

新聞情報データベースの分析方法の成立条件について

On the Analytical Method of Newspaper Information Database

三 石 博 行

新聞情報データベースの分析方法の成立条件について On the Analytical Method of Newspaper Information database

情報処理技術を活用した社会学調査方法のための検討課題

Examination subject for the sociology investigation method utilized information processing technology

三石博行

E-mail h-mitsuishi@kinran.ac.jp

キーワード 阪神大震災生活情報、新聞全文検索、デジタル社会情報分析方法、

The key word: Living information at Great Hanshin-Awaji Earthquake Whole sentence reference of the newspaper, Analytical method of the digital social information

We did investigation about the daily living information appeared in newspaper news about the Hanshin-Awaji Great Earthquake a way to utilizing a digital database in a newspaper publishing company. Generally we have already traditional sociological investigation method of newspaper information, but doesn't still establish sociological investigation and analysis to utilize these digital data. Therefore we need to argue about an approval condition of this analytical method. We must argue about three models of daily living information supposed inside to study the formation historical process of life structure theory before doing statistical analysis. Because, we put up an idea called historical theory of the daily living information, and by which we developed a model about a tendency of generation change and a distribution change of the daily living information. In other words, it is supposed that the generation and distribution of those daily living information, that we defined last year, show each characteristic one tendency after life structure destroyed by the Great Hanshin-Awaji Earthquake. We need the scientific inspection about it. And, the reliability of the data is furthermore demanded so that it begins to be taken as data if an article includes reference term by one in the whole sentence reference. Finally, this subject is described to have the problem submission about the sociological method of the digital social information and the crisis management or the development of the search engine technology of the daily living information.

はじめに

阪神・淡路大震災時の震災情報に関する研究は災害時社会情報の危機管理体制の課題を提起している¹。すでに、ラジオ²、新聞³、インターネット、携帯電話⁴などの震災情報の分析がなされ⁵、震災時生活情報の伝達問題も取り上げられてきた⁶。多くの震災時生活情報に関する研究結果が出されている中で、我々はデジタル化された新聞情報に焦点を当て、その分析を行った。しかし、この分析方法は社会的な理論としてまだ成立していない。そのために、生活情報構造のモデルに基づく統計的分析の論理、検索データの信憑性、検索ソフトの内容検討と多くの課題を抱えている。さらにこの課題は、高度情報化社会におけるデジタル社会情報の活用方法、生活情報の危機管理やサーチエンジン技術開発等の問題提起として発展していく可能性をもつ。

1、デジタル社会情報の分析は可能か

1-1、伝統的ニュース情報分析の有効性とその課題

ニュース情報の社会的分析は、観察者が直接ニュース記事を読みその内容を理解し、記事の文字数を調べ紙面当りに占める文字情報の割合を算出し情報の重要性を定量的に割り出す伝統的な方法として、すでに確立している。

例えば、中森広道が関東大震災時に出された村の新聞の23項目に関する記事の分析で用いた方法では、全体の記事数の中で占めるその記事の文字数の割合を記事の重要度として算出している⁷。また、三上俊治の行った神戸新聞と毎日新聞の生活情報の分析でも、同じ方法で情報の質や量の分析を行っている。さらに、三上俊治は、情報量の期間的变化を求め、情報発生傾向や調査期間を通じて発生したそれらの情報量を比較する分析もおこなっている⁸。

このように社会学の伝統的方法によるニュース情報の調査方法では、まず観測者が記事に直接触れその解釈を行い、記事と関連する他の社会情報を把握しその背景を理解し、さらに情報量を算出しミクロ的分析を加える等のことができる。現在のところニュース記事分析としてはこの方法が最も確かな手段であることは疑えない。

しかしこの方法では、記事情報のあるなしの判断が研究者個人に委ねられている。例えば、記事の中の一つの調査対象単語を情報と観ることができるかどうかは個人的に解釈によっては意見を異にする場合が発生しかねないため、情報の存在に関する判定が個人的判断に委ねられている限り、個人的主観が入り込む余地は大きい。

また伝統的方法では、莫大な情報を素早く調べるのに多くの労力が必要である。そのため予め調査対象や調査範囲を限定しなければならない。しかし社会情報の現実を正確に分析するためにはできるだけ多くの資料の分析が必要であるので、この伝統的な方法論が今後も有効であり完全なものであると評価されるには限らない。

