

VI-203

建設工事労働災害のテキスト情報を用いた災害リスク要因別の事例判別

労働省産業安全研究所 正員 鈴木芳美
同 上 正員 花安繁郎

1. まえがき

労働災害の発生事例に対してその原因やリスク要因に関して作成される災害調査記録は、類似災害の再発防止ばかりでなく積極的に災害予防策を講じて行くにあたって有用な参考資料となる。したがって、これらの資料から得られる情報の有効活用については以前から多方面で要望のあるところでもある。しかし、これらの資料の有効活用のためには、まずこれらから得られる情報が有している基本的な性質や構造を把握しておく必要がある。

今回は、建設工事で発生した災害事例に関して作成された記録文書の記述内容が有している情報の基本的な性質や構造の把握を試みた。その結果として、これらの情報は災害タイプの差異に基づく情報構造を有していること、またそれに基づいた災害事例の判別がかなりの程度で可能なことなど、幾つかの知見を得ることができた。

今回はさらに、災害タイプ以外の他の要因、すなわち工事種類や被災者の属性などに着目して同様の情報解析を行った結果について報告する。

2. 分析対象

分析対象は前回¹⁾の報告と同様に、労働災害発生に伴って労働省内で作成される「災害調査復命書」と呼ばれる資料 807件である。これはある1年間（正月元旦から12月大晦日まで）で発生した建設工事における全労働災害のうち約9割に相当し、建設工事全般で発生している労働災害をほぼ網羅していると考えられる。

本資料の記載内容のうち「災害発生状況の概略」は300字程度のテキスト情報であるが、災害事例の様々なインデクスの情報としても活用出来得るものと考えられる。今回の情報解析は主にこの「災害発生状況の概略」に着目して下記に示すような分析を行った。

3. 分析内容

これまでの分析^{1), 2)}と同様に分析資料に関するデータベースを作成し、データベース機能を用いて「災害発生状況の概略」の記述中に使用されているフリータームの抽出・切出しを行った。特殊な用語や言い

回しに関して、辞書メンテナンス等を含めたチェックの実施は前回までと同様である。チェック後のフリータームをキーワードとし、数量化Ⅲ類をはじめとする多変量統計解析などを実施した。

4. 分析結果と考察**4・1 キーワードの出現頻度分布の傾向**

既に前回報告¹⁾したように、抽出された全キーワード（9059種類、平均 11.22 種類／事例）についての出現頻度およびその頻度分布の傾向に関しては、以下のことが判明している。すなわち、キーワードの出現頻度（当該キーワードを使用している事例数）とある出現頻度を有するキーワードの種類数との双方の間には、両対数紙上で直線（傾きは約-2）の関係が存在すること、したがって労働災害に関する情報の抽出にあたって多くのキーワードのうち比較的高頻度のものを選択分析しこれらを活用し得ると考えられること、などである。

4・2 キーワードの選択

今回の分析では、前項に述べたようなキーワードの頻度分布状況の検討結果を基に、キーワード全種類のうち頻度順位で上位3%以上に相当（累積頻度で見た場合は全体の50%に相当）する比較的高頻度のキーワード293語を分析対象のキーワードとして選択した。

4・3 数量化Ⅲ類による分析

これらのキーワードを用いて数量化Ⅲ類による分析を行ったところ、大きな固有値を得るまでには至らなかったものの相関係数 >0.4 なる解は3つを得ることができ、労働災害に関する情報は災害種類や工事種類に基づく構造的なことが窺えた。

さらに全807件のサンプル（災害事例）スコアに基づく散布図（図1～3）を作成し、図中には各サンプルの散布範囲を細分類のカテゴリ毎に確率楕円（危険率10%）で表示した。

図1～2などの確率楕円の重なり具合などから判る通り、労働災害に関する情報は「災害パターン」に加えて「工事区分や工事種類の差異」に関する事項が比較的抽出し易い形で包含されていると考えられ

た。一方、図3に示されるように、「被災者の経験・年齢」などの被災者属性や「災害の発生時刻・曜日」などに関する事項などに関しては、明瞭な差異を読み取ることはできなかった。

4・4 災害事例の判別可能性の検討

前項までの結果を確認するために、上記のキーワードを用いて各災害事例の判別分析を行ったところ表1に示すような判別結果が得られた。例えば、「工事区分」や「工事種類」に関しては、「災害種類」ほどではないにせよかなり良好な判別の可能性を見ることができる。しかし、他の要因に関するものについては、誤判別の割合が大きく事例相互の判別は難しいと考えられる。

