

## くらしの安全掲示板と安全情報マイスター

### くらしに役立つ情報教えてください～安全なくらしのために～

ふだんのくらしの中で、危ないと思ったことってありませんか？

例えば、電化製品を使っていてケガをしそうになったこと、日曜大工のときに道具の具合が悪くてケガをしたことはありませんか？

そうした、危ないと思ったことや、ケガをした状況などの情報があれば、ぜひ教えてください！

みんなで知恵を出し合って、安全なくらしを実現しましょう！

#### くらしの安全掲示板

身の回りで見かけた危険や、経験したケガなどの情報を、自由に書き込める二つの掲示板をつくりました（写真、動画による投稿も可能です）。

##### 1. 製品あんしん掲示板

生活の中で使用するいろいろな製品の安全性に関する苦情などが掲載されます。

##### 2. 空間あんしん掲示板

道路や公園など公共の場所の安全性に関する苦情や注意点が掲載されます。

\* 掲載される内容は「[安全性にかかわるもの](#)」が対象となります。製品性能やサービスに関連する苦情等は対象となりません。

#### 投稿いただく皆様

だれでも自由に投稿いただけます。（事業者の皆様からの投稿は御遠慮頂いております。）

消費者の皆様に加えて、福祉・保育・教育機関など公益に関わるお仕事に従事されている方々からの投稿もお待ちしております。

なお掲示板へは、ニックネームでの公開となります。（初回書込みに際して御氏名、御連絡先、ニックネームを登録いただきます。）

#### ポイント制度

くらしの安全掲示板に投稿いただくと“投稿ポイント”が付与されます。

さらに、Webサイトに公開された投稿に「役に立った！」ボタンが1クリックされるごとに“読者ポイント”が付与されます。

また、ポイントが貯まると“安全情報マイスター”に認定され、さまざまな特典をお受けいただけます。（詳細は次のページを御覧ください。）

## 安全情報マイスター

安全情報マイスターは、くらしの中の事故について、高い感性と発信力を認められた証です。

安全情報マイスターに認証されると次の特典がございます。

- 1) ヒヤリ・ハット投稿功労者に表彰されます  
獲得ポイント数や事故の未然防止への寄与の大きさなど総合的に判断し、「ヒヤリ・ハット投稿功労者」を決定します。  
※表彰者は、表彰検討委員会にて決定します。
- 2) ヒヤリ・ハットフォーラムに参加いただけます。  
ポイント獲得者の中から参加者を募集し、フォーラムを開催します。  
※参加者は、募集の上、フォーラム実行運営委員会にて決定します。
- 3) ヒヤリ・ハット講習会に優先参加いただけます。  
ポイント獲得者の中から参加者を募集し、事故未然防止の為に講習会を開催します。  
※参加者は、募集の上、講習会運営委員会にて決定します。  
講習会講師には、事業者、研究者等を招へいし、様々なテーマで開催します。

## 情報の登録方法

- ① Webサイトへの直接の書込み、
- ② 専用アドレスへのメール
- ③ 専用窓口への電話またはFAX、にてお受けいたします。

### 【 くらしの安全掲示板での情報の取扱いについて 】

#### 情報精度の確認と掲載の判断

1. 「製品あんしん掲示板」において、具体的な商品やサービス等の名称が関わる情報については、消費者等からの書込みを受けて、企業への情報の事前提示を即日行う。その情報を受理した後、企業は5営業日以内に内容確認を行い、受理日より10日以内に意見申入れを行うことができる。
2. 前項1の企業からの意見申入れが行われた場合、その内容を●●●が審査検討し、修正並びに掲示板への公開を判断する。ただし、著しく事実と異なるあるいはその恐れがある場合を除き企業の利益保護を優先して掲載の判断を歪めることはない。

#### 個人ならびに関連する情報の保護

1. 掲示板への公開に際しては、事前に登録されたニックネームを使用し、閲覧者等から個人の特定が行われないようにする。
2. 掲示板への書込みに際しては、初回の書込み時にニックネーム、氏名、連絡先の登録を行いポイント制に連動させてゆく。全ての個人情報法規に則り●●●の責務のもと管理する。

