

Specific Research Reports of the National Institute  
of Industrial Safety, NIIS-SRR-NO.26 (2002)  
UDC 614.8:007:331.05:519.23

### 3. 化学プラントを対象としたヒューマンエラーに起因する災害事例の抽出手法の研究 (第2報：キーワード及び事例検索結果の数量化解析)

藤本康弘\*, 大塚輝人\*, 水谷高彰\*, 韓 宇燮\*\*

#### 3. Study on Search of Human-Error Cases in Accident Database of Chemical Plant (Second Report : Quantification Analysis on Human-error Keywords and Accident Reports)

by Yasuhiro FUJIMOTO\*, Teruhito OTSUKA\*,  
Takaaki MIZUTANI\* and Ou-Sup HAN\*\*

**Abstract:** In this paper, the results of the categorization of human-error keywords and accident reports with quantification analysis are reported. The procedures are follows;

1) Summarization of the human-error search keys with regular expressions.

The search keys were summarized from 817 to 747 words by regular expressions. For example, a chinese character and their Japanese phonetic characters are described using one regular expression.

2) Quantification analysis on the summarized keys.

The quantification analysis with the keys and the 1487 human-error keywords gave many axes with almost same eigen values. Because of the difficulty of the selection of effective 2 axes by eigen values, the alternative method was tried, which method was to pallarelize each samples (summarized keys) to the axis which has the maximum score for the sample. Four axes which pallarelized decades of samples were found, and these axes could be shown more effective than the two axes with the maximum eigen values.

3) Quantification analysis on the human-error reports searched with the summarized keywords.

The quantification with the keys and human-error accident reports could show that the effective two axes for the reports. Each axis was defined as 'Internal-factor axis' and 'Interaction-factor axis'.

Futher analysis to categorize the keys and/or the accident reports using cluster analysis would be proceeding with the result of this report which shows the improved selection of effective axes for categorize.

**Keywords;** Explosion and fire accidents, Accident database, Human error

---

\* 化学安全研究グループ Chemical Safety Research Group

\*\* 科学技術振興事業団 重点研究支援協力員 Supporting Staff for Priority Research, Japan Science and Technology Corporation

## 1. はじめに

化学プラントを対象とした危険性評価技術の検討においては、過去の災害事例について、その発生状況や原因等の貴重な情報が記録されている災害データベースの利用は欠かすことができない。例えば化学プラント災害の発生理由を検討することで、ヒューマンエラーが原因となった事例が60%にのぼることが報告されている例もある<sup>1)</sup>。

しかし、情報が蓄積され重要度を増すデータベースそのものに対して、そこから必要な情報を抜き出す手法には、いくつかの問題点が存在する。その問題点とは、例えば、上記のヒューマンエラーが原因となった事例の比率を算出した事例について考えると、ヒューマンエラーの抽象的な定義はいくつか提案されていても<sup>2), 3), 4)</sup>、それぞれの事例について、ヒューマンエラーに起因するかどうかを判定するための客観的、具体的な基準が見当たらないことに起因する。これは、すなわち、判断を担当する者の知識、経験などの能力により、ヒューマンエラーに関連しているかどうかの判断が異なりうることであり、またもうひとつは、判断する担当者が同一であったとしても、数多くの事例を判断せざるを得ない場合などには、全体を通じてその判断が一定であることを保証できないことである。

このため、現状のままでは例えばヒューマンエラーに関連した統計的な数値<sup>1)</sup>などは、再現性に疑義を持たざるを得ない。そこでまず我々は、ヒューマンエラーに関連した事例をより客観的にそして再現性を持って検索する手法の開発に取り組み、その成果を前報<sup>5)</sup>で報告した。ここでは、和和辞書を利用してヒューマンエラー関連キーワードをほぼ自動的に抽出する手法を提案し、また、得られたキーワードを用いて、災害情報データベースからヒューマンエラー関連の事例を検索する手法も提案した。これらの手法により、恣意的な要素をほぼ排除したかたちで、ヒューマンエラーに関連した事例を検索することができる。また、ヒューマンエラー災害事例の件数についての年次推移について、全災害件数が減少傾向にあるなか、ヒューマンエラー関連の災害事例件数がほぼ一定値のままであることを客観的に示すことができた (Fig. 1)。

