーキング協議会が開催される

3. イスラーム研究におけるデジタルヒューマニティーズ

3-1. 国外の動向

本節では、筆者の専門であるイスラーム研究に焦点を当てたデジタルヒューマニティーズの現況について紹介したい。筆者はデジタルヒューマニティーズという研究分野の存在を認識した2014年、アメリカの「(北米)中東研究学会 Middle East Studies Association (of North America)」(MESA)に参加した³¹⁾。その前年にコンピュータを利用したアラビア語伝記集の読解でミシガン大学より博士号を取得したマキシム・ロマノフ Maxim Romanov³²⁾の発表や、アーガー・ハーン財団による歴史的建築物のデジタル化プロジェクトに触れ、大いに刺激を受けた。

その後、ロマノフはイスラーム研究におけるデジタルヒューマニティーズの第一人者のひとりとして、「アーガー・ハーン大学イスラーム教徒文明学研究所 Aga Khan University's Institute for the Study of Muslim Civilisations」(AKU-ISMC)の「知識・情報技術・アラビア語書籍 Knowledge, Information Technology, and the Arabic Book」(KITAB)プロジェクト³³⁾などに参加している。KITABプロジェクトの「イスラーム化テキスト公開計画 Open Islamicate Texts Initiative」(OpenITI)³⁴⁾は、アラビア語の写本や刊本のテキストファイルを公開している。

イスラーム研究に特化したデジタルヒューマニティーズ関連の書籍としては、2016年出版のド・グリュイター社『デジタルヒューマニティーズとイスラーム・中東研究 The Digital Humanities and Islamic & Middle East Studies』が挙げられる。同社は2019年以降、「デジタルヒューマニティーズの入門書：宗教編 Introductions to Digital Humanities - Religion」叢書を刊行している。L・W・C・ヴァンレット L.W.C. van Lit とジェイムズ・ハリー・モリス James Harry Morris が編者を務めた『アジアにおけるデジタルヒューマニティーズと宗教：入門編 Digital Humanities and Religions in Asia: An Introduction』は、この叢書の1冊として2024年に出版された。両者がイスラーム研究に関わっていることから、イスラーム関連写本

のデジタル化プロジェクト「東南アジアの危機に瀕した写本のデジタル保管庫 Digital Repository of Endangered and Affected Manuscripts in Southeast Asia」(DREAMSEA)³⁵⁾などが取り上げられている。編者のひとりヴァンレットは、2013年にオンライン雑誌『デジタル東洋学者 The Digital Orientalist』³⁶⁾を創刊した人物で、2020年にはブリル社より『デジタル化された写本の間 Among Digitized Manuscripts』を出版している。

ここで関連するプロジェクトや書籍をすべて列挙することはできないが、イスラーム研究においても、現在のところデジタルヒューマニティーズの中心は欧米である。しかし、近年ではトルコ発の研究成果も見逃せない³⁷⁾。また、アラブ首長国連邦もアラビア語コーパス開発などを進めるニューヨーク大学アブダビ校や、2019年に創立したムハンマド・ビン・ザイド人工知能大学が注目される。2024年にはシャルジャ・アメリカン大学が、アラブ・イスラーム世界を対象としたデジタルヒューマニティーズの国際会議³⁸⁾を開いており、今後国際的な研究拠点のひとつとなっていく可能性がある。

聖書や仏典と同様に、聖典クルアーンもデジタル化されて久しい。イスラーム研究者でなければ、このことを意外に思うかもしれない。しかし、デジタルヒューマニティーズが注目される以前より、イスラーム諸国では電子書籍やオンライン辞書が活用されていた。著作権による制限が厳しい日本より進んでいた面もあったように思う。また、サウジアラビアは2017年に世界で初めてAIロボットに市民権を与えており、イランも人型ロボットを小児科病院や宗教の授業に導入している³⁹⁾。デジタルヒューマニティーズが隆盛する素地は充分にある。

3- 2. 国内の動向

さて、話を2014年のMESAに戻す。筆者は国内にも同好の士を求め、翌2015年に日本中東学会で「17-18世紀ハラマインにおける師弟関係：シャー・ワリーウッラーの伝記情報のネットワークによる視覚化とその分析」(萩原淳との共著)、JADHで「17-18世紀におけるイスラーム改革主義の勃興：学者ネットワーク分析とGephiによる視覚化」を発表し、研究の成果を2016年の論文⁴⁰⁾にまとめた。日本のイスラーム研究においても、図や統計を利用した研究は行われてきた。しかし、デジタルヒューマニティーズという自覚をもった学術活動は、筆者を嚆矢とする自負している。