## (2) 事故情報データの品質向上に向けて (村田 磨理子)

### ア 目的

事故情報データバンクに登録されている事故情報に関して、収録されている項目や分類の状況を確認することにより、データの品質を検証する。検証結果からデータの有用性を高める方策を検討する。

データの品質を向上させることにより、

- これまで、(項目指定) 検索・集計に利用されていなかった項目について、適切に分類して検索・集計を可能とすることにより、利用者ニーズの適合性が向上し、利用の拡大が期待される。
- 項目の定義や分類基準を明確にすることによって、事故情報集計の明確性や比較可能性を向上させ、さらに統計分析を促進することが期待される。
- 一定の標準的な項目や分類基準を示すことにより、より多くの消費者からの情報提供が容易になることや、情報入力担当者の負担軽減が期待され、データバンクの効率性の向上に寄与することができる。

### イ 分析

データの収録内容について、度数分布表、クロス集計表やグラフを用いて、項目や分類に関するニーズ適合性や整合性・比較可能性を検証した。分析対象のデータは次のとおり。

- ・ 83項目 (一般向けウェブサイトでは非表示の項目を含む)
- ・ 198,135レコード (登録年月日が2009年10月1日から2016年12月31日まで)

本報告は、既存の完成されたデータファイルの内容を利用した分析の範囲にとどまるが、分析結果からは、データのコード化及び編集に直接関係するもののほかに、設計やデータの収集などに関する示唆も得られた。

#### (ア) 項目の収録状況

事故情報データバンクは、各機関で持っているデータベースの項目に基づいて事故情報を入力することを基本としているため、参画機関ごとに入力項目に差異が生じている。このため、参画機関をとおして共通して入力することとされている項目が限定されている。そこで、研究者、行政機関等利用者、一般利用者などのニーズを把握して、入力を推奨する項目を明確にすべきと考える。

本専門調査会において、自由記述の有用性を高める枠組みなどが提案されており、項目の選択にも役立つと考えられる。

分析に使用したデータにおいては、項目によっては、空白が非常に多く、全レコードの95%以上が空白である項目は40項目、そのうち13項目はすべて空白であった。空白の原因は、無記入、該当なしなどの理由が考えられるが、利用にあたっての説明が必要だと考える。

#### (イ) 収録の形式

発生日時を例にとり、収録された内容の度数分布表を作成して確認すると、次のように様々な形式になっていることが分かる。

- ・ 年のみ
- ・ 年月
- ・ 年月日
- ・ 年月日時
- ・ 年月日時分
- ・ 年月時など
- ・ 複数の時点を併記
- ・ 季節
- ・ 期間
- ・ 頻度
- ・ 発生条件（例：「夏 気温30度以上」）
- ・ 現在または基準となる時点からの経過・遡及時間（例：「購入から1年」）

事故情報のトレンドをみるとき、データの収録項目として、発生日点、受付時点、データバンクへの登録時点の3つを基準として時系列分析を行うことが想定される。しかし、発生日点は空白が約3割あることに加えて上記のように形式がまちまちであるため、使い勝手がよくない。

