

情報国文学について考える

安永尚志

国文学研究資料館

yasunaga@nijl.ac.jp

<http://www.nijl.ac.jp/>

あらまし (講演の趣旨)

情報国文学という語は藤原鎮男会長による創作である。1994年、情報学の手法を国文学に導入したらどういう形になるか、それを情報国文学と呼び、情報国文学の構成を図られた[1]。その後の国文学研究資料館などでの諸活動は、ご趣旨に沿っているか甚だ疑問であるが、古典文学に関わる目録、全文、原本画像、演能などのデータベースとシステムの開発研究はかなり進んできている。本講演では、主として国文学研究資料館などにおける研究事例を紹介しつつ、蓄積されてきている様々な情報資源を活用する文学研究とは何かについて、報告し、ご批評を得たい。

人文科学とコンピュータという新しい領域の問題解決に向けての研究活動が、国内外において盛んである。例えば、最近のシェークスピアの作品の真贋鑑定などの目覚ましい成果をあげることができる。もちろん、日本文学研究へのコンピュータ利用も進んでいる。とりわけ、1999年は極めて画期的な年であった。3種の源氏物語のデジタルテキスト、索引等のCD-ROMの市販、および国文学研究資料館からの日本古典文学本文データベース (実験版) のインターネットによる試験公開 (約600作品) などである。

今後、益々重要な日本文学研究資料のデジタル化とその公開が、インターネット環境並びに電子出版環境において加速されよう。しかし、我々に課せられる課題は多いし、大きい。例えば、IT革命の進行の中で、新しい日本文学研究の展開はあり得るのだろうか。不幸にして、未だ然したる研究成果は得ていないように思われる。今後、日本文学研究がどのように情報技術を利用していけるか、あるいは参画することができるかについて、私案を述べ、ご批評を仰ぎたい。

なお、これからのキーワードと思われる電子資料館システムと、コラボレーションの展開についても触れる。

1. まえがき

文学、歴史学、考古学、美術、音楽などの多くの人文科学分野でコンピュータが活用

されてきている。例えば、古典和漢学の研究においても、研究成果を入手したり、成果物の確認を行うにはコンピュータを利用しなければならないという事態が起こり始めている[2]。また、文学研究を進展させるためのコンピュータ利用の方法論および手法の研究も進展している。

日本文学研究においても、原本、テキスト、書誌、目録、索引などのデジタル化が進み、インターネットやCD-ROMで、容易に利用できるようになってきた。インターネットに接続可能なパソコンさえあれば、デジタル化された多様な学術情報を利用して研究を進め、検証し、あるいはプロダクトを構築することなども可能となってきた。

日本文学研究におけるコンピュータ利用の目的を整理してみる。一応、インターネットを介在するものとしよう。

- (1) 研究課題の探索と調査研究
- (2) 研究論文（雑誌、図書）の探索と入手
- (3) 資料、材料（原本も含む）の探索と入手
- (4) 電子化原本（画像）、テキスト（全文、索引）の探索と入手
- (5) 語彙索引などの作成
- (6) 語彙並びに用例検索と辞書（コーパス）の作成
- (7) テキスト校訂（異本テキストの比較、差異の検出と検証）および解析
- (8) 研究成果の発表（インターネットでの情報発信）
- (9) コラボレーションの推進（問題提起、討議、評価、検証）

などであろうか。従来には無かったあるいは考えられなかった新しい展開が始まっていると言えよう。基より、コンピュータ利用の目的は、大量の学術情報を網羅的に整理し、処理し、新たな知見を得ることである。最近では、情報の量は1人の人間が取り扱うことのできる範囲を超えている。

もう少し体系的にまとめてみれば、日本文学におけるコンピュータ利用は次のような3課題を考えることである[3]。ここでの情報の意味は広くとらえ、モノである文学資料、原本など各種資料を含むものとする。

- (1) 情報をみつける
- (2) 情報をひもとく
- (3) 情報をつくろう

(1)は、研究活動に必要な情報をどうやって探すかという課題である。対象とする日本文学の本、文献、テキストを探し出す方法について、データベースにたどり着き、実際に検索を行い、手に入れるまでを考えることである。今後、情報の探し方の技術そのものも情報基盤の発展環境に合わせて大きく変わる。新しい情報検索の機能を探る必要がある。