2020年には、文部科学省科学研究費・学術変革領域研究(A)「イスラーム的コネクティビティにみる信頼構築：世界の分断をのりこえる戦略知の創造」(イスラーム信頼学)が始まり、研究班「デジタルヒューマニティーズ的手法によるコネクティビティ分析」⁴¹⁾には、筆者も研究分担者として参加している。さらに、筆者を含むイスラーム信頼学の研究者たちが多くの章を執筆した、須永恵美子・熊倉和歌子編著『イスラーム・デジタル人文学』(人文書院、2024年)も出版されている。

一方、2020年以降『西洋史学』誌上の連載企画「Digital History Insights (デジタルヒストリー識見)」において、歴史学に特化したデジタルヒューマニティーズを紹介している日本西洋史学会⁴²⁾、2023年にフォーラム「西洋古典学とデジタル・ヒューマニティーズ」⁴³⁾を開催し、西洋古典に特化した生成AIシステム「ヒューマニテクス」を2024年にウェブサイト上で限定公開⁴⁴⁾した日本西洋古典学会などの他分野と比較すると、国内イスラーム研究におけるデジタルヒューマニティーズへの関心は、一部の研究者に限定されたものといえるかもしれない。その要因のひとつとして、イスラーム研究の多くがアラビア文字と関連するため、デジタルヒューマニティーズを取り入れる際に障壁があることが挙げられよう。この点については、次節で詳細を述べる。

4. アラビア語テキストデータを利用したデジタルヒューマニティーズ

4- 1. アラビア語の特徴

イスラーム教は、天使ガブリエルが預言者ムハンマドへ神の言葉を伝えたことによって始まった。この時にアラビア語が使われたことから、イスラーム教文化圏では現代においてもアラビア語の地位が高い。研究対象による程度の差はあれども、イスラーム研究にはアラビア語の知識が欠かせない。日本においてアラビア語は、「ミミズが這ったような」⁴⁵⁾と形容されることが多いが、これは英語やフランス語の筆記体のように、文字と文字を連結するという特徴のためであろう。アラビア語で用いられるのは28文字だが、文字単体の独立形、文字が単語の最初にある時の語頭形、単語の途中にある時の語中形、単語の最後にある時の語末形の4つの形がある。しかし、28文字すべてがこれら4つの形をもっているわけではない。

このように複雑な規則が、アラビア語用OCRの開発を

妨げてきたと思われる。OCR の仕組みは、テキストが含まれる画像より、テキストの部分を見つけ出し、そこに書かれたテキストを出力するというものである。ルールベース型人工知能では、「A の音であるアリフは、独立形と語頭形が「ا」で、語中形と語末形が「ـ」である」といった具合に、人間がひとつずつ規則を設定していかなければならない。人工知能自身がデータから規則を学習していく機械学習の時代になり、アラビア語用 OCR の精度も向上した。しかし、ラテン文字を使用する欧米諸語に比べれば、依然として精度は低いままである⁴⁶⁾。アイコンなどの使用で視覚的に操作できる OCR 及び「手書き文字認識 Handwritten Text Recognition」(HTR)を提供している「トランスクリバス Transkribus」⁴⁷⁾の試用画面では、英語やドイツ語はもちろんのこと、文語であるラテン語や、話者が数百万人のスロヴァキア語やスロヴェニア語を含む 15 言語の選択肢があり、その精度を確認することができる。しかし、話者が 2 億人を超え、国連公用語ともなっているアラビア語は選択肢にない。利用者申請が必要であるものの、トランスクリバスと同様の機能を提供している「イースクリプトリウム eScriptorium」⁴⁸⁾や「カルファ Calfa」⁴⁹⁾にはアラビア語モデルがある。

アラビア語のもうひとつの特徴として挙げられるのは、右から左へ書くことだろう。右揃えにして文末に句点を打つと、句点が文頭の右端に移動してしまうなど、未だにワードでさえも根本的な不具合が残っている。デジタルヒューマニティーズの各種技術においても、左から右向きのテキストが標準となっており、たとえばテキストデータの国際標準である「テキスト符号化計画 Text Encoding Initiative」(TEI)でアラビア語を入力する際の注意点については、拙稿⁵⁰⁾を参照されたい。

以上のように、イスラーム研究と不可分なアラビア語の特徴のため、「欧米のデジタルヒューマニティーズ研究で用いられているソフトウェアを試してみたら使えなかった」ということは珍しくない。アラビア語だけではなく、欧米諸語とは異なる言語的特徴をもつアジアの諸言語は、多かれ少なかれ同様の問題に直面する。欧米の基準をそのままアジアに適用することで格差が生じている現状について、インド工科大学インドール校のニルマラ・メノン Nirmala Menon が「第二の植民地化」と評したこともある。ただし、この問題が無視され放置されているわけではない。ここ数年の ADHO の国際会議では、多言語について議論するワークショップが開かれており、2024 年の

