

データベースは研究に影響を与えるか
—国文学研究資料館「日本古典文学本文データベース」利用者調査—

○山西史子
安野一之
安永尚志

What is the use of databases for Japanese literary researches?

Fumiko YAMANISHI
Kazuyuki YASUNO
Hisashi YASUNAGA

Abstract

The questionnaire has been carried out to the frequent researchers for the full-text database of National Institute of Japanese literature. The aim of this investigation is two, that is, 1) to know the activities of the researchers on using the database, and 2) to analyze how to be available to their researches by the database.

1. はじめに

「日本古典文学本文データベース」は、岩波書店刊行の旧版『日本古典文学大系』の本文データベースである。国文学研究資料館ホームページからアクセスが可能で、『日本古典文学大系』全100巻の約600作品を収録し、語彙を検索したり、全文をダウンロードすることなどができる。

2002年2月現在、利用登録者数は約1400名で、うち1割強(150名)は海外からの利用者である。1999年から試験公開されており、一日当たり平均100件強の利用がある。利用が多い作品は『古事記(原文)』『日本書紀(上)』『日本書紀(下)』『源氏物語(一)』『今昔物語(一)』『平家物語(上)』などである。

「日本古典文学本文データベース」は、本文の検索機能等を有するシステムであるが、その公開目的は、研究者に自由に利用できる本文データを提供することであり、そのデータの利用方法は個々の研究者に委ねられている。

本稿では、「日本古典文学本文データベース」を利用して、研究者たちがどのように研究を行っているか、という目的でアンケートによる調査を行った。その集計課程で、興味深い結果が得られたので、報告する。本調査は次の2点を明らかにすることを目標としている。

- (1) 「日本古典文学本文データベース」の利用者の利用目的と利用状況を探る。
- (2) 「日本古典文学本文データベース」の利用により、研究に影響が出てきたかを探る。

2. 調査概要

2. 1 調査方法

調査には電子メールを用いた。

調査対象: 「日本古典文学本文データベース」の利用登録者中、利用回数上位の約300名

調査手順: 2002年2月20日に電子メールにより質問を送付。3月10日を回答期限とした。

回収結果: 不着(16件)、回答数(97件)、有効回答数(94件)であった。

2. 2 調査内容

質問は、10項目の問で構成されており、内容は以下の通りである。

- (1) ～ (4) 「日本古典文学本文データベース」の利用に関する質問
- (5) ～ (7) 「日本古典文学本文データベース」の利用と研究動向の関連の質問
- (8) 「日本古典文学本文データベース」に期待すること
- (9) 同種のデータベースでよく利用するもの
- (10) その他の意見（自由記述）

3 調査結果

3. 1 「日本古典文学本文データベース」の利用目的と利用方法

(1) ～ (4) の回答結果からは以下の4点が明らかになった

- ・回答者の多くが「日本古典文学本文データベース」を「研究目的」に利用している。
- ・利用方法は、「二次資料」（レファレンスツール）としての回答が多いが、「一次資料」（原資料）としてを回答も全体の約3分の1を占めた。
- ・主な利用機能は「全文表示・ダウンロード」と「文字列検索」であった。
- ・利用目的との関連では「用例収集のための文字列検索」「特定の場所を読むための文字列検索」「加工目的の本文ダウンロード」の回答件数がそれぞれ過半数を占めていた。

表1 (3)主に使う機能 (N=93)

全文表示・ダウンロード	55
検索	33
文字頻度分析	1
底本情報の確認	0
決まってない	4

表2 (4)利用目的と使用方法 (N=94)

用例収集の為の文字列検索	73
場所特定の為の文字列検索	50
本文ダウンロード（自分で加工する目的）	51
文字頻度分析	10
諸本底本確認	21
試用段階（研究には未使用）	4
その他	0

3. 2 「日本古典文学本文データベース」の利用と研究動向

(5) ～ (8) の回答結果からは以下の3点が明らかになった。

- ・「日本古典文学本文データベース」は研究の役に立つという評価が大多数であった。
- ・研究動向への変化については、データベース「コンピュータの利用が増えた」と「研究の成果が出しやすくなった」という回答が多い。
- ・研究時の資料利用の変化の有無を質問した表4では、「変化あり」が過半数であるが、「冊子体の資料よりもデータベースを使うようになる」という回答が52件ある一方で、「今までよりも冊子体を使うようになる」との回答が12件あった。また、「影響なし」の回答も28件あり、全回答数の約3分の1を占める点から注目される。

表3 (6)研究動向への変化 (N=92)

研究の成果が出しやすくなった	37
異なったアプローチの研究が可能になった	27
コンピュータの利用が増えた	40
変化なし	11
その他	3

表4 (7)資料利用への変化 (N=92)

冊子体よりもデータベースを使うようになる	52
今までよりも冊子体を使うようになる	12
影響なし	28

3. 3 「日本古典文学本文データベース」へ期待することなど

アンケートの自由記述欄に記入された意見の中から、多かったものをいくつかとりあげてみる。データベースに対する評価としては、「用例集めが容易にできるようになった」という記述が12件あった。要望については、旧漢字の取り扱いなどを含めて、「データの質を向上させてほしい」という意見が最多で23件。次いで、「データベースの収録作品数を増やしてほしい」が13件、「複数の作品間一括検索を出来るようにしてほしい」が11件あった。収録作品の数については、設問(8)の表5にも「より多くの作品の収録」の選択肢があり、こちらでも89件の回答があった。