(2)は、日本文学研究の立場からの情報とは何かという課題である。研究活動は情報の処理の仕方にあるということを基本にして、情報の様々な側面を分析、評価し、新しい

モデルを提案する必要がある。そのモデルに基づいて、実際のテーマを分析し、処理し、考察し、組み立てる。

(3)は、日本文学研究における情報の処理の仕組みや、データの作り方を考える課題である。とくに、古典テキストの処理方法、標準的なデータの定義法や、文書としての構造の認識法などの研究が必要である。また、原本（写本）は本来マルチメディアであるから、マルチメディアデータベースとしての取り扱いが不可欠である。さらに、研究成果の取りまとめと公表を行うことも含む。

なお、このような課題は他の学問分野においても同様と考えられる。

2. 情報国文学[4][5]

2.1 国文学研究とは

以下、文献[4][5]を引用しつつまとめる。

国文学研究はわが国の文学全体に渡る文学論、作品論、作家論、文学形態論、文学史などを対象とする研究分野である。また、広く書誌学、文献学、芸能学、国語学などを含み、歴史学、民俗学、宗教学などに隣接する。研究対象は上代の神話から現代の作品まで、全ての時代に渡り、地域的にも歴史上のわが国全土を網羅する。

古典文学は千数百年に渡る歴史を持ち、ジャンルも多様である。強いて分ければ、散文、韻文、戯曲のジャンルがある。散文には物語、論評、説話、随筆、日記、紀行などがあり、韻文には和歌、連歌、俳諧、漢詩、歌謡などがあり、戯曲では能、狂言、歌舞伎、浄瑠璃などがある。

文学は人の感性の言語による表出であるから、国文学は日本人の心の表現であり、日本語を育んだ土壌であると言える。すなわち、国文学研究は現代日本人の考え方と感じ方を育てた土壌を探る学問であると言える。国文学研究は文学作品を通じて、すなわち文字によるテキストを主体として、思潮、感性、心理を科学する。

テキストは単なる文字の羅列ではなく、作者の思考や感情などが文字の形で具象化されたものであるから、研究者は書かれた文字を「ヨム」ことによって、作者の思考や感情を再構築しようとする。「ヨム」こととは、読む、詠む、訓むなどの意味である。換言すれば、文学作品を鑑賞し、評論し、その作品を通しての作者の考え方を知ることである。

元来、国文学にとってコンピュータは最も縁遠い存在と見られてきた。国文学者からみれば、コンピュータに文学が分かるかとか、日本語のコンピュータが無いなどの理由である。一方では、コンピュータへ寄せる大きな期待と落差から来る批判もある。

そこで、文学研究にコンピュータは役立つかということの検証が必要である。例え、コンピュータは単に「ハサミとノリ」の役割であるとしても、答はYESと言えるであろうか。研究過程における膨大な資料から生成される大量多種多様なデータや情報の取扱いには、コンピュータは欠かせないに違いない。

2.2 国文学情報の分析[4]

国文学研究の研究対象である資料は、文献資料および本文（ほんもん）資料である。文献資料には写本や刊本などの原本や写真資料があり、また翻刻され、印刷された活字本もある。本文資料は本文、語彙索引、用例索引などがある。また、研究論文の必要なことは言うまでもない。

文献[4]に、国文学情報の分析を行い、その種類と特質をまとめている。ここではその詳細は割愛するが、取扱うべき情報の質的な違いを区分するものとして、5レベルの区分をしている。すなわち、国文学情報は階層構造を持つと考えられ、これを階層モデルと呼んでいる。情報形態は、各レベルにおいてそれぞれ独自に表現され、文字だけではなく、数値（符号を含む）、画像、音声などのマルチメディアによる表現である。

さらに、階層モデルに基礎を置き、作品をヨムためにはその作品の諸本の関連と研究の経緯を知る必要がある。諸本の伝来の系譜モデルが必要であり、提案している。詳細は割愛する。

