

Research Reports of the Research Institute
of Industrial Safety, RIIS-RR-92, 1993
UDC 614.8:681.3

建設工事労働災害に関するテキスト情報の解析*

鈴木芳美**

An Information Analysis by Statistical Method for the Descriptive Statements of the Occupational Accident Report Concerning Construction Work*

by Yoshimi SUZUKI**

Abstract; In this study, driving a method for making practical and direct use of descriptive statements of the "Occupational Accident Report" concerning Construction Work, statistical analysis methods are applied to examine free-terms used in descriptive statements of these reports. Especially, statistical quantification analytical method are applied to study free-terms selected from descriptive statements of the "Accident Situation and Circumstances".

Brief outline of this study and main results are as follows; Firstly, author has developed information retrieval database system about original occupational accident reports, and 9059 kinds of free-terms in total are extracted from the Japanese sentence item of the "Accident Situation and Circumstances" by means of the utility in the Database. The frequency of these extracted free-terms were distributed according to the Bradford's law.

Secondly, 293 free-terms of high frequencies are selected as keywords for arranging information concerning many factors of occupational accidents occurred in construction work sites. In this procedure, the relationship between these keywords and accident cases are analysed by the quantification method of third type. As results of this analysis, scatter diagram of keywords (category weight, Fig. 3 ~ Fig. 7) and scatter diagram of accident cases (sample score, Fig. 8 ~ Fig. 9) are obtained and the information about "Accident Situation and Circumstances" is arranged by kind of accident types and kind of construction works.

Thirdly, using these 293 keywords, discriminant analysis was made to examine the probability for exploration of accident cases according to various factor of occupational accidents occurred in construction work sites. Classification results of discriminant analysis are obtained as classification result tables (Table 5 ~ Table 7), and each accident cases are sufficiently classified by accident types or by kind of construction works.

Then, after these statistical analysis of free-terms extracted from the Japanese sentence item of the "Accident Situation and Circumstances", information structure concerning occupational accident in construction work can be arranged according to a few key concept; accident types, kind of construction works and so on. For distinguishing various information from occupational accident data, statistical methods using free-terms in descriptive statements of "Accident situation and

*平成 4 年 9 月 28 日土木学会年次学術講演会において一部について発表

**土木建築研究部 Construction Safety Research Division

Circumstances” of occupational accident report can be used practically.

Keywords; Occupational accidents, Construction work, Information analysis, Statistical analysis, Data-base, Key-word, Quantification method of third type, Discriminant analysis.

1. ま え が き

労働災害事例に関する諸記録から得られる情報は、労働災害の発生傾向の分析のためのデータやその後の類似災害防止のための資料としてきわめて貴重かつ有用なものとなっている。また、これらの情報の有効活用を図ることについては、従前より各方面から望まれてきたところでもある。

本研究は、労働災害事例に関して記述された記録を適切かつ有効的に活用を図ってゆくための基礎資料を得ることを目的として行ったものである。本報告では、労働災害の発生に伴って作成される災害記録資料のオリジナルの記述内容を取り上げ、これらが有している情報としての性質や構造を明らかにするために行った情報解析の結果を述べたものである。

前報¹⁾では、特にトンネル建設工事における労働災害を取り上げ、労働災害の災害発生状況に関する記録・記述から得られる情報が有している性質についての解析を行った。今回は、分析対象を特定の工事に限らず建設工事全般に拡大し、また分析対象事例数も増大させて分析を試みたものである。

2. 労働災害の記録資料

2.1 労働災害の記録資料が有する特性

2.1.1 オリジナル資料の特性

労働災害の事例に関しては様々な記録資料がある。これらの記録資料は、発生件数などの数値統計を目的とした資料を除くと、各事例ごとに災害発生日時や災害発生状況などを記録したものが一般的である。すなわち、これらは個別の災害事例に関して文書として作成された記述資料と言えるものである。

現状では、これらの多くの記述資料が収集された上で、分析用のデータや資料として提示されたり引用されていることが多い。

ところで、これらの記録資料のオリジナルデータの記述内容は、当然のことながら当該事例に固有の事柄や条件あるいは背景などが詳述されている。したがって、個々の事例によって、それらの記述内容の

細部は千差万別にならざるを得ない。換言すると、労働災害事例に関するオリジナル資料の記録内容は、本質的には情報検索や情報活用などを念頭に置いたものではなく、雑多な情報が混在して記述されたいわば未整理な「情報源」に過ぎない性向を有している。

2.1.2 二次的資料の特性

一方、労働災害事例の記録資料の中にはオリジナル資料の記述内容に基づいて必要事項を取捨選択し整理を行ったものも多い。これらの言わば二次的資料では、情報整理のフィルタリングが施されているため、本来のオリジナルデータには記録・保持されていた情報内容の細部が最終的な段階では欠落してしまうものも少なくない。逆に、このような二次的資料では、例えば災害タイプの分類などいわゆる二次的情報（あるいは付加情報）が加えられているのが特徴である。

災害タイプの分類に例を挙げると、これらの資料では「墜落」「土砂崩壊」等の従来から行われている災害種類の分類や区分けに基づいた情報付加が施されている。これらの付加情報の加えられた記録資料は、取扱がきわめて便利であるなど付加価値が加わった優位性を有する情報と成り得る。

しかし反面、これらの二次的情報では前述した通り、オリジナルデータには記録されていた細かな記述内容などはかえって取捨抹消されてしまう傾向もある。例えば、同じ墜落災害でも「足場から」か「梯子から」かなどの細分類までは通常の二次的資料の災害分類では言及しない場合が多い。したがって例えば、従来の災害分類区分とは異なった角度からの分析を必要とする場合やより細分化した分類を必要とする場合などには、このような記録資料では対応できないケースも多々生じている。このような場合にはオリジナル資料の記述内容にたちもどって災害種類に関する細かな記述内容などを再吟味する作業が必要となる。

2.1.3 オリジナル資料活用の可能性

前項までに述べたように、労働災害事例に関する「情報」には、2つの相反する性質すなわち「オリジナル資料の記述内容の細部の保持・留保」と「情報の

取扱い易さ（すなわち情報内容の整理・取捨選択・簡潔化）」との両面が要求される場合が多い。

そこで、労働災害に関する記録資料を最大限有効的に活用していこうとする場合、未整理なオリジナル資料の記述内容が「情報源」としてどのような形で労働災害に関する「情報」を包含しているのかを明らかにしておく必要がある。