5. あとがき

労働災害事例の記録文書を用いて労働災害発生状況に関するテキスト情報の解析を行った。結果として、災害種類や工事区分に関する情報の抽出やこれに基づいた事例判別は充分可能であることが判った。一方、被災者の属性に関する情報等に関してはその抽出はかなり難しいと判断され、原資料の記述内容にはこれらに関する情報内容が不十分あるいは欠如していると考えられることなどが判明した。

労働力の高齢化などの問題を抱える今日、これまでの労働災害記録では情報内容に不足する点があるとも考えられ、今後この種の記録文書のあり方を再考する必要があると思われる。

表1 高出現頻度を有するキーワード293語を用いた事例判別の結果（判別正答率）

災害種類別		工事区分別		被災者年齢別	
墜落	91.7%	土木	95.5%	20未満	93.3%
建機	87.0	建築	88.8	20才代	75.0
車両	90.9	設備	86.0	30 "	62.7
クレーン	91.0	他	100.0	40 "	59.9
土砂	88.9	工事種類別		50 "	61.2
飛来	88.0			60以上	72.5
倒壊	84.4			不明他	84.8
他	89.7			被災者経験別	
発生時刻別		トンネル	89.7%		
		橋梁	91.4		
		道路	76.8		
		河川	76.7		
		砂防	87.9	5年未満	57.9%
		水道	76.5	5～	56.7
		他土木	80.4	10～	60.0
		ビル	77.7	15～	57.6
午前	86.8%	木造	75.3	20～	59.4
午後	72.2	建築設	77.4	25～	58.2
深夜	68.3	設備	82.2	不明他	59.3
不明	100.0				

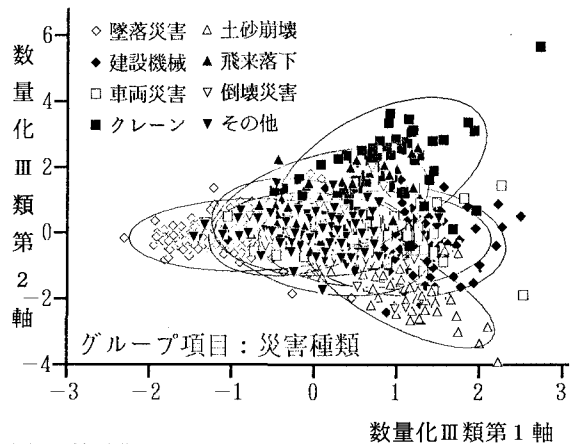


図1 数量化Ⅲ類による分析結果（災害種類別、サンプル散布図）

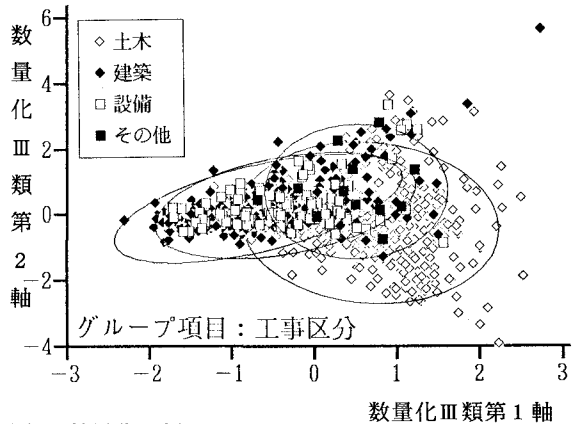


図2 数量化Ⅲ類による分析結果（工事区分別、サンプル散布図）

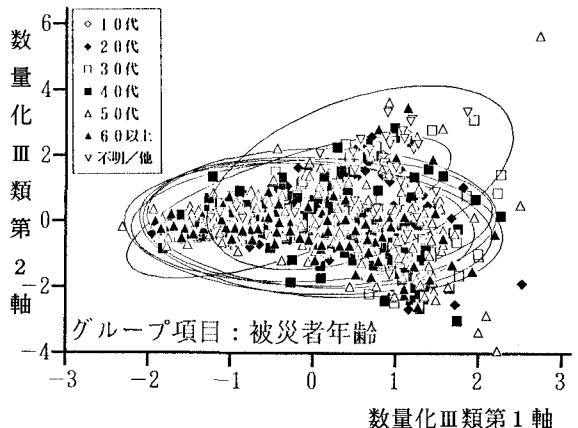


図3 数量化Ⅲ類による分析結果（被災者年齢別、サンプル散布図）