例えば、三上俊治は調査対象にした阪神大震災時の情報を神戸新聞「がんばろう……震災関連情報」と毎日新聞「希望新聞」のコラムに記載された記事に限定した。時間をかけ莫大な労力を使って分析結果を出すことが、現実社会で問われている生活情報の危機管理を問題にしたこの研究の目的に合うとは思われない⁹。そのために分析範囲を限定する方法で伝統的方法により調査を行ったのであるが、実際にはコラム以外のニュース記事の至る所に震災関係の情報が存在している。新聞が記載した阪神大震災の生活情報を問題にする場合にはそれら全てが対象となるのだが、それを調べるためには大変な労力が必要である。それに比べて中森広道の対象とした資料は、「村の新聞」という極めて限定された情報量を基にしている。そのため莫大な情報分析の作業はなく、伝統的分析方法の限界を指摘される問題は生じていない。

1-2、ニュース記事のデジタルデータベース化

現在、全国的な商業新聞社の殆どが過去のニュース情報をCD-ROMで提供している。このCD-ROM商品が生み出されるには新聞社の情報化が既に進んでいなければならない。記事のデジタル化は、編集、整理、校正や印刷などニュース作成過程の情報化の入り口であり、それらのデータベースが情報処理検索システムを可能にし、記事及び情報作成の資材としてさらに社内用で活用される。このデータベースはCD-ROM版にして消費者に販売される。また、読売新聞社のようにデジタルデータベースを分散型管理にすることでデータベースの危機管理体制を確立している所もある。

毎日新聞の場合、1988年から東京本社で1992年から大阪本社で社内のデータベース化がはじまった。この段階では、記事をキーワードで検索するだけのシステムであったが、その後、社内のニュース記事及び情報作成の資材として活用するためにソース辞書と全文検索のソフト開発がなされた。さらに松下電器株式会社の「バナサーチ」全文検索ソフトを改良し、毎日新聞社の全文検索とキーワード検索を兼用した「シンドバット」検索システムが共同開発され、1995年1月1日から実用化されている。現在、東京本社では東京本社、中部本社と北海道支社から発行された記事の、大阪本社では大阪本社と西部本社から発行された記事のデータベースを作り、東京本社と大阪本社で相互にデータベースサーバを置きそれぞれのデータを送り合っている。

我々は、毎日新聞社の社内用データベースを活用して、1995年1月17日から1996年1月16日までの阪神大震災に関する生活情報の記事の全文検索を行った。

1-3、デジタル情報検索結果の統計処理方法とその問題点

我々は1997年に朝日新聞CD-ROMで20項目の生活情報の記載記事を一定期間毎にキーワード検索し記事件数を調べ、それらを理論的に仮定されたモデルに従って統計処理を行った。

生命維持に直接関係する生活構造から生じる情報・第一次生活情報、豊かな生活を維持するための社会システムや生活構造から生み出される情報・第二次生活情報と、余暇やレジャーなど生活の余裕を享受する生活スタイルや生活構造から生じる情報・第三次生活情報の三つの生活情報を仮定し、それらの生活情報の生活情報発生パターンと生活情報分布パターンをさらに仮定した¹⁰。そのパターンモデルに即した幾つかの資料を得ることが出来たが¹¹、キーワード検索方法では新聞情報の隅々に潜む情報を全て拾い出すことが出来なかつたので、毎日新聞社大阪本社や読売新聞社大阪本社の社内用データベースを使い全文検索の調査を行った。しかしその全文検索の調査の中で、我々の用いたデータベース検索による社会情報分析方法が科学的成立すると言うことを検証しなければならない問題が発生した。それらに関して以下に示す。

第一点目は社会学的方法論に関する基本的問題である。この新聞情報の分析では観測者が直接に記事を読んで判断する作業はなく、情報処理器によって調査したい用語の件数を検索し分析するため、調査対象の社会情報の背景にある別の社会情報を理解することは出来ない。社会情報の分析を進めるための経験的な感覚の不在が生じ、社会学的分析方法としての資格が問われる。そこで検索された記事の全文を出力することによって、直接観察では見逃してしまつた小さい情報も一つ残らず拾いだし、それを伝統的な方法で調査することもできる。そうなるこの方法は伝統的方法の補足に留まると言わざるを得ない。

