歴史地震史料集電子データベースの現状と課題

岩 崎 伸 一*

Current Status and Future Remarks about an Electronic Data-Base System of Collected Historical Documents on Earthquakes and Related Phenomena

Sin-Iti IWASAKI*

Abstract

It is well known that, in Japan, very high seismic activity can be traced back to more than one thousand years ago based on old written documents. Many seismologists and tsunami researchers have been receiving benefits from collected and published documents on historical earthquakes and related phenomena (hereafter, collected documents). However, it had been stated for long time that it is difficult to use the collected documents properly because of their huge amount. In fact, only a few "experts" could use them. Besides, it was also a fact that some studies made mistakes with careless reading of the collected documents. To improve this situation, a committee was established in the National Research Center for Disaster Prevention (at present, National Research Institute for Earth Science and Disaster Prevention) to construct an electronic data-base system of the whole collected documents. The committee is now sleeping after constructing a small data-base system and its handling software. Recently, a new movement occurred to make a data-base. Then, it is not meaningless to summarize the activities of the committee about 10 years ago and clarify the problems that arose through the constructing processes of the small data-base system. Major problems pointed out were; 1) huge amount, 2) no indices except time, 3) mixture of documents of different qualities.

Whereas the seismic (geological, geophysical) time scale is long compared to that of a human life, we much owe to historical records when we discuss seismicity in some regions. In other words, the collected documents are a privilege resources of our Japanese scientists. Accomplishing easy access and easy usage of collected documents based on a data-base system is a key to future developments in seismology and related studies.

Key words : data-base, historical documents, earthquakes, tsunamis

キーワード : データベース, 史料集, 地震, 津波

* 防災科学技術研究所

* National Research Institute for Earth Science and Disaster Prevention

I. はじめに

もう10年以上も前のことになってしまったが、石橋(1987)が、地震予知研究における歴史地震史料研究の現状と問題点と題する、いわば歴史地震研究促進のアピールとでも言うべき論文を発表した。これを受ける形で、歴史地震史料データベース化委員会が国立防災科学技術センター(現在は改称されて、防災科学技術研究所)内に作られた。この委員会の活動は、小規模な電子データベースを作成した後、現在その活動を停止している。最近、静岡大学の小山等によって増訂大日本地震史料(武者, 1941, 1943, a, b: 以下武者史料)のデータベース化の話が持ち上がった(たとえば、小山・倉澤, 1998)。小山等が作ろうとしているデータベースと我々が当時考えていたものは、いろいろな面で異なる。特に挙げるべき点は、史料集を(歴史)地震研究に必要な不可欠な“道具”であるとする、小山等は、道具が手に入りやすくなっている、多くの研究者に配布するという目的でデータベースの作成を考え、我々は、道具が使いにくいので使いやすくする、さらには道具そのものを改良し、多くの研究者が簡単に使える道具を新たに作るためのデータベース作成を考えたいと言えるであろう。しかしながら、両者とも、地震およびその関連した研究において歴史地震研究が重要な位置を占めていて、史料集をなるべく多くの研究者に使って欲しいとの認識が根底にある。また、電子データベース化に関する共通した問題も多々あり、ここで10年前の活動を総括し、歴史地震史料集の電子データベースあるいは、それを用いた研究について論じておくのも無駄ではないだろう。

II. 歴史地震史料データベース化委員会の活動

石橋(1987)は、地震予知研究における歴史地震(史料)研究の重要性を指摘した上で、そのためには専門の研究拠点の設立と新収日本地震史料(東京大学地震研究所, 1981~1994: 以下新収史料)、および武者史料のデータベース化が不可欠であると論じた。委員会は、そのうちのデータベ

ス化の実現のために設立された。最初に決まった方針は、小規模なデータベースを作成し、史料集全体のデータベース化、さらには専門の研究拠点設立の地盤を確保するというものであった。

そのためには、様々な分野からデータベースに関わっている方の意見を聞いて、できる限り使いやすく、魅力的な、かつ、後々も使える(今回の努力が無駄にならない)データベースを作成することが肝要であると考えた。そこで、歴史地震史料データベース化委員会拡大委員会が歴史地震研究会の一環として、東京大学地震研究所で開かれた。そこに集まった方々は、地震、津波の研究者のみでなく日本史、史料編集、図書館学、情報関係の研究者、ソフトウェアの作成・販売をする情報・出版関係の企業からの参加者、地方自治体防災担当者等まさに産官学がそろっていた。そこで明らかになった史料集の問題点、データベースの理想像については、歴史地震第4号(歴史地震研究会, 1988) p. 179以下に掲載した「歴史地震史料のデータベース化に関する討論会要旨集」を参照していただきたいが、要点を列記すると以下1)、2)のようにまとめられる。

