

人文学においてデジタル技術はどう活用されてきたか： CH 研究会研究報告を手がかりとして

永崎研宣^{†1}

CH 研究会のこれまでの研究報告に関するサーベイをもとに、我が国の人文学においてデジタル技術がどのように活用されてきたかを報告する。特に、人文学のなかでも、どのような分野において、どのような技術や手法が適用されてきたか、ということについても目配りしつつ、海外の同種の動向とも適宜対比しながら検討する。

How Digital Technologies Have Been Used: Through the History of the SIG Computers and the Humanities

Kiyonori Nagasaki^{†1}

Through the 827 presentations of the SIG Computers and the Humanities, this presentation describes the way of adoption of digital technologies in the humanities from 1989 onward. Especially, it focuses on what fields have been treated in the humanities and what digital methods have been adopted, comparing with the similar activities in other countries.

1. はじめに

人文学におけるデジタル技術の活用とは、情報処理学会人文科学とコンピュータ研究会（以下、CH 研究会）のテーマそのものである。CH 研究会は、コンピュータの技術的発展と普及が人文学において大きな変革をもたらすだろうという見通しの下ではじめられたものであった[1]。当初は計算処理による自動化や効率化に関する様々な試みが行われる一方で、紙に代るメディアとしてのデジタル技術というあり方についても理論としての様々な検討はあったものの、その圧倒的な有用性について具体的に議論するには、Web の普及や大規模ストレージの発達を待たねばならなかった。21 世紀に入って数年、ようやく、デジタルメディアはメディアとしての現実的な利用に耐え得るものとして一般化し、実際に広く社会的役割を担うに至った。その結果、その情報を保管し探し引き出す量・正確性・速度は、紙メディアを相対化するに十分な存在であると広く認識されるに至り、具体的な議論と実践が可能となったのである。

2. 科研費新規採択課題からみえる

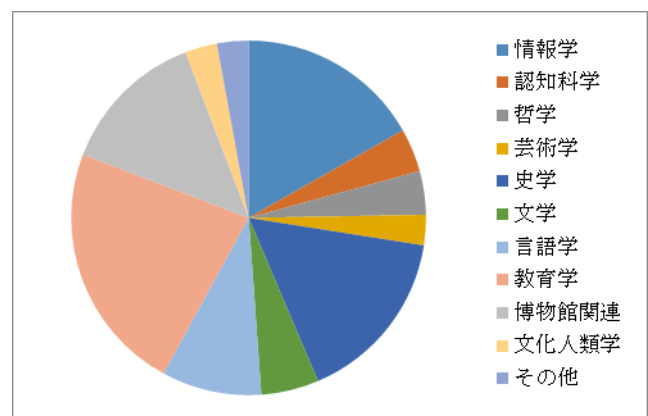
そして、今や、人文学においてもデジタル技術の活用は、大きく広がりつつある。たとえば、科研費データベース[a]を用いて、この数年の科研費の配分状況において、新規採択課題の課題名から、CH 研究会に関わり得る（＝人文学とデジタル技術の両方に関わることが課題名からして明らか）研究課題を集計してみたところ、表 1 のような結果

となった。

年	件数	総額
2009	99	251,862,000
2010	126	385,300,000
2011	172	505,390,000
2012	203	654,160,000

（表 1：科研費新規採択課題の推移）

すでにじんもんこん 2010 において筆者らが発表したもの[2]の続きということになるが、引き続き、件数・総額ともに着実に増えていることは注目に値するだろう。



（グラフ 1：2012 年度の関連新規採択課題の分野）

また、内訳を見てみると、2012 年には人文系の分野で採択された課題の総額が情報系を上回っている（グラフ 1）。特に、人文学分野における審査を通じて採択されたものが半分以上を超えたということは人文学におけるデジタル化への態度の明らかな変化をみて取ることができる。もちろん、こ