ヒューマンエラー災害事例に関する数量化の手法を用いた報告には、例えば建設工事における墜落災害に関する報告<sup>6)</sup>があるが、化学工場の爆発、火災災害に関しては見当たらないので、本報では、前報で検索に用いられたヒューマンエラー検索キーに検

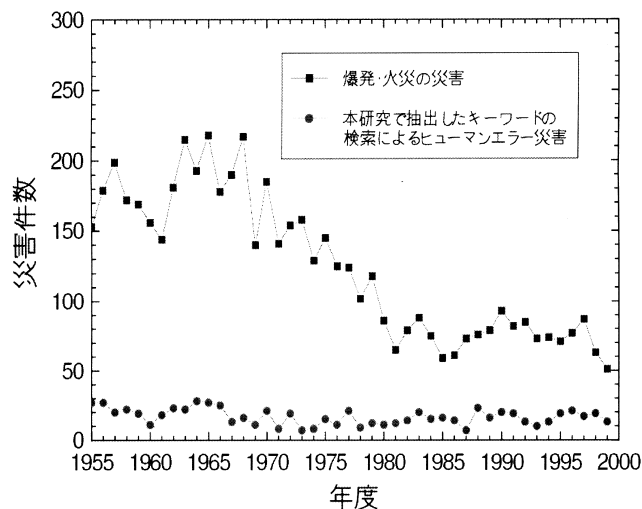


Fig. 1 Variation of human-error accidents to whole accidents in the fire and explosion database. 火災爆発災害におけるヒューマンエラー災害件数の推移

討を加え、また検索された災害事例についても数量化の手法を用いてその分類を試みたので、以下に報告する。

## 2. 語句の定義

本報で使用する語句について、以下のように定義、整理しておく。

### ヒューマンエラー (関連) 事例

形態素分解<sup>7)</sup>されたデータベースに対して、以下に定義されるヒューマンエラー検索キーを用いて検索し、ヒットした事例。ここで、検索にあたって、検索キーが複数の単語で構成されている場合、データベース内のひとつの文中(“。”で文は終了するものとする)に該当する各単語が順番に現れた場合に、ヒットしたものとする。

### ヒューマンエラー災害 (事例)

上記検索における対象データベースが、当所の(災害調査復命書情報を中心とした)爆発火災災害情報データベースである場合に検索される事例。

### (ヒューマンエラー (関連)) キーワード

前報において、和和辞書から抽出されたヒューマンエラーに関連する 1487 語

### (ヒューマンエラー) 検索キー

前報において、ヒューマンエラーキーワードを

形態素分解し、共通要素をまとめた 817 語

#### (ヒューマンエラー) 正規表現化検索キー

上記検索キーについて、かなとカナ、かなと漢字などを正規表現を用いて、まとめた 747 語。  
(詳細は次章参照)

### 3. ヒューマンエラー検索キーの正規表現化

前報において、得られたヒューマンエラーキーワードを形態素分解して共通要素をまとめて、817 語の検索キーを得ている。その数量化解析を行なうにあたり、検索キーを見直してみると、

1. 同じ読みで、同じ意味 (カタカナとひらかなと漢字) のもの  
“落ちつく ぬ”と“落ち着く ぬ”  
“脅える”と“おびえる”  
....等
2. 部分集合がキーワードとして別にあるもの  
“だめだ”と“駄目だ する”と“駄目だ なる”....  
等
3. 読みが微妙に違うが同じ意味、用途のもの  
“がっかり”と“がっくり”  
“めちゃくちゃ”と“めちゃめちゃ”
4. 否定形 “ない/ぬ” の統一

例えば、“落ちつく ぬ”と“落ち着く ぬ”は、“落ち (つ|着) く ぬ”というひとつの表現でまとめることが可能である。また、データベースを形態素分解する際の問題点として、文の構造により否定形の要素が、“ない”と表現される場合と“ぬ”で表現される場合とがある。ここでは、検索キーの方は“ぬ”で統一しておき、データベースについては、形態素分解したものをさらに“ない”“ぬ”と変換して、検索に用いることとした。ただし、“足りない”など、分解されない否定形については、“足り (ない|ぬ)”として、検索キーの方で正規表現を用いて対応した。