2.3 国文学データベース[4]

国文学情報の階層モデルに基づく様々なデータベースが開発研究され、実用に供されてきている。原本やマイクロフィルム資料などの書誌目録、研究論文の目録などの基本的なデータベースは既に実用化されている。文献[4]に詳しいので割愛する。最近では、原本画像データベースや全文データベースの実用化が始まっている。代表的なものを後ほど触れる。

2.4 文字セット[6]

文字の問題がある。とくに、JIS 外字の問題である。情報交換のための漢字は、JIS 規格で定まっている。しかし、この枠外に必要な文字が数多く発生してきており、これをコンピュータで取り扱えなければ、国文学研究の推進はないと言い切る国文学者もいる。しかし、JIS 外字を1組織体で努力して作成しても、その組織体の外への流通はほとんど不可能に近い。

最近、5万字を越える大量電子化漢字のパソコン利用が進んでいる。インターネットによる情報交換も可能となってきた。これらについても後述する。

2.5 コンピュータの役割

そもそもの期待は文房具の役割である。村上教授[7]はコンピュータと国文学を対比して、「コンピュータは絶対実在世界の存在を前提として存立しているのに対して、文学研究は世界とは主体によって認知されるときに始めて存立するもので表象された虚構である」と言う。そして、国文学研究では「研究対象への思い入れが必然的に伴って、研究対象への価値付与・意味付与がその最終目的になっている」と言う。このことを前提

として、コンピュータは「文房具化」の役割と規定した。

すなわち、「与えられた命令に従って（意味を除外して），原資料（作品テキスト・史料・作品論文リストなど）から必要なデータをフィルタをかけ抽出し，加工（整序）し，それを出力したり，そっくり仕舞っておいたりすることが出来る道具．それ以上でも以下でもない．ただし，量・速度・機械的無謬性（網羅性）において極めて強力になれる．」と言う。

コンピュータが特定の分野の仕事を進めるための文房具足り得るならば，その文房具の機能や使い勝手の善し悪しは，その仕事の進捗に大きな影響を与える．一方，国文学の研究態様は甚だ個別的でかつ多様であるため，一般化は難しい．

すなわち，両者の現状から見て，コンピュータは未だ十分な文房具に成り得ていないのではないかと思われる．確かに，残念ながら現在のコンピュータは国文学研究においては「ハサミとノリ」程度の文房具でしかないのかも知れない．

ところで，上記の論考には人間の思考を援用するという観点が欠けているように見える．「ハサミとノリ」という知的生産のための道具から，発見的知見を得る道具への進展が強く期待される．

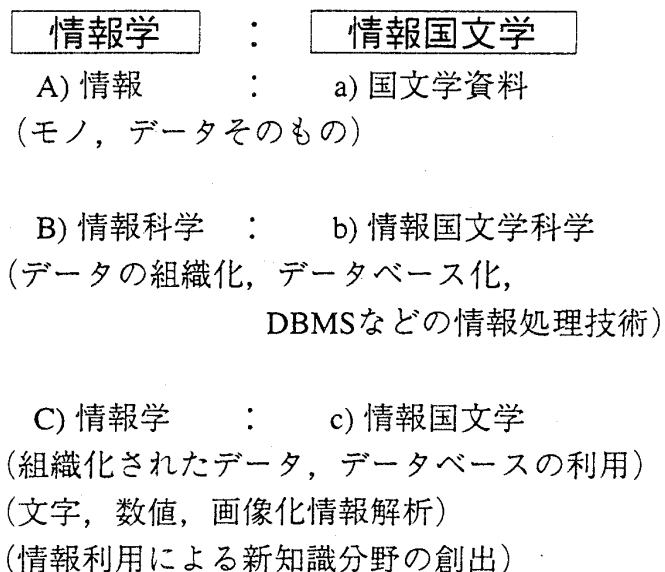


図1 情報学との対比による情報国文学の構成[1]

2.6 情報国文学論[1]

一方，藤原鎮男会長によれば，情報学の手法を国文学に導入したらどういう形になるか，それを情報国文学と呼び，図1のように情報国文学の構成を情報学との対比により行っている．