第二点目は入手した情報の信憑性の問題である。新聞記事の一語が全く無関係に使われることはほとんど無いという記者の経験的な発言もあるが、検索用語が一語でもあれば検索結果が出力されるため、全く関係ない情報が混入する可能性もある。そこで検出された情報の信憑性が問われる。その確率を問題にしなければならない。

第三点目は、第一点目と関連する課題であるが、データを理論やそのモデルによって解釈することになるため、社会学学説史の中で前提にされた理論やモデルに関する点検が必要となる。我々が仮定した生活情報のモデルの定義や、その発生や分布パターンに関して厳密な学説上の議論が問題になる。¹²

このように、データベース検索による社会情報の分析方法は社会学的方法論としての欠陥を指摘されている。しかし他方で、この方法は情報処理機器の力を借り短時間に多くのデータを取り出すことができるため、数多い生活情報の検索には便利な手法である。今後、高度情報化社会ではデジタルデータベースは益々増え、それと共にそのデータベースを活用する社会学的方法も問題にされる。その意味で、この方法を検討する作業は評価しなければならないだろう。この方法を確立するために、以上の課題についての検討を行う。

2、デジタルデータ統計分析方法の理論的点検

2-1、全文検索過程の論理

a、期間集合族とその部分集合の表現

具体的な検索過程では、検索要件が年月日情報と記事情報に分けられ、全ての記事情報は年月日情報、朝刊夕刊情報、誌面情報、地方版情報などが記載され、各記事に時間と場所、発信地名が記録されて、それらが記事情報の基本情報を構成している。つまり時間情報は記事の基本情報である。

我々は、阪神大震災と1995年1月17日から1996年1月16日までの期間を設定した。それをここでは区間集合 B として、その区間集合 B をさらに二週間ごとに区切って、それぞれの期間を B_p で示す。区間集合 B は小区間集合 B_p を要素とする集合族であり、この部分集合 B_p の集合和を以下の関係に示すことが出来る。また、それぞれの区間は重ならないので、以下の条件が成り立つ。

$$B_p \cap B_{p+1} = \phi \quad (1)$$

この条件からそれぞれの部分集合 B_p の集合和は代数的な加算和となるので、以下の関係が成り立つ。

$$\bigcup_{p=1}^n B_p = \sum_{p=1}^n B_p = B \quad (2)$$

但し、部分集合 B_p は 1995 年 1 月 17 日から 1996 年 1 月 16 日までの区間集合 B ・集合族の要素であるので、以下の関係が前提となる。

$$B_p \in B$$

この区間集合 B を従属変数と考え、それぞれの小区間集合 B_p に対する「阪神大震災」という用語を持つ情報と、さらにその阪神大震災の、例えば「安否」についての情報を独立変数と考える。

また、区間集合 B での全ての新聞記事を Ω とし、区間集合 B の部分集合 B_p での全ての記事件数を C_p とすると、その関係は以下のように書ける。

$$B_p \cap \Omega = C_p \quad (3)$$

但し、 $C_p \in \Omega$ である。

b、阪神大震災という用語をもつ記事集合の表現

今、ここで阪神大震災という用語をもつ全ての情報を独立変数 A として考える。阪神大震災という用語をもつ集合 A は紙面に記載された全ての阪神大震災という用語を一つ以上もつ記事からなる部分集合 A^i からなる集合族である。部分集合 A^i の集合和として集合族 A は表現できるので、その関係を以下に示す。

$$\bigcup_{i=1}^m A^i = A \quad (4)$$

さらに、検索条件に当てはまる記事の件数が示される時、同じ記事を一回以上数えることはない。つまり全文検索では記事集合の部分集合である A^i の一つ一つの記事が互いに重なり合うことはないので、部分集合 A^i は以下の条件を充たさなければならない。

$$A^i \cap A^{i+1} = \phi \quad (5)$$

従って、阪神大震災という用語をもつ集合族 A は以上の条件を充たす時に成立する部分集合 A^i のベキ集合と考えられる。

それぞれの小区間集合 B_p に対する「阪神大震災」という用語を持つ情報を独立変数と考え、それらの小区間ごとに阪神大震災の記事件数が求められる。

そこでこの関係は以下に示すことができる。

$$A^i = \{i | 1 \leq i \leq 1, A^i \in C_p, C_p \in \Omega, A^i \subseteq A\} \quad (6)$$

