

「人文科学とコンピュータ」の発展にむけた課題 — 境界を広げる基礎教育の必要性 —

原正一郎^{†1}

概要: 情報処理学会「人文科学とコンピュータ研究会」(じんもんこん)は発足からまもなく30年が経過する。その間に達成できた成果は多く、フロンティア領域としては成功した研究会の一つである。しかし、次の10年間の「人文科学とコンピュータ」を考えると、既存の殻をやぶる努力と方策が必要であると感じている。

キーワード: 人文科学とコンピュータ, デジタルヒューマニティーズ, 教養教育

Problems for the Development of “Humanities Computing” -The Importance of Liberal Arts to Extend the Border-

Shoichiro HARA^{†1}

Abstract: The “IPSI SIG Computers and the Humanities” (SIG-CH) has achieved much and fruitful results during its 30 years activities. The SIG-CH is one of successful SIGs in the Frontier Research Groups. However, when we think about the next 10-year “Computers and the Humanities,” more efforts and strategies would be necessary to make a new breakthrough.

Keywords: Computers and the Humanities, Humanities Computing, Digital Humanities, Liberal Arts

1. そこまで進んだ「人文科学とコンピュータ」

異分野融合研究を推進するメリットは、他分野の手法を導入して自分分野の研究を促進する場合か、新しい分野を創出する場合である。

「人文科学とコンピュータ」における人文学分野の場合、情報学の手法を導入することにより、テキストのデジタル翻刻、画像などのデジタル化、データベース構築、デジタルデータの交換・共有、可視化、テキストマイニングなどにおいて著しい前進が見られた。この分野に足を踏み込んだ30年ほど前、施設内LANを敷設して電子メールを使えるようにした際には「私はそんなものは使わないし、無くとも研究はできる」とか、資料整理や資料交換のためにスプレッドシートやワープロの利用を進めた際には「1万字程度で研究ができるか、何のためにデータ交換をするのだ、不要だ」とか言われた経験を持つ者にとって、この変化は隔世の感がある。

情報学分野においては、人文学研究者より提供される多様なデータと取り組むことにより、データ処理に関わる多くの技術において着実な発展を遂げることができた。自分のモデルに当てはまらないデータは無視するか、モデルに当てはまるように都合良く変えてしまいがちな情報学研究者にとって、人文学分野のデータは「異世界」あるいは「目から鱗」であり、「柔らかなサイエンス」の前進にとって、

よい訓練になったのではないかと思う。以上は肯定的な評価であるが、否定的な側面も存在する。

2. それでも進まない「人文科学とコンピュータ」

30年前に比べると隔世の感があると言ったものの、それはデジタル化の部分である。図1は研究プロセスをモデル化したものである。人文学分野における画像のデジタル化やテキストのデジタル翻刻などは、図の最初の「データ収集」に相当する。「人文科学とコンピュータ」黎明期のころに、人文学研究者が自力で作成したテキストデータの多くは、プレーンテキストか正規文法レベルのマークアップであった。しかし最近では、XMLを駆使した複雑な構造化テキストも珍しいものではなく、TEIのような標準マークアップを利用するケースも見られるようになった。さらに、IIIFのような最新技術をいち早く導入する事例も散見されるようになった。また、ここ数年の特徴として人文学分野でもGISの手法を取り入れるケースが増えている。通常のメタデータに緯度・経度、場合によっては時間属性を追加することにより、データを地図上やタイムライン上で可視化できるようになった。緯度、経度、時間属性の多くは、非数値データである地名や暦名から推論する必要があり、これは一種の計量化操作が必要である。これは図1の「初期分析」に該当する。

このように、人文学分野においてデジタルあるいはコンピューティングが進んでいるのは、「データ収集」から「初期分析」のレベルであり、そこから先への展開事例は多いと言えない(もちろん、人文学には多くの研究科目が存在

