

## アトリエに吹く風

— デジタル・ヒストリーと史料 —

櫻田宗紀

小風尚樹

## 序論

## 背景

今からおよそ20年前、『史学雑誌』のコラム「歴史の風」にデジタル時代の到来に警鐘を鳴らす2本のエッセイが掲載された<sup>1)</sup>。二宮宏之と三鬼清一郎は、それぞれフランスと日本の図書館におけるコンピュータ検索システムの導入が引き起こしうる問題への危機感をあらわにしている。

同じく「歴史の風」に2015年に掲載された深沢克己のエッセイを読むと、我々は隔世の感をおぼえる。深沢は、国際学会で発表し、欧米語の学術雑誌に論文を掲載し、インターネットを用いて外国人研究者と日常的に交流する研究者を「新人類」と称した<sup>2)</sup>。人類の進化は、しばしば道具の作製・使用と結びつけて説明されるが、深沢の言う「新人類」が手にした最大の道具は、間違いなく情報通信技術だろう。ベルリンの壁が崩壊した年に生まれた筆者らにとって、情報端末は日常生活に浸透していたため、2010年代に研究を始めるにあたってインターネットから得られる研究関連の情報との向き合い方を積極的に考えようとしたのは自然なことであった。ただし今後の義務教育の対象となる世代は、小学校におけるプログラミング教育の導入などによって、さらに高度な道具を手にして我々の前に現れるだろう。

本特集の企画にあたって筆者らに白羽の矢が立ったのは、このような時流を捉えてのことと思われる。2017年、我々は小風を代表とする Tokyo Digital History (以下 ToDH) というコミュニティを立ち上げた。ToDH は、歴史家・アーキビスト・エンジニアの協働に基づき、情報通信技術を歴史研究に導入することの可能性や留意点を実践的に模索し、国内外で成果を発信している<sup>3)</sup>。本稿の筆者は、中世教皇史を専門とする櫻田と近代イギリス海事史を専門とする小風の2名ではあるが、その内容は、ToDH のメンバーとの議論に基づいている<sup>4)</sup>。筆者らが拠って立つ専門分野の拘束は免れえないものの、専攻を異にする様々なメンバーとの議論を通して、読者

1) 二宮宏之「ビブリオテークに風が吹く」『史学雑誌』108(9)、1999年、40-42頁；三鬼清一郎「コンピューター歴史学の光と影」『史学雑誌』108(11)、1999年、40-42頁。

2) 深沢克己「「あまりに遠し」から西洋史研究の「新人類」まで」『史学雑誌』124(1)、2015年、40-42頁。

3) ToDH の立ち上げについては、小風尚樹「ここからはじめる日本のデジタル・ヒストリー — Tokyo Digital History の立ち上げによせて —」図書出版文学通信ブログ、2018年5月。業績一覧は、<https://historiansworkshop.org/2018/07/09/> (アクセス確認日時：2019年9月7日、以下同様)

一人ひとりに訴えかける内容を持つように努めた。

### 議論の射程

本論に先立ち、デジタル技術と人文学の関わりをめぐる議論を引き合いに出しつつ、本稿が扱う範囲について触れたい。その前に、本稿で何を論じないのか、先に断っておくのがよいだろう。本稿は、すでに欧米では研究・教育活動が定着しつつあるデジタル・ヒューマニティーズ（以下DH）について論じるわけではない。たしかにDHは、歴史学・文学・言語学・考古学・書誌学・図像学など人文系諸学問とデジタル技術の邂逅から生まれる新たな可能性と課題を、分野横断的に広く議論し共有する「方法論的共有地」や「大きなテント」と形容される知の枠組みである点で<sup>5)</sup>、歴史研究にも示唆するところは大きい。しかし本稿は、ステーブン・ロバートソンが述べるように、DHは大きなテントというよりも、ひとつの家のなかに分野ごとの部屋が分かれているものと理解し、デジタル技術の活用を議論するには各分野の特質を考えるべきだとする姿勢に共感し<sup>6)</sup>、歴史研究に軸足を置く。その意味で本稿は、歴史学のデータを活用して情報学の研究を行う歴史情報学や<sup>7)</sup>、歴史コンテンツを社会に発信する際にデジタル技術を活用するデジタル・パブリック・ヒストリーと<sup>8)</sup>、議論の射程が重なり合う部分もある。しかし、むしろ本稿は、歴史家の日常的な研究実践、すなわち「どのように史料を集め／どう読むのか」に着目し、そこにデジタル技術がどのように影響するのかを考察したい。

### 本論の構成

本論では、デジタル技術が歴史研究にもたらす変化について、3つの観点、すなわち研究環境の変化／集め方の変化／読み方の変化に分けて考察する。これらの考察対象は、ジョージ・メイソン大学の歴史とニューメディア研究センターが2015年に公開した『デジタル・ヒストリー白書』で紹介された「歴史家の基本的な営み」を参照しつつ<sup>9)</sup>、デジタル技術がもたらす影響が大きいと考えられる論点に絞って設定した。本論では、まず1章で研究者をとりまく環境の変化を確認したのち、2章・3章で史料の集め方と読み方それぞれにおける最近の動向および課題と対応策を取り上げる。史料と言っても、絵画や地図、貨幣や彫刻といった非文字史料ではなく<sup>10)</sup>、

- 
- 4) 本稿の執筆に関わる議論に参加してくださった方々（敬称略、五十音順）、小川潤、金甫榮、小風綾乃、小林拓実、佐治奈通子、中村覚、永崎研宣、長野壮一、福田真人、槇野翔、山崎翔平、山中美潮に謝意を表したい。なお、本稿の文責は筆者2名が負う。
- 5) DHの理念に関する概説として、Burdick, A. et al., *Digital Humanities*, Cambridge, MA, 2012, pp. 122-135（邦訳は<http://21dzk.lu-tokyo.ac.jp/DHI/index.php?sg2dh>）。
- 6) Robertson, S., 'The Differences between Digital Humanities and Digital History', in M. K. Gold & L. F. Klein, eds., *Debates in the Digital Humanities 2016*, Minneapolis, 2016, pp. 289-290.
- 7) たとえば、後藤真・橋本雄太編『歴史情報学の教科書——歴史のデータが世界をひらく——』文学通信、2019年。
- 8) 国内のデジタル・パブリック・ヒストリーの最新の成果として、橋本雄太「市民参加型史料研究のためのデジタル人文学基盤の構築」博士論文（京都大学）、2018年。
- 9) Arguing with Digital History Working Group, 'Digital History and Argument', White Paper, Roy Rosenzweig Center for History and New Media, 2017, pp. 3-4. 同白書は、歴史家の基本的な営みとして、史料の選択／とりまとめ／順序立て／文脈化／比較／時期設定／想定読者との議論、を提示した。
- 10) デジタル技術は非文字史料との結びつきも強めている。とくに図像については、永崎研宣「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)、2017年、61-66頁。