以上 1 ~ 4 で示される、別の検索キーとして扱うよりは、同じものとして扱うべきと思われるものについて、正規表現を用いて統一することで、検索に使われるキーは、747 語にまとめることができた。これらをヒューマンエラー正規表現化検索キーと呼ぶ。

### 4. 数量化三類による解析

#### 4.1 正規表現化検索キーの数量化解析

正規表現化検索キーとヒューマンエラーキーワードとの対応表に対して数量化三類を用いて、正規表

現化検索キーの分類について検討した。対応表は、ヒューマンエラーキーワード 1487 語を形態素分解したものについて、正規表現化検索キーを用いて検索をかけ、正規表現化検索キーそれぞれが、どのキーワードと対応しているかを調べ、検索頻度数を値として作成した。

### 正規表現化

検索キー  $\rightarrow A \Rightarrow A1 \Rightarrow A2 \Rightarrow A3 \Rightarrow$   
(検索) キーワードの  
和和辞書での展開  
 $1/2! \quad 1/3! \quad 1/4!$

Fig. 2 Parallelization of human-error keywords and human-error regularized search keys.  
正規表現化検索キーとキーワードの対応

ここで、1487 語のキーワードは、それぞれが、その抽出に用いた和和辞書を媒体として、互いに意味的に対応関係を持っている。すなわち、ある正規表現化検索キーを、キーワード 1487 語に対して検索をかけた時に、ヒューマンエラーキーワード A 1 にヒットしなかったとしても、ヒットした別のキーワード A が和和辞書を介してキーワード A 1 と関連していることがありうる。その場合、この正規表現化検索キーには、キーワード A だけでなく A 1 も対応させたほうが、検索キーの分類、解析には適していると考えられる。さらに、この和和辞書を介した関連づけは、何段階も繰り返すことが可能である。この場合、段階が増すごとに、最初のキーワードと得られたキーワードとの意味的な関連性は徐々に弱くなっていくので、各段階毎に適当な係数を与えて関連性が徐々に弱くなるようにする必要がある。最適な係数の与えかた等についての議論は省くが、ここでは各段階での正規表現化検索キーが持つ検索頻度数の最大値の大きさを目安として、段階が進む毎に最大値が  $(1/n!)$  の比率で減少していくような係数を与えることにした (Fig. 2)。すなわち、各段階での検索頻度数の最大値を、1 段階目の展開で  $1/2$ 、2 段階目で  $1/6$ 、3 段階目で  $1/24$ 、4 段階目で  $1/120$  とする。展開前の検索頻度数の最小値はその最大値の  $1/100$  より大きかったので、このようにすると、4 段階目で追加される検索頻度数の値はいずれも展開前の最小値以下となり、解析に与える影響は充分小さくなる。そこで、検索キーの展開を 3 段階目までで打ち切ることにした。このようにして、検索キ

Table 1 Top 4 axes parallelized with search keys (A result of quantification analysis using the expanded table)  
**検索キーと、検索される憲作されるキーワードを3段階まで和和辞書で展開したキーワードからなる拡張対応表を数量化三類で解析し、そこで得られた軸とその軸に対応する検索キーの数（検索キーの対応数多いもの4軸）**

対応検索キー	軸	固有値
26語	9軸	0.889
16語	15軸	0.818
14語	7軸	0.931
10語	150軸	0.384

一とキーワードとの拡張された対応表の作成を行ない、数量化三類で解析した。また、比較として展開をしない場合についても解析した。

3段階まで展開した対応表を用いて数量化解析した場合に得られる1487個の固有値は、1を除いて大きい順に、0.961, 0.945, 0.941, 0.937...となる。サンプルの分類を行なう際、固有値の大きな2軸を選んでその平面上での分布を調べることが一般に行われているが、このようにほとんど大きさの変らない固有値を持つ軸が多数ある場合には、どの2軸を選択してもその寄与率には大差がないことから、必ずしも固有値の大きな上位2軸が有効であるとは言い難いと考えた。