1) 史料集自体の問題点、データベース化に当たっての問題点、注意すべき点など

1. 史料集自体の問題点

- (1) 量が多い。(約1,500万字(当時)、印刷本の厚さの合計1m以上)
- (2) 読みにくい、あるいは、読めない。(旧字、武者史料第2巻など印刷が不鮮明)
- (3) 両史料とも編年体で書かれており、場所別のインデックスなどが付けられていないので使いにくい。(確かに、安政東海地震等大地震は、量が膨大なため、地域毎にまとめられている。それでも目当ての記事を探すには大変な労力がある。さらに困った問題は、その地震のあった日付だけでなく「先の地震では」等の記述で後の地震の記述の中に以前の地震の記事がでてくる場合があることである。また、地震を感じた地域名にしても、新収史料では「諸国」「全域」など、扱いに躊躇してしまう場合がある。さらに、出版初期のもの

と後期のものとでは記述の仕方が異なる。)

- (4) 新収、武者史料は相補的であり、両者がないと完全なものにならないが、記述の仕方が両者で異なる所がある。
- (5) 信頼性の確定した史料と、そうでないものが、区別なく混在している。
- (6) 親史料、子史料、孫史料が何らの区別なく混在している。
- (7) 印刷された当時との社会情勢が異なるため、現在では、慎重な取り扱いを必要とする記述がある。

2. データベース化に関しての問題点

1. で挙げた問題点は、当然データベース化に際しても問題となるが、特に、史料集をパソコンで読めるテキストファイル化し入力する際の問題点を挙げると、

- (1) JIS 第2水準以外の字、異体字、半行改行、太字、2行割書、ふりがな等、印刷媒体として読む場合には意味がわかるが、パソコンのテキストファイルでは再現できない。そのため付加情報が必要になる。現在はhtmlタグを用いれば表現できるものもあるが、これは当時の話である。また、当然のことながらhtmlタグを用いれば、入力に手間がかかることとなる。
- (2) 著作権問題。特に、近代史料に関しての文献所蔵者の承諾。

3. その他の問題点

史料集を用いた研究が数多くなされ、その中には史料集の記述に疑義を唱えるものもある。しかしながら、これらの研究成果がフィードバックされ、史料集の改訂作業に結実していないのが現状である。そのため、史料集を使った研究を行なう際に、前述した問題点を含む膨大な量の史料集を読み、関連文献を調べ、信頼性を確定しなくてはならない。史料集が研究の素材であることを考えると、活用したいが躊躇してしまう研究者も多いのではないかと考えている。

石橋(1987)の主張した専門の研究拠点の設立は、史料集の改訂作業を行なうことを念頭に置いているが、このことは史料集を使った研究の底辺

を拡大することを意図としたものであると考えている。このような改訂作業を効率的に進めるには、様々なインデックスがあり、個々の記述の比較、検討が容易にできる“電子史料集”が必要であり、その視点に立つとデータベースの理想像と言うべきものが浮かび上がってきた。

2) 電子データベースの理想像

- (1) イメージ入力でなく、文字コードによって入力されたデータベース。
- (2) 様々なインデックスで表示、検索ができ、データの一部取り出し、保存、印刷ができるソフトウェアが必要。特にキーワード検索機能は不可欠であり、後からいくらでもキーワードが付けられる形式のデータベースが望ましい。

拡大委員会では、都司(東京大学地震研究所)から究極のデータベース理想像が紹介された(図1: 都司, 1988, 第Ⅲ章参照)。

3) 作成したデータベース

1. 文字コード化入力した内容

武者史料: 第1巻 605-808頁

年代, 西暦1596年9月5日-1658年5月5日

地震件数 632件

新収史料: 第2巻本巻, 58-195頁

年代, 西暦1596年9月5日-1657年6月21日

地震件数 757件

2. コード化に当たっての整形加工

詳しい内容については、歴史地震史料データベース化委員会(1989)を参照していただきたいが、基本的には印刷本をそのままの形で入力することとした。読み出しの手がかりとして、様々なインデックス(印刷本の巻号、頁、本文の内容の分類、日付等)を付けた。Ⅱ章1)に挙げたように、様々な問題があった。図2、図3は少々見にくいだが、業者にコード化入力を依頼した時の原稿とコード化入力された結果である。インデックスに関する全部の説明は省略するが、たとえばS200188Pは、最初のSは新収史料、次の2桁20は第2巻本巻(たとえば、第4巻別巻2であれば42となる)、次の4桁0188は頁数、最後のPはこれが頁インデックスであることを示す。図2の肉筆による書