筆者らの力の及ぶ範囲で文字史料を中心に考察する。基本的に櫻田が1・2章、小風が3章を執筆した。

## 1章 研究環境の変化

外国史研究に従事する研究者たちは、常に諸外国の研究動向を追っていることだろう。日常的な作業のなかで、デジタル技術を活用した歴史研究を行う個人やプロジェクトの存在に目を留める機会は確実に多くなっている<sup>11)</sup>。本章では、歴史学界における既存の制度、すなわち研究機関と学会に視点を限定し、デジタル・ヒストリーが我々の研究環境にどのように入り込んでいるのかを、ごく簡単に眺めてみよう。

まず研究・教育拠点の状況を確認する。DH 研究・教育の国際ネットワーク centerNet に加盟している機関は、2019年9月現在202を数える<sup>12)</sup>。欧米では、DH の課程・講座を設置する大学が増えており、DH の入門やデータ構築の実践などさまざまなレベルのサマースクールが開催されている<sup>13)</sup>。しかし、デジタル・ヒストリーに限って言えば、正規課程を置くのは稀であり、研究機関附属のセンターが講義やサマースクールを開催する機会が多い<sup>14)</sup>。そのような拠点としては、アメリカの Roy Rosenzweig Center for History and New Media、フィンランドの Helsinki Computational History Group、ルクセンブルクの The Luxembourg Centre for Contemporary and Digital History が有名だが、数ある組織のなかでも開かれたコミュニティ形成の点で際立っているのが、ロンドン大学歴史学研究所の Digital History Seminar である。デジタル・ヒストリーに関するさまざまなレッスンを提供するとともに、月一回程度の頻度で定例会を開催する同セミナーは、デジタル・ヒストリー研究のひとつの発信源となっている。過去のセミナー報告の動画とスライドは、セミナーの YouTube チャンネルおよび SlideShare で公開されている<sup>15)</sup>。

次に、日本の歴史家がこれまでも参加してきた学会において、デジタル技術を取り入れた研究発表がどの程度浸透しているのか確認しよう。たとえば中世学の国際学会として双璧をなすリーズの International Medieval Congress と<sup>16)</sup>、カラマズーの International Congress on Medieval Studies の過去のプログラムをみると<sup>17)</sup>、最近10年間で DH 関連の発表が増えていることがわかる。近世史研究者が数多く参加する国際18世紀学会では、2019年大会において DH 関連のセッション23件、ペーパー77件を数えた<sup>18)</sup>。デジタル・ヒストリーに力を入れる歴史学専門の学会としては、アメリカ歴史学協会 (AHA) が有名である。2018年の大会では、24パネルがデジタ

11) 関連情報の把握には、菊池信彦「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、2013年、55-68頁を参照。

12) centerNet 登録機関一覧、<https://dhcenternet.org/centers>

13) ヨーロッパ DH 協会の HP (<https://eadh.org/education>) で、DH の課程を置く大学、DH の研究センターなどの一覧をみることができる。

14) 組織運営や資金運用の柔軟さの点で、デジタル・ヒストリーの研究・教育拠点は大学の学部よりもセンターが適切だとされている。Cohen, D. J. et al. 'Interchange: The Promise of Digital History', *The Journal of American History*, 95(2), 2008, p. 465. 国内の DH の研究・教育インフラの現状については、後藤真「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、2018年、1-16頁を参照。

15) 詳しくは IHR Digital History Seminar の HP (<https://ihrdighist.blogs.sas.ac.uk/>) を参照。

16) <https://www.imc.leeds.ac.uk/about/programme-archives/>

17) [https://scholarworks.wmich.edu/medieval\\_cong\\_archive/](https://scholarworks.wmich.edu/medieval_cong_archive/)

ル・ヒストリー関連のものだった<sup>19)</sup>。

日本国内の状況はどうだろうか。日本西洋史学会においては、2016年の第66回大会で「デジタル資源を活用した資料の共有化とこれからの西洋研究への展望」と題する小シンポジウムが組まれたことがある。現在、日本国内でデジタル・ヒストリー研究を発表・議論するのであれば、情報処理学会の研究会である「人文科学とコンピュータ研究会」（通称「じんもんこん」）がよい選択肢となる<sup>20)</sup>。だが、後藤真が指摘するように、日本のDH研究は研究基盤構築や情報探索のシステム設計に議論が偏りがちで、分析に着目した研究は欧米に比べて少ない<sup>21)</sup>。分析系の研究を多く蓄積している欧米の情報にアクセスしやすい研究者には、分析手法や分析視角の議論に参入する可能性が十分に開かれていると言える。

## 2章 集め方の変化

デジタルアーカイブやオンライン史料目録の充実により、史資料へのアクセスが容易になった、という文言が常套句と化して久しい。今となってはアクセスの容易さに言及するよりも、ウェブ上の情報に全くアクセスすることなく現在求められている水準の研究を行うことは不可能となった、と言ったほうが適切だろう<sup>22)</sup>。ウェブ上には研究材料に関する情報があふれており、研究対象・時代・地域によってそれらに関する情報の質・量・アクセスの容易さは均一ではない。そこで本章では、中世ドイツ史研究に関わる文書史料を収集する過程を例として<sup>23)</sup>、デジタル化された史料を扱う際に留意すべきことを検討してみたい。

### 最近の動向

西洋中世の文書史料を収集しようとするとき、まずは求める史料がすでに批判校訂版に採録されているかどうかを確認することになる。ドイツ史の場合、各地域・都市・司教区ごとに文書集(Urkundenbuch)が編纂されている。これらの文書集がウェブ上で閲覧可能か調べる場合、RI-

18) 小風綾乃「18世紀研究におけるDHの広がり——第15回国際十八世紀学会(ISECS 2019)に参加して——第1回：個別発表にみるデータ可視化」『人文情報学月報』97、2019年8月。

19) AHAのデジタル・ヒストリー関連の活動については、山中美潮「アメリカ史研究とデジタル・ヒストリー」『立教アメリカンスタディーズ』40、2018年、12-13頁。