情報学はA)情報そのものつまり実体と，B)それを処理する科学技術と，さらにA)とB)により体系化された結果C)という3つの部分から成ると考えられる．そこで，国文学の

実体である資料a)と、それをデータベースにしたり、検索したりする資料処理の科学技術b)があり、その上でそれを利用する情報国文学c)があるとする。重要な指摘は、情報国文学は情報利用による新知識の獲得であるという点である。

2.7 コンピュータへの期待[8]

福田秀一教授等による代表的な指摘を以下にまとめる。以下の課題はなかなか難しい。この指摘は、情報システムへの要求と言うよりも、むしろ国文学におけるデータの作り方の問題のように思える。しかし、情報システムから見た場合の課題もあると考えられる。

(1) 情報の即時利用

テキスト同定のための諸本の検討とそれに先立つ探索が必要で、冊子体目録ではデータ更新が望めない。文献資料目録データベースの最新性の維持と、即時性への対応が要求される。例えば、国書総目録[9]はすでに絶版であり、その補遺は古典籍総合目録[10]に引き継がれてはいるが、次の改訂版の出版は未定である。データベースの早期公開が望まれている（一部試行サービス中）。

(2) 典拠コントロール

異名同書や同名異書などの典拠コントロールが望まれる。異名同書では、「更級日記」を「更科日記」，「十六夜日記」を「いさよひ日記」「不知宵記」「不知夜記」などとする表記の違い程度はあまり問題はない。しかし、「吾妻問答」と「角田川」は同書であるが、これなどは目録作成に専門的知識が必要である。

一方、同名異書は難しい問題がある。例えば、「とはずがたり」には、国書総目録に17項目の異書があると指摘されている。

このことは著者の場合にも通じる。同名異人は、論文目録データベースの著者に多く、例えば「伊藤博」氏は正確に何人いるであろうか。

(3) キーワードの統制

例えば、論文目録データベースの場合は、使用される用語が学術用語として確立していないので、キーワードの統制は不可能である。概念は著者により、また読み手により異なり、キーワードの策定に苦勞が多い。タイトルからの自動切り出しも行うが、用語が人により異なる。さらに、分類キーワードも持っていない。

(4) 語彙認定

古典作品は分かち書きされていないので、不用意な語彙区切りは危険であるとの指摘がある。例えば、「無名草子」に「五月十日よひの程」とあり、この「よひ」は古来「宵」と解釈されてきた。しかし、文脈上「余日」つまり「余り」とした方がよいと新説が最近出され、現在こちらの方が正しいとされているという例もある。

2.8 課題の整理

国文学研究資料は、文字を主とするものであるが、写本はマルチメディア情報として、画像として表現される。あるいは、演能、浄瑠璃、歌舞伎などでは、テキストも必要であるが、その演じられることが重要であるから、動画であり、音曲、音楽、朗読、謡などが用いられる。国文学研究資料はマルチメディアであり、これらの電子化と組織化が不可欠である。

すなわち、文字だけを取り扱うのではなく、文字を基に、テキストや本の多様な情報表現を情報工学として考察することが肝要である。

一方、先に触れたように、国文学研究資料の解析により得られた階層モデルに基づき、国文学データベースが数多く構築され、実運用も進んできている。しかし、これらは個別対応のデータベースとして作られ、利用されている。国文学のように思弁的な学問分野においては、様々な研究資料、情報の総合的で多角的な活用が必要とされる。複数の異質なデータベースを渡り歩いて、発見的知見が得られるような仕組みを必要とする。例えば、原本と活字本のテキストにおいて、文や字句などの直接的な比較が可能なシステムが必要である。

インターネット環境を基礎とした電子資料館システムの双方向活用が必要である。あるいはコラボレーションの推進である。

3. 展開と課題

最近の日本文学研究へのコンピュータ利用の展開について、主な成果を以下にキーワードとしてあげる。プロダクトとしてはデータベースである。インターネット型とパッケージ型に分かれる。現在までの傾向では、データベースは作ることに主眼が置かれている。それを活用した研究への展開はこれからのようである[11]。