国際会議 (DH2024) のための査読評価項目には「多様性 Diversity」が登場した。

また、言語の特徴が不利になる場合もあれば、新しい技術の開発につながる場合もある。人文学オープンデータ共同利用センターの「くずし字 OCR (AI くずし字認識)」⁵¹⁾、人文情報学研究所の「TEI 古典籍ビューワ」⁵²⁾などは、くずし字や縦書きといった日本語の特徴に対応しようとする試みである。さらに、近年デジタルヒューマニティーズの技術は次々と開発されており、何を選べばよいか迷うことがある。アラビア文字を表示できるか否かを選択の基準のひとつとすることで、情報の洪水から身を守ることもできる。

4-2. OpenITI を活用した伝記資料の比較

最後にデジタルヒューマニティーズ研究の具体例として、OpenITI が公開しているアラビア語テキストを活用した研究事例を示したい。筆者の研究対象は、シャー・ワリーウッラーに連なる 17-18 世紀アラビア半島を中心とした師弟関係である。この時代の師弟関係を調べたアズラ Azra の先行研究⁵³⁾が主な典拠としていた資料のうち、テキストファイルが OpenITI で公開されているものを表 2 にまとめた。左端から、整理のための番号、典拠資料の略称、その資料を典拠とする学者の数、資料の著者名、著者の没年である。

表 2 師弟関係の典拠資料

番号	資料略称	学者数	著者名	著者没年
1	<i>Fahrāsa</i>	57	al-Kattānī	1962
2	<i>Khulāṣa</i>	48	al-Muḥibbī	1699
3	<i>Hadīya</i>	32	al-Baghdādī	1920
4	<i>al-A'lām</i>	26	al-Ziriklī	1976
5	<i>'Ajā'ib</i>	13	al-Jabartī	1822
6	<i>Silk</i>	10	al-Murādī	1791

著者の没年より、表 2 において 17 - 18 世紀の同時代人だったといえるのは、2 番目のムヒッビー al-Muḥibbī と、6 番目のムラーディー al-Murādī である。出身地を示す単語に着目し、彼らの作品のアラビア語テキストデータを検索してみる。図 2 の縦軸は、テキストに出現する出身地を示す単語全体に対する割合である。横軸は学者の出身地となる地名であり、左からマグリブ（現在のモロッコなど