20) 国内のDH関連の学会・研究会については、後藤真「人文情報学と歴史学」後藤・橋本編、前掲書、15-18頁。

21) 同論文、16-17頁。

22) 近藤和彦・後藤はる美・伊東剛史「ECCOやNCCOを使わない研究は国際基準にならないという段階に来ています」GALE『IN THE WORDS OF USERS——GALEユーザーインタビュー集——』センゲージラーニング株式会社、2017年、127-146頁。

23) 西洋中世史料のデジタル化の状況に関する日本語での情報発信の例として、鈴木道也「記録管理の過去と現在——サン＝ドニ修道院証書集<Cartulaire blanc>の電子テキスト化について——」『埼玉大学紀要教育学部』55(1)、2006年、29-41頁；同「歴史資料デジタル化の現状 中世フランス史の場合」『ヨーロッパ・グローバル化と諸文化圏の変容 研究プロジェクト報告書』東北学院大学オープン・リサーチ・センター、2012年、209-220頁；同「〈特集 デジタルメディアと歴史学〉ヴァンサン＝ド＝ボーヴェ編著『大いなる鑑(Speculum Maius)』の構造分析におけるテキスト・データベースの活用——現状と課題——」『ヨーロッパ文化史研究』14、2013年、5-27頁；菊池信彦「デジタル化の「その先」へ——ヨーロッパにおける中世写本のデジタル化の現状とデジタルヒューマニティーズ——」『情報の科学と技術』65(4)、2015年、156-163頁。

OPAC を利用するのが定石である<sup>24)</sup>。RI-OPAC では、検索された出版物が各種のデジタルアーカイブで公開されていれば、検索結果のページにそのリンクが貼られている。公開されているデータにアクセスすることができれば、文書集の紙面をスキャンした画像を目にすることになるだろう。テキストデータを公開している史料データベースの代表例としては、西洋中世史料の収集・編纂・研究機関として最も有名なモヌメンタ・ゲルマニアエ・ヒストリカ (Monumenta Germaniae Historica, MGH) が発行する史料集のデジタル版、dMGH が挙げられる<sup>25)</sup>。dMGH では、紙面のスキャン画像と HTML バージョンが切り替え可能となっており、利用者は史料の文面をテキストデータとして扱うことができる。

一方、まだ転写・校訂されていない文書史料を探索する、あるいは校訂版が手に入るにもかかわらずオリジナルを閲覧する必要がある場合は、所蔵文書館が史料の画像を公開しているかどうか確認することになる。文書館のオンライン史料目録で史料を検索するとき、多くの人は、まず年代や日付の下限と上限を設定した上で、キーワード検索を行うだろう。もし画像が公開されていれば、この時点で史料の画像を入手することができる。近年では多くの研究機関が、史料の画像データベースの構築・公開を進めている。たとえば、1928年から中世史料の画像を作製・保存してきたマールブルク大学の研究所 Lichtbildarchiv älterer Originalurkunden は、2008年にデータベースを公開した<sup>26)</sup>。ここでは、原則的に神聖ローマ帝国内で受領された1250年までの文書の画像を公開し、研究者のみならず広いオーディエンスの利用に供している<sup>27)</sup>。機関によって進度はさまざまだが、今後、所蔵史料の画像を公開する流れが逆行することはないと予想される。

### 課題と対応策

かつて、コンピュータ検索システムの導入により容貌を変えたフランス国立図書館の書棚を目にした二宮は、「一点一点その顔だちをもち、初めから終わりまで読まれるものとしてあった書物は、部分的に情報を引き抜くための素材と化しつつある」と述べ、変化をこうむる史料との接し方への懸念を示した<sup>28)</sup>。この懸念を少し具体的に換言すれば、デジタル化が進むことによって歴史家と史料のあいだが点と点の関係になり、史料の持つコンテキストへの注意がおろそかになってしまうのではないかと、ということだろう。史料のコンテキストを重視することは、史料批判の基本でもあり、今世紀最初の10年に日本で高まりをみせた西洋中世史料論研究によって深く議論されてきた<sup>29)</sup>。しかし、ウェブ上で史料の画像を入手する機会が増えていく状況においては、史料のコンテキストへの視線を改めて重視させられることになる。

デジタル化によって史料がコンテキストから切り離されうるという点に関しては、近現代史研

24) RI-OPAC, [http://opac.regesta-imperii.de/lang\\_de/](http://opac.regesta-imperii.de/lang_de/). パッサウ大学の旧歴史補助学講座のHPでは、対象地域のアルファベット順に既刊文書集の一覧が公開されている。

25) dMGH, <https://www.dmgH.de/>

26) Lichtbildarchiv älterer Originalurkunden, <http://lba.hist.uni-marburg.de/lba/pages/>

27) このような画像データベースの存在を把握するには、ドイツの大学の中世史講座・歴史補助学講座のHPを参照するのが有益である。

28) 二宮、前掲論文、42頁。

29) 日本の西洋中世史研究における史料論の動向については、岡崎敦「〈学界展望〉西洋中世史料論と日本学界——いま何が問題か——」『西洋史学』223、2006年、43-56頁。

究で利用される新聞データベースを題材としてよく議論されてきた。ボブ・ニコルソンは、人が関心に沿った新聞記事を見つけようとする際、紙の新聞とデータベース上の新聞ではまったく逆方向の過程を経ると指摘している<sup>30)</sup>。すなわち、紙の場合は誌名⇒号⇒見出し⇒記事という具合に、視野に入る多くの情報を取捨選択しながらトップダウンで特定の記事に到達するのに対して、キーワード検索機能をそなえたデータベース上ではボトムアップの方向で情報を得る傾向にある。データベースで検索する際には、キーワードを含む記事を、記事の外側にあるコンテキストを飛び越えて特定することになるのである。

同様の観点が前節で述べた中世の文書史料の収集方法にもあてはまることは、容易に想定できるだろう。紙であれ、PDFであれ、文書集を手にしようとするとき、我々は、まず対象とする地域・都市・修道院・教会をタイトル名に含むものを探す。その後、調査する年代を絞りつつ、各文書の梗概を順次読みながら研究課題にとって重要な文書を特定する。このときの視点は、より広いコンテキストからひとつの史料に絞っていくというトップダウンの過程を経ている。それに対して、文書館のオンライン史料目録をはじめとするデータベースの検索機能を使って史料の画像を見つけた場合、その史料の外側にあるコンテキスト情報を得るには、ボトムアップの過程を経ることになる。手書き文字を解読することができたとしても、我々がそのコンテンツから歴史的事実を引き出すためには、史料が含まれている文書群の総体、文書作成の年代・日付、作成の主体や目的、あるいは文書の支持体やフォーマットなどを確定させる必要がある。この作業を怠ると、異なる背景・異なるバイアスを持つ史料の証言を、同じ性格のものとして並べ、自身の主張を支える論拠として提示するという誤謬に陥ってしまう。