^{†1} 一般財団法人人文情報学研究所
International Institute for Digital Humanities

a) <http://kaken.nii.ac.jp/>

れだけでなく、課題名からはうかがえないが研究内容としてデータベース化やデジタル技術の活用を挙げている研究課題も少なくないことにも留意しておく必要があるだろう。

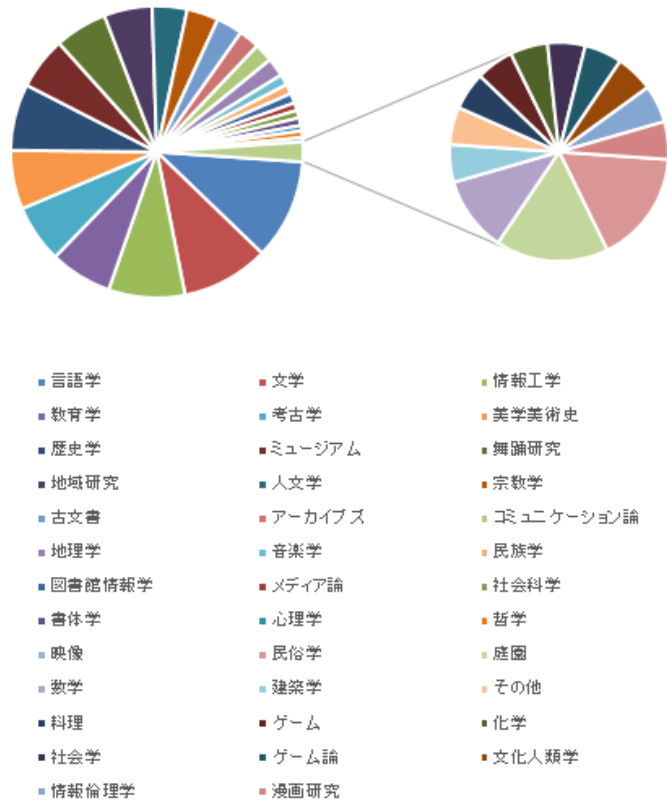
3. CH 研究会研究報告を手掛かりとして

ここに至るまでに CH 研究会の果たしてきた役割が決して小さなものではなかったことは、1989 年の研究会発足以来、毎年 4 回の研究会を継続的に開催し、総数 827 件という累計発表数からしても明らかである。では、CH 研究会では、これまでどのような研究がなされてきたのか。これについてはすでに最初の 10 年間のサーベイが公開されており[3]、第 81 回研究会においても同様にサーベイが発表されている[4]。さらに、筆者自身も、「デジタル・アーカイブ」という観点から CH 研究会の流れを振り返り、それが技術的制約を強く受けつつもその変化し拡大しつつある枠の中で常に発展し続けてきているということを明らかにした[5]。本発表では、これらの流れを踏まえつつ 2012 年度までの研究発表について検討するとともに、他のいくつかの要素も併せて、今後のこの種の研究の方向性についての検討材料を提供したい。なお、827 件にのぼる研究報告は、発表に際しての査読がないためにやや玉石混淆の感否めないが、それぞれの時期の状況を把握する材料の一つとしては一定の有益さがあると考えられる。