Table 2 Top 2 axes parallelized with search keys (A result of quantification analysis using the normal table)  
**検索キーと、検索される憲作されるキーワードを3段階まで和和辞書で展開したキーワードからなる拡張対応表を数量化三類で解析し、そこで得られた軸とその軸に対応する検索キーの数（検索キーの対応数多いもの4軸）**

対応検索キー	軸	固有値
15語	11軸	0.868
13語	7軸	0.918

そこでここでは、軸の選択手法として次のような手順を試みた。まず、それぞれのサンプル（検索キー）毎に各軸上のスコアを比較し、スコアの絶対値が最大値を取る軸にそのサンプルを対応させる。そして、各軸に対応したサンプル数の大小に注目して、

一番多くのサンプルが対応した2軸を選択する。この手法では軸の選択時点でそれぞれの軸に対応してサンプルが分布するので、特定の数本の軸にサンプルの多くがうまく対応した時には、それだけで分類ができてしまうという利点もある。得られている1487軸の中で固有値が1及び0.1以下に対応するものについては対象外とし、それ以外の軸に対応させた結果、二桁の検索キーが対応した軸が4本見つかった（Table 1）。

ここで各軸に対応したサンプル（正規表現化検索キー）の内容を眺めてみると、9軸に対応する単語は、一例では以下のようなものであり、ここでは、“怒る軸”と名づけておく。

“(向|む) かつ腹”, “堪忍袋”, “とさか くる”, “頭 (来|く) る”, “キレる” ...

同様に15軸は“ぼんやり軸”,

“成り行き”, “無頓着”, “頓着 ぬ”, “注意 足りない|ぬ”, “注意 欠ける” ...

7軸は“びっくり軸”,

“ギョッ”, “度肝 抜く”, “(怖|恐|こわ) がる”, “(脅|おび) える”, “あがる” ...

最後に、150軸は“機嫌悪い軸”と名づける。

“機嫌 悪い”, “お冠”, “虫 居所 悪い”, “ご機嫌 斜め”, “苛立”, “癩だ 障る” ...

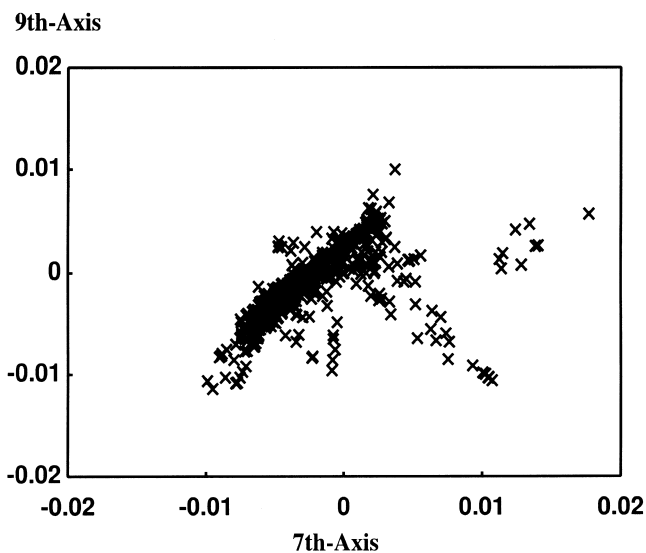


Fig. 3 Regularized search keys distribution with the 7th axis and the 9th-axis.  
**7軸と9軸を用いた正規表現化検索キーの散布図**

検索キーを展開しない対応表を用いて数量化解析した場合についても、同様に各軸に最大スコアを持

つサンプルを対応させてみると、二桁の検索キーが対応した軸が2本見つかった (Table 2)。

各軸に対応する単語の内容を評価すると、11軸には、先に“怒る軸”に含まれていた検索キーが、また7軸には、“びっくり軸”に含まれていた検索キーが集まっていた。しかし、先の検索キーを展開した対応表を用いた解析結果にくらべると、それぞれのグループの単語数は少く、関連語を集めきれていない。例えば、11軸に集まった検索キーを、先に“怒る軸”に集まった検索キーと比較すると“とさかくる”、“あたりちらす”、“イライラ”、“理性失う”などの検索キーが抜けており、今回行なった和和辞書による検索キーの展開で、本手法による軸の選択がより有効に行なわれていると考えられる。