デジタル時代において、歴史家が史料の持つコンテキストへの意識を正しく高める方策として、アーカイブズ学の基本を知ることが提案したい。西洋中世史料へのまなざしとアーカイブズ学を結びつける視点は、全く新しいものではなく、すでに岡崎敦が、アーカイブズ理論と史料論的研究の融合を試みている<sup>31)</sup>。これらの研究を参照しつつ、ここでは史料のコンテキストに関するアーカイブズ学の基本的な考え方を紹介し、アーキビストと歴史家が互いの研究領域を隣接学問として尊重し合うことの意義を述べたい。

まずは、アーカイブズの定義を確認しておこう。「アーカイブズ」という用語は、国際アーカイブズ評議会の話彙集によって次のように定義されている。「ある法人あるいは個人が、その活動の過程で作成、受領し、さらに組織固有の必要のために、それを形成させる主体あるいは後継者によって保管されるか、あるいはアーカイブズ上の価値ゆえに、適正な史料保管組織に移管される資料の総体で、日付、形態、物的支持体の如何を問わない」<sup>32)</sup>。加えて、アーカイブズ学における資料整理の原則として「出所原則」と「原秩序尊重の原則」が存在することを考慮すれば、アーカイブズの性質が明確になるだろう。つまりアーカイブズ資料は、一点一点で存在しているのではなく、業務や法関係管理を反映する「有機的な総体」であり、本来的に持っているその構

30) Nicholson, B., 'The Digital Turn: Exploring the Methodological Possibilities of Digital Newspaper Archives', *Media History*, 19, 2013, pp. 66-67.

31) 岡崎敦「現代アーカイブズ理論と西洋中世史料論研究」『西欧中世文書の史料論的研究 平成23年度年次活動報告書』2012年、82-89頁。

32) Walne, P., ed., *Dictionary of Archival Terminology*, 2nd ed., München/New York/London/Paris, 1988, p. 22. 3点ある定義のうち1点のみを挙げ、日本語訳は岡崎に従った。岡崎敦「アーカイブズ、アーカイブズ学とは何か」『九州大学附属図書館研究開発室年報』2012年、2頁。

造を崩すことなく残されるべきものとみなされているのである<sup>33)</sup>。

このようなアーカイブズ学の知見は、文書館のオンライン史料目録にも活かされている。たとえばニーダーザクセン・ブレーメン州とヘッセン州の州立文書館が共同で開発・導入した検索システム Arcinsys では<sup>34)</sup>、各文書の記述目録の画面上に「コンテキストを表示する (Kontext anzeigen)」というリンクがあり、その史料が属する「まとまり (Bestand)」の一覧をボトムアップの方向で辿ることができるようになっている。また一点一点の目録には、作成者、作成場所、作成年月日、梗概、さらには史料の外的特徴 (支持体の材料、タテ・ヨコの長さ、印璽の有無・材質・大きさなど) が記述されている。史料を利用する歴史家には、この記述目録をアーキビストの研究成果として丹念に読み、学問的に正当な方法で引用・批評することが求められている<sup>35)</sup>。もちろん、ある組織体の機構やある個人の履歴そのものは歴史研究の対象でもあり、一義的に確定できるほど自明ではないため、中世史料一点一点に詳細な目録記述を付与することが難しい場合もある。目録記述を充実させるためには、アーカイブズ学と歴史学の学問的な相互批判を活性化させることが望ましい。

日本においては、アーカイブズ学が、関連学問である歴史学や情報学といかなる関わり方をしているか明確にされてこなかったことが反省され、歴史家や情報学者との協力の必要性が説かれている<sup>36)</sup>。同様の反省は、ドイツ語圏の歴史学にもみられる。ドイツ語圏においては、アーキビストと文書形式学者 (Diplomatiker) の仕事の領域が重なり合うことを自覚しつつも、むしろ両者の目録記述の基準や作業条件の差異が強く意識されていた<sup>37)</sup>。しかしそれでも、情報通信技術の発展にともない、両者が共通の方法と目的のもとに史料を扱う仕組みが模索されてきた。その一例として、デジタルアーカイブ Monasterium.net を紹介したい<sup>38)</sup>。

Monasterium.net は、文書館のデジタル関係事業をサポートする相互扶助ネットワーク International Centre for Archival Research (ICARUS) によって運営されている。ニーダーエスターライヒの修道院に保管されていた11、12世紀の文書群の画像を作成することから始まったこのプロジェクトには<sup>39)</sup>、現在ヨーロッパ10か国から60以上の機関が参加し、所蔵する中近世文書の画像公開を進めている。ICARUS の2018年3月の事業報告書によれば、Monasterium.net で画像が公開されている文書の数は636,658点にのぼる<sup>40)</sup>。

Monasterium.net の特徴は、データの提供者と利用者が対話する仕組みが設けられている点に

33) 青山英幸『記録から記録史料へ——アーカイバル・コントロール序説——』岩田書院、2002年、264頁。

34) Arcinsys Hessen, <https://arcinsys.hessen.de/arcinsys/start>

35) 安藤正人 (聞き手: 小川千代子・宮間純一) 「〈シリーズ 歴史家とアーキビストの対話 第3回〉歴史学とアーカイブズ学の課題」『歴史学研究』967、33頁。

36) 青山、前掲書、287頁。

37) Vogeler, G., 'Das Verhältnis von Archiven und Diplomatik im Netz. Von der Archivischen zur Kollaborativen Erschließung', in J. Kemper & G. Vogeler, eds., *Digitale Urkundenpräsentationen*, Norderstedt, 2011, pp. 61-82.

38) Monasterium.net, <https://www.monasterium.net/mom/home>

39) Heinz, K., 'Monasterium.net - Auf dem Weg zu einem Europäischen Urkundenportal', in T. Kölzer et al., eds., *Regionale Urkundenbücher*, St. Pölten, 2010, pp. 139-145.