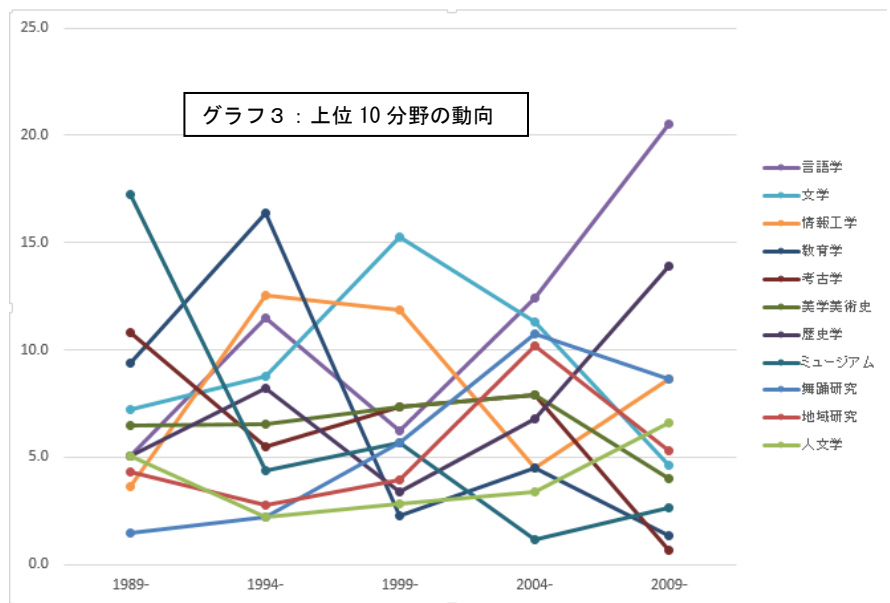
3.1 過去の研究発表を対象分野からみている

この種の研究の多くはある程度まで類型化することが可能である。たとえば以下の通りである。「人文科学(のある分野)についてこのようなデジタル技術を適用((してこのような結果を得られた)|(した|する)|(ことにはこのような意義がある)|するべきである)。」また、査読なしの研究会であることから、特に対象としての人文科学ということ強く意識せずに技術的研究を提示してみるというものも散見される。そのような事情を踏まえつつ、とりあえず、「どの分野に対して」という観点でこれまでの研究発表を分類してみるとグラフ 2 のようになった。さらに、上位 10 分野に限ってその動向を見てみたものがグラフ 3 である。ただし、この分類の仕方はあくまでも筆者

グラフ2: 対象分野の割合



グラフ3: 上位10分野の動向



の主観によるものであり、若干正確性を欠いているところもあるかもしれない。また、文字（特に漢字）の研究に関してはほとんどを「言語学」に分類したことをお断りしておく。

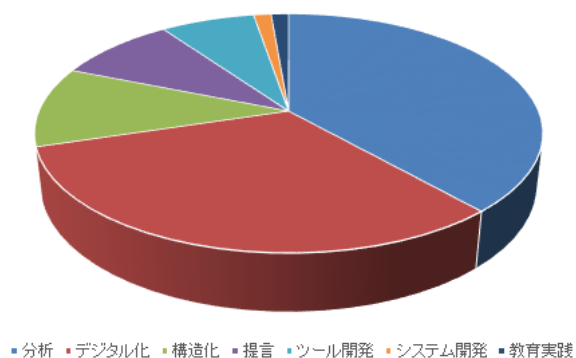
グラフ 2 によれば、期間の全体を通してでは言語学がトップ、文学が 2 位、となっているが、それもそれぞれ 1 割程度であり、全体として様々な分野がまんべんなく扱われていることがわかる。

さらに、グラフ 3 をみても、当初はミュージアムに関わる発表が多かったが徐々に落ち着き、言語学と歴史学

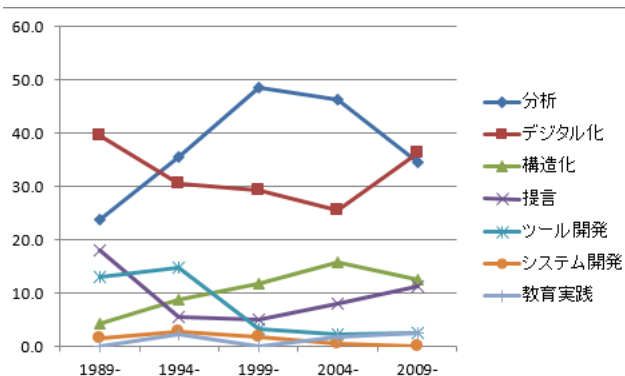
が特に近年大きな伸びを示していることがわかる。また、ここでは細かく分類していないが、英語文献を扱う研究発表が少ないという指摘[4]は筆者も同意するところであり、また、研究者の人数に比べて国語学からの研究発表が少ない、といったこともあるが、それらはいずれも、他に学会があるため、敢えて CH 研究会まで来て発表をする必要がないという面もある。英語文献に関しては、英語コーパス学会に加えて英語圏の学会で発表をするという選択肢も有力であり、実際のところ、欧州や米国の関係学会にて日本人の名を目にすることは少なくない。国語学に関しては、計量国語学会が 1950 年代より活動を続けている。ミュージアム関係についてもアート・ドキュメンテーション学会が活発に活動している。いずれもそれほど大きな人数ではないため、個人的な交流のレベルでほぼ解消してしまうとは言え、連携することによるメリットも場合によってはあるかもしれない。CH 研究会の側からは、特に技術面における情報提供等の形での貢献が可能であると思われる。