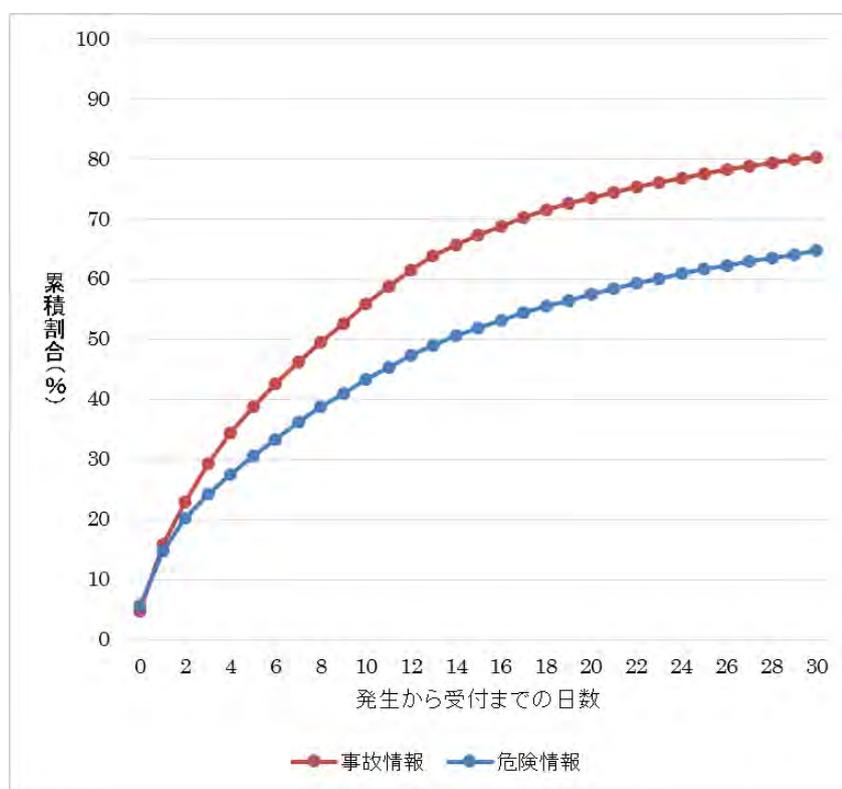
改善策として、まず、年月日と時刻は定型フォーマットとし、その他の自由記述文と分けて収録されることが望ましい。また、空白を減らすためには、最初の情報提供を受ける際に、少なくとも「年」を必須とするなどの対応が併せて必要だろう。

なお、現状のデータに対して形式統一処理をしたとすると、半数は「年月日」が利用でき、約6割は「年月」の情報が利用できると思われる。

ここでは、発生日時と受付日時の両方から年月日を取り出せた一部のデータ（82,698件）を使って、事故等の発生から受付までの日数を集計した結果を示す。（ただし、日数が利用できないものが約10万件あるため、本当の経過日数の分布を正確に推測することは困難である点に留意する必要がある。）

事故情報は発生から30日までに8割が受付され、危険情報は30日で6割程度が受付される。中央値と比較すると、事故情報は9日、危険情報は14日である。

【図表 2-6】 発生から受付までの日数別件数の累積割合



この結果は、受付時点からみると、例えば、今受付の件数が増えているということが、今発生が増えていることと必ずしも同じではないことを示唆している。特に危険情報は、4割近くは1か月以上前に発生した案件であった。なお、発生から時間が経っている案件は、類似案件の報道発表に触発されて通報するといった行動が考えられる。受付件数の増減は、報道発表との関連を併せてみるのが重要と思われる。

次に、事故情報に限定して、情報提供元（参画機関）ごとにみると、

日数の分布に違いがみられる。違いは、事故の被害者が情報提供しているのか、それとも、製品・サービスに係る事業者からの情報提供なのかといった、通報者の属性、通報の手段・経路に起因すると推察するが、現状のデータの範囲では明らかにできなかった。違いの要因となる項目をデータに追加することで、分析が深まると期待される。少なくとも、データの利用者に対して、参画機関ごとの特徴を説明することが望ましいのではないか。

#### (ウ) クロス集計

被害者の年代、性別といった属性と、商品や事故内容のクロス集計によって、事故の特徴を分析したいが、現状のデータでは、項目ごとの空白の多さや収録形式の課題があり、クロス集計が容易ではない。

一例として、被害者性別の空白が比較的少ない情報提供元に限定して、被害者性別と商品の分類とのクロス集計を試みた。

【図表 2-7】被害者性別と商品分類とのクロス集計

	男性	女性	不明、空白
食料品	6,975	12,535	7,793
家電製品	1,133	1,628	8,616
住居品	3,082	5,789	8,287
文具・娯楽用品	1,339	1,517	1,718
光熱水品	187	260	1,125
被服品	875	2,247	595
保健衛生品	4,831	26,468	2,136
車両・乗り物	1,618	1,355	8,051
建物・設備	2,194	4,362	2,860
保健・福祉サービス	3,881	17,445	667
他の商品・サービス	3,959	6,659	5,117
総数	22,993	60,750	37,620