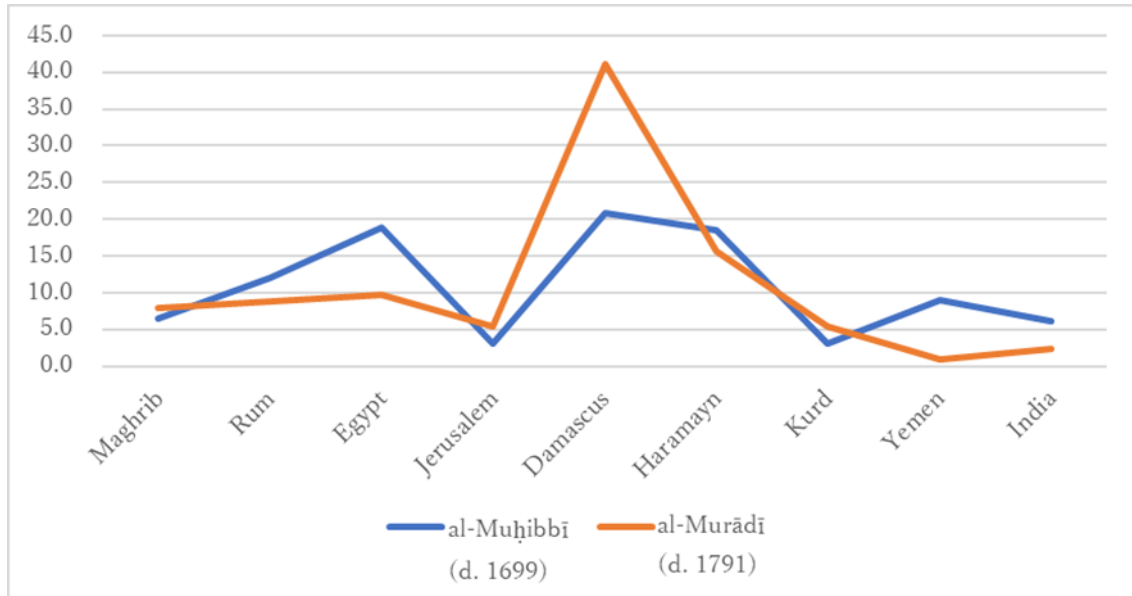


図2 出身地を表す単語の各地の割合（同時代の2資料）

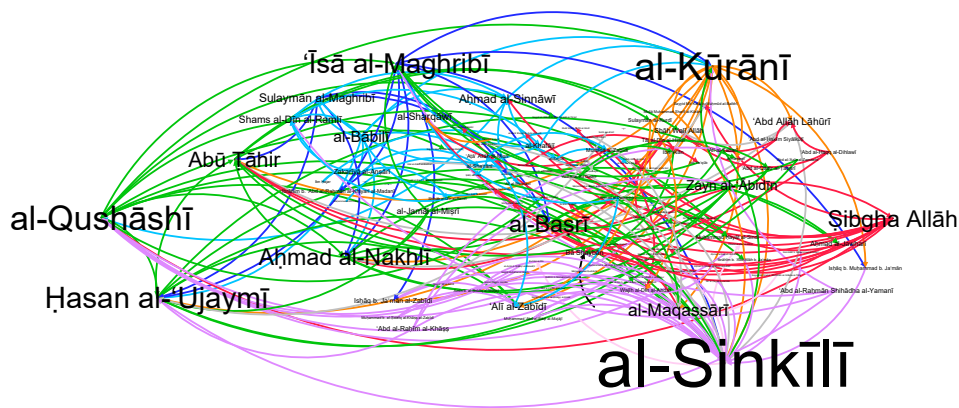


図3 主な学者たちの師弟関係と出身地

北西アフリカ)、ルーム（アナトリア半島）、エジプト、エルサレム、ダマスкас、ハラマイン（マッカとマディーナ）、クルド、イエメン、インドである。

一方、図3は主な学者たちの師弟関係と出身地を視覚化したものである。師弟関係が結ばれた学者同士が辺（エッジ）で結ばれ、それぞれの学者の出身地が色で示されている。マグリブの青色、エジプトの水色、ハラマインの緑色、クルドの橙色、インドの赤色に加え、インドネシアの桃色が目を引く。これは、アズラが研究の焦点をインドネシアとハラマインの関係に当てていたためである。入力したデータに偏りがある点に注意が必要であるものの、出身地を異にする学者たちの間の師弟関係を確認することができる。

ムヒッピーもムラーディーもダマスкас出身であるため、ダマスкас出身者の情報は集めやすかったのであろう。しかし、図2でムヒッピーと比べると、ムラーディーの作品の方がダマスкас出身者への言及が多いようである。また、ムラーディーの方が、エジプトやイエメンの出身であることを示す単語の数が少ないようである。17世紀と比べて18世紀においては、エジプトやイエメン出身の学者たちの影響が小さくなったことを示しているのかもしれない。17世紀の学者であるイブラーヒーム・クーラーニーはエジプトやダマスкасを経てマディーナへ定住し、その兄弟子であるアフマド・クシャーシー *Aḥmad al-Qushāshī* (1583-1661年) はマディーナ出身であるがイエメンに留学している。一方、イブラーヒーム・クーラーニーの息