3.2 過去の研究発表を対象分野からみてる

では次に、手法についてみてみよう。グラフ 4 は 827 件全体を分類したもので、グラフ 5 はそれを 5 年ごとに分けて推移を表しているものである。



(グラフ4: 手法別分類)



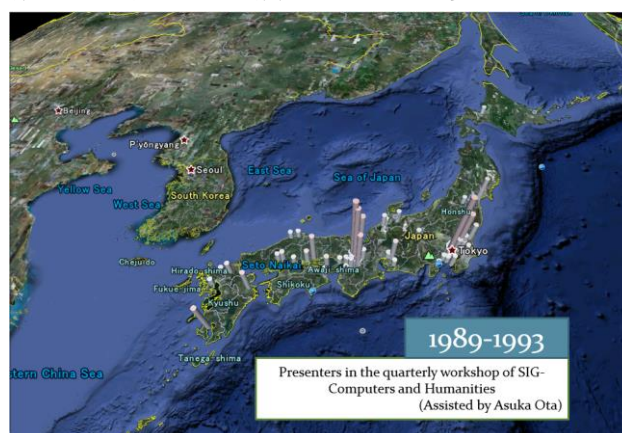
(グラフ5: 各手法の推移)

手法についてはかなり大ざっぱなくくりでまとめている。「分析」にはテキストの分析も画像の分析も入っており、「デジタル化」には Web データベースの構築のようなもの

も含まれている。「構造化」は、デジタル化にあたっての情報の構造について議論しているものを含んでいる。この分け方では「分析」が全体としては最大となっているが、「デジタル化」もかなり近接しており、2009-2012 年では「デジタル化」の方が多くなっている。これは特に 2012 年にデジタル化に関わる発表が多かったことが強く影響している。このことは、冒頭に述べたように、近年、デジタルメディアの有用性がきわめて高まり、この研究会のみならず、国会図書館・近代デジタルライブラリーや青空文庫の隆盛に端的に見られるように、社会全体として文化資料をデジタル化して共有しようとする大きな流れができつつあることが深く関係しているだろう [b]。もちろん、一般に比べると人文学における利活用というのはパイが小さいように見えるが、人文学の役割の一つとして、文化 (資料) の価値や意義を一般市民につなぐ役割、あるいはつなぐための礎を築き上げていくというものがある。それを考えるなら、人文学において利活用されることは、単に研究者が活用するというだけにとどまらず、それを通じてその文化 (資料) の価値や意義をより広く社会に浸透させていくことも十分に期待できるだろう。その方向性での研究としては、利活用の担い手や受益者についてよく検討しつつ、用途に応じて情報を取捨選択できるインターフェイスを用意するなど、実践面からも様々なアプローチが可能だろう [7]。

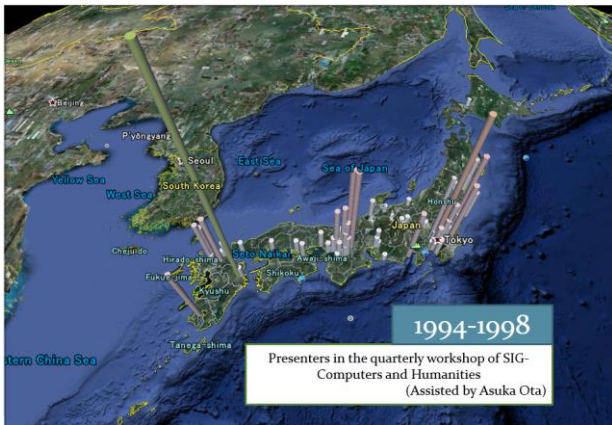
4. 地域分布

CH 研究会では、初期の段階から、全国各地での研究の掘り起こしを掲げ、全都道府県での研究会開催を目指して活動してきたという経緯があった。これは 2006 年中に達成され、2007 年 1 月には記念シンポジウムが開催された [c]。以下の図は、2011 年までのものだが、発表者の当時の所属研究機関を 5 年ごとに地図上にプロットしたものである。棒グラフの高さは同一機関からの発表の多さを表している。