(注) 商品など分類は複数に該当するため、内訳の和は総数に一致しない。

総数では、男性22,993人、女性60,750人、不明・空白37,620人であり、女性が男性の2.6倍となっている。商品など分類ごとにみると、男女比の偏りが小さいものから大きいものまでさまざまであり、保健衛生品及び保健・福祉サービスにおける女性の多さが際立っている。

一方で、この結果において、女性が男性の2.6倍となっていることが、

全体で発生している事故の比率と同じとみることは危険である。実際に発生している事故等のうち、情報提供される割合が偏っている可能性が高く、バイアスがあると考えられる。

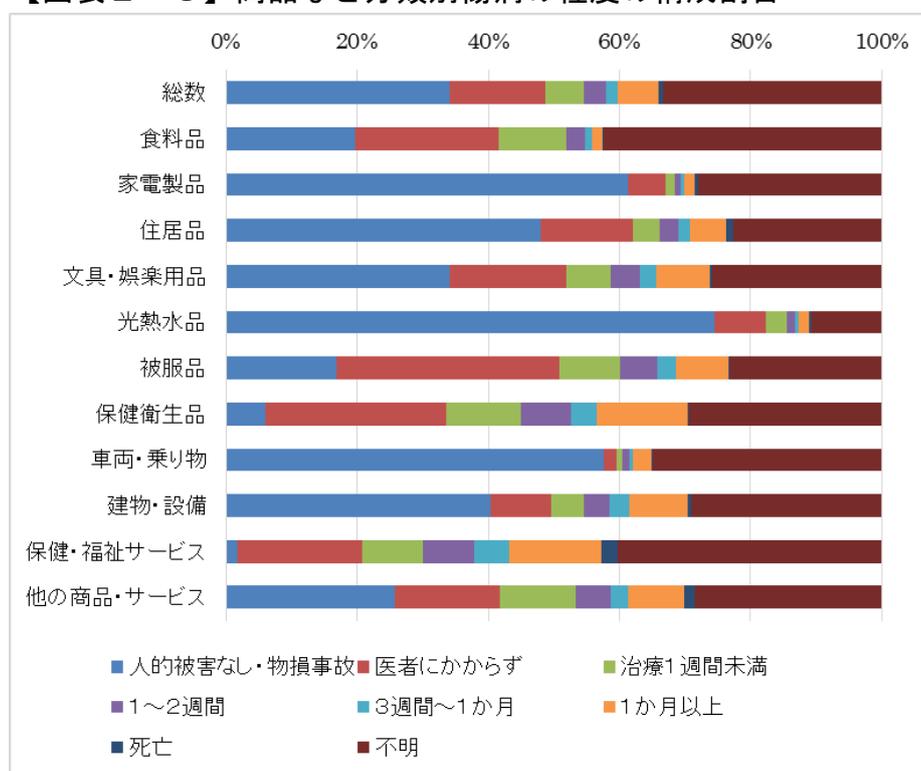
### (エ) 項目の区分

事故情報には傷病の程度の収録が比較的多いが、危険情報ではほぼ空白となっている。危険情報の定義は「けが人が発生していないものの、発生するおそれがあった事案等に関する情報」であるため、傷病の程度が空白であると考えられるが、「人的被害なし・物損事故」といった区分を追加して、事故情報と一貫した順序尺度をもつ区分とすることによって、傷病の程度の分析の幅が広がると思われる。

ここでは、危険情報の傷病の程度をすべて「人的被害なし・物損事故」とみなして、事故情報と併せて、商品など分類別に集計した結果を示す。

保健・福祉サービス、保健衛生品、被服品、食料品は人的被害が出ている割合が高く、保健・福祉サービス、保健衛生品は1か月以上・死亡の割合も比較的高いこと、光熱水品、家電製品、車両・乗り物は人的被害なし・物損事故の割合が高いことが示される。