40) ICARUS Annual Report 2017/2018, pp. 5-6, <https://documents.icar-us.eu/documents/2018/06/annual-report-2017-2018.pdf>

ある。利用者には、みずから検討した史料の目録記述の加筆・修正を申請する手段が提供されており、その申請は専門家のレビューを経て既存の記述に反映される。このシステムは、文書館が Monasterium.net に提供した史料のデータを研究者が利用し、その史料を使った研究成果が目録記述に反映される、という共同作業を可能している。さらに、国際的に共有されている目録記述やテキスト符号化の標準を組み合わせることによって、データの交換や保存の持続性を保つ工夫も施されている。Monasterium.net は、アーカイブズ学と歴史学の対話によって成立したデジタルアーカイブの好例と言える。

### 3章 読み方の変化

ここでは、デジタル技術の進歩が歴史研究にもたらす変化のうち、史料の「読み方の変化」に着目したい。ララ・パットナムは、ウェブ上でアクセス可能な学術資源がますます増える一方で、多くの歴史家にとっての研究実践は情報の検索にとどまっていると指摘した<sup>41)</sup>。本章では、ウェブの世界に広がるテキストデータを、検索を越えてどのように活用し得るのか、どのような点が課題になっているのか検討したい。

#### 最近の動向

デジタル技術がもたらした最も顕著な「読み方の変化」は、「読まずに分析」できるようにした点である。よく知られているように、たとえば Gale 社の運用する Early English Books Online や The Times Digital Archives など、ウェブ上には数多くのテキスト・データベースが存在し、キーワードを検索することで自分が関心を寄せる情報に容易にアクセスできる。このような検索性の向上に加えて、テキスト・データベースの利点は、単語の出現に関する数量的事実を把握できる点である。たとえば長谷川貴彦は、Google Books 全体の中での単語の出現頻度を基に、政治史・経済史の減退と社会史の隆盛、そこから文化史へと続く20世紀歴史学の研究潮流を論じた<sup>42)</sup>。

テキストを分析する歴史家としては、デジタル文学研究がもたらし得る示唆も把握しておきたい。イタリアの文学者フランコ・モレッティは、伝統的なテキスト精読は「あまりに少ない分量の正典に依拠せざるを得ない」と批判し、大量のテキストを対象に、単語の使用状況に基づくジャンル推定やテキストに登場する人物の相関関係を統計的に分析する遠読の実践を提唱した<sup>43)</sup>。遠読はあくまで文学研究の話であって、歴史研究ではあまり有益と考えられていないとも言われるが<sup>44)</sup>、遠読がもたらす示唆にも注目したい。それは、語の使用パターンや登場人物の社会的関係性など精読からは得られにくい大量の情報を、短い時間でおおまかに把握するのに用いられる視覚化データ群の役割である。たとえば図1は、シェイクスピア『ハムレット』における登場人物の相関図を、作中の会話の様子を基にモレッティが描画したものである<sup>45)</sup>。

41) Putnam, L., 'The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast', *The American Historical Review*, 121 (2), 2016, pp. 377-402.

42) 長谷川貴彦『現代歴史学への展望——言語論的転回を超えて——』岩波書店、2016年、157頁。

43) Moretti, F., *Distant Reading*, London & New York, 2013 (秋草俊一郎ほか訳『遠読：〈世界文学システム〉への挑戦』みすず書房、2016年), p. 48.

44) Blevins, C., 'Digital History's Perpetual Future Tense', in M. K. Gold & L. F. Klein, eds., *op. cit.*, pp. 308-324.



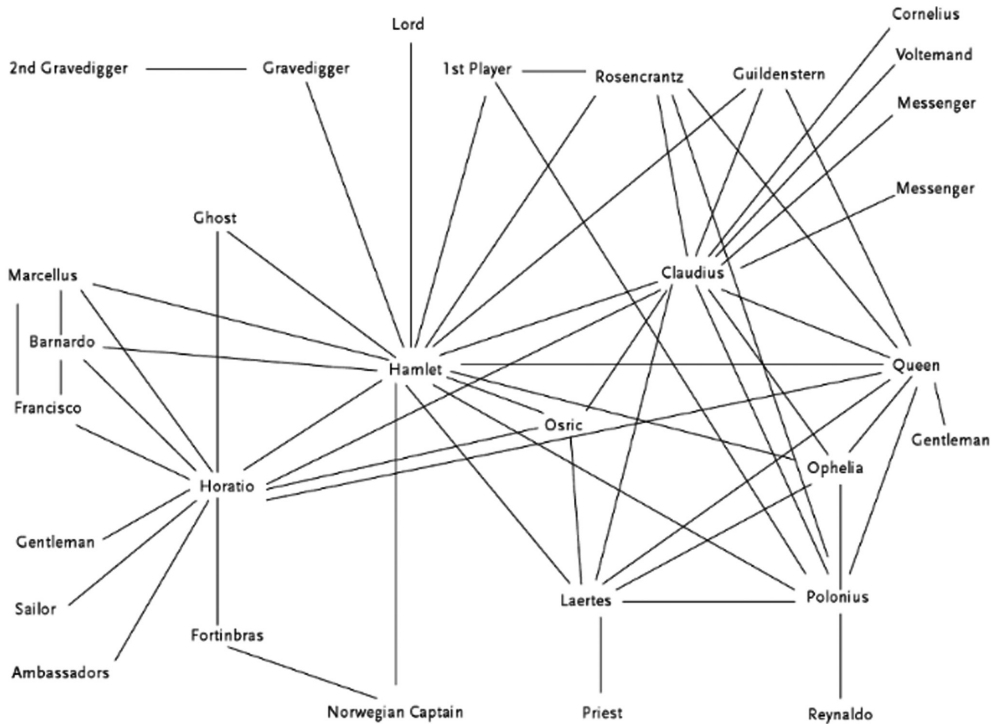


図1 『ハムレット』登場人物相関図

社会ネットワーク分析として知られるこのアプローチは、「作品中で言葉を交わした」など何らかの指標でもって人物間のつながりを定義し、そのつながりを集積した結果生成されるネットワーク構造を数学的に分析することにより、社会的関係性を把握しようとするものである<sup>46)</sup>。

このような視覚的アプローチは、テキストをデータに変換することで可能になる「読み方」のひとつである。これは、書物を手に取り、目で追い、線を引き、余白にメモを取るといった身体的な「読む」行為とは本質的に異なるものだ。ただし、「読む」行為の本質が異なれば、得られるものも異なる。遠読をはじめとする統計的テキスト分析は、テキストの精読を否定するものではなく、どの箇所あるいはどの人物を集中的に読むかを判断するための情報を提供するもの、あるいは経験的に蓄積・共有されてきた「読み」の是非を定量的に検証するものであると理解されたい<sup>47)</sup>。

### 課題と対応策

このようなテキストのデジタル化に伴う変化がもたらす課題として、物質性の変容と、コンテキストからの逸脱を検討したい。それぞれの課題と、その考えられる対応策について順に検討しよう。

45) Moretti, *op. cit.*, p. 213.