REC:3166				S200188P
REC:3167	八月四日丁卯	天晴	未刻地少動	酉刻微雨
REC:3168	(神原藩日記(江戸))			01AD
REC:3169	八月四日丁卯	天晴	之刻少地震	
REC:3170	◎承応二年八月十五日(一六五三・一〇・六)	盛岡→三田		1653100601AZ
REC:3171	(雑書)◎盛岡			02AD
REC:3172	八月十五日	朝霧	巳刻ヨリ雨	
REC:3173	寅刻地震	戌刻地震	亥刻地震	
REC:3174	◎承応二年八月二十二日(一六五三・一〇・一三)	盛岡→二十三田		1653101302AZ
REC:3175	(雑書)◎盛岡			03AD
REC:3176	八月廿二日	晴	戌ノ刻地震	丑ノ刻ヨリ雨
REC:3177	八月廿三日	甚雨	子刻地震	
REC:3178	◎承応二年九月五日(一六五三・一〇・二六)	盛岡		1653102603AZ
REC:3179	(雑書)◎盛岡			04AD
REC:3180	九月五日	曇	午ノ刻地震	
REC:3181	◎承応二年十月二十二日(一六五三・一一・一一)	江戸		1653121104AZ
REC:3182	(稲葉氏永代日記)◎江戸			05AD
REC:3183	十月廿二日	勿刻地震		
REC:3184	◎承応二年十一月十六日(一六五四・一一・四)	盛岡		1654010405AZ
REC:3185	(雑書)◎盛岡			06AD
REC:3186	十一月十六日	晴	午刻地震	
REC:3187	◎承応二年十一月二十九日(一六五四・一一・一七)	江戸		1654011706AZ
REC:3188	(江戸幕府日記(西丸))			07AD
REC:3189	十一月廿九日辛酉	天晴	亥刻地震	
REC:3190	(神原藩日記(江戸))			08AD
REC:3191	十一月廿九日辛酉	亥ノ下刻地震		
REC:3192	◎承応二年十二月十五日(一六五四・一二・二)	五所河原		1654020207AZ
REC:3193	(平山日記)◎五所河原市湊			09AD
REC:3194	十二月十五日	大地震、当年も冬雪降不申候		
REC:3195	承応二年十二月十八日(一六五四・一二・一五)	津軽		1654021508AZ
REC:3196	(新注、あるいは十五日の地震のことか)			
REC:3197	◎承応二年十二月二十七日(一六五四・一二・一四)	盛岡→二十八田		1654021409AZ
REC:3198	(雑書)◎盛岡			10AD
REC:3199	十二月廿七日	晴	丑刻地震	
REC:3200	十二月廿八日	曇	午刻地震	
REC:3201	◎承応二年十二月二十八日(一六五四・一二・一五)	江戸		1654021510AZ

図 3 図 2 の原稿からコード化入力された内容の例。

き込みからわかるようにかなりの整形加工が必要であった。

3. ソフトウェア

表示、検索のためのソフトウェアのプロトタイプは吉井(東京大学地震研究所)によって作成された。これを基礎として改訂版を作った。その機能を簡単に紹介する。このプログラムは PC98 シリーズのパソコン用に書かれたものであり、DOS-BASIC を使っている。MS-DOS, Ver 5.00 までの動作は確認済みである。

- (1) 表示機能：行番号，日付，頁を入力することにより，本文が表示される。本文のスクロール，印刷，抽出保存が可能である。
- (2) 検索機能：本文およびインデックスの任意の文字列での検索が可能である。

本文，インデックスとも 2 つまでの文字列を選択することができ，それぞれについて and/or の条件検索ができる。たとえば，地震，津波のどちらかを含み，1500 年から 1600 年までの記事を検索する場合，「地震 or 津波」「15 or 16」と入力すれば，すべて検索し表示する。検索文字列を反

転表示して，検索行のうち本文のみが表示される。また，検索行を先頭として，それに続く本文を表示することができ，やはりここでもスクロールが可能である。検索結果および，それによって表示された本文は，印刷，保存が可能である。