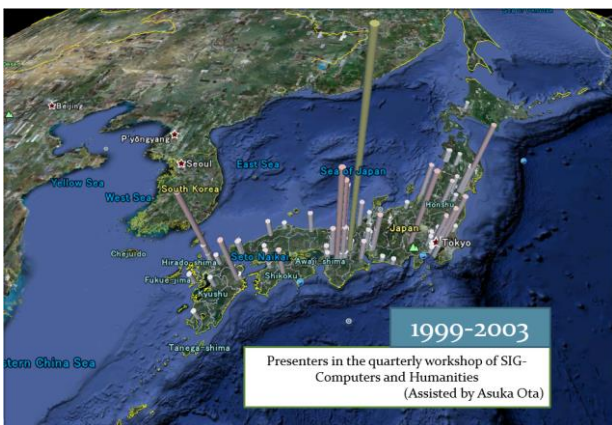


(図 1 : 1989-1993 年発表者分布)

b) たとえば、勉誠出版をはじめとして様々な出版社からデジタル文化資料に関する書籍が出版され [6]、関連する会合も盛んに行われている。
 c) <http://jinmoncom.jp/prev/2007.01/>



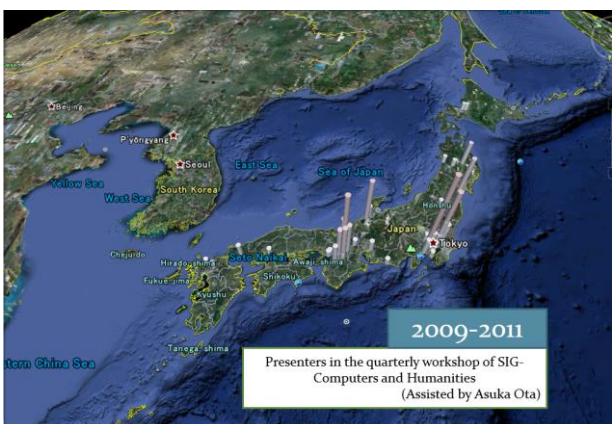
(図 2 : 1994-1998 年発表者分布)



(図 3 : 1999-2003 年発表者分布)



(図 4 : 2004-2008 年発表者分布)



(図 5 : 2009-2011 年発表者分布)

CH 研究会による全国行脚の成果が如実に表れていると言えるが、一方で、急激に発表者が増えた後にいなくなってしまう組織・地域も少なくない。これも今後の課題として考えていく必要があるだろう。

5. 海外の動向との関連で

一方、海外の動向に関しては、1970 年代に設立された欧州の ALLC: European Association for Digital Humanities (当初の名称は Association for Literary and Linguistic Computing) (ALLC) [d] や米国の Association for Computers and the Humanities (ACH)[e]といった学会がこうした動きを担ってきており、文学と言語学の研究者が中心となってきた。さらに、Modern Language Association (MLA)[f]でも近年はデジタル化に関わる分科会を設けるようになっており、やはり、文学・言語学研究者が中心となって活発に議論が行われている。また、1980 年代にはテキスト資料デジタル化の共通フォーマットとして Text Encoding Initiative 協会[g]が設立され、上記の ALLC, ACH 等が中心となってガイドラインの制定に取り組んできた。これは現在も改訂が続けられており、XML を利用した基本フォーマットは多種多様な人文資料に研究用途としての利用に耐えるものとして広く利用されるようになっている。また、この種の分野は長らく Humanities Computing と呼ばれてきたが 21 世紀に入ってから Digital Humanities (以下、DH) と名前を変えることでさらに大きな広がりを見せるようになり、欧米、そしてすぐにカナダも交えた DH の世界連合、Association of Digital Humanities Organizations (ADHO)[h]が設立され、現在はオーストラリア[i]、日本の DH 学会[j]も加盟している。

財政面では、米国では人文学専門の助成団体 National Endowment for the Humanities (NEH) [k]があり、人文学の将来について専門的に検討することが組織的に行われている。この NEH が DH を重要な戦略の一つと位置づけたことにより DH をターゲットとした助成金や DH を媒介とした国際的な共同研究を助成する枠組みなどが提供されるようになってきている[l]。これにより、特に米国では DH が大いに盛んになりつつある。カナダでも、人文社会科学向けの助成団体 [m]があり、DH に徐々に力を入れつつある。

また、Public Humanities という流れも出てきておりそれが DH と連携し、一般市民と人文学がどういう関わりを持つかということについて実践を伴う研究プロジェクトも行