すなわち、これらの情報源が本質的に有している性質や構造が前もって的確に把握されておれば、フリーターム検索等を駆使することにより、利用目的に対応した事例検索やデータ抽出を可能とすることが出来る。

このことは、オリジナルデータの記録内容を無理なくまた無駄なくそのまま直接活用することにも通じ、労働災害に関する適切な情報の提供・労働災害に関する記録の有効活用に結び付くとも考えられるからである。

2.2 分析対象資料

本報では分析対象とした労働災害のオリジナル記録資料として、労働省の安全衛生行政の中で取り扱われる「災害調査復命書」と呼ばれる文書資料を取り上げた。当該資料は、死亡労働災害あるいは重大災害（一時に3名以上の被災者を伴った災害）などの発生時に、所轄の労働基準監督署の係官が現地調査を踏まえて作成するものである。

本資料の概略の構成は以下の様になっている。表紙に相当する第1頁（正式には「災害調査復命書甲紙」と呼ばれている）はフォーマットの定まった表形式になっており、災害発生日時や発生場所をはじめ、災害発生状況の概略などが記載され²⁾、個々の災害事例のインデックス的な役割を果たしている。また第2頁以降（正式には「災害調査復命書乙紙」と呼ばれている）には、災害発生状況の詳細や災害発生原因・要因など係官の調査した内容が図表・写真等を含めて記載されている²⁾。第2頁以降の記載量は不定で、およそ数頁（多いものでは数十頁にわたるものもある）程度である。

本資料はその性格上、記載内容等については公表できる性質のものではない。しかしながら、例えば第1頁目に記載されている「災害発生状況の概略」に関する記述などは、各種の労働災害の実際の発生状況やその背景などを「災害事例のインデックス」的情報として簡潔な文章情報として得ることができる。

Table 1 Detail classification of 807 accidents
分析事例（807件）の内訳

月別件数		工事種別発生件数			
1月	59件	土木工事	410件	建築工事	322件
2	62	水力ダム	2	ビル	161
3	76	トンネル	23	木造	74
4	65	地下鉄	1	建設設備	39
5	47	鉄道	10	その他	48
6	42	橋梁	29	設備工事	75件
7	73	河川	29		
8	58	砂防	21	電気通信	29
9	94	土地整理	33		
10	75	上下水道	43	機械設備	28
11	58	港湾	21	その他	18
12	98	その他	24	分類不能	83件
曜日別件数		被災者の年齢別			
月	122件	19才以下	15件	45～49才	120件
火	122	20～24才	34	50～54才	135
水	121	25～29才	46	55～59才	110
木	112	30～34才	54	60～64才	57
金	122	35～39才	72	65～69才	21
土	126	40～44才	97	70才以上	13
日	61				

これらは、労働災害の傾向についての分析や災害防止のための教育的資料などとしても充分に応用できる極めて有効で貴重な資料となり得るものである。

したがって今回の分析では、本資料の「災害発生状況の概略」をオリジナルデータとして情報分析を行う対象として取り上げた。

前述したように、これらの記述内容の情報としての性質や構造を把握しておくことにより、将来的にはこれらの資料を埋もれさせずに十分に活用し得るような様々な手段や手当を講じあるいは施してゆくことも可能と思われる。

2.3 分析対象事例の概要

今回の分析では、ある1年間（正月元旦から12月大晦日までの間）で発生した建設工事における死亡労働災害など807件（災害調査復命書によるもの）の事例をデータベース化し、そこから得られるデータに基づいて行った。これは、この年に発生した全建設工事死亡労働災害のうち約9割に相当し、その内訳はTable 1に示す通りである。

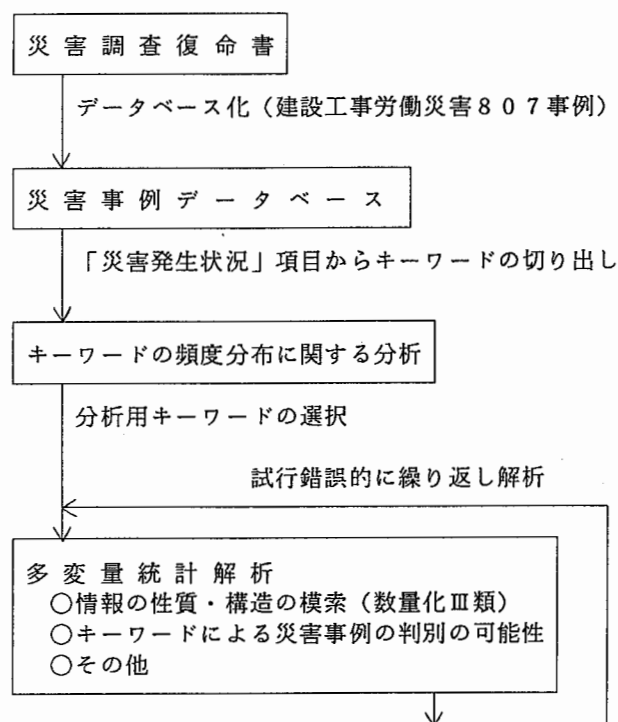


Fig. 1 Brief procedure of analysis
分析内容のフロー概略

Table 1 から判るとおり、これらの労働災害事例は建設工事全般を漏れなく全体的に網羅したものである。残りの約 1 割の労働災害については、たまたま諸般の事情により本解析実行時点でデータベース化されていなかったため除外されたもので、選択的に意味があるわけではない。

Table 2 Comparative table about number, kind and cumulative frequency of keywords
キーワードの切り出し結果

	建設工事労働災害	トンネル工事労働災害 ¹⁾	全産業分野労働災害
オリジナルデータ および 事例数	災害調査復命書 807 事例	死傷病報告書 161 事例	災害調査復命書 3,592 事例
キーワード総種類数	9,059 種類	3,070 種類	26,824 種類
キーワード総のべ語数	40,285 語	10,714 語	220,081 語
1 事例あたりの 平均キーワード種類数	11.22 種類/事例	19.07 種類/事例	7.47 種類/事例
1 事例あたりの 平均のべ語数	49.92 語/事例	66.55 語/事例	61.47 語/事例

注：トンネル工事労働災害の数値については参考文献 1) より引用

3. 分析内容と分析方法

3.1 分析内容のフロー

今回の分析内容は前報¹⁾に報告した分析内容と基本的には異なるものではなく、全体の分析内容の概略のフローを示すと Fig. 1 の通りである。

すなわち、まず労働災害に関する記録資料のデータベース化を行い、特に災害発生状況に関する項目（最大文字数 350 字の日本語で記された文章項目）からデータベースが有するユーティリティ機能を用いてフリーターム（キーワード）の切り出しを行った。