データベースの提供はCD-ROMなどによる個人レベルでの提供が増えてきている。ホームページでの情報発信は増加しているが、研究にたる情報資源（例えば、校訂テキストなど）は少ない。むしろ、リンク集などに敬服に値する良いものがある[12]。信頼度の高い全文データベース、貴重書などの原本画像データベースなどの提供は、やはり専門の研究機関などによることになるだろう[13]。

良い研究用のデータベースの作成は、1研究者、1研究機関の枠内では限界がある。コラボレーションによる学際的な領域まで広がった専門的かつ総合的な情報基盤の整備が望まれる。すなわち、人、システム、ネットワーク、コンテンツの4カテゴリの総合化で捉えなければならないであろう。

3.1 フルテキストデータベース

3.1.1 源氏物語のフルテキストデータベース

1999年、源氏物語に関する3つの画期的なプロダクトが集中的に出版された。日本文

学研究にとって重要なプロダクトである。

- (1) 伊井春樹他編：CD-ROM 角川古典大観 源氏物語，角川書店，1999
- (2) 村上征勝他編：CD-ROM 勉誠出版 源氏物語，勉誠出版，1999（遅れて，2000年の予定）
- (3) 中村康夫他編：CD-ROM 国文学研究資料館 絵入源氏物語，岩波書店，1999

従来，源氏物語本文の電子化と流通は様々に展開されてきている。しかし，文学研究の観点からの評価はそれほど高いものではなかった。上記3点のCD-ROMは文学研究者の手を経たもので，研究者に活用されることを目標にした完成度の高いプロダクトと言える。それぞれ代表的な写本系列の校訂テキストを中心に，諸本，校異，文法，事例，語彙等々の様々な関連する資料，情報に対応するデータベースを指向している[11]。3つとも，それぞれに大きな特徴を有するが，詳細は割愛する。

3.1.2 和歌集のフルテキストデータベース

新編国歌大観は45万首のほとんど全ての和歌を網羅し，検索機能も多様である。正保版本による勅撰二十一代集は新たな校訂テキストである。これらによって，和歌の引用が極めて容易となった。

- (1) 角川書店編：CD-ROM 新編国歌大観，1996
- (2) 中村康夫他編：CD-ROM 国文学研究資料館 二十一代集，岩波書店，1999

3.1.3 グローバルな古典テキストのデータベース

岩波書店刊行旧版「日本古典文学大系」全100巻のフルテキストデータベースである。古典の代表的な作品約600の校訂テキストからなる。

- (1) 国文学研究資料館：日本古典文学本文データベース（実験版），1999，
<http://www.nijl.ac.jp/>

3.1.4 課題

CD-ROMによる研究支援は画期的である。しかし，問題がない訳ではない。例えば，CD-ROMという限られたメディア空間であること，プラットフォームの限定，価格の問題，新たな電子出版という知的財産権の発生に伴う類似の研究の展開の困難性などが指摘されている。これらの閉じた情報資源の効果的な開かれた相互運用が望まれる。この解の方向としては，双方向運用可能な開かれた電子資料館システムが考えられる。

現在のところ，以上のプロダクトならではの（を活用した）日本文学研究の新しい展開，および新たな研究成果は得られていない。このようなデータベースを形成することは進んでいるものの，これによる文学研究の展開が今後の大きな課題と思われる。

3.2 大量漢字の電子化

日本語処理にとって重要な要因は大量漢字の処理である。周知の通り、現在の日本語処理のための情報交換用符号として、JISが定められている。よく使われる文字集合として、JIS-X0202-1983があり、これは約6000文字である。現在、大きく4つの漢字処理が現れ[14]、これからの展開は予測がつかない状況にある。JISを除いて、大量漢字処理では共通するものの、その目的、用途はかなり異なる。利用者には正しい情報を伝え、啓蒙し、かつ利用後の評価が真摯に行われなければならない。

(1) GT書体プロジェクト	：東京大学	約5万字
(2) e漢字プロジェクト	：京都大学	約9万字
(3) 今昔文字鏡プロジェクト	：学会など有志	約9万字
(4) JIS規格：第3、第4水準		約1万1千字