なお，文字列の検索機能は学習機能を持っている。本文 1 行を 1 レコード文番号を付け，検索文字列(キーワード)によって検索された本文の文番号の情報を付けたキーワードインデックスを作成するようになっている。たとえば「地震」という文字列で検索すると，キーワードインデックスに「地震」の項が付加され，そこには本文中に「地震」という文字列を含むすべての文番号の情報が入力される。これにより，検索に際し最初にキーワードインデックスを参照するので，一回でも検索された文字列は短時間で結果が表示される。また，キーワードはいくらでも増やせることになる。

III. 今までの史料集を用いた研究の問題点と史料集のデータベース化

ほぼ同一の地域内で大規模な地震の発生が、ある程度の決まった周期を持っているという知見は確からしいが、これは史料集を用いた研究の大きな成果の一つである。また、地震、津波の研究者で史料集を利用したものは少なくない。これは、地震発生に関係する地学的、地質学的時間スケールが、研究者の活動期間の時間スケールより遙かに長い場合、ある地域での地震、津波発生の根拠を論じる際に過去の記録に頼らざるを得ないからである。いわば、史料集は遙か過去に置いてきた観測機器と考えることができる。しかしながらこの観測機器はまさに玉石混淆であり、その種類もいろいろであるし、精度も不確定である。

たとえば同じ津波の記述にしても、その津波が起きた直後、経験した人物によって書かれたもの、何十年、あるいは何百年後に人づてに聞いて書かれたもの、藩の正史に当たるもの、旅人がそのとき気ままに書いた覚え書きのようなもの、上質の史料とそうでないものが混在している。

史料に書いてあることをそのまま鵜呑みにして研究を進めるといったことはしないまでも、誤りの入り込む余地は大きい。津波の研究を例にとると、痕跡高、到達時刻の記録等を史料から抜き出し、現地での測量、伝播時間の計算等から、津波を起こした地震の規模、位置を推定する、という手法を取る場合が多い。これについては羽鳥の一連の研究があり（たとえば、羽鳥、1975等）、その成果はめざましく、これほどまでに過去の地震の規模、位置がわかっている国は日本しかない。しかしながら、痕跡高については同一地域でなるべく多くの記録を発掘し、測量結果から合理的である（コンターが滑らか、数値計算と調和的など。たとえば、羽鳥ほか、1981）と考えられる値を取る場合が多いが、たとえば、日本海中部地震の時の峰浜村（東北大学工学部津波実験所、1984）、北海道南西沖地震の時の奥尻島藻内で見られたように、異常に大きな痕跡高がある1点でのみ観測されるという事実（都司ほか、1994）があり、実

際のところ「異常な結果、即、この記述は信用できない」とは結びつかないのである。また到達時刻にしても、江戸時代の時間感覚は岩崎（1987）にもあるように、それほど正確なものではない。加藤・木村（1983）のように、地質学的傍証を手がかりとして史料集の記述を評価できる例はまれである。つまり、史料集の記述を手がかりとしてその後に科学的フィルターをかけて判断する、という現在の手法は、史料に不慣れなものが行なうには危険性が伴う。

史料集の記述そのものを考慮したものとして石橋（1998）を挙げておきたいが、この作業でやっと研究素材の素性がわかるのである。地震関連の他の分野の研究に比して、かなり異なった分野の知識と労力を必要とすることがよくわかる。

史料集が大変な苦労のもとに作成された貴重なものであることは間違いないが、史料集は研究の素材であり、関連分野の研究の進展に伴い何回も読み直さねばならない性格のものである。しかし、Ⅱ章で挙げた問題点、特にその膨大な量を考えると、現在の形式のままの史料集で、今までのような活用の仕方を続けることが可能なのであろうかと考えざるを得ない。史料集自体の信頼性を確定し、精選史料集の作成という前段階の仕事が必要になってきているのである。これには、理工学系の研究者だけでなく、日本史、文献史学の研究者の参加が是非とも必要である。石橋（1987）の言うように、この際には、理工学系の研究者はほぼ一方的に恩恵を受ける立場にある。従って、研究基盤整備、インフラ整備は我々が行なわなくてはならない。そのための歴史地震専門研究拠点の設立であり、史料集のデータベース化である。

IV. 最後に

「現状と課題」ということで書かせていただいたが、ここに書かれていることは、ほとんどが10年前の委員会で話あったこと、あるいはその成果である。それ以降データベース化はあまり進展しているとは言えないので、現在でも「現状」と言ってもいいのかもしれない。石橋（1987）は、歴史地震研究を大きく3つの分野、1. 過去の地震現