切り出しの行われたフリータームについてその頻度分布状況をチェックし、その結果を踏まえて分析に用いる比較的高頻度のフリータームをキーワードとして選定した。次に選定されたキーワードを用いて、数量化 III 類をはじめとする多変量統計解析を実施した。数量化 III 類による分析ではキーワードと災害事例との間の相関に関して試行錯誤的に解析を何回か繰り返し実施した。さらにこれらの結果を踏まえてキーワードによる災害事例の判別可能性などについても検討を加えた。

4. 分析結果と考察

4.1 キーワードの頻度分布

今回分析した全 807 事例について切り出されたキーワードの総種類数は 9059 種類（1 事例あたりの平

Table 3 Cumulative frequency of keywords (807 labour accidents in construction work)
切り出しキーワードの頻度・種類・累積頻度分布表 (建設工事労働災害 807 事例)

	頻 度 (キー ワード使用 事例数) ①	キーワードの種類数と頻度		頻度数 ①×②	累積頻度 (累積率%)	
		種類数 ②	頻度順位 (累積率 %)			
1 作業	559	1	1 (0.01 %)	559	559	1.73 %
2 工事	503	1	2 (0.02)	503	1062	3.28
3 被災	417	1	3 (0.03)	417	1479	4.57
4 者	401	1	4 (0.04)	401	1880	5.81
5 死亡	385	1	5 (0.06)	385	2265	7.00
6 m	346	1	6 (0.07)	346	2611	8.07
7 現場	343	1	7 (0.08)	343	2954	9.13
8 被災者	339	1	8 (0.09)	339	3293	10.17
9 町	273	1	9 (0.10)	273	3566	11.02
10 墜落	243	1	10 (0.11)	243	3809	11.77
11 市	239	1	11 (0.12)	239	4048	12.50
12 下	221	1	12 (0.13)	221	4269	13.19
13 高さ	202	1	13 (0.14)	202	4471	13.81
14 内	185	1	14 (0.16)	185	4656	14.38
15 メートル	159	1	15 (0.17)	159	4815	14.87
16 コンクリート	152	1	16 (0.18)	152	4967	15.34
17 名	151	1	17 (0.19)	151	5118	15.81
18 用	139	1	18 (0.20)	139	5257	16.24
19 側 , 他	128	2	20 (0.22)	256	5513	17.03
20 使用	127	1	21 (0.23)	127	5640	17.42
21 階 , 他	121	2	23 (0.25)	242	5882	18.17
22 運転 , 他	118	2	25 (0.28)	236	6118	18.90
23 部	117	1	26 (0.29)	117	6235	19.26
24 労働者	116	1	27 (0.30)	116	6351	19.62
25 新築	114	1	28 (0.31)	114	6465	19.97
26 地上	113	1	29 (0.32)	113	6578	20.32
27 移動 , 他	111	2	31 (0.34)	222	6800	21.01
28 床 , 他	109	2	33 (0.36)	218	7018	21.68
29 長さ	99	1	34 (0.38)	99	7117	21.98
30 道路	97	1	35 (0.41)	97	7214	22.28
85 ショベル , 他	22	13	218 (2.41)	286	15008	46.36
86 埋設 , 他	21	6	224 (2.47)	126	15134	46.75
87 接触 , 他	20	19	243 (2.68)	380	15514	47.92
88 合図 , 他	19	14	257 (2.84)	266	15780	48.74
89 先端 , 他	18	6	263 (2.90)	108	15888	49.08
90 感電 , 他	17	14	277 (3.06)	238	16126	49.81
91 操作 , 他	16	16	293 (3.23)	256	16382	50.60
92 改築 , 他	15	22	315 (3.48)	330	16712	51.62
93 崩れ , 他	14	24	339 (3.74)	336	17048	52.66
94 昇降 , 他	13	37	376 (4.15)	481	17529	54.15
95 ポンプ , 他	12	26	402 (4.44)	312	17841	55.11
96 側壁 , 他	11	34	436 (4.81)	374	18215	56.27
97 ブラケット , 他	10	35	471 (5.20)	350	18565	57.35
98 金具 , 他	9	43	514 (5.67)	387	18952	58.34
99 連絡 , 他	8	60	574 (6.34)	480	19432	60.03
100 すべり , 他	7	82	656 (7.24)	574	20006	61.80
101 スラブ , 他	6	137	793 (8.75)	822	20828	64.34
102 移送 , 他	5	168	961 (10.61)	840	21668	66.93
103 不安定 , 他	4	232	1193 (13.17)	928	22596	69.80
104 ミス , 他	3	446	1639 (18.09)	1338	23934	73.93
105 被覆 , 他	2	1019	2658 (29.34)	2038	25972	80.23
106 ペイント , 他	1	6401	9059 (100.0)	6401	32373	100.0

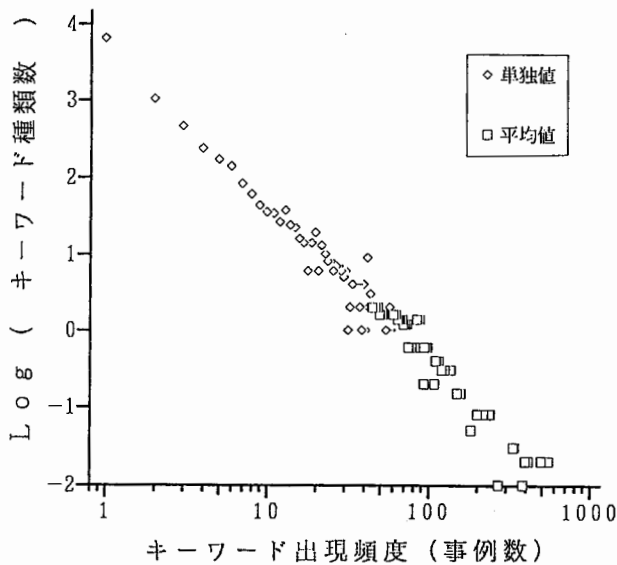


Fig. 2 Frequency distribution of keywords
キーワードの頻度分布図

均のキーワード種類数では 11.22 種類/事例), またキーワードの繰り返し使用を考慮した総のべ語数は 40285 語 (1 事例あたりの平均ののべ語数では 49.92 語/事例) であった。