3.3 情報国文学の展開

コンピュータが文学研究に役立つかという課題に本格的に挑戦する必要がある。文学研究過程をコンピュータ上にシミュレートし、新たな知見を得ることができるかどうかの研究が展開されている[5]。研究過程を通じて生成される大量の多種多様の灰色情報を、再活用し、データベースに展開するシステム環境を提案し、具体的な文学テーマの推進への応用を試みている。

以下の研究例は、文学の問題を越え、国語学の問題としても興味深い。

藤原鎮男会長[15]は、連歌の使用語彙の頻度統計の分析から、語彙の頻度が指数分布をすることを発見し、また普遍性と個別性を持つものへの偏りのあることを発見されている。この発見により、耳慣れた普遍性の言葉を多用することによって、連歌の基調が作られ、意味の限定した個別性の言葉から句の独自性、特色が作られることを実証された。

近藤みゆき、泰弘夫妻[16]は、平安和歌（古今和歌集など）を歌語を始めとする表現の分析に適した形でデータベース化する方法を研究している。さらに、和歌表現、和歌語彙におけるジェンダー性の分析に取り組んでいる。その手法に和歌語彙のnグラム統計分析を用いている。nグラム統計は、手法としては簡単であるが、文字表記の統一性、掛詞の処理、長い文字列の慣用表現や類句の抽出の容易性に効果的で、日本文学研究における強力なツールに成りうる。事実、nグラム統計による女性使用語彙の偏りの発見によるジェンダー研究への展開の足がかりを得ている。

その他、村上征勝氏による使用語彙、品詞などのクラスター分析によるテキストの真贋判定などへの応用研究がある。有名な源氏物語の宇治十帖の作者判別などの研究があるが、詳細は文献[17]などに譲る。

3.4 コラボレーションの推進

インターネット環境におけるコラボレーションは、最も単純なシステム環境として

は、電子メールとWWWの共有によるものである。例えば、テキスト校訂などを行う共同研究が考えられる。遠隔地の研究者や専門家などが共同して、データベースなどのプロダクトを作るという課題を、グループウェアとして、高度の情報交換を可能とし、随時に結果を共有ファイルに形成しつつ、直ちにその評価を行い、より精緻な結果を生成することを考えている。

最終的な目標は、共同利用型の電子資料館システムの構築である。例えば、近接した分野（人文科学）の研究機関の持つ個々の特色ある情報資源のネットワークによる共有の実現である。具体的な実験が始まっている。ある概念で検索すれば、国立歴史民俗博物館データベースの歴史情報が、国際日本文化研究センターデータベースの文化比較情報が、国文学研究資料館データベースの本、テキスト等の文学情報が、研究者の机のパソコンに自動的に揃うことの実現である。

3.5 電子資料館システムの展開

情報発信型と情報受信型の電子資料館システムが必要である。情報発信型は通常の電子図書館システムとして機能する。メタデータ、Z39.50プロトコルなどによる検索システムを装備し、例えば国文学研究資料館の多様なデータベースへの共通アクセスを可能とする。ただし、上述のように共同利用機関のコラボレーションを指向する必要がある。

一方、情報受信型は次のようなシステムである。これは今までにない発想であるが、インターネットの発展を考慮すると、1人または1機関の力は限りがあるため、多くの人々のプロダクトを有効に活用し、場合によって再編し再発信するなども含め、新たな個人向きの情報空間を形成しようというものである。ダイナミックに変化する個人環境を柔軟に作り上げる個人環境向き電子資料館の構築である。必要な情報、データを世界中から瞬時にアクセスし、取得し、自分の知的空間を再編、作り上げていくシステムを目指すものである。言わば、個人の机が世界に向かって開かれた電子図書館として機能する。

4. あとがき

情報国文学という立場の意味について考察した。文学研究へコンピュータが役立つかという第一の課題は重要である。が、ここではむしろ文学研究の情報学的展開という第二の課題への挑戦を意識している。しかし、勉強不足のため重要な事例の研究に乏しく、かつ考察が浅いままに止まっている。