象の復原と分析, 2. 過去の地震災害の復原と分析, 3. 過去の社会現象の復原と分析, に分け, 論文の最後にこう結んでいる。少々長いが引用する。「歴史地震の研究が大きな成果を生みだし, それによって地震学や地球科学全般に大きく貢献することの出来る国は, (中略) 今のところ日本しかない。一方, 今世紀末から来世紀初めまでには, 日本列島のどこかで大地震が発生する可能性が高いと考えている研究者は少なくないようだが, それまでにあと 10 数年しかない。その前にここで提案した 2 点が早急に実現されるか, それともこれまで通り, 少数の「物好き」が暇を見つけて, 手作業でこつこつ行なう研究だけを頼りにするか, その選択は日本の地震予知の将来にかなりの影響を与えるだろう。」。この後, 北海道南西沖, 北海道東方沖, 阪神と大地震が起き, 地震(予知)研究を取り巻く体制も変質しつつある。しかしながら, 未だに大地震が日本周辺で起きる可能性は高いと考えられ, 石橋(1987)の3分類のうち, 1. についてはもちろん, 2. についても歴史地震研究の必要性は重要性を増している。まだ, 間に合うのである。

文 献

- 羽鳥徳太郎(1975): 三陸沖歴史津波の規模と推定波源域, 東京大学地震研究所彙報, **50**, 397-414.
- 羽鳥徳太郎・相田 勇・岩崎伸一・日比谷紀之(1981): 尾鷲市街に遡上した津波の調査—1944 年東南海・1960 チリおよび 1854 安政津波—. 東京大学地震研究所彙報, **56**, 245-263.
- 石橋克彦(1987): 地震予知研究における歴史地震研究の現状と問題点. 地震予知シンポジウム, 129-142.
- 石橋克彦(1998): 実在しない天福元年二月五日(ユリウス歴 1233 年 3 月 17 日)の南海巨大地震. 地震 **2**, **51**, 335-338.
- 岩崎伸一(1987): 東海地方の歴史津波. 月刊地球, **9**, 200-204.
- 加藤祐三・木村政昭(1983): 沖縄県石垣島のいわゆる津波石の年代と起源. 地質学雑誌, **89**, 471-474.
- 小山真人・倉澤康二(1998): 歴史時代の天変地異記述データベース(火山活動編)の試作. 平成 8・9 年度特定研究「広領域分野における学術・教育資料の情報体系分析と情報資源化に関する研究」(研究代表者: 八重樫純樹)研究成果報告書, 57-62.
- 東京大学地震研究所(1981~1994): 新収日本地震史料, 第一巻(1981, 193 p.), 第二巻と別巻(1982, 575+290 p.), 第三巻と別巻(1983, 961+590 p.), 第四巻と別巻(1984, 870+582 p.), 第五巻と別巻一・二(1985, 599+403+1931 p.), 第五巻別巻三・四(1986, 293+667 p.), 第五巻別巻五(1987, 2528 p.), 第五巻別巻六(1988, 1834 p.), 補遺と別巻(1989, 1222+992 p.), 続補遺(1993, 1054 p.), 続補遺別巻(1994, 1228 p.). 東京大学地震研究所.
- 武者金吉(1941): 増訂大日本地震史料 第一巻. 文部省震災予防評議会, 945 p.
- 武者金吉(1943a): 増訂大日本地震史料 第二巻. 文部省震災予防評議会, 754 p.
- 武者金吉(1943b): 増訂大日本地震史料 第三巻. 文部省震災予防評議会, 933 p.
- 東北大学工学部津波防災実験所編(1984): 第 2 編 調査資料. 津波防災実験所研究報告, 第 1 号, 88-196.
- 都司嘉宣(1988): データベースの利用にさいしての問題点. 歴史地震, **4**, 185-191.
- 都司嘉宣・加藤健二・荒井賢一・山中佳子(1994): 北海道南西沖地震津波の特徴. 月刊海洋, 号外, 110-122.
- 歴史地震研究会編(1988): 歴史地震史料のデータベース化に関する討論会 要旨. 歴史地震, **4**, 179-197.
- 歴史地震史料データベース化委員会(1989): 歴史地震史料のデータベース化. 歴史地震, **5**, 151-155.

(1999 年 4 月 23 日受付, 1999 年 6 月 21 日受理)