46) 歴史研究におけるネットワーク分析の先行研究リストとして、Historical Network Research のHP (<http://historicalnetworkresearch.org/bibliography/>)などを参照。

47) Arguing with Digital History Working Group, *op. cit.*, pp. 13-16.

## 物質性の変容

テキストのデジタル化がもたらした変化は、本に対する認識と行動の変化である。ケヴィン・ケリーによれば、紙の本を手にとって読む場合と、デジタル・テキストをスクリーンで読む場合とでは、我々は異なる知の体系にアクセスしている。紙の本は、物理的な制約から、世界のある部分を切り取った姿を読者に提供しているのだから、註記された情報をほかとリンクさせることを読者に委ねている。一方、ウェブにつながったスクリーンで読むデジタル・テキストは、そこに記された単語や概念について、ほかで書かれた関連情報と即座にリンクさせることができるので、読者はウェブ上に広がる知的資源を共時的に参照しながら読み進めることができる<sup>48)</sup>。たとえば Wikipedia の記事を見てみれば、知識が相互にリンクする様子が明らかである。Wikipedia は情報の信頼性についての懸念はあるが、誰もが知識を容易に発信・訂正・追補・共有できる集合知のプラットフォームである点で<sup>49)</sup>、紙の本と比較した場合のデジタル・テキストの特徴を端的に表している。

このように、史料をデジタル化してそこに記されたテキストをウェブとリンクさせながら読解を進めることもできるが、それで史料分析は事足りりとするのは片手落ちである。モノとしての史料の特性が失われてしまうことに自覚的でなければならない。エレイン・トレハーンは、中世写本のデジタル化事業に携わった経験から、史料のデジタル化の欠点を考察した。彼女は、重さ・寸法・紙やインクの材質・装飾の立体感といった物質的情報を、史料を扱う人間の身体との関係で把握できないことを問題視した。すなわち、持ち運べたかどうか／どのような体勢だと読みやすいように設計されていたのか、といった物理的な情報は、原物と対峙してはじめて体感できるものであって、撮影された史料の写真とそれを基に翻刻した電子テキストを平面的なスクリーン上で閲覧したとしても、現実世界の中で史料と人間がどのように関わり合っていたのかを知ることはならないと批判したのである<sup>50)</sup>。

一方で、このような批判に対するデジタル技術の応答例もある。ドット・ポーターは、VisColl というソフトウェアを開発し、中世写本の冊子体の綴じ方についてのデータを記述し、綴じの様子をスクリーン上で視覚的に把握できるようにした。一般に史資料のデジタル化プロジェクトにおいては、史料画像の撮影や書誌情報の記述、テキストの電子化に関心が集まりがちだが、VisColl の活用によりモノとしての手稿本への眼差しが養われることが期待される<sup>51)</sup>。これはあくまでスクリーン上でのデジタル再現の試みだが、文化遺産を原寸大で複製するプロジェクトもいくつか見られる。たとえば、2015年にイスラム系過激派組織のテロによって破壊されたシリアのパルミラ遺跡の 3D 複製や<sup>52)</sup>、凸版印刷が手がけた国宝「洛中洛外図屏風（舟木本）」な

48) ケヴィン・ケリー（服部桂訳）『〈インターネット〉の次に来るもの ― 未来を決める12の法則 ―』NHK 出版、2016年、4章「SCREENING」。

49) 日下九八「ウィキペディア ― その信頼性と社会的役割 ―」『情報管理』55(1)、2 - 12頁。

50) Treharne, E., 'Fleshing out the Text: The Transcendent Manuscript in the Digital Age', *Postmedieval: A Journal of Medieval Cultural Studies*, 4(4), 2013, pp. 465-478.

51) Porter, D. et al., 'VisColl: A New Collation Tool for Manuscript Studies', in H. Busch et al., eds., *Codicology and Palaeography in the Digital Age 4*, Norderstedt, 2017, pp. 81-100.

52) Thompson, E. L., 'Legal and Ethical Considerations for Digital Recreations of Cultural Heritage', *Chapman Law Review*, 20(1), 2017, pp. 153-176.

どの高精細デジタル複製の展示などである<sup>53)</sup>。文化財のデジタル複製の懸念と可能性はすでに美術史家により議論されているが<sup>54)</sup>、文字史料を扱う歴史家は、仮に原物と見間違えるような、重さと手触り、質感を兼ね備えた史料の複製をいつか手にした時、果たして何を感じ、どのような行動を取るだろうか。

「読む」ことを想定した文字とテキストは、歴史の記録媒体として今後も残るだろう。しかし、3D デジタル複製などの登場によって、歴史は語り手による叙述とそれを受け取る側の「読み」によって再構成されるだけでなく、人間の五感を使って没入的に「体感する」ものにもなってきた<sup>55)</sup>。もちろん、これは歴史家の語り、すなわち史料を選別し、議論の枠組みを定め、歴史の断面図を活写する営みを否定しない。情報伝達媒体の多様化に伴って人間同士のコミュニケーション方法が多様化し、情報を得るための「読む」行為そのものも多様化しているのだ。時空をこえて過去の声に耳を傾ける歴史家も、その影響を受ける者として例外ではないだろう。

### コンテキストからの逸脱

改めて「読み方の変化」に立ち戻って検討したい二つ目の課題は、テキストのコンテキストからの逸脱である。ここで言うコンテキストとは、前章で検討したアーカイブズのコンテキストではなく、史料に残された文字情報としてのテキストの内容について、そこで使われている単語が「いつ／どこで／誰によって」発されたのかを示す情報である。前述の統計的テキスト分析の特徴は、ひとりの人間が精読できないほど大量のテキストを対象に、語の使用傾向や人物同士の関係性についての数量的事実を、視覚化データの力も借りて把握することであった。しかし、もう一步深くテキストを解釈する場合、まずは「いつ／どこで／誰によって」発された言葉なのかを検討するのが良いだろう。