やはり、「コンピュータ活用は、人間の知的活動の延長にあるもので、旧来の手法やノウハウも、次世代への大きな糧となって継承されるべきで、そこにこそ人間の文化的生活のあり方があると言える[2].」とまとめるべきであろう。

サイバースペースにおける知的財産権が大きな関心を呼んでいる。インターネットの爆発的な普及、電子出版、DVDの出現など近年のデジタル化、ネットワーク化の進展は

著しい。著作権の面から見たデジタル化の特徴はデジタル化自体は複製にすぎない。しかし、デジタル化により著作物の利用の仕方が多様になった。本稿ではこの問題については言及できなかった。

サイバースペース上には、同じ情報資源の異種、複製など（とりわけ、同一作品の異テキストなど）が大量に存在する。電子版写本の横溢がある。世界中にインターネット上に源氏物語のテキストはどのくらいあるであろうか。研究に値する資源を見極めることが極めて重要な研究者の資質条件となっている。一方、20世紀後半にして、どの分野においても文献数は激増し、その寿命は短くなり、極端にはWWW上の情報はその安定性、寿命、信頼性を保証することができない[18]。

セキュリティの確保など、情報化の進展に伴って深刻な問題も発生している。このような社会的問題に対処するための方策を、国文学研究と言えども探求することが求められている。

謝 辞

本稿をまとめるに当たって、しばらく体調を崩していたため、主として文献[5]から引用しつつまとめた。内容的にも、最新の研究事例にあまり触れることができなかった。ご寛恕をたまわりたい。本会田畑孝一理事、五所吉哉氏にはご迷惑をおかけした。深くお詫び申し上げますと共に、御礼申し上げます。

藤原鎮男会長には本稿のテーマのきっかけを作っていただき、研究の位置づけなどにご配慮をいただいている。また、藤原譲副会長には終始研究の方向、進め方などのご指導をいただいている。深く感謝申し上げます。

参考文献

- [1] 藤原鎮男：情報国文学の一路程，国文学とコンピュータシンポジウム(6)，1994
- [2] 相田満他：シンポジウム「コンピュータの和漢学」，東洋文化，84，pp.59-71，2000
- [3] 安永尚志：文科系のための情報検索入門，平凡社，1996
- [4] 安永尚志：国文学における情報の考察とデータベースの構築，情報知識学会誌，5.2，pp.1-15，1995
- [5] 安永尚志：国文学研究とコンピュータ，勉誠出版，1998
- [6] 安永尚志：国文学研究資料館におけるJIS 外字文字作成の状況，人文学と情報処理，10，pp.26-34，1996
- [7] 村上学：エンドユーザの視点から，国文学とコンピュータシンポジウム(1)，1989
- [8] 福田秀一：国文学における大量多種データ運用管理のための知識ベースシステム，

科学研究費報告書 #60810009, 1988

- [9] 市古貞次他編：国書総目録, 1～8, 岩波書店, 1970
- [10] 国文学研究資料館編：古典籍総合目録, 第1巻～第3巻, 岩波書店, 1990
- [11] 伊藤鉄也：源氏物語受容環境の变革, 「伊勢と源氏」臨川書店, pp.159-234, 2000
- [12] 例えば, 後藤斉 (東北大学) : <http://www.sal.tohoku.ac.jp/~gothit/zinbun.html>
日本文学研究所 : <http://koudan.com/link/linkd.html>
- [13] 例えば, 京都大学電子図書館 : <http://ddb.libnet.kulib.kyoto-u.ac.jp/minds.html>
- [14] 情報処理学会：次世代の漢字環境, 全国大会 (60) パネル討論集, 2000
- [15] 藤原鎮男：連歌の語彙にみる普遍性と個別性, 国文学研究資料館紀要, 22, pp.213-262, 1996
- [16] 近藤みゆき：n グラム統計による和歌集の分析, 人文科学とコンピュータ最終報告書, 1998
- [17] 村上征勝：真贋の科学, 朝倉書店, 1995
- [18] 名和小太郎：変わりゆく情報基盤, 関西大出版部, 2000