- d) <http://www.allc.org/>
- e) <http://www.ach.org/>
- f) <http://www.mla.org/>
- g) <http://www.tei-c.org/>
- h) <http://www.adho.org/>
- i) <http://www.aa-dh.org/>
- j) <http://www.jadh.org/>
- k) <http://www.neh.gov/>
- l) <http://www.neh.gov/divisions/odh>
- m) <http://www.sshrc-crsh.gc.ca/>

われるようになりつつある。とりわけ、功利主義者 Jeremy Bentham の手稿を Web 上で共同翻刻するプロジェクト Transcribe Bentham プロジェクトは、英語ということもあり、世界中から数百人の参加者を集め、それまでは極めて困難であり時間がかかっていた翻刻の作業のスピードを一気に速め、Bentham 研究の進展[n]に大きく貢献するとともに、手稿を所蔵する University College London (UCL)の図書館におけるコンテンツの充実化、すなわち、手稿を全文検索できるサービスを提供できるようにするという点でも貢献しているとのことである[8]。また、この種のことは、特に英米では図書館情報学と DH の距離が近いことから、学術情報流通という観点からアプローチされることもある[o]。

DH 的な考え方や手法を広めるためのワークショップを世界各地で開催するという取り組みも ALLC を中心に行われた時期があり、日本デジタル・ヒューマニティーズ学会 (JADH) もこの流れから生まれてきたものであった。さらに近年は、1 週間程度のサマースクール的なセミナーを行うという流れも作られつつある。ヴィクトリア大学の Digital Humanities Summer Institute[p]、オックスフォード大学の Digital Humanities @Oxford Summer School[q] が先行する取り組みだったが、これにライプツィヒ大学(ドイツ)、ベルン大学(スイス)、メリーランド大学(米国)等が続き、世界中に広がりを見せつつある。また、一方で、欧米圏のみならず、その外に DH の連携の輪を広げていこうという試み、Global Outlook::Digital Humanities が主にカナダの DH 研究者のリードにより、ADHO の取り組みとして開始されている[r]。今のところは、スペイン語圏・中国語圏に特に注力されているようである。

もう一つ特筆すべき点を挙げておくと、特に欧米先進国において近年目立つアプローチとして、単にデジタル技術に応用してみせるにとどまらず、その手法を誰もが採れるようなツールを開発し、CC ライセンスや GPL ライセンスなどで公開するという動きが目立つようになってきている。Web 上で共同翻刻をするための Web CMS[s]、グーグルマップやオープンストリートマップ上に年表とともに情報をマッピングして時空間情報を提示できるようにする Web CMS[t]、映像も含めた様々なメディアを統合的に扱いつつ統計機能なども兼ね備えた Web アーカイビングシステム

n) Bentham は高校教科書にも出てくる程の著名な歴史的人物だが、その著作の出版においては適切な編集がなされず、Bentham 自身の思想が適切に表明されていないということが明らかとなっており、手稿に基づく Bentham の思想の再構成が UCL を中心に進められているが、手稿が膨大な上に読みづらいためになかなか研究が進まず、そこに Transcribe Bentham プロジェクトが強力な援軍として現れたとのことである。

o) なお、日本でも図書館情報学と人文学におけるデジタル技術の応用に関わる研究との連携については、北海道大学にて開催されたじんもんこん 2012 と九州大学にて開催された日本図書館情報学会がテレビ会議システムで中継を行うなど、端緒についたところであると言えるだろう。

p) <http://www.dhsi.org/>

q) <http://digital.humanities.ox.ac.uk/dhoxss/>

r) <http://www.globaloutlookdh.org/>

s) <http://scripto.org/>

t) <http://neatline.org/>

[u]など、かなり充実したものが公開されつつある。我が国では、デジタル技術を応用するところまでは多く発表されているものの、ここまで踏み込んだものはまだきわめて数少なく、今後のアプローチの一つの方向性として考慮していく価値はあるだろう。

なお、このような共有ツールを開発し公開していこうとする流れには、Methodological Commons[v]と呼ばれる DH の基本的概念の在り方が反映されているものとみてよいだろう。Methodological Commons は、人文学各分野が持つ明示的・暗黙的な方法論を、デジタル技術を媒介とすることで明らかにしつつ、それを Commons に提供することで、人文学各分野がそれらを共有し、よりよい研究成果を得られるようにしていこうとする考え方である。これは CH 研究会が今後目指していく在り方を考える上でもおおいに参考になると思われる。