では、大量のテキスト群の概要を数量的に掴みながら、テキストの背景にあるコンテキストも加味して分析するにはどうしたら良いか。その対策のひとつは、テキストを単なる文字列の集合体としてではなく、それぞれ意味を持った単語や文のまとまりとしてデータ化することである。たとえば、テキスト中に new という単語があったとする。これをアルファベット3文字の連なりとしてではなく、ある段落では形容詞として、ほかの場所では New York など地名の一部として、あるいは New Deal など政策名の一部として記述されたものであるというように、単語がどのような意味を持っているかをコンピュータが解釈できるようにデータ化すれば良い。このように、電子テキストに意味を付与してデータ化することをマークアップ／符号化／タグ付けと呼ぶ<sup>56)</sup>。

このように、解釈可能なデータとして人文学のテキストを記述し、かつ関心を同じくする研究者の間ではある程度共通のルールを用いてデータを共有・交換しようとする動きは、すでに1980年代末から欧米で見られ、Text Encoding Initiative (TEI) と呼ばれる国際的な枠組みとして定

53) トップパン VR・デジタルアーカイブ、<http://www.toppan-vr.jp/bunka/>

54) 高岸輝「デジタル技術と文化財の複製」『人文情報学月報』65【前編】、2016年。

55) Cohen, D. J. et al., op. cit., pp. 454, 467-479; 川村佳男ほか「歴史資料を展示する博物館の未来」歴史学研究会編『歴史を社会に活かす ― 楽しむ・学ぶ・伝える・観る ―』東京大学出版会、2017年、245-256頁。

56) Coombs, J. H. et al., 'Markup Systems and the Future of Scholarly Text Processing', *Communications of the ACM*, 30(11), 1987, pp. 933-947.

着している<sup>57)</sup>。TEIは、詩・文学作品・戯曲・写本・辞書・表・楽譜といったさまざまな形式のテキストを対象に、単なる文字列の集合としてではなく、人文学的解釈を付与した精緻なデータとして記述することができる。すでに欧米では、史料への註釈や版ごとの異同をひとところに示す批判校訂版を電子テキストとして提供するには、TEIに準拠することが事実上の標準的実践となっており<sup>58)</sup>、西洋史研究との関連では Perseus Digital Library、Regesta Imperii Online、Old Bailey Online などの著名なデータベースが TEI マークアップを採用している。一般に、史資料のテキスト本文を TEI でマークアップする際に重視されるのは、暦や日付といった時間の記述、人物・地名・出来事といった固有表現の記述である<sup>59)</sup>。つまり、歴史研究の文脈における TEI マークアップは、コンテキスト情報を取り出せるようにすることに重きが置かれているので、速読に見られるようなテキスト分析よりも、対象を絞った比較的精緻な分析を可能にしているのである。

TEI ファイルを用いてコンテキスト情報を重視したテキスト分析の事例として、アメリカのカールトン・カレッジで2015年冬に開講された DH の入門講座 Hacking the Humanities を紹介したい。ここで例示されているのは、Folger Shakespeare Library が提供する TEI ファイルを用いて、シェイクスピア『ロミオとジュリエット』における登場人物ごとの発話内容を分析するレッスンである<sup>60)</sup>。口絵1と2は、同レッスンに沿って出力した結果で、それぞれロミオとジュリエットがどのような発話をしたのかを視覚的に把握しようとするものである。

この図は、ワードクラウドと呼ばれる視覚化データで、出現回数の高い単語が中心付近に大きく表示されるようになっており<sup>61)</sup>、同レッスンにしたがえば、ロミオは愛と死を、ジュリエットはロミオへの呼びかけと愛を伝える発話構成になっていることが窺えるようである。

もとよりシェイクスピアの作品世界を云々することは筆者の力の及ぶところではないし、分析の上澄みを紹介したにすぎないが、このようなアプローチは文学作品だけでなく歴史史料のデジタル・テキストにも適用可能である。たとえば、議会議事録や裁判記録のテキスト・データベースが似たような形式のデータとして保存されていれば<sup>62)</sup>、発言や証言の内容について「いつ／誰が／どのような言葉と共に」発していたのかを分析することができる。つまり、大量のテキストを対象に広範な情報処理を行う統計的テキスト分析の利点を活かしながら、コンテキスト情報を

57) TEI のまとまった解説として、永崎研宣「歴史データのさまざまな応用 — Text Encoding Initiative の現在 —」後藤・橋本編、前掲書、131 - 154頁。

58) 日本をはじめとする東アジア圏の史料を TEI でマークアップする研究として、後藤真ほか「構造化記述されたテキストの基盤整備に向けて — 延喜式の TEI マークアップを事例に —」『じんもんこん 2018論文集』2018年、243 - 248頁。

59) Vogeler, G., "The "Assertive Edition": On the Consequences of Digital Methods in Scholarly Editing for Historians", *International Journal of Digital Humanities*, 1(2), 2019, p. 314.

60) Mason, A., "Text Analysis and the TEI", *Hacking the Humanities*, 2015.

61) この図を出力した統計的テキスト分析のための Voyant-tools (<https://voyant-tools.org/>) については、Sinclair, S. & G. Rockwell, "Text Analysis and Visualization: Making Meaning Count", in S. Schreibman, et al., eds., *A New Companion to Digital Humanities*, 2nd ed., Chichester, 2016, pp. 274-290.

62) Old Bailey Online からは裁判記録の TEI マークアップファイルを入手できるし (<https://www.oldbaileyonline.org/static/Data.jsp>)、19世紀初頭から現在にかけてのイギリス貴族院・庶民院における議事録をまとめた Hansard Archive でも、TEI ではないが、コンピュータで処理しやすく構造化された XML 形式のファイルを入手できる (<http://www.hansard-archive.parliament.uk/>)。

重視した比較的粒度の高い分析が可能なのである。改善点として、それぞれの単語が否定的な文脈で使われていないかどうか／どのような助動詞を伴ってどのようなニュアンスを表しているか、といったより精緻な分析ができる方が良い。このような観点での開発は、主に自然言語処理系の研究者に委ねられており<sup>63)</sup>、今後も分析ツールの精度や完成度は上がるだろう。歴史家としては、自らの研究プロセスや分析のあり方を顧みながら一般化し、他分野の研究者との批判的かつ建設的な対話に臨むことで、史資料を題材とした情報処理技術の発展、ひいては歴史研究への新しいアプローチの導入に寄与できるだろう<sup>64)</sup>。