このキーワード切り出し状況を, トンネル工事労働災害に限ったもの (161 事例) あるいは対象産業分野が建設工事以外の労働災害事例を含むもの (3592 事例) と比較^{1), 2)}すると, 前述の Table 2 のような結果となった。

建設工事における労働災害では種々の資機材や様々な作業工程が関連するため, その記述内容に多様性のあることは前報¹⁾でも述べた通りであるが, 今回の分析でもその傾向が, 1 事例あたりの平均キーワード種類数の数値などに窺うことができる。

Table 3 には, 今回分析を行った建設工事での労働災害事例 (全 807 事例) の中で使用されていたキーワードに関して, キーワードの頻度 (そのキーワードを使用している事例数)・当該頻度を有するキーワードの種類数・高頻度順の順位 (種類数の蓄積数) および累積頻度などをまとめたものを示した。

Table 3 からわかるとおり, 今回分析した 807 事例の中での最大頻度のキーワードは 559 事例の「作業」であった。以下, 503 事例の「工事」, 417 事例の「被災」などが続き, 1 事例のみで使用された出現頻度 1 回のもの (キーワード種類数 6401 種類) まで 106 段階に分布した。

Table 4 Model table of correspondence between keywords and accidents
災害事例とキーワードの対応表 (模式表)

	KW ₁	KW ₂	KW ₃		KW _{x-1}	KW _x
事例 1	v		v		v	
事例 2		v				v
事例 3	v	v				
⋮						
事例 n-1			v			v
事例 n					v	

1 事例のみでしか使用されていないキーワードの数はキーワード全種類の中の 70.66% (全業種労働災害では, 56.08%, トンネル建設工事では 70.46%) であった。また頻度順位で上位から 3%あるいは 5%までに相当するキーワードを使用している事例数は累積頻度で各々, 49.81%, 57.35% (全業種労働災害では 60.95%, 68.65%, トンネル建設工事では 40.52%, 46.42%) に達していることが判った。

4.2 キーワードの頻度分布図

上記 106 段階のキーワードの頻度分布データを両対数紙上にプロットしたものが Fig. 2 である。

横軸はキーワードの出現頻度 (そのキーワードを使用している事例数) を表し, 縦軸はキーワードの種類数 (当該頻度を有するキーワード数) を示す。

但し, 本図の右側部分にプロットされるデータ (頻度の大きなキーワードに関するプロット) については, 頻度数で前後に欠測値があるものが多い。したがって, 欠測値のある場合に限って, ある頻度範囲ごと (5 ~ 50 頻度) のキーワード種類数の平均値を算出し, その値をプロットの形状を変えて (□印) 示してある。

本図の左側部分にプロットされるデータ (頻度の小さなキーワードに関するプロット) には前後の頻度数に欠測値は存在しないので, 観測されたキーワード種類数および頻度数はいわば単独値としてそのまま図中に表示 (◇印) してある。

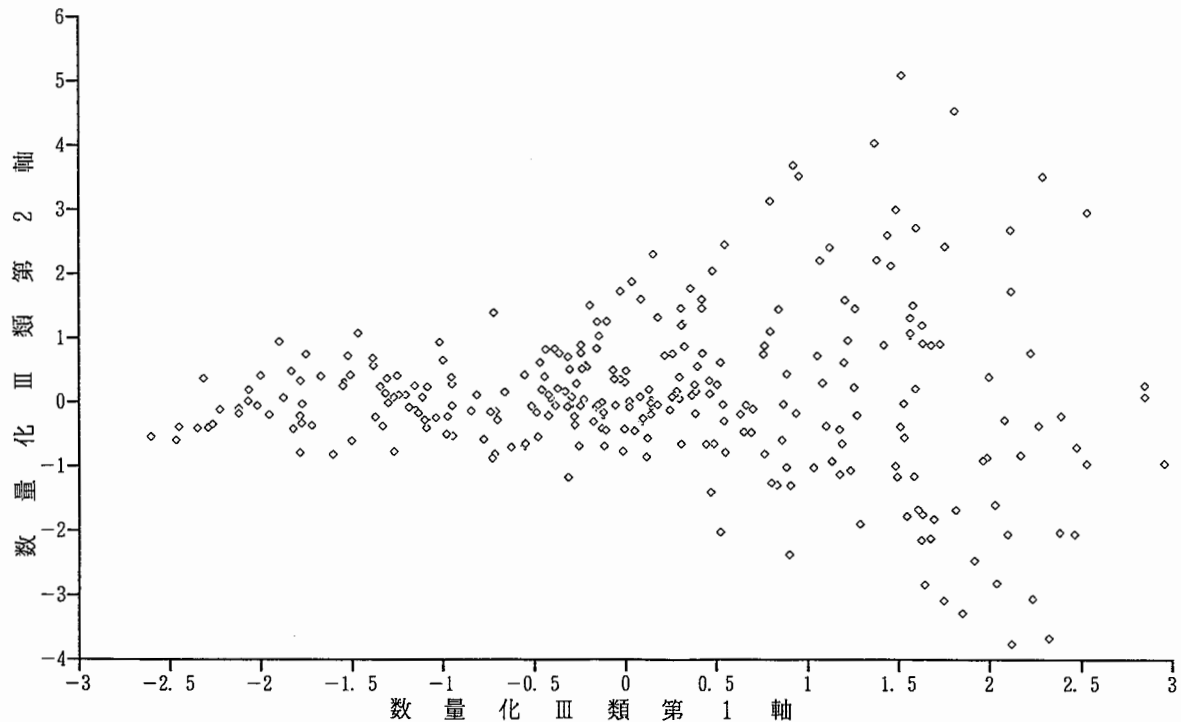


Fig. 3 Statistical analysis of keywords by the quantification method of the third type (Scatter diagram of category weight)
数量化Ⅲ類の結果に基づくキーワードの散布図 (カテゴリーウェイト散布図)

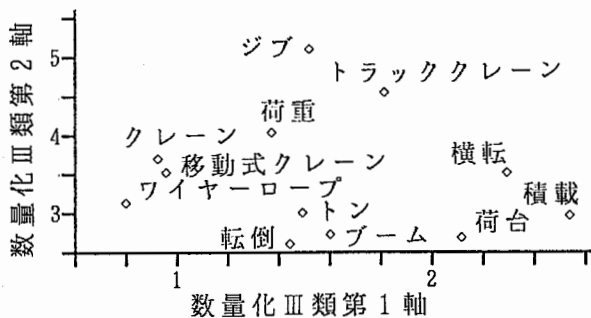


Fig. 4 Statistical analysis of keywords by the quantification method of the third type (Scatter diagram of category weight: magnified diagram No. 1: the end of upper side of Fig. 3)
数量化Ⅲ類の結果キーワード散布図
(拡大図その1: 上端)