【図表 2-8】商品など分類別傷病の程度の構成割合



ただし、この結果についても、前述のとおり、全体で発生している事故の傷病の程度と同じとみることは危険であると考えられる。情報提供される割合が商品など分類によって偏っている可能性がある。

#### (オ) 自由記述の扱いについて

情報提供元によっては、商品や事故内容の分類ルールを詳細に設定していることを確認した。しかし、複数の情報提供元の間で、分類ルールが統一されているかは不明であり、データをそのまま一括して使うと、分類の揺らぎが表面化していると推察する。

そこで、事故情報データベースシステムにおいて、統一的なアフターコード化を実施することを提案する。(公的統計の作成では、例えば、独立行政法人統計センターにおいて、経済センサスの産業分類や社会生活基本調査の生活行動分類などについて、自由記述文から自動的にコーディングする方法が研究され、実用化されている。)本専門調査会においても、言語解析技術や機械学習を活用した分析を行っており、同技術を活用することで、統一的なアフターコード化を確立できると思われる。

分類の揺らぎのほかに、適切な分類区分を網羅しているかという点からみると、例えば、事故内容は、「その他」が4割を占める。「その他」の割合は、情報提供元によって、大きく異なる。そこで、情報提供元による分類の違いを考慮に入れて、統一的なアフターコード化により、「その他」の細分化を検討する余地があると思われる。

また、現状では空白が多い事故原因、措置状況に相当する内容が、別の項目の自由記述文に含まれる場合や、関連する2つの項目が必ずしも整合していない場合などが散見されており、統一的なアフターコード化の対象とすることにより、空白を減らし整合性のある有用性の高いデータとすることが期待される。

#### (カ) データ収録内容に関するその他の課題

次に挙げる点を解消することで、データの有用性を高めることができると思われる。

- 「商品など分類」のように、複数の区分に該当するような場合のデータ構造を工夫する。(集計用のデータでは、区分ごとに「該当する、しない」のフラグを入れる方法がよく使われている。)
- 「商品など分類」の大分類、中分類のように異なるレベルで分類す

るものは、大分類と中分類をそれぞれ別の項目として収録する。

- 「被害者の情報 年代等一年代」の年齢階級は、利用者ニーズと適合しているか確認する。
- 「被害者の情報 年代等一属性」の使い方を明確にする。（現状の収録内容から推察すると、「要介護者」といった情報を補足するものであることを明確にするなど）
- 「原因究明等一原因調査状況」と「原因究明等一事故原因」の項目間の関連を明確にする。（例えば、「原因究明等一事故原因」に記載があれば、「原因究明等一原因調査状況」は空白にしないなど）
- 「発生場所」、「発生場所一施設用途」、「発生場所一施設名」、「発生場所一発生場所」は、空白がかなり多く、4つの項目の使い分けが難しいと感じる。特に、「発生場所一施設名」は、「発生場所一施設用途」との混同が多い。
- 「発生場所」に番地までの住所が必要かどうか。

#### ウ 今後の課題

分析結果から得られた、データの品質向上のための課題を下記にまとめる。

- 自由記述文とコード化された項目を明確に分ける。
- 自由記述文からの分類、コード化は情報提供元による定義の違いを考慮する。可能ならば、事故情報データバンクシステム独自のアフターコード化を行う。
- 項目間の関連性を明確に定義する。
- 利用者のニーズに応じた情報提供として、事故原因、措置状況、被害者の性別など、現状では実質的に利用できない項目について、ニーズに応じてさらなる提供を検討する。
- データの説明文書を充実する。（事故情報登録作業向けと、データ利用者向け）

### (3) 事故情報の活用等のあり方について 言語解析技術（相澤 彰子）

#### ア 目的

事故情報データベースに登録されている事故情報に関して、言語解析技術を活用した分析を行うことにより、分類項目の網羅性や整合性を調査する。また、自由記述項目に対する言語解析の適用可能性を調べ、効率的なデータ収集・管理・分析に向けた課題の整理を行う。