^{†1} 京都大学
Kyoto University

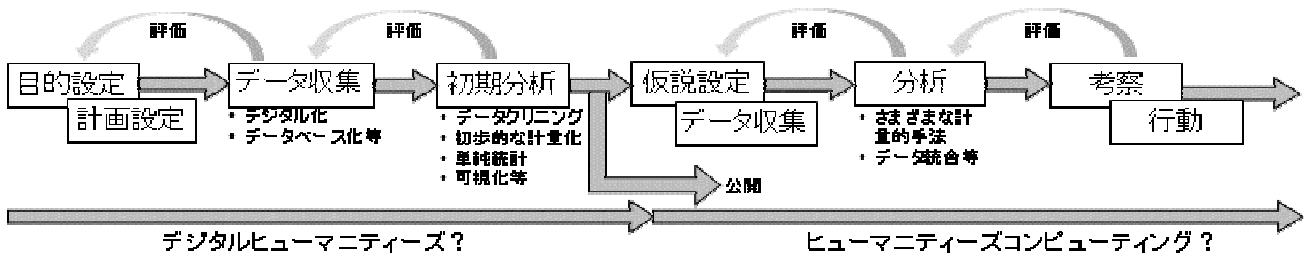


図1 研究プロセス

Figure 1 Research Process

し、その状況は一様ではない)。「じんもんこん」を振り返ってみると、データベースの構築事例は毎年のように報告されているが、構築したデータベースを駆使して新しい知見を得たという報告はそれほど多くない。一方で、たとえば国文学関連の紀要などを見ても、コンピュータを利用した分析事例は皆無に等しい。

情報学分野において多くの分析が試みられているのは当然であろう。しかし、その内容を見ると、人文学研究者からの Given データや Web から入手可能なデータに興味あるまたは手元で利用可能な手法を適用している、という印象を強く受ける。実際、技術的な事項に関する考察は厚いが、分析結果に対する考察は相対的に薄い。データそのものに対する批判が皆無、あるいは分析結果を人文学研究者にフィードバックしなかったり評価を受けたりしていない研究事例も少なからず見られる。

「人文科学とコンピュータ」の現在は、図1の左側を「デジタルヒューマニティーズ」と称してデータリテラシの高い人文学研究者がデジタル化を行い、右側を「ヒューマニティーズコンピューティング」と称して情報「処理」リテラシの高い情報学研究者が分析を行っている、というようにも見える。分離はせずにそれなりに共存している。つまり、最初に述べた異分野融合研究を推進するメリットのうち、他分野の手法を導入して自分分野の研究を促進するという側面について、「人文科学とコンピュータ」は成功しているとの印象を受ける。

3. その先の「人文科学とコンピュータ」

発足から30年を迎える「人文科学とコンピュータ」は、フロンティア領域という性格上、体系的というよりも試行錯誤的あるいは事例的に研究を進めてきたと言えよう。さいわい、30年にわたる長期的かつ濃密な相互交流(衝突?)を維持し、その経験を蓄積してきた。その成果は前述のように大きなものであったと評価できる。また、研究協力や知識交換のインタフェースがそれなりに機能しており、研究の棲み分けもできている。ある意味、居心地の良い状況である。しかし、フロンティアのままで30年が経過しSIG-CHの会員数も漸減しつつある現実を直視すれば、「人文科学とコンピュータ」が人文学からも情報学からも魅力

を失いつつあることを示しているように思われる。

「人文科学とコンピュータ」に参加している人文学研究者や情報学研究者は全体の一部である。人文学研究者の大多数は、情報学研究者の支援がなくとも、多くのすばらしい業績をあげており、情報学研究者も同様である。そのような研究者に、「人文科学とコンピュータ」は面白いあるいは役に立つから参加したいと思わせるには、新しい研究領域の創出が不可欠である。もし『その先の「人文科学とコンピュータ」』があるとすれば、新しい研究領域の創出がフロンティア領域のあるべき姿であると考えられる。もちろん、社会あるいはマスコミが興味を引くようなトピック的な研究と情報発信も重要であるが、それだけで他分野から研究者が参入して来ることはないであろう。

従来のように、人文学研究者は分析を「お任せ」、情報学研究者は分析を「お助け」だけで、新しい研究領域を創出することは困難である。人文学研究者は、自分の持つさまざまなアイデアを自分で試みる程度は試みることができるレベルの情報リテラシを習得する必要がある。情報学研究者も、データリテラシを習得し、内容に踏み込んだ分析ができるようになる必要がある。このような環境が整えば、「人文科学とコンピュータ」においてもイノベーションを起こす可能性が生まれる。

情報リテラシやデータリテラシを習得するための即効薬は存在せず、人文学分野においては基礎数理や情報技術に関する教育と実習の充実、情報学分野においてはリベラルアーツ教育の充実を地道に進めることが肝要と考える。しかし、必要な基礎科目を並べるだけであれば(おそらく多くの大学では、そのようにカリキュラムを組んでいるであろう)、単に講義時間が増えるだけで、学生の興味やモチベーションの向上を期待することはできない。1998年に「講座 人文科学研究のための情報処理(全5巻)」が刊行された。これ自体は画期的であったが、それによりSIG-CHの登録会員が増えたわけではなかった。対象と目的をより明確にして、より「魅力的(面白い!)」かつ「効率的(総講義時間が増えない!)」かつ「有用(使える!)」カリキュラムを工夫する必要がある。国内外の事例調査やSIG-CHにおける経験の体系化、さらには類似の問題を抱えている他のフロンティア領域との連携が重要であると考えられる。