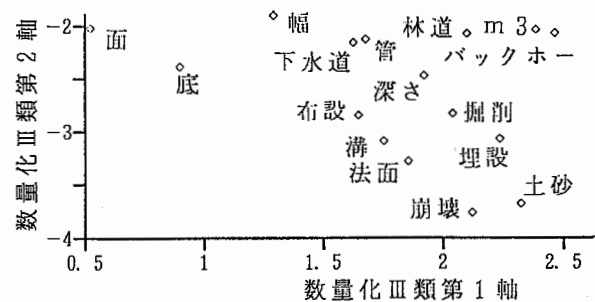


Fig. 5 Statistical analysis of keywords by the quantification method of the third type (Scatter diagram of category weight: magnified diagram No. 2: the end of bottom side of Fig. 3)
数量化Ⅲ類の結果キーワード散布図
(拡大図その2: 下端)

各プロットは前報¹⁾の場合と同様に傾き約-2の直線の近傍に位置し、そこから大きく外れるようなプロットは見あたらない。すなわちこれらのキーワードの頻度分布状況を表すプロットはブラッドフォー

ドの法則に従った分布を示していることが明らかになった。このことから、上記の切り出されたキーワードを用いてこれらの情報が有している性質等の分析・検討を進めても差し支えないと判断した。

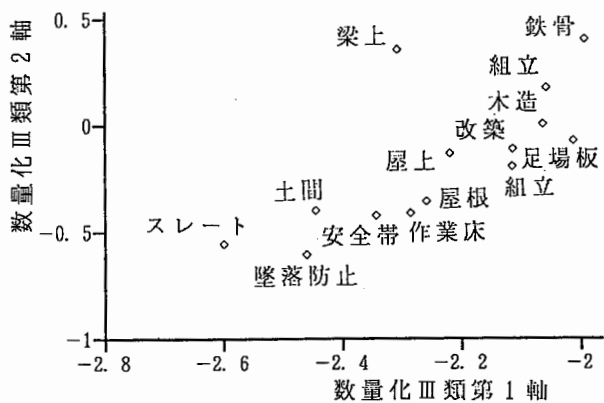


Fig. 6 Statistical analysis of keywords by the quantification method of the third type (Scatter diagram of category weight: magnified diagram No. 3: the end of left side of Fig. 3)
数量化 III 類の結果キーワード散布図
(拡大図その 3: 左端)

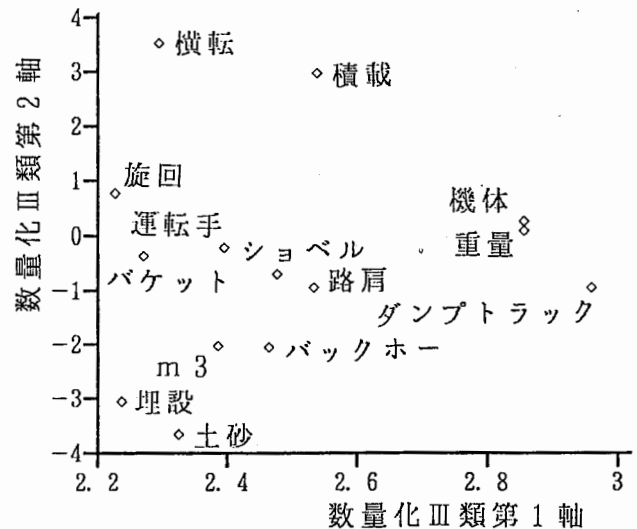


Fig. 7 Statistical analysis of keywords by the quantification method of the third type (Scatter diagram of category weight: magnified diagram No. 4: the end of right side of Fig. 3)
数量化 III 類の結果キーワード散布図
(拡大図その 4: 右端)

4.3 災害事例とキーワードとの対応表の作成

キーワードの頻度分布の状況が明らかになったところで、これらのキーワードが実際に各事例の記述の中ではどの様に使用されているのかを示す対応表を作成した。Table 4 はその対応表を模式的に示した。

この対応表の作成に関しては、キーワードを変数パラメータとして与えた上でデータベースを検索し、検索結果を別の変数パラメータとして対応表データの格納されたデータファイルを書き換えるプログラムを作成して実施した。

本表に示されたような関係からキーワードと災害事例との間の何等かの相関を得ようとする場合、数量化 III 類による分析が用いられる³⁾。

4.4 数量化 III 類による分析結果

数量化 III 類による分析に先だって、比較的高頻度のキーワード 293 語を分析対象として選択した。これらのキーワードは Table 3 から判るとおり、全キーワード種類数のうち頻度順位で上位 3% 以上に、また累積頻度で見た場合では全体の 50% 以上に相当するキーワードである。

Fig. 3 に示したのは、このように選択された累積頻度上位 293 種類のキーワードについて、数量化 III 類による分析の結果から得られたカテゴリーウェイトに基づいて示された各キーワードの散布図である。

この解析例では、大きな固有値を有する解は得ら

れなかったものの、相関関数 > 0.4 なる解は以下に示す λ_1 から λ_3 までの 3 つを得ることができた。

$$\begin{cases} \lambda_1^2 = 0.282 & (\lambda_1 = 0.531) \\ \lambda_2^2 = 0.218 & (\lambda_2 = 0.467) \\ \lambda_3^2 = 0.173 & (\lambda_3 = 0.416) \\ \lambda_4^2 = 0.150 & (\lambda_4 = 0.387) \\ \text{以下省略} \end{cases}$$

Fig. 3 の左右上下端に位置したキーワードを分かりやすくするために範囲を限って多少拡大した 4 つの図 Fig. 4 ~ 7 の中に各々位置するキーワードを明示した。

これらの図から、Fig. 3 の左端 (参照 Fig. 6) には墜落災害に関係の深いキーワードが、また図の右上 (参照 Fig. 4, Fig. 7) 方向にはクレーン関連のキーワード、図の右下 (参照 Fig. 5, Fig. 7) 方向には土砂崩壊関連のキーワードが位置することなどが判明した。すなわち、これらのキーワードの分布位置の違いから推定できるように、オリジナル資料の記述内容では災害種類の差異に関する情報を比較的明瞭に抽出できると考えられる。