次に、検索キーを展開した対応表から得られた数量化解析の考察で得られた4本の軸の内、大きな固有値に対応している7, 9軸について注目し、このふたつの軸を用いた平面上へ各検索キーをプロットして (Fig. 3) その分布を見た。この散布図と、1以外で最も固有値の大きな軸 (すなわち、2, 3軸) で描いた散布図を比較してみると、2, 3軸で描いた散布図では、スコア値の絶対値が大きな単語が数語あって、そのため固有値が高くなっていると思われた (Fig. 4)。実際2, 3軸からの散布図上の極端に大きなスコア値を持つ単語以外は、7, 9軸からの散布図の分布状態と比べると、ずっと小さな領域の中に集まっており、それらの単語を分類するのは困難であ

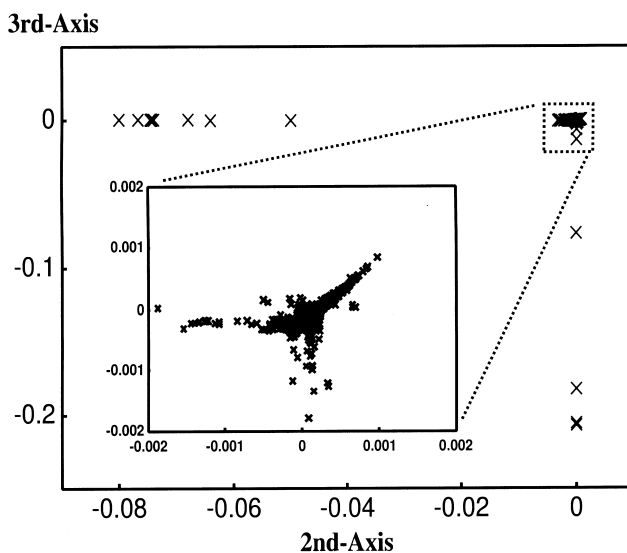


Fig. 4 Regularized search keys distribution with the 2nd-axis and the 3rd-axis.  
2軸と3軸を用いた正規表現化検索キーの散布図

る。すなわち今回選ばれた軸は、単語全体を均一に分散させており、単に固有値の大きな2軸を選びだして検討するよりも、今後の検索キーの分類の検討に有効である。

#### 4.2 災害事例検索結果を用いた数量化解析

次に、検索キーを用いて、当研究所で稼動している爆発火災災害データベースからヒューマンエラー関連の事例を検索した結果について議論する。

まず、形態素分解し、さらに“ない→ぬ”変換をした災害事例データベースに対して、747語の正規表現化検索キーを用いて検索を行なった。そこでヒットした事例について前報と同様に年次推移を調べたが、Fig. 5のようにほとんど同じ傾向を得た。この推移を5年毎にまとめてみると、1970年代以降はヒューマンエラー関連災害の件数は減少どころか、