#### イ 分析する事故情報の分野

まず、事故情報データベースに登録されたすべてのレコード198,135件について、俯瞰的な調査を実施する（以下、「分類項目ごとのデータ登録状況に関する俯瞰的な調査」と呼ぶ）。次に、レコード件数や自由記述項目に登録されたテキストの量を勘案し、《情報提供元》が《国民生活センター》で、《事故種別》が《事故情報》である85,778件をサンプルとして抽出し、言語解析を含む詳細な分析を行う（以下、「言語解析の適用可能性に関する調査」と呼ぶ）。

#### ウ 分析に用いられた技術の概要

##### (ア) 分類項目ごとのデータ登録状況に関する俯瞰的な調査

全事故情報 198,135件を Unicode (UTF-8) に変換して、83個の分類項目について、「辞書サイズ」、「網羅率」、「平均バイト数」を調べた。ここで「辞書サイズ」とは、項目ごとの登録文字列の異なり数とする。たとえば分類項目《種別》に対しては、《事故情報》と《危険情報》の2つの値が登録されていることから、辞書サイズは2となる。分類項目《事故内容詳細》のように自由記述項目である場合には、登録文字列は原則として互いに異なるため、辞書サイズはレコード数に近くなる。また、「網羅率」とは、値が登録されていない、いわゆる欠損値を除くデータの割合と定義する。「平均バイト数」は、項目ごとに登録されているテキストの長さの平均で、欠損値を除いて計算する。調査は UTF-8 で符号化されたテキストを対象に行ったため、日本語については1文字3バイトで計算される。

調査の結果、分類項目ごとのデータ登録状況は、情報提供元によって大きく異なることが判明したため、情報提供元URLが入力されていて登録件数が1,000件以上ある5つの情報提供元について、それぞれの数値を求め、図表2-9のとおり比較可能な形に整理をした。