また、(5)式の条件で(4)式は以下のように示すことができる。

$$\bigcup_{i=1}^m A^i = \sum_{i=1}^m A^i = A \quad (7)$$

c、阪神大震災に関する記事に含まれている生活情報用語の件数

阪神大震災に関する生活情報の期間毎の変化傾向を調べるのが、この課題の目的である。そこで阪神大震災という用語をもつ全ての情報のベキ集合と「阪神大震災 and 生活情報」の集合との関係について理解しなければならない。

ここで言う「阪神大震災 and 生活情報」とは何らかの具体的な生活情報を意味する。つまり、それらは「阪神大震災 and 一次生活情報」、「阪神大震災 and 二次生活情報」や「阪神大震災 and 三次生活情報」に関する用語を持つ情報に大きく分類される。そして生活情報を示す用語 j を含む情報をここで A_j^i と考えると、阪神大震災という用語をもつ全ての情報 A も A_j^i を要素とするベキ集合であると考えられる。

集合 A の部分集合 A^i の中に、同様に生活情報の用語 j を含む集合が存在する。その集合を A_j^i と考える。つまり、下に図示したように集合 A の部分集合 A^i の中の生活情報の用語 j を含む要素をもつ集合 A_j^i を取り出す。

この場合、集合 A_j^i は j を構成する数に関係なく、その存在が認められれば空集合ではない。つまり、集合 A_j^i は以下のように示すことができる。

$$A^i \cap A_j = A_j^i \quad (8)$$

従って、 A_j^i がいま(0)の場合、つまり、生活情報の用語 j を含む集合が存在しない場合には以下ようになる。

$$A^i \cap A_j(0) = \phi \quad (9)$$

つまり集合 A_j^i は空集合となる。

ここで任意の一つの記事集合 A^i を以下のように示す。

$$\{j | j \neq 0, 1 \leq j \leq n, i = \text{constant}, A_j^i \in A^i\} = A^i \quad (10)$$

以上の議論を以下の図に示すことができる。

生活情報の用語 j を含む要素をもつ集合 A_j^i は、集合 A の部分集合 A^i の中に用語 j を持つ場合のみ存在する集合であり、用語が存在するとき A_j^i は 1、用語が存在しないとき A_j^i は 0 を示すので、その関係は単位集合(0.1)で示すことができる。

$$A_j^i = (0, 1) \quad (11)$$

この関係から、任意の一つの記事集合 A^i の中に、生活情報の用語 j を含む要素が一つでもある場合、 A_j^i は集合 A の部分集合 A^i に含まれ、次の関係を満たす。しかし A_j^i が 0 を取る場合もありうるので、以下のように示す。

$$\bigcup_{j=1}^q A_j^i \subseteq A^i \quad \text{但し、} A_j^i \in A^i \quad (12)$$

この式から、 A_j^i は単位集合(0.1)であるから、 A_j^i の存在は(9)式の条件を満たすので単位集合の加算形式で計量化できる。例えば、いま生活情報ゼロでないつまり $A_j^i(1)$ である A_j^i 集合が h 個ある場合には記事集合 A^i は以下のように示すことができる。

$$\bigcup_{j=1}^q A_j^i = \sum_{j=1}^q A_j^i = h \quad (13)$$

つまり、「阪神大震災 i 」に関する記事の中に、 h 個の異なる「生活情報の用語」が存在している。

d、ある任意の生活情報語 j を含む要素をもつ記事集合

さらに(6)式に示したように、ある B_p 期間での阪神大震災に関する全ての記事集合 A^i の中で、生活情報 j を含む記事集合 A_j^i の発生する件数をここで B_p 期間での生活情報 j を含む全ての記事集合 A_j と考える。その集合は以下に示すことができる。

$$\{i | i \neq 0, 1 \leq i \leq m, j = \text{constant}, A_j^i \in A_j\} = A_j \quad (14)$$