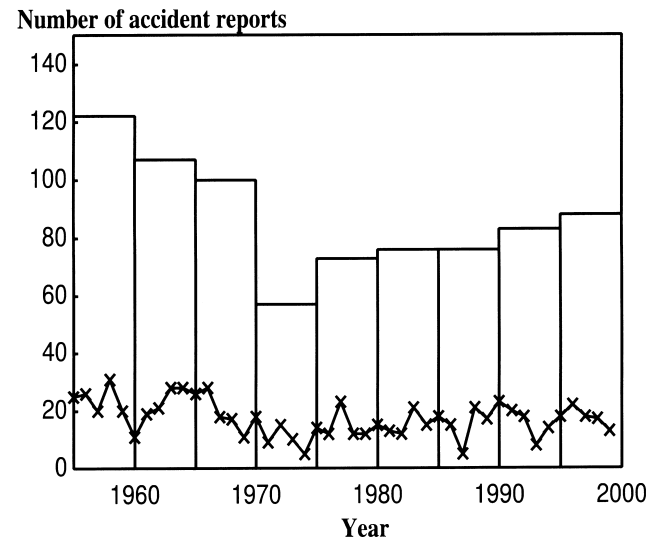


Fig. 5 Variation of human-error accidents  
ヒューマンエラー災害件数の年次推移と5年毎の件数

増加傾向にあると言ってもよいことがわかる。

このように、全体の災害件数が大幅に減少しているにもかかわらず、ヒューマンエラー関連災害の件数が横ばい、もしくは増加傾向を見せているということは、これまで行なわれてきた災害防止対策が、ヒューマンエラー災害の災害防止には必ずしも有効ではなかったことを示している。

ちなみに、正規表現化検索キーそれぞれのヒット件数を見ると、以下のような順番となり、エラーという単語に直結する誤り系以外では、“するぬ”、

Table 3 Search hit number of regularized searching keys for the fire and explosion database.  
爆発火災災害データベースに対する、正規表現化検索キーのヒット件数上位一覧

ヒット回数	正規表現化検索キー
366	する ん
166	誤る
53	転倒
23	誤り
23	ミス
22	誤認
22	急ぐ
20	する 過ぎる
19	忘れる

Table 4 Top 3 axes parallelized with accident reports. (A result of quantification analysis using the search keys and reports table)  
検索キーと、検索される災害事例からなる対応表を数量化三類で解析し、そこで得られた軸とその軸に対応する災害事例の数（災害事例の対応数多いもの3軸）

対応検索キー	軸	固有値
315事例	10軸	0.957
144事例	14軸	0.927
36事例	25軸	0.766

“転倒”，“急ぐ”などのキーが数多く災害事例にヒットして (Table 3)。したがって、例えば“転倒防止”などの具体的な対策が、ヒューマンエラー災害の防止対策として有効である可能性が高い。

ここで、数量化解析における議論で各軸に対応した検索キーの集合別に災害事例を検索を試みたが、“怒る軸”など各軸に含まれる検索キーの数が少いうえに、実際に検索に使われる単語はその内のごく一部なので、統計的に議論できるほどの数の検索結果を得ることができなかった。実際、747語の正規表現化検索キーの内、実際に当研究所nの災害事例データベースでの検索で利用された単語は、わずか87語である。そこで、この87語について検索事例との対応表を作成して数量化三類でもう一度解析を試みた。得られた各軸に対応する固有値からは、やはり軸の選択が困難であったので、これまでの解析と同様にスコアが最大になる軸に各検索キーを対応させて軸の選択を試みた。しかし、87語の検索キーは得

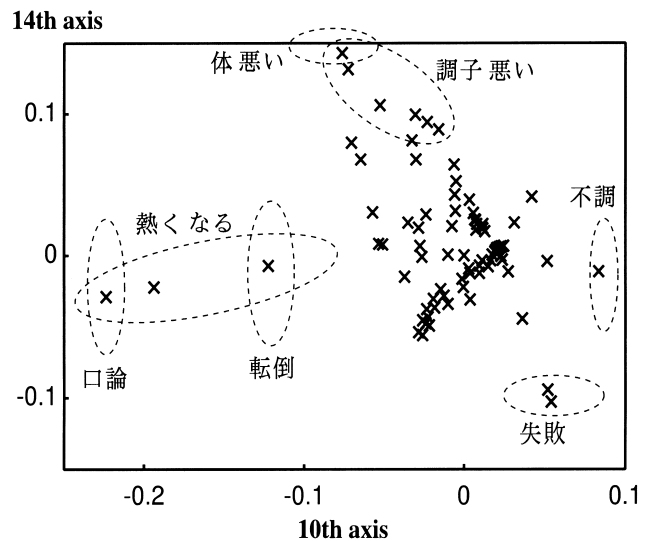


Fig. 6 Human-error accident reports distribution with the 10th-axis and the 14th-axis.  
ヒューマンエラー災害事例の散布図と検索でヒットした正規化検索キーの例