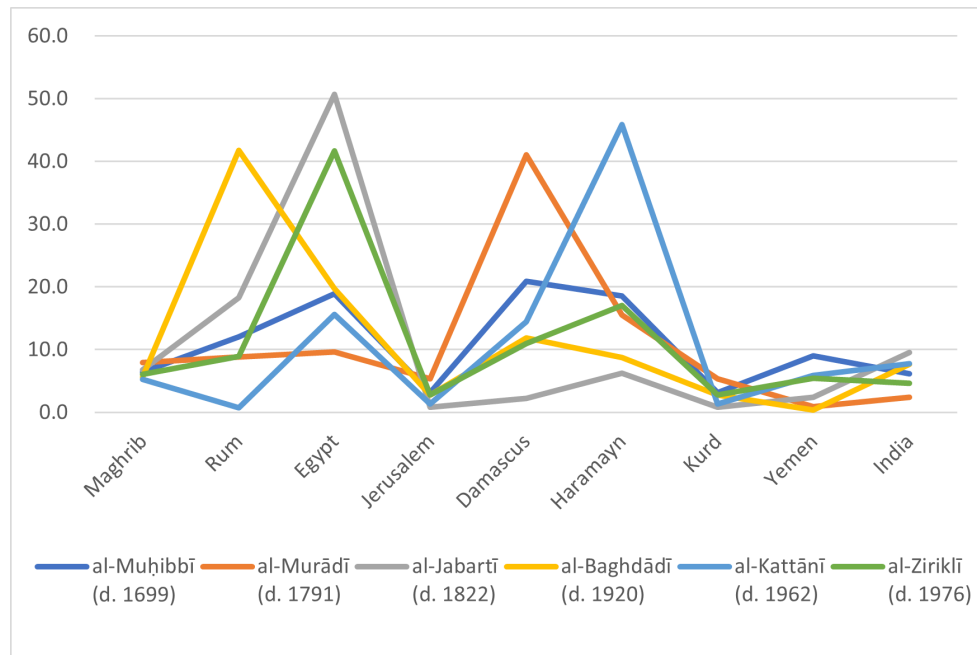


図4 出身地を表す単語の割合

子であり弟子であるアブー・ターヒルには、マディーナを離れて遊学したという記録は現在のところ見つかっていない。アブー・ターヒルの弟子となったシャー・ワリーウッラーも、エジプトやダマスカスなど他の都市には寄ることなく、インドからハラマインのみを目指した。各地を移動しながら学問を重ねるという学者たちの周遊的な移動が、ハラマインへの一極集中と往復だけの硬直的で直線的な移動へと変化した可能性がある。

続いて、ムヒッピーとムラーディーを含む全6作品の全体的な傾向を考察したい(図4)。各作品の出身地を表す単語における割合からすれば、やはりムラーディーのダマスカス出身者への言及は多く、エジプト出身のジャバルティー al-Jabartī とベイルート生まれダマスカス育ちのズィリクリー al-Ziriklī はエジプト出身者、マグリブ出身のカッターニー al-Kattānī はハラマイン出身者への言及が多い。名前からバグダード出身であることが分かるバグダーディー al-Baghdādī は、ルーム出身者について多く言及している。これは、彼がイスタンブルにいたことと関係しているかもしれない。

写本や刊本の画像データだけではなく、テキストデータまで公開されることにより、上記のような考察が可能となる。たとえばズィリクリーの『偉人 al-A'lām』は8巻本であり、他の伝記も基本的に複数巻本である。何百ページもある各作品を1ページずつバグダード出身者を示す「バグダーディー al-Baghdādī」は何個、マグリブ出身者を示

す「マグリビー al-Maghribī」は何個と数えていたら、数週間はかかる作業であっただろう。一方、OpenITI のテキストデータには、OCR の出力結果そのままのファイルや、そのファイルを人の手で修正してもなお誤りが残ったままのファイルもある。今回の分析では、厳密なテキストの整形を行っていないことに注意されたい。また、人名に出身地を示す単語が含まれていない人物や、人名の出身地が先祖の出身地で本人の出身地とは異なる人物も存在する。たとえば、カッターニーはフェズの名家の出身であるが、マグリビーと書かれていなかったり、ジャバルティーはカイロの名家の出身であるが、エジプト出身者を示す「ミスリー al-Miṣrī」と書かれていなかったりする。正確な集計結果を出すには、やはりテキストの精読は不可欠である。

おわりに

本稿では、筆者のこれまでの経歴とともに、デジタルヒューマニティーズの歴史を振り返ってきた。色とりどりの美しいネットワーク図などを示した方が、読者のデジタルヒューマニティーズへの関心を高めることができたかもしれない。しかし、デジタルヒューマニティーズは手段であって目的ではない。大量のデータを最新のデジタル技術で視覚化しても、そこに学術的発見が伴っていなければならない。

写本や刊本の画像だけではなく、テキストファイルも公開されることによる利点については、本稿では「遠読」⁵⁴⁾ 的

手法を可能にする点にしか言及できなかった。しかし、たとえば TEI を使って複数人によるテキストの読み方を示すことで、学術的な検証可能性を高めることもできる。また、学術的な継承という観点でも、研究の過程を含めて資料のデジタル化を進めていくことが望ましい。人文学研究においては、或る書体の写本を読むことのできる研究者が世界に一人しかいないという事態が起こりうる。このような場合、画像と対応したテキストデータを後世に残しておくことの意義は計り知れないだろう。とりわけ少子化の進む日本においては、あらゆる学問分野で後継者が不足することになるので、切迫した問題であるといえる。従来の人文学を継承したいと願う研究者こそ、研究のデジタル化に取り組むべきではないだろうか。筆者としては、デジタルヒューマニティーズという言葉が使われなくなるほど一般化することを願っている。