また、それぞれの生活情報 j に関する記事は決して重ならないので、以下の式が成り立つ。

$$A_j \cap A_{j+1} = \phi \quad (15)$$

さてこの条件を充たす B_p 期間での阪神大震災に関する全ての記事集合 A^i の中で、生活情報 j を含む記事集合 A_j^i の発生する総件数 g は以下のように示すことができる。

$$\bigcup_{i=1}^p A_j^i = \sum_{i=1}^p A_j^i = g \quad (16)$$

e、ある期間の「阪神大震災 and 生活情報 j 」を含む記事集合件数

今、小区間集合 B_p を要素とする集合族 B を考えた。それぞれの小区間において、阪神大震災に関する全ての記事集合 A^i の中で、調査したい生活情報を含む記事係数は(13)式に示したように、それぞれの記事に含まれている生活情報を全ての記事に関して加算して求めることもできる。また、一つ一つの生活情報に関して全ての記事についての件数を出し、それを加算することによって求めることもできる。

$$\begin{aligned} \{A_j^i | 1 \leq i \leq m, 1 \leq j \leq n, A_j^i \in A^i, A_j^i \in A_j, A^i \in C_p, C_p \in \Omega, A^i \subseteq A\} \\ = \sum_{j=1}^q A_j^i \cup \sum_{i=1}^p A_j^i \end{aligned} \quad (17)$$

この(17)式を以下のように表現できる。

	A^1	A^2	...	A^i	...	A^m	
A_1	A_1^1	A_1^2	...	A_1^i	...	A_1^m	$\bigcup_{i=1}^m A_1^i$
A_2	A_2^1	A_2^2	...	A_2^i	...	A_2^m	...
...
A_j	A_j^1	A_j^2	...	A_j^i	...	A_j^m	$\bigcup_{i=1}^m A_j^i$
...
A_q	A_q^1	A_q^2	...	A_q^i	...	A_q^m	$\bigcup_{i=1}^m A_q^i$
	$\bigcup_{j=1}^q A_j^1$	$\bigcup_{j=1}^q A_j^i$...	$\bigcup_{j=1}^q A_j^m$	(18)

(13)式も(16)式もその集合和は互いにベキ集合の形をしているので、(18)式は(13)式から導きだされる h か、もしくは(16)式から導きだされる g のいずれかで表現できるのだが、ここでは混乱を避けるために x という数を設定し、この(18)式を以下のように示す。

$$\sum_{j=1}^q A_j^i \cup \sum_{i=1}^p A_j^i = \sum_{j=1}^q \sum_{i=1}^p A_j^i = x \quad (19)$$

2-2、新聞情報の統計的分析の論理

a、記事数割合変化傾向分析

小区間集合 B_p を要素とする区間集合族 B は、それぞれの部分集合 B_p が互いにベキ集合であるので、(3)式に示したようにその加算和として区間集合族 B を表現した。それぞれの部分集合 B_p が1995年1月17日から1996年1月16日までをそれぞれ2週間毎に区分けしたものであるため、その区間 B_p 毎の阪神大震災に関する全ての記事集合 A^i の中で、生活情報 j を含む記事集合 A_j^i の発生する総件数を g とした。

そこで、区間集合族 B の部分集合である区間 B_1 の場合の生活情報 j を含む記事集合 A_j^i の発生する総件数を g_1 とし、さらに区間 B_2 の場合の生活情報 j を含む記事集合 A_j^i の発生する総件数を g_2 とすると、これらの情報件数の区間毎の移行過程は(16)式から以下のように示すことができる。

$$\sum_{i=1}^p A_j^i = \sum_{i=1}^p g_1 = g \quad (20)$$

そこで、区間全体での発生件数に対する各区間での発生の割合を求めると以下のように示すことができる。これは情報の発生割合を示す。

$$g_1 / g = y_1 \quad (21)$$

この(21)式をそれぞれの区間で求め、以下のように表現できる。

$$y_1 \Rightarrow y_2 \Rightarrow y_3 \Rightarrow y_4 \Rightarrow \dots \Rightarrow y_p = f(y_i) \quad (22)$$