また Fig. 8 および 9 には前述した数量化 III 類による分析の結果から、807 件の災害事例のサンプル

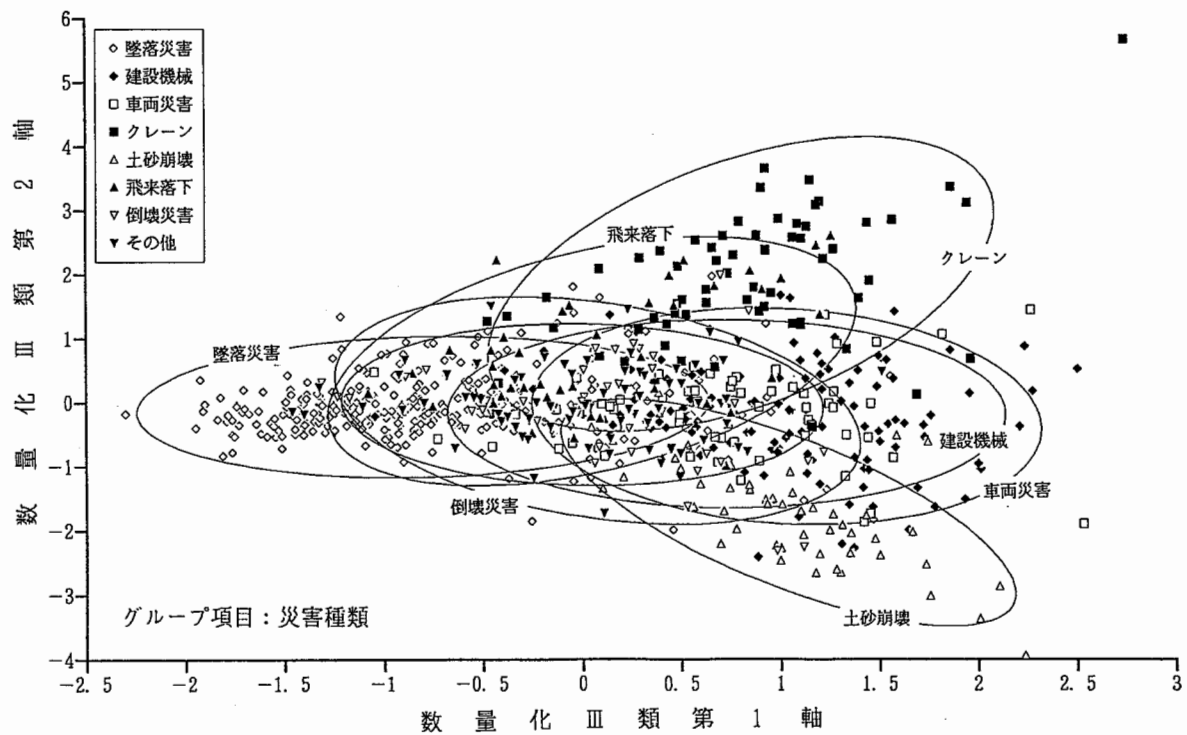


Fig. 8 Statistical analysis of keywords by the quantification method of the third type (Scatter diagram of samples No. 1, According to types of accident)

数量化 III 類の結果（その 1：災害種類別に見た災害事例のサンプルスコアに基づく散布図）

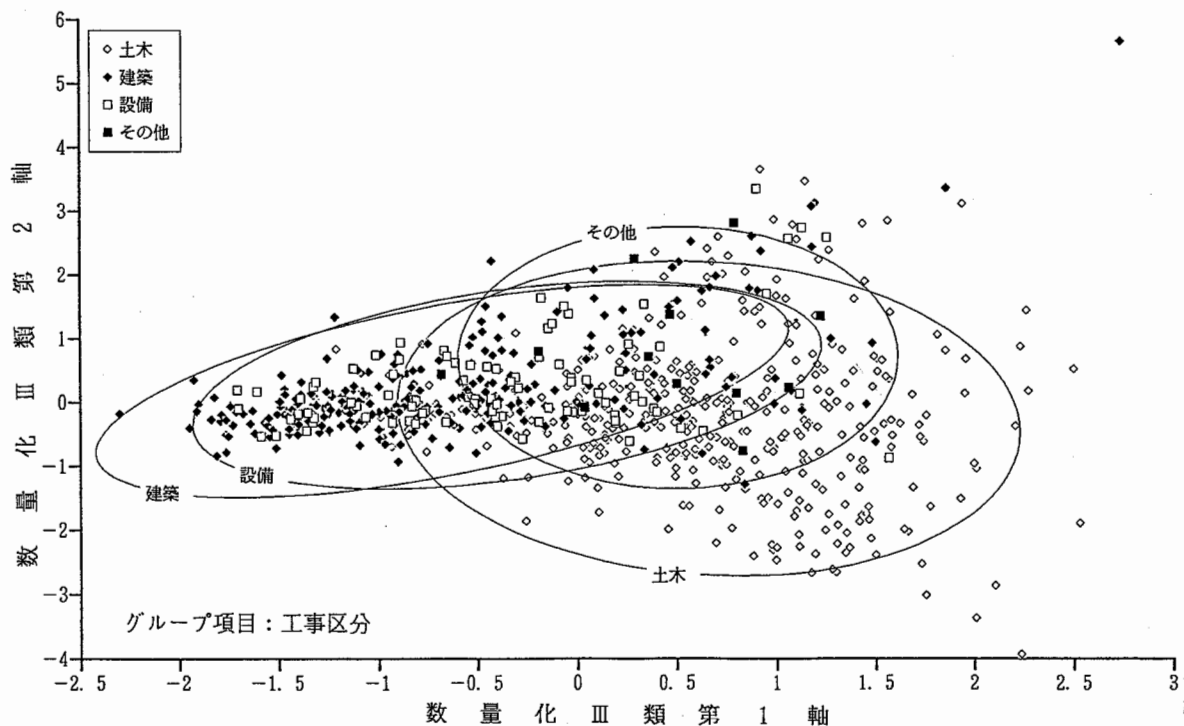


Fig. 9 Statistical analysis of keywords by the quantification method of the third type (Scatter diagram of samples No. 2, According to kinds of construction work)