これらの記述からは、「日本古典文学本文データベース」が用例収集のツールとして高い評価を得ているということが分かる。また、より多くの作品の良質な本文データを入手したい、という要望があることも分かる。

表5 (8)「日本古典文学本文データベース」に期待すること (N=94)

より多くの作品の収録	89
機能の充実	25
論文、図書等とのリンク	46
外字が扱える様にする	45
特になし	1
その他	13

4. 考察

4. 1 「日本古典文学本文データベース」の利用者の利用状況

「日本古典文学本文データベース」の利用方法に関する結果からは、回答者の多くが、「研究目的」で「用例収集」をするために、「文字列検索」や「全文表示」の機能を使用していることが分かる。また、「本文データのダウンロード」をした上で自ら加工する者も回答者中約半数いることが分かる。

一方、回答者の要望をまとめた表5において、回答者ほぼ全員が「より多くの作品の収録」を選択している。さらに、

「市販では入手しにくい作品のDB化を期待する」

「個人ベースで作成しているテキストデータの集約をしてほしい」

といった意見もある。

これらの意見は二通りの見方が可能である。前者は電子化された本文データを自由に使用したいということである。これは「日本古典文学本文データベース」を研究のための原材料（一次資料）としてとらえていると解釈できる。もう一つの見方は、より多くの作品から用例収集をしたり、検索をしたいということである。この場合、検索等の機能が必須になる。これは「日本古典文学本文データベース」を索引等（二次資料）として捉えていることになる。「複数の作品間の一括検索を出来るようにしてほしい」の要望は後者の利用者から出てきたものと解釈するのが自然であろう。

「日本古典文学本文データベース」の利用目的を「一次資料」か「二次資料」かとたずねた質問では、回答者の3分の1が「一次資料」、3分の2が「二次資料」を選択していた。

この結果から回答者は「日本古典文学本文データベース」に冊子体の索引等の機能を期待していると即座に結びつけることは不可能だが、自由に加工できる本文データだけでなく冊子体資料の索引に対応するような、別の付加価値を期待する傾向があることは指摘できる。

4. 2 「日本古典文学本文データベース」の利用により、研究に影響が出てきたか

研究への影響を効率面から考察すると、影響が出ていることは明らかである。

表3の(6)「研究動向に変化は出たか」では、「コンピュータの利用が増えた」「研究の成果が出しやすくなった」ことを認める研究者は多い。また、表4の(7)「資料利用に変化はあるか」に対して、「冊子体よりもデータベースを使うようになる」を回答者の過半数が選択している点は、「日本古典文学本文データベース」が研究を効率よく進めるための道具として定着していると解釈できる。それ以外の影響は無いのだろうか。具体的な事例をあげて考えてみたい。

事例1：「古代の公文書中で紫菜と表記されていたアマノリが平安時代に源順にとって甘苔と書かれ、次いで、うつほ物語の中で甘海苔と表記されて今日に至ったこと、同時代あるいはそれ以前に記録がないことが確認できた」

事例2：「例えば、日本で中世の自殺の研究をしているので、「切腹」だけではなく「捨身」「生害」「投水」などの表現でこの現象が現れているので、「日本古典文学本文データベース」に載っているテキストでは関係のある表現で簡単に調べられることは、非常に役に立つ」

事例3：「膨大な資料の中から特定の言葉を探すのはとても時間のかかる作業ですがこれなら検索にかけるだけなので時間がかかり短縮される。そのぶん考える（考察する？）時間にまわせる。」

事例4：「最後は、人間が読み判断するという（研究の）根幹に揺るぎはない」←()内は著者注

事例1、2は、古典文学の本文を網羅的に調査することによって可能な研究であり、「日本古典文学本文データベース」の存在を前提とした研究テーマが出てきていることが明らかになった。

事例3からは、作業の効率化によって、データの考察がより深くなる可能性を示唆している。

一方、事例4の回答者は、研究の考察段階や解釈にはデータベースの影響を認めていない。

これらを統合して考えると、「日本古典文学本文データベース」を高頻度で利用しながらも、研究テーマの設定から影響を受けるという利用者と、影響は最小限であるという利用者がいることが分かる。つまり「日本古典文学本文データベース」利用者の多様性を示すしている。

5 今後の展望

本調査では、研究のテーマや作業段階への「日本古典文学本文データベース」の影響は確認できた。作業の効率化による研究成果の出しやすさ、「日本古典文学本文データベース」の存在を前提とした研究テーマの設定、などが具体的な影響である。一方、研究の根幹部である、考察や解釈への影響は認めないとする回答者も無視できない。現時点では「日本古典文学本文データベース」による量的な変化が質的な変化には直結していない。しかし、利用者数の増加、利用者層の多様化が進み、それに伴って「日本古典文学本文データベース」側でもシステムの改良を重ねることにより、多くの新しい研究が生まれることを期待できる。

参考文献

安永尚志. 日本古典文学データベース（実験版）の試験公開. 学術月報. pp. 51-54. Nov. 2000

安永尚志. 日本文学研究とコンピュータ. 文学, 2000年7, 8月号, pp. 85-93. 2000. 7

安永尚志. 情報国文学について考える. 情報知識学会第8回研究報告会講演論文集, pp. 33-44. 2000. 5

山西史子（愛知淑徳大学大学院 文学研究科 図書館情報学専攻）

安野一之（国文学研究資料館）

安永尚志（国文学研究資料館）