この関数 $f(y_i)$ を記事数割合の変化傾向関数と考えることができる。つまり、ここで $f(y_i)$ で示された傾向は期間全体の中での「生活情報 j 」の発生の割合を一定期間毎に取り、その変化傾向を求めたものである。実際に、出力したデータからこれらの記事数割合の変化傾向グラフを求めることができる。

b、記事数割合変化傾向グラフの成立条件

この記事数割合の変化傾向のグラフは全期間の生活情報量の総数に対しそれぞれの期間毎に発生した生活情報量の割合の変化を示すものである。しかし、期間毎の母集団である全ての記事数に大きなばらつきがある場合は、相対的に記事数つまり情報発生を比較する事は出来ない。従って、期間毎の全ての記事数とその平均記事数とのばらつき度・標準偏差値を測定しなければならない。その標準偏差が極めて小さいという条件が成立してはじめて、記事数つまり情報発生を比較する事が出来る。また、この記事数割合の変化傾向グラフによって、生活情報の一年間の発生傾向を観ることができる。

区間集合 B を1995年1月17日から1996年1月16日までの期間に設定し、それを二週間ごとに区切った期間を B_p で示した。ここで区間集合 B における全ての新聞記事を Ω とすると、期間 B_p での記事数 C_p は(3)式

で示されている関係を使って、記事数分散を以下のように示す。

$$S^2 = \frac{1}{n} \sum_{p=1}^n (C_p - \overline{C_p})^2 \quad (23)$$

但し、期間毎の記事の平均数は以下に示される。

$$\overline{C_p} = \frac{1}{n} \sum_{p=1}^n C_p \quad (24)$$

標準偏差は以下に示される。

$$S = +\sqrt{S^2} \quad (25)$$

この標準偏差値が0.1以下である場合は、期間毎の母集団である全ての記事数に大きなばらつきはないと考えられる。

c、情報の相対的重要度変化傾向分析

単に情報件数の期間毎の変化を示すだけでは阪神大震災という課題の中のある情報の関心度を他の情報との関係で相対的に比較することはできない。そこで、生活情報 j を含む記事集合 A_j^i の発生する総件数 g_1 を、「阪神大震災」の数 x で割ることによって、「阪神大震災 and 生活情報 j」の情報量が阪神大震災に関する記事の中で占める比率 z_1 を求めることができる。その z_1 は、区間 B_1 での「阪神大震災 and 生活情報 j」の相対的重要度の計量的表現である。この方法から、それぞれの生活情報に関する発生情報件数のばらつきに関係なく、関心度を計量的に求めることができる。

$$g_1/x = Z_1 \quad (26)$$

さらに、(19)式から区間 B_1 での「阪神大震災」に関する情報件数を x_1 として、さらに区間 B_2 での情報件数を x_2 とし、区間集合族 B でのそれぞれの部分集合である区間 B_p の場合の「阪神大震災」に関する情報件数を x_p とすると、(26)式からその情報件数は以下のように示すことができる。

$$Z_1 \Rightarrow Z_2 \Rightarrow Z_3 \Rightarrow Z_4 \Rightarrow \dots \Rightarrow Z_p = f(z_i) \quad (27)$$

この(27)式は「阪神大震災 and 生活情報 j」の阪神大震災に関する全ての情報の中での相対的重要度変化傾向を示す関数である。

実際に、記事数割合の変化傾向グラフと同様に、出力したデータからこれらの相対的重要度変化傾向のグラフを求めることができる。

2-3、信憑性点検の理論と実験課題

a、データの信憑性確率の理論

全文検索では、記事の中に一つでも検索用語があれば、検索結果にカウントされる。阪神大震災時の飲料水に関する調査を行うため「阪神大震災」and「水 or 飲」と入力した場合、「阪神大震災」and「水島某」or「飲酒」or「飲食店」という記事の場合でも、検索結果として出力される。このように、全文検索では検索結果の信憑性を問題にしなければならない。

3、社会科学的方法論の課題

3-1、デジタル社会情報分析と学際的研究

ホームページ情報検索に有効なサーチエンジンが必要とされているように、デジタル化してゆく新聞記事・社会情報の活用も実用的な検索ソフトの開発が問題になっている。検索者の意図に即して働く検索ソフトの技術を開発しなければ、高度情報化された社会情報データベースを活用したマネジメントは不可能である。この技術の確立によって、デジタル化した社会情報を社会学的分析の材料にしようとする我々の課題いずれにしても解決されるであろう。

高度情報化社会における社会情報の調査は、今回の調査のように新聞社が開発したシステムを前提にして行われている。データベースの作成過程では、データの入力条件や検索過程での検索ソフトの性質等の影響を受けているため、デジタル社会情報を活用する場合にはデータベースシステムや検索ソフトの論理構造に関する理解が必要となる。社会情報の研究は、他方で高度情報化社会のシステムを動かす情報処理の理論や技術に関する知識が要求され、それらの先端科学技術との学際的研究が課題になってくる。