6. 終わりに

本発表では、CH 研究会の研究報告を一つの手がかりとしてこの種の分野でどのような研究がなされてきたかということ概観してみた。もちろん、これまでの動向をより正確に把握するには、関連する各学会研究会における動向や人文学分野の学会研究会におけるデジタル化の動向、研究機関の動向などもきちんと調査しなければならないが、それは今後の課題としたい。また、今後の動向については、2011 年からメールマガジン『人文情報学月報』[w]が発刊され、この種の情報がある程度まとめて把握できるようになってきており、さらに、<http://www2.dhii.jp/dhm-nl/> にて、関連研究集会等のレポートを時空間マッピングによって閲覧できるようになる等、この種の動向の把握は徐々に容易になりつつある。しかしながら、そこからただ自分自身の研究に役立つものを探そうとするだけでは CH 研究会としての発展は限定的だろう。むしろ、単に自分が直接関わる分野にデジタル技術が応用可能かどうかというだけでなく、他分野においてどうかということについても目配りしていくことで、CH 研究会における活動がより大きな意味を持つてくることになるだろう。そのように考えたとき、繰り返しになるが、Methodological Commons という枠組みは、まさに CH 研究会の今後の取り組みの指針の一つになり得るのではないだろうか[x]

謝辞 本研究は、2012 年 3 月のオーストラリア DH 学会、中国人民大学、及びヴァージニア大学スカラーズ・ラボで

u) <http://shanti.virginia.edu/>

v) <http://www.allc.org/node/188> 及び <http://www.allc.org/node/189>

w) <http://www.dhii.jp/DHM/>

x) なお、本発表はこれまでの状況を概観し把握することが主な目的であるため、今後の具体的な展開については敢えてあまり言及していない。それについては、じんもんこん 2012 でのアンカンパレンスの報告[9]が特に参考になるだろう。

の発表を通じて検討されてきた内容を下敷きとしている。その過程で、国内外の多くの関連研究者に様々なご協力をいただいたことを感謝とともに記しておく。また、特に、CH 研究会の過去の発表者の組織の調査に当たっては、筑波大学大学院（当時）の太田あす香氏に全面的にご協力いただいたことを記しておきたい。なお、本研究の成果には、JSPS 科研費 30343429 の助成を受けたものが含まれている。

参考文献

- 1) 杉田繁治「人文科学とコンピュータ」『情報処理学会研究報告』(1989-CH-001), pp. 1-8.
- 2) 永崎研宣, 中村雄祐, 後藤真「人文学におけるデジタル化に関するオープンなメタ議論の意義—じんもんこん/Humanities Computing/Digital Humanities の将来に向けて—」『人文科学とコンピュータシンポジウム論文集』(社) 情報処理学会(2010年12月), pp. 1-6.
- 3) 山田奨治「人文科学とコンピュータの10年」『情報処理学会研究報告』(1998-CH-42), pp. 1-6.
- 4) 後藤真「人文科学とコンピュータ研究会—80回の歴史と課題」『情報処理学会研究報告』2009-CH-81, pp. 47-52.
- 5) 永崎研宣「デジタルアーカイブの弁証法」『情報処理学会研究報告』2005-CH-68, pp. 17-24.
- 6) NPO 知的資源イニシアティブ編『デジタル文化資源の活用：地域の記憶とアーカイブ』(2011), 勉誠出版.
- 7) 永崎研宣「人文科学のためのデジタル・アーカイブにおけるステイクホルダー —仏教文献デジタル・アーカイブを手掛かりとして—」『人文科学とコンピュータシンポジウム論文集』(社) 情報処理学会(2007年12月), pp. 347-354.
- 8) Tim Caser, Justin Tonra and Valerie Wallace, Transcription maximized; expense minimized? Crowdsourcing and editing *The Collected Works of Jeremy Bentham, Lit Linguist Computing* (2012), 27(2): 119-137.
- 9) 関野樹「「人文科学とコンピュータ」をとりまく状況と将来展望」『情報処理学会研究報告』(2013-CH-97(8)), pp. 1-4.