数量化 III 類の結果（その 2：工事種類別に見た災害事例のサンプルスコアに基づく散布図）

Table 5 Result of discriminant analysis for types of accident
キーワードを用いた災害種類の判別結果

件 数		判 別 さ れ た 災 害 種 類							
(%)		墜落災害	建設機械	車両災害	クレーン	土砂崩壊	飛来落下	倒壊災害	合計
実 際 の 災 害 種 類	墜落災害	297 (91.7%)	2 (0.6%)	2 (0.6%)	0 (0.0%)	2 (0.6%)	3 (0.9%)	10 (3.1%)	324 (100.0%)
	建設機械	2 (1.6%)	107 (87.0%)	4 (3.3%)	1 (0.8%)	2 (1.6%)	2 (1.6%)	2 (1.6%)	123 (100.0%)
	車両災害	2 (3.0%)	0 (0.0%)	60 (90.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (1.5%)	66 (100.0%)
	クレーン	2 (3.0%)	0 (0.0%)	0 (0.0%)	61 (91.0%)	0 (0.0%)	2 (3.0%)	1 (1.5%)	67 (100.0%)
	土砂崩壊	0 (0.0%)	3 (5.6%)	0 (0.0%)	0 (0.0%)	48 (88.9%)	0 (0.0%)	2 (3.7%)	54 (100.0%)
	飛来落下	2 (4.0%)	1 (2.0%)	1 (2.0%)	1 (2.0%)	0 (0.0%)	44 (88.0%)	0 (0.0%)	50 (100.0%)
	倒壊災害	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (4.4%)	2 (4.4%)	38 (84.4%)	45 (100.0%)
	その他	3 (3.8%)	2 (2.6%)	1 (1.3%)	0 (0.0%)	0 (0.0%)	1 (1.3%)	1 (1.3%)	78 (100.0%)
	合計	308	115	68	63	54	54	55	807

スコアに基づいた散布図を示した。Fig. 8 は災害種類別にまた Fig. 9 は工事種類別にプロットの形状を区別して表示するとともに、マハラノビス汎距離 D^2 に基づいた確率楕円 (危険率 10%)^{3,4)} で各分類区分ごとの散布範囲を合わせて表示した。

Fig. 8 に示されるとおり墜落災害・クレーン関係災害・土砂崩壊災害はサンプル (災害事例) の散布位置や散布範囲が相互に比較的明瞭に異なっていることが判り、Fig. 3 ~ Fig. 7 から得られた結果を裏付けることができた。しかしそれ以外の事例については図中心付近に散布されるためこの図のみからでは明瞭に区別することは難しい。また Fig. 9 に示されるとおり、工事種類別に見たサンプル散布図では、建築工事・設備工事と土木工事では散布位置に差異がみられるものの、Fig. 8 に示された災害種類ほどには工事種類の違いを図中に明確に見い出すまでには至っていない。

また、ここでは図を省略したが、災害種類や工事種類以外の例えば、被災者の属性 (年齢・経験) あるいは災害発生日時 (曜日・時刻) などの要因別にもサンプル散布図を作成した。これらのサンプル散布図では、各々のカテゴリの細分類に基づくサンプルの散布位置や散布方向にはほとんど差異は認められず、各々の確率楕円もほとんどが互いに重複するように描かれる。すなわち、災害種類や工事種類以外の要因に関する情報を明瞭な形で抽出することはかなり難しいと判断された。

4.5 災害事例の判別結果

災害事例とその発生状況の記述中に用いられているキーワードとの関係が明らかになったところで、これらのキーワードを用いて災害事例の判別の可能性を検討した。前項までに示したように各事例の災害種類の差異についてはかなりの程度で判別が可能と

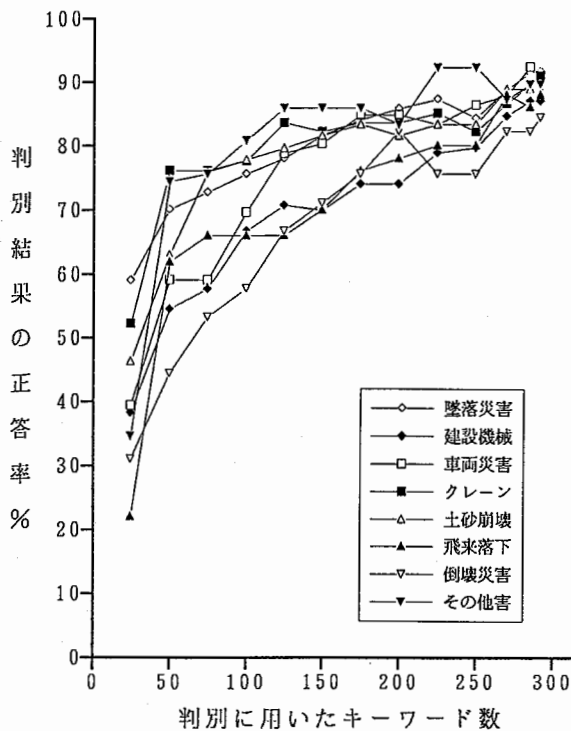


Table 6 Classification result of kinds of construction work (roughly classification) by discriminant analysis
キーワードを用いた工事種類 (大分類) の判別

件 数 (%)		判 別 さ れ た 工 事 種 類				
		土木工事	建築工事	設備工事	不明／他	合計
実 際 の 工 事 種 類	土木工事	378 (95.5%)	10 (2.5%)	8 (2.0%)	0 (0.0%)	396 (100.0%)
	建築工事	17 (5.6%)	270 (88.8%)	15 (4.9%)	2 (0.7%)	304 (100.0%)
	設備工事	7 (7.5%)	5 (5.4%)	80 (86.0%)	1 (1.1%)	93 (100.0%)
	不明／他	0 (0.0%)	0 (0.0%)	0 (0.0%)	14 (100.0%)	14 (100.0%)
	合計	402	285	103	17	807

注：高頻度キーワード 293 語（うち 168 語選択）による分析結果

Fig. 10 Change of classification results by discriminant analysis according to number of keywords
判別に用いたキーワード数と正答率の推移

考えられた。そこで、試行錯誤的に高頻度のキーワードを順次選択し、これらのキーワードを用いて災害事例の災害種類の差異についての判別の可能性を検討することとした。

Table 5 は高頻度 293 ワードを選択した際の分析結果をまとめたものである。293 ワードのうち判別効果に意味のある 172 ワードが選択された。災害種類ごとの判別の正否の程度は表中の左上から右下への対角線に位置する外に示された数値で見ることができる。各災害種類ともおしなべて 90% 前後の正答率を得ることができた。