毎日新聞社のデータベースの全文検索の問題は単純な論理で問題を検討したが、毎日新聞社と松下電器産業が開発した「シンドバット」ソフトのプログラムの内容に触れるなら、その検索ソフトの理論的問題にまで立ち入らなければならない。例えば今後我々が計画している全文検索データの信憑性の調査によって、もし仮にその信憑性確率が低い場合には、この実験結果を新聞社データベース部やソフト開発研究者に還元し、検索ソフトの内容を問題にしなければならないだろう。

この社会情報データベース分析を試みた社会学的研究は、有効なサーチエンジンを開発する情報科学と共通の課題を持つことになり、そこに共同研究の可能性が生まれ出されるだろう。しかし、有効な検索ソフトを開発しようとする情報工学分野との共同研究や学際的研究は情報社会学の基本的な課題ではない。社会学の課題は、あくまでもデジタルデータベースを活用しようとする社会学的方法論のために必要な学際的知識である。

3-2、統計分析の前提をなす社会認識の課題

新聞情報の統計分析方法論は、新聞記事を新聞記者個人の行為として捉える考え方からは生まれてこない。統計的分析の社会的土壌は、記事が社会文化システムの中で新聞記事作成過程と呼ばれる新聞記者の意識を越えた意思決定過程として発生すると考えることによって成立している。

記事の作成過程を構成する要素は、記事の背景となる社会的事象、新聞社の社会的機能、経営的要素や記者の個人的要素が考えられる。新聞社の社会的機能とは市場での新聞社の社会的役割・セールスポイントと呼ばれるもので、その機能を実現するために記事作成や印刷、さらにはデータベース作成など新聞紙面作成過程の組織が作りだされる。経営的要素はその紙面作成過程のみではなく、広告や販売などの企業体としての組織を構成するものである。さらに個人的要素に関しては、記事と呼ばれる商品生産に従事しているジャーナリズム企業の一員として位置づけなければならない¹⁴。

言い換えると、新聞記事の統計分析を行うためには、記事の作成過程を意思決定過程として位置づけ、記者個人の行為が社会文化的環境の中で決定され、個人の行為を支配する精神構造が社会文化構造と相似的な関係にあり、社会文化現象として新聞記事が発生するという、社会的存在者としての記者個人の在り方を社会分析するための前提条件が社会認識論として問題になっている。その上で、新聞記者の記事作成に関する個人的行為を、彼らの新聞社共同体の中での社会的役割行為、つまり企業の集団表象を土台にする行為として、またその精神構造を共同主観や間主観として語る、社会文化の現象としての新聞記事分析つまり社会情報学の言及が必要とされる。

3-3、構成的方法、演繹的・帰納的方法

朝日新聞の分析では、キーワード検索資料の統計分析を行う前に、生活情報史観と生活情報構造のモデルを仮定し、そのモデルから情報・記事件数割合の変化傾向や情報の相対的重要度変化傾向を予測し、その予測されたモデルに即して、調査対象となる生活情報の全文検索資料の統計処理をするという演繹的方法で分析を行った。

しかし、現実のデータからは理論的に予測したパターンに当てはまらないケースが出てきた。予測に当てはまるものから仮定された生活情報史観と生活情報構造のモデル¹⁰を立証することはできるが、予測に当てはまらなかったパターンに関する分析は演繹的方法の延長では展開できない。それらを現実の資料を用いて、前提にされた理論を点検する方法である帰納的方法が問われる。この方法で毎日新聞社の71項目の資料の分析を行った。その中で、生活情報の相対的重要度変化傾向のパターンや記事件数割合変化傾向のパターンが、はじめに仮定したものから別のものに移行している現象が仮定できた。その原因に関しては今後厳密な検討が必要であるが、現在のところ我々は生活構造の変化にともなう生活情報の質的な移行過程によって生じる現象であると理解している¹⁵。