注

- 1) digitalnagasaki. 2020. 「Methodological Commons: デジタル人文学で昔から定番の話」『digitalnagasaki のブログ』 <https://digitalnagasaki.hatenablog.com/entry/2020/12/20/182659> (2024 年 9 月 3 日閲覧)。
- 2) 石田友梨. 2016. 「シャー・ワリーウッラー・ディフラウィー著『か弱き下僕の生涯における優雅なる一編』」『大阪経済法科大学論集』 111: 96–97. https://www.keiho-u.ac.jp/assets/pdf/pages/research/association/daigakuronshu_111.pdf (2024 年 9 月 3 日閲覧)。
- 3) Benjamin Kiessling. 2015–2024. “kraken.” The Kraken Website. <https://kraken.re/main/index.html> (2024 年 9 月 3 日閲覧)。
- 4) python.jp. 2001–2023. 「プログラミング言語 Python の紹介」『python Japan』 <https://www.python.jp/pages/about.html> (2024 年 9 月 3 日閲覧)。
- 5) Gephi.org. 2008–2022. “Home.” Gephi. <https://gephi.org/> (2024 年 9 月 3 日閲覧)。
- 6) dh-notes. 2022. “dhnotes.” GitHub. <https://github.com/dh-notes/dhnotes/blob/master/pages/dh-programs.md> (2024 年 9 月 3 日閲覧)。
- 7) 東京大学大学院人文社会系研究科・文学部附属次世代人文学開発センター人文情報学部門. n.d. 「東京大学大学院横断型教育プログラム デジタル・ヒューマニティーズ」『東大デジタル人文学』 <https://dh.tu-tokyo.ac.jp/home> (2024 年 9 月 3 日閲覧)。
- 8) Chiba University. 2010–2024. 「プログラム紹介」『千葉大学卓越大学院プログラム アジアユーラシア・グローバルリーダー養成のための臨床人文学教育プログラム』 <https://jinbun-taknetsu.chiba-u.jp/top/introduction.html> (2024 年 9 月 3 日閲覧)。
- 9) 後述するグローバル COE プログラムにより、立命館大学が 2007–2011 年度に「日本文化デジタル・ヒューマニティーズ教育プログラム」で修了証を授与していたようである。Ritsumeikan University. 2008. 「教育プログラム一覧」『日本文化デジタル・ヒューマニティーズ拠点』 https://www.arc.ritsumei.ac.jp/lib/GCOE/guideline_j.html (2024 年 9 月 3 日閲覧)。
- 10) 人間文化研究機構デジタル・ヒューマニティーズ促進事業. n.d. 「NIHU DH について」『NIHU DH』 <https://dh.nihu.jp/about> (2024 年 9 月 3 日閲覧)。
- 11) 人間文化研究機構デジタル・ヒューマニティーズ促進事業. n.d. 「プロジェクト一覧」『NIHU DH』 <https://dh.nihu.jp/projects> (2024 年 9 月 3 日閲覧)。
- 12) ヨーロッパでは 2014 年設立の「芸術と人文学のためのデジタル研究基盤 Digital Research Infrastructure for the Arts and Humanities」(DARIAH) に、現在 23 カ国が加盟している。DARIAH-EU. n.d. “Members and Partners.” DARIAH-EU. <https://www.dariah.eu/network/members-and-partners/> (2024 年 9 月 3 日閲覧)。
- 13) 人間文化研究機構デジタル・ヒューマニティーズ促進事業. n.d. 「DH 若手の会で育む」『NIHU DH』 <https://dh.nihu.jp/projects/dh-new-members-union> (2024 年 9 月 3 日閲覧)。
- 14) 浜名恵美. 2019. 「デジタル・ヒューマニティーズの本格的導入の提案：日本の英語文学研究の resilience のために」『英文学研究支部統合号』 11: 95. https://doi.org/10.20759/elsjregion-al.11.0_95 (2024 年 9 月 3 日閲覧)。
- 15) Susan Schriebman, Ray Siemens, and John Unsworth eds. 2004. A Companion to Digital Humanities. Malden; Oxford; Carlton; Blackwell. <https://companions.digitalhumanities.org/DH/> (2024 年 9 月 3 日閲覧)。2015 年には同じ編者たちによる新版 (Susan Schriebman, Ray Siemens, and John Unsworth eds. 2016. A New Companion to Digital Humanities. Malden; Oxford; Chichester; John Wiley & Sons) が出版されている。
- 16) Matthew G. Kirschenbaum. 2010. “What Is Digital Humanities and What’s It Doing in English Departments?.” ADE Bulletin 150: 56–57. <https://www.maps.mla.org/content/download/155871/file/ade.150.55.pdf> (2024 年 9 月 3 日閲覧)。
- 17) Google. n.d. Google Books Ngram Viewer. <https://books.google.com/ngrams/> (2024 年 9 月 3 日閲覧)。