Fig. 10 は上記の正答率の推移を分析に用いたキーワードの数ごとに追跡した結果である。キーワード数が多くなれば災害種類ごとに多少の差異はあるものの正答率は高くなる傾向が見られる。ただし正答率はキーワード数 100 ~ 150 ワード程度でほぼ一定 (80 ~ 90%) となりそれ以上キーワード数を増やしても正答率はさほど上昇しない。

またキーワードには災害種類に固有のものが存在することから、それらのキーワードの選択・非選択に

よってその災害種類についての正答率はかなりの影響を受けることがグラフから読み取れる。

以上のように、各災害事例の災害種類の差異については、高頻度の比較的少数のキーワード (293 種類、全体のおよそ 3% 程度) だけでもかなりの程度で判別が可能との結論を得た。

そこで参考までにこれらのキーワードを用いて、工事種類について全く同様に災害事例の判別を行ったところ Table 6 および Table 7 に示したような結果を得ることができた。Table 6 は土木工事・建築工事などおおまかな分類によるもの、また Table 7 はさらにそれを細分化した場合の結果で、判別に際して判別効果に意味のあるキーワードは 293 ワードのうち、各々 168 ワード、181 ワードが選択された。

工事種類については、一般的に必要とされる分類項目が通常は Table 7 に示されるように多岐にわたることもあって、これらの結果は災害種類の判別結果に見いだされたほどの正答率を得るまでには至っていないことが判った。しかし Table 6 に示されたように判別すべき項目数やキーワードの選択法次第

Table 7 Result of discriminant analysis for kinds of construction work (detail classification)
キーワードを用いた工事種類（細分類）の判別

件 数		判 別 さ れ た 災 害 種 類											
		(%)	トンネル	橋梁工事	道路工事	河川／ 港湾	砂防／ 土地整理	上下水道	他の 土木工事	ビル建設	木造家屋	建築設備 他の建築	設備工事 ／その他
実 際 の 工 事 種 類	トンネル	35 (89.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (7.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (2.6%)	39 (100.0%)
	橋梁工事	0 (0.0%)	32 (91.4%)	0 (0.0%)	0 (0.0%)	3 (8.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	35 (100.0%)
	道路工事	1 (1.0%)	3 (3.0%)	76 (76.8%)	1 (1.0%)	5 (5.1%)	1 (1.0%)	7 (7.1%)	2 (2.0%)	0 (0.0%)	0 (0.0%)	3 (3.0%)	99 (100.0%)
	河川港湾	1 (2.3%)	1 (2.3%)	3 (7.0%)	33 (76.7%)	1 (2.3%)	0 (0.0%)	1 (2.3%)	1 (2.3%)	0 (0.0%)	0 (0.0%)	2 (4.7%)	43 (100.0%)
	砂防／他	2 (3.0%)	0 (0.0%)	1 (1.5%)	0 (0.0%)	58 (87.9%)	1 (1.5%)	3 (4.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (1.5%)	66 (100.0%)
	上下水道	1 (1.5%)	0 (0.0%)	1 (1.5%)	0 (0.0%)	2 (2.9%)	52 (76.5%)	4 (5.9%)	2 (2.9%)	0 (0.0%)	1 (1.5%)	5 (7.4%)	68 (100.0%)
	他の土木	1 (2.2%)	1 (2.2%)	1 (2.2%)	2 (4.3%)	3 (6.5%)	1 (2.2%)	37 (80.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	46 (100.0%)
	ビル建築	0 (0.0%)	0 (0.0%)	2 (1.3%)	4 (2.5%)	1 (0.6%)	0 (0.0%)	3 (1.9%)	122 (77.7%)	2 (1.3%)	12 (7.6%)	11 (7.0%)	157 (100.0%)
	木造家屋	1 (1.2%)	0 (0.0%)	1 (1.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (3.5%)	6 (7.1%)	64 (75.3%)	5 (5.9%)	5 (5.9%)	85 (100.0%)
	建築設備	1 (1.6%)	0 (0.0%)	0 (0.0%)	2 (3.2%)	1 (1.6%)	0 (0.0%)	1 (1.6%)	4 (6.5%)	2 (3.2%)	48 (77.4%)	3 (4.8%)	62 (100.0%)
	設備工事	2 (1.9%)	0 (0.0%)	2 (1.9%)	0 (0.0%)	1 (0.9%)	3 (2.8%)	0 (0.0%)	4 (3.7%)	0 (0.0%)	7 (6.5%)	88 (82.2%)	107 (100.0%)
合計		45	37	87	42	75	58	62	141	68	73	119	807

注：高頻度キーワード 293 語（うち 181 語選択）による分析結果

では、これらの判別精度をもっと高めることは充分可能と考えられる。

5. あとがき

各種の建設工事で発生した労働災害の記録資料から得られる情報の有効的な活用を図る為の一環として、当該情報が有している性質・構造についての解析を試みた。

個々の災害事例の記録資料の「災害発生状況」に関する記述内容を取り上げ、オリジナルデータからフリーターム（キーワード）の切り出しを行った。これらのフリータームの頻度分布（使用頻度状況）を調べたところ、災害発生状況の記述の中でこれらのフリータームはブラッドフォードの法則に従った頻度分布を示していることが確認できた。

そこで、比較的高頻度のフリータームをキーワードとして、各災害事例とキーワードとの相関関係を数量化 III 類を用いて分析したところ、災害種類の差異に基づく情報を比較的明瞭に抽出することができた。またこれらのキーワードを用いることにより、各事例の災害種類・工事種類などの判別がかなりの程度で可能であることが判った。しかし、災害種類や工事種類以外の要因に関する情報については、これ

らを明瞭な形で抽出することはかなり難しいと判断された。

今後は、キーワード相互の親近性等の検討を加えて検索効率を考慮したキーワードの選定法やそれが判別結果に与える影響などを吟味し、労働災害の記録資料の有効活用を図るための基礎資料の充実をさらに図ってゆきたいと考えている。

（平成 3 年 3 月 30 日受理）

参 考 文 献

- 1) 鈴木芳美：トンネル工事労働災害情報の性質と構造について，産業安全研究所研究報告 RIIS-RR-91, pp. 49-58, 1991.
- 2) 鈴木芳美，前田 豊：産業安全研究所における労働災害および産業安全情報検索システムの試行開発（第 3 編労働災害調査復命書データベース），産業安全研究所安全資料 RIIS-SD-90-2, 1991.
- 3) 例えば，田中 豊，脇本和昌：多変量統計解析法，現代数学社，pp. 161-171, 1983.
- 4) 富士通：統計処理パッケージ ANALYST コマンド説明書 (V11 用).