このようなテキスト分析の改善可能性を念頭に置くことに加えて、デジタル技術を利用した分析であるからと言って、真に科学的客観性が保証された手法に則っているなどという幻想は抱かないようにすべきである。なぜなら、データ／ツール／解析プログラム／視覚化モデル、といった分析の諸段階で立ち現れてくるあらゆる取捨選択の局面において、分析者の主観的判断が介在するからである<sup>65)</sup>。しかし、史料の選別／解釈の枠組み／ナラティブの構成に関しても、歴史家の判断は介在する。言語論的転回がもたらしたひとつの教訓が、歴史を再構成する者ひとりひとりに内在する語り手としてのスタイルに自覚的になることだったとすれば<sup>66)</sup>、デジタル・ヒストリーの実践がもたらす教訓のひとつは、歴史データを扱う分析プロセスひとつひとつに潜むバイアスに自覚的になることだろう。

## 結 論

本稿では、研究者をとりまく環境、史料の収集、史料の読解それぞれにデジタル技術がもたらす影響を検討してきた。1章では、外国史研究者が議論の相手とする研究者の間で、デジタル・ヒストリーの実践が確実に浸透し始めていることが確認できた。続く2章、3章では、史料の外側と内側にあるコンテキスト、および史料が帯びる物質性がデジタル化によって不可視化されてしまう課題への対応として、アーカイブズ学への接近 (Monasterium.net)、デジタル技術による物質性の再現 (VisColl や 3D 複製)、人文学的解釈を付与した電子テキストの作成 (TEI) を論じた。

歴史研究に従事する者に要求される仕事の質は、デジタル時代においても大きな変更をこうむることはない。また、一人ひとりの作業場にもデジタル技術の影響が及んでいるとはいえ、本稿で歴史家全員がプログラミングなどのコンピュータ技術を身につけるべきだと主張するつもりはない。我々に求められるのは、デジタル技術が利便性とともに関々のバイアスを生むことを自覚し、改めて歴史学の作法や成果発信方法をかえりみることなのである。

そのような状況のもとでは、本稿で扱いきれなかった諸問題もまた議論される必要がある。アウトリーチ活動<sup>67)</sup>、若手研究者のキャリア形成<sup>68)</sup>など相互に関連する論点は数多くあるが、人文

63) たとえば、金山博ほか「日本語 Universal Dependencies の試案」『言語処理学会第21回年次大会発表論文集』2015年。

64) このような対話のあり方として、西村陽子・北本朝展「デジタル史料批判と歴史学における新発見」『人工知能』31(6)、2016年、769-774頁などを参照。

65) Boyd, D. & K. Crawford, 'Critical Questions for Big Data', *Information, Communication & Society*, 15(5), 2012, pp. 662-679.

66) 二宮宏之「歴史の作法」上村忠男ほか編『歴史を問う (4) 歴史はいかに書かれるか』岩波書店、2004年、33-34頁。

学における共著論文の査読・評価方法もとりわけ重要な問題である。AHA が発表したデジタル研究の評価に関するガイドラインには、歴史学科は共著論文をはじめとする共同作業の成果を評価するための規約を改善すべき、との項目が盛り込まれた<sup>69)</sup>。共著で論文を書く例がきわめて少ない日本の歴史学界でも<sup>70)</sup>、今後この問題に向き合う必要があるだろう<sup>71)</sup>。

もう一点、デジタル・ヒストリーを史学史に位置づける可能性も指摘しておきたい。数量分析と親和的なデジタル・ヒストリーは、大量のデータを統計的に分析して長期的時間軸のなかでの変化を跡づけようとした「系の歴史学」の系譜を継ぐものとみなすこともできると思われる<sup>72)</sup>。ただし、デジタル・ヒストリーは、これまで歴史学が邂逅してきた方法論のなかでも、きわめてグローバルな現象という点で特異でもあり、しかし各国の学問的伝統に各々少しずつ異なるかたちで接ぎ木されようとしていることにも注目すべきであろう。

最後に、デジタル・ヒストリーと研究者の心理面の相関に言及し、稿を閉じたい。デジタル技術は、専門や母語を異にするアカデミア内外の人々との対話を促す共通言語でもある。歴史家がデジタル技術を学ぼうとすれば、必然的に歴史学以外の専門家から学ぶ機会が増えることになる。このように、交流する人々が変化する、つまり「依存先が増える」ことによる精神面の変化は、研究者の日常により影響を与えるのではないだろうか<sup>73)</sup>。

(アーヘン工科大学・東京大学大学院／東京大学大学院)

67) 歴史家によるアウトリーチ活動の一例として、大谷哲「インターネット歴史学コンテンツと社会への発信——「せんだい歴史学カフェ」の活動より——」『史苑』79(2)、2019年、164-172頁。

68) 博士号取得後の若手研究者のキャリアを考えるための言葉として、alt-ac (alternative academic careers) という DH 研究者による造語が広まった。ノースカロライナ大学チャペルヒル校図書館の HP (<https://guides.lib.unc.edu/c.php?g=216166&p=1623647>) では、人文社会系分野における alt-ac 関連の情報を発信している。

69) AHA, 'Guideline for the Professional Evaluation of Digital Scholarship by Historians', 2015 (菊池信彦ほか訳「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、2016年)。

70) 小風は、『史学雑誌』123(9)、124(1)、124(5)に掲載された巻末文献目録(西洋史)を対象として、研究文献の単著・共著の別を分析したことがある。その結果、単著の割合は98.5%(1304/1323)であった。小風尚樹「歴史学系院生によるオープン・コミュニティの構築——Tokyo Digital Historyの挑戦——」Japan Open Science Summit 2018(2018年6月18日)発表資料。

71) 人文学の研究成果の評価方法をめぐる議論については、後藤真「研究の量的評価は人文学に対して可能なか——人間文化研究機構の試み——」『学術の動向』23(10)、2018年、42-49頁。

72) 長野壮一「海外 DH 特集：フランス歴史学における DH の伝統」『人文情報学月報』58【後編】、2016年；Mullen, L., 'A Braided Narrative for Digital History', in M. K. Gold & L. F. Klein, eds., *Debates in the Digital Humanities 2019*, Minneapolis, 2019, pp. 606-617.

73) 当事者研究で知られる熊谷晋一郎は、「自立とは、依存先を増やすこと」と主張している。当事者研究については、熊谷晋一郎編『みんなの当事者研究』(『臨床心理学』増刊第9号)金剛出版、2017年。