グーグル社が収集した書籍のデジタルテキストのデータ (コーパ

- ス) から、特定の単語の出現頻度を検索し、図表として視覚化する機能をもつ。開発の経緯や分析例については、エレッツ・エイデン、ジャン＝バステイト・ミシェル『カルチャロミクス: 文化をビッグデータで計測する』(草思社、2016 年) を参照のこと。
- 18) URL は前注 15 を参照のこと。
- 19) Alliance of Digital Humanities Organizations. n.d. "About." Alliance of Digital Humanities Organizations. <https://adho.org/about/> (2024 年 9 月 3 日閲覧)。
- 20) Alliance of Digital Humanities Organizations. n.d. "Conference." Alliance of Digital Humanities Organizations. <https://adho.org/conference/> (2024 年 9 月 3 日閲覧)。
- 21) European Association for Digital Humanities. n.d. "About." European Association for Digital Humanities. <https://eadh.org/about> (2024 年 9 月 3 日閲覧)。
- 22) Alliance of Digital Humanities Organizations. n.d. "About." Alliance of Digital Humanities Organizations. <https://adho.org/about/> (2024 年 9 月 3 日閲覧)。
- 23) Ritsumeikan University. 2008. 「拠点概要」『日本文化デジタル・ヒューマニティーズ拠点』https://www.arc.ritsumei.ac.jp/lib/GCOE/guideline_j.html (2024 年 9 月 3 日閲覧)。
- 24) Japanese Association for Digital Humanities. 2012. "Purpose." Japanese Association for Digital Humanities. <https://www.jadh.org/purpose> (2024 年 9 月 3 日閲覧)。
- 25) Japanese Association for Digital Humanities. 2012. 「デジタル技術を活用した人文学研究に携わっている日本の研究者のみなさまへ」『Japanese Association for Digital Humanities』<https://www.jadh.org/node/25> (2024 年 9 月 3 日閲覧)。
- 26) 永崎研宣. 2011. 「人文情報学月報への招待」『人文情報学月報』創刊準備号 <https://www.dhii.jp/DHM/dhm00> (2024 年 9 月 3 日閲覧)。
- 27) 鹿内菜穂. 2020. 「[CH] 人文科学とコンピュータ研究会」『情報処理学会 60 年のあゆみ』<https://www.ipsj.or.jp/60anv/60nenshi/3-3-3-6.html> (2024 年 9 月 3 日閲覧)。
- 28) じんもんこん 2009 実行委員会. n.d. 「ホーム」『人文科学とコンピュータシンポジウム「じんもんこん 2009」』<https://www.jinmoncom.jp/sympo2009/> (2024 年 9 月 3 日閲覧)。
- 29) 計量国語学会. n.d. 「学会概要」『計量国語学会』<https://www.math-ling.org/> (2024 年 9 月 3 日閲覧)。
- 30) digitalnagasaki. 2020. 「「デジタル人文学」以前の日本の人文系デジタルテキスト研究を探訪してみる」『digitalnagasaki の プ ロ グ 』<https://digitalnagasaki.hatenablog.com/entry/2020/11/23/021820> (2024 年 9 月 3 日閲覧) など参照されたい。
- 31) Middle East Studies Association of North America. n.d. "Previous Meetings." Middle East Studies Association. <https://mesana.org/annual-meeting/previous-meetings> (2024 年 9 月 3 日閲覧)。
- 32) Regents of the University of Michigan. 2024. "Computational Reading of Arabic Biographical Collections with Special Reference to Preaching in the Sunni World (661--1300 CE)." University of Michigan Library. <https://hdl.handle.net/2027.42/102300> (2024 年 9 月 3 日閲覧)。
- 33) KITAB. 2024. "About KITAB." KITAB. <https://kitab-project.org/about/> (2024 年 9 月 3 日閲覧)。
- 34) Open Islamicate Texts Initiative. 2024. "About." Open Islamicate Texts Initiative. <https://openiti.org/about.html> (2024 年 9 月 3 日閲覧)。
- 35) dreamsea.co. 2018–2023. "About." DREAMSEA. <https://dreamsea.co/about/> (2024 年 9 月 3 日閲覧)。
- 36) The Digital Orientalist. n.d. "About The Digital Orientalist." The Digital Orientalist. <https://digitalorientalist.com/about/> (2024 年 9 月 3 日閲覧)。
- 37) Digital Ottoman Studies. n.d. "Home." Digital Ottoman Studies. <https://www.digitalottomanstudies.com/> (2024 年 9 月 3 日閲覧)。
- 38) American University of Sharjah. 2024. "The Praxis of Digital Humanities: Expanding Horizons and Transforming Scholarship in the Arab/Islamic World." American University of Sharjah. <https://info.aus.edu/ara-2024> (2024 年 9 月 3 日閲覧)。
- 39) 石田友梨. 2021. 「現代イスラーム社会における AI の位置づけ」『宗教研究』94 (別冊): 23–24. https://jpars.org/journal/bulletin/wp-content/uploads/2021/01/vol_94-1.pdf (2024 年 9 月 3 日閲覧)。
- 40) 石田友梨. 2016. 「イスラーム研究におけるデジタル・ヒューマニティーズの活用に向けて: シャー・ワリーウッラー『ハラマインの師たちの瞳孔』に基づく 17–18 世紀ハラマインの学者ネットワーク分析」『イスラーム地域研究ジャーナル』8: 25–36. <http://hdl.handle.net/2065/48314> (2024 年 9 月 3 日閲覧)。
- 41) 学術変革領域研究 (A) 「イスラーム的コネクティビティにみる信頼構築」. 2024. 「研究体制」『イスラーム信頼学』<https://connectivity.aa-ken.jp/outline/#sec02> (2024 年 9 月 3 日閲覧)。
- 42) The Editorial department of Seiyoshigaku. 2022. 「ホーム」『日本西洋史学会『西洋史学』編集部』<https://sites.google.com/view/seiyoushigaku/> (2024 年 9 月 3 日閲覧)。
- 43) 日本西洋古典学会. 2024. 「日本西洋古典学会第 73 回大会 (2023 年度)」『日本西洋古典学会』<https://clsoc.jp/>