この作業から、演繹的方法と帰納的方法の弁証法的関係に気付いた。資料の分析には理論が必要で、その理論的解釈やそのモデルから実験結果を分析できる。と同時に実験結果から得た資料から理論的仮説を点検することで、科学的考察がさらに展開していく。つまり演繹的方法と帰納的方法が相互に展開され、現実的な理論の点検作業が可能になる。社会現実から生活情報の危機管理という問題提起を受けている阪神大震災時の生活情報分析やデジタル社会情報分析の課題では、実践的課題と理論的課題の両義性を相互に確認しなければならない方法論が問われている。この演繹的方法と帰納的方法の両立を課題にする方法論を吉田民人は構成的方法と呼んだ¹⁶。今後、高度情報化社会での社会情報分析の科学的手段としてこの構成的方法の展開を試みる必要がある。

引用文献

- 1 廣井 脩「阪神・淡路大震災と災害情報」 in 『東京大学社会情報研究所調査研究紀要』第6号、1995.12、pp85-106
廣井 脩「震災時の生活情報をどう伝えるか」 in 『国民生活』1995年5月号、1995.5、pp64-69
- 2 川端信正 廣井 脩「阪神・淡路大震災とラジオ放送」 in 『東京大学社会情報研究所調査研究紀要』第10号、1997.12、pp83-95
- 3 岡田亞弓 野田 隆「阪神・淡路大震災における安否情報・生活情報とマスメディア」 in 『家政学研究』奈良女子大学家政学会々誌 Vol43 No2、1997.3 pp47-54
- 4 中村 功 廣井 脩「兵庫県南部地震時の携帯電話の役割と問題点」 in 『東京大学社会情報研究所調査研究紀要』第7号、1996.12、pp1-17
- 5 中森広道 廣井 脩「阪神・淡路大震災と初動情報」 in 『1995年阪神・淡路大震災調査報告書-1-』東京大学社会情報研究所「災害と情報」研究会、1997.3、pp66-80
- 6 中村 功 廣井 脩「災害時の安否情報とメディアミックス」 in 『東京大学社会情報研究所調査研究紀要』第10号、1997.12、pp155-180
- 7 中森広道「関東大震災と狭域情報-『村の新聞』からの考察」 in 『社会学論叢』第131号 日本大学社会学会 1998年3月、pp19-36
- 8 三上俊治「都市直下地震発生時のメディア環境と情報行動に関する研究-阪神・淡路大震災における生活情報へのニーズと流通の実態-」平成9年度科学研究費補助金 重点領域研究(2)成果報告書 平成10年3月、pp12-17
- 9 三上俊治「阪神・淡路大震災における生活情報へのニーズと流通の実態」 in 『第2回都市直下地震災害シンポジウム』東京 1997.11.19 pp1-8
- 10 三石博行「生活情報構造モデルと生活情報史観」 in 『社会・経済システム学会 1997年第16回大会報告要旨集』社会経済システム学会 関西大学 1997.11.1 p3-6
- 11 三石博行「阪神大震災以後の生活情報発生時の調査と生活情報構造分析」 in 『第5回情報文化学会全国大会講演予稿集』、情報文化学会 東京工業大学、1997.11、pp20-23、ISSN 1341-593X
- 12 三石博行「社会システム論的生活構造論学説史批判と現代生活情報論の科学性批判」 in 『社会・経済システム論学会 1998年第17回全国大会報告要旨集』社会・経済システム論学会、京都精華大学、1998.10、pp4-5
- 13 R.A. フィッシャー『統計的方法と科学的推論』渋谷政昭、竹内啓訳 東京、岩波書店、1962.11、pp8-28
- 14 三石博行「生活構造論から考察される生活情報と生活情報史観の概念について」 in 『情報文化学会誌』第6巻第1号 1999.10、pp57-63

- ¹⁵ 稲葉三千男「マス・コミュニケーションの制作過程における主要な矛盾」in 佐藤毅、細谷昴、竹内郁郎、藤竹暁編集『社会心理 マス・コミュニケーション社会学セミナー4』東京、有斐閣、1973. 10、pp225-237
- ¹⁶ 三石博行「生活情報パターン移行減少と情報文化パラダイム」in『第6回情報文化学会全国大会講演予稿集』、情報文化学会、明治大学、1998. 11.
- ¹⁷ 吉田民人「生活空間の構造－機能分析」in 作田啓一編『人間形成の社会学』現代社会学講座 有斐閣 1964年、pp137-196