られた各軸にはほぼ均等に対応し、この手法でも適当な軸を選びだすことは困難であった。

次に視点を変えて、検索された災害事例についても、同様の手法（カテゴリスコアが最大値を持つ軸に対応させる）で軸の選択を試みると、ふたつの軸が3桁の事例に対応することが判った (Table 4)。

10軸と14軸をそれぞれX、Y軸とする平面にヒューマンエラー災害事例をプロットしたものが、Fig. 6である。

それぞれの軸の意味付けについては、今後の議論を待ちたいが、図中のスコア値の大きいいくつかの事例にヒットした正規表現化検索キーを調べてみると Fig. 6 に一部記入しているように、10軸には、“口論”，“熱くなる”，などのキーが対応しており、また14軸では、“体悪い”，“調子悪い”などのキーが対応しているので、とりあえずここでは、10軸を“外部要因-軸”，14軸を“内部要因-軸”と名づけた。

ここまでの検索結果を吟味していくと、検索事例の中には想定した意味とは違う意味で使われている事例が検索されてくる場合も少なくない。例えば、“熱くなる”という検索キーは、本来感情を表わすものとして、検索語として入っているが、“(温度が高くて)熱い”という事例にもヒットしてくる。このようなあきらかなミスヒットをどのようにフィルタリングしていくか、今後さらなる検討が必要である。

## 5. まとめ

- 1) 正規表現化検索キーとヒューマンエラーキーワードとの対応表から 3 段階まで和和辞書でキーワードを展開して得られた拡張対応表を用いて、数量化三類による解析を行なった。得られた 1487 軸について、固有値から有効な軸を選びだすことは困難であった。しかし、それぞれの検索キーについて各軸上のスコアが最大値を取る軸を対応させる手法を用いることで、検索キー分類に有効な軸の選びだしが可能になった。
- 2) 当所でまとめている災害事例データベースに対して、正規表現化検索キーを用いて検索を行った。用いた 747 語の正規表現化検索キーの内、実際の検索で利用された単語は 87 語であった。その 87 語について検索事例との対応表を用いて数量化三類による解析で分類を試みた。その結果、固有値の値からも、また最大スコアを持つ軸を対応させる手法でも検索キーを分類する有効な軸を選びだすことはできなかった。しかし、この数量化の解析の結果を用いて、災害事例に最大スコアを持つ各軸に対応させることで、災害事例を分類するためのふたつの軸を見つけだすことができた。それぞれ事例の内容から“外部要因-軸”，“内部要因-軸”と名付けた。  
今後は、ミスヒットをどのようにフィルタリングしていくかの検討が必要である。また、検索キーや災害事例に対して、分類に利用できる軸を選びだすことができたので、クラスター分析手法などを用いて分類方法をさらに検討していく。

## 参考文献

- 1) 井上紘一, 高見勲; ヒューマンエラーとその定量化, システムと制御, Vol.32, No.3, pp. 152-159 (1988).
- 2) Meister, D.; Human Factors Theory and Practice, Wiley Sons, Inc. (1971).
- 3) Swain, A.D. and Guttman, H.E.; Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Application, U.S. NRC-NUREG/CR-1278, April (1980).
- 4) Reason, J.T. ; Human Error, Cambridge University Press, Cambridge (1990).
- 5) 韓宇燮, 大塚輝人, 水谷高彰, 藤本康弘; 化学プラントを対象としたヒューマンエラーに起因する事故事例の抽出手法の研究 (第 1 報: ヒューマンエラーに起因する事例の抽出手法の開発), 産業安全研究所特別研究報告, NIIS-SRR-No.26 (2002).
- 6) 鈴木芳美, 白井伸之介, 江川義之, 庄司卓郎; 建設工事における墜落災害の人的要因に関する多変量統計解析, 産業安全研究所研究報告, NIIS-RR-97 (1998).
- 7) 松本裕治, 黒橋禎夫, 宇津呂武仁, 妙木裕, 長尾真; 日本語形態素解析システム JUMAN 使用説明書 version2.0, NAIST Technical Report, NAIST-IS-TR94025 (1994).

(平成14年 8 月 9 日 受理)