- [meetings/2023/meeting.html](#) (2024 年 9 月 3 日閲覧)。
- 44) 田中一孝. 2024.「西洋古典に特化した AI 対話システム「ヒューマニテクスト」の開発」『日本西洋古典学会』<https://clsoc.jp/news/2024/240616.html> (2024 年 9 月 3 日閲覧)。
- 45) 本来は汚い筆跡を形容する言葉。国立国会図書館「NDL Ngram Viewer」(<https://lab.ndl.go.jp/ngramviewer/>) で検索すると、「ミミズがはったようなヘブライ文字」(京都新聞社編『日本人の解体』駸々堂出版、1974 年)や「ミミズがはったようなテルグ語」(川又一英『さすらいびとの唄：インドの大地と放浪者たち』音楽之友社、1979 年)など、アラビア語やアラビア文字以外を形容する例もある。
- 46) アラビア語を読み取ることのできる各種 OCR の精度を比較した Ishida Yuri and Shinoda Tomoaki. 2021. "A Study on the Accuracy of Low-cost User-friendly OCR Systems for Arabic: Part 1." The Digital Orientalist. <https://digitalorientalist.com/2021/09/17/a-study-on-the-accuracy-of-low-cost-user-friendly-ocr-systems-for-arabic-part-1/>; "A Study on the Accuracy of Low-cost User-friendly OCR Systems for Arabic: Part 2." The Digital Orientalist. <https://digitalorientalist.com/2021/09/24/a-study-on-the-accuracy-of-low-cost-user-friendly-ocr-systems-for-arabic-part-2/> (2024 年 9 月 3 日閲覧)を参照のこと。
- 47) READ-COOP. n.d. "Try it out." Transkribus. <https://www.transkribus.org/> (2024 年 9 月 3 日閲覧)。
- 48) eScriptorium. n.d. "Home." eScriptorium. <https://escriptorium.inria.fr/> (2024 年 9 月 3 日閲覧)。
- 49) CALFA. n.d. "OCR." CALFA. <https://calfa.fr/ocr> (2024 年 9 月 3 日閲覧)。
- 50) 石田友梨. 2024. 「TEI と RDF を用いたアラビア文字伝記資料のデータ化の試み：イブラーヒーム・クーラーニーの伝記を例に」『岡山大学大学院社会文化科学研究科紀要』57: 12-15. <http://doi.org/10.18926/66825> (2024 年 9 月 3 日閲覧)。
- 51) ROIS-DS. n.d. 「くずし字 OCR (AI くずし字認識)」『人文学オープンデータ共同利用センター』<http://codh.rois.ac.jp/char-shape/OCR/> (2024 年 9 月 3 日閲覧)。
- 52) TEI 研究会. n.d. 「TEI 古典籍ビューワ TEIviewer4EAJ ver. 2024-08-28」『TEI 研究会』<https://tei.dhii.jp/teiviewer4ej> (2024 年 9 月 3 日閲覧)。
- 53) Azyumardi Azra. 2004. The Origins of Islamic Reformism in Southeast Asia: Networks of Malay-Indonesian and Middle Eastern 'Ulamā' in the Seventeenth and Eighteenth Centuries. Crows Nest: Allen & Unwin; Honolulu: University of Hawai'i: Press.
- 54) フランコ・モレッティ. 2016. 『遠読：〈世界文学システム〉への挑戦』みすず書房。