【図表2-9】分類項目ごとのデータ登録状況比較

番号	フィールド名	全体			参画機関①			参画機関②		
		辞書サイズ	網羅率	平均 バイト数	辞書サイズ	網羅率	平均 バイト数	辞書サイズ	網羅率	平均 バイト数
0	No	9,939	1.00	3.8	9,939	1.00	3.8	5,824	1.00	4.0
1	事故情報ID	198,135	1.00	9.0	121,363	1.00	8.9	8,308	1.00	9.2
2	事故概要	42,985	1.00	22.5	30,651	1.00	17.3	5,442	1.00	52.5
3	種別	2	1.00	12.0	2	1.00	12.0	1	1.00	12.0
4	発生日時	19,250	0.65	16.6	14,097	0.49	16.4	2,994	1.00	16.8
5	発生場所	219	0.46	9.4	48	0.49	9.3	47	0.87	9.3
6	商品など分類	2,177	0.98	42.1	2,170	1.00	49.5	1	1.00	27.0
7	商品など名称	55,808	1.00	29.5	43,355	1.00	24.2	5,459	1.00	70.1
8	事故内容	20	0.92	13.0	17	1.00	13.3	1	1.00	12.0
9	事故内容詳細	167,882	0.94	610.2	121,316	1.00	865.6	0	0.00	0.0
10	傷病内容	28	0.52	18.2	20	0.71	19.6	1	1.00	9.0
11	傷病の程度	8	0.45	15.3	7	0.71	15.4	0	0.00	0.0
12	措置状況	3	0.06	9.1	0	0.00	0.0	0	0.00	0.0
13	情報提供元	13	1.00	27.4	1	1.00	24.0	1	1.00	15.0
14	情報提供元URL	7	0.88	37.8	1	1.00	39.0	1	1.00	38.0
15	登録年月日	2,432	1.00	16.7	2,387	1.00	16.7	48	1.00	16.8
16	備考	125,713	0.74	424.0	116,475	0.99	506.9	0	0.00	0.0
17	通知方法	6	0.65	6.1	3	0.97	6.0	0	0.00	0.0
18	被害者の情報-人数	277	0.10	4.3	0	0.00	0.0	247	1.00	4.5
19	被害者の情報-氏名	5	0.00	9.6	0	0.00	0.0	0	0.00	0.0
20	被害者の情報-所属-職業	20	0.62	11.9	10	1.00	11.8	0	0.00	0.0
21	被害者の情報-所属-住所	197	0.42	9.4	49	0.69	9.3	0	0.00	0.0
22	被害者の情報-年代等-年代	51	0.41	8.2	11	0.64	8.2	0	0.00	0.0
23	被害者の情報-年代等-属性	6	0.00	9.3	0	0.00	0.0	0	0.00	0.0
24	被害者の情報-性別	3	0.45	6.0	3	0.70	6.0	0	0.00	0.0
25	発生場所-施設用途	41	0.53	12.5	9	0.67	11.8	24	1.00	20.2
26	発生場所-施設名	62	0.00	14.0	0	0.00	0.0	0	0.00	0.0
27	発生場所-発生場所	141	0.08	9.4	10	0.12	7.3	0	0.00	0.0
28	型式・ロット・生産国	68,282	0.48	57.2	32,571	0.36	32.1	0	0.00	0.0
29	傷病内容詳細	340	0.01	24.4	0	0.00	0.0	0	0.00	0.0
30	傷病の程度詳細	296	0.09	18.7	0	0.00	0.0	0	0.00	0.0
31	事業者-事業者名1	47,913	0.69	21.7	39,887	0.85	18.0	0	0.00	0.0
32	事業者-事業者区分1	1,304	0.12	18.4	1,294	0.02	44.2	0	0.00	0.0
33	事業担当部署-所在地1	5	0.00	47.0	0	0.00	0.0	0	0.00	0.0
34	事業担当部署-電話番号1	5	0.00	11.6	0	0.00	0.0	0	0.00	0.0
35	事業担当部署-担当部署1	3	0.00	20.0	0	0.00	0.0	0	0.00	0.0
36	事業担当部署-担当者役職1	2	0.00	7.5	0	0.00	0.0	0	0.00	0.0
37	事業担当部署-氏名1	3	0.00	8.0	0	0.00	0.0	0	0.00	0.0
38	製造・輸入時期1	2	0.00	15.5	0	0.00	0.0	0	0.00	0.0
39	購入・販売時期1	5,087	0.45	13.0	5,084	0.73	13.0	0	0.00	0.0
40	設置期間1	1	0.00	39.0	0	0.00	0.0	0	0.00	0.0
41	使用期間1	425	0.23	10.8	37	0.22	13.5	0	0.00	0.0
42	事業者-事業者名2	9,468	0.14	17.1	9,052	0.22	16.9	0	0.00	0.0
43	事業者-事業者区分2	119	0.00	19.2	111	0.00	37.9	0	0.00	0.0
44	事業担当部署-所在地2	3	0.00	45.7	0	0.00	0.0	0	0.00	0.0
45	事業担当部署-電話番号2	1	0.00	12.0	0	0.00	0.0	0	0.00	0.0
46	事業担当部署-担当部署2	0	0.00	12.0	0	0.00	0.0	0	0.00	0.0
47	事業担当部署-担当者役職2	1	0.00	15.0	0	0.00	0.0	0	0.00	0.0
48	事業担当部署-氏名2	1	0.00	15.0	0	0.00	0.0	0	0.00	0.0
49	製造・輸入時期2	0	0.00	15.0	0	0.00	0.0	0	0.00	0.0
50	購入・販売時期2	0	0.00	15.0	0	0.00	0.0	0	0.00	0.0
51	設置期間2	0	0.00	15.0	0	0.00	0.0	0	0.00	0.0
52	使用期間2	1	0.00	11.0	0	0.00	0.0	0	0.00	0.0
53	事業者-事業者名3	16	0.00	20.8	0	0.00	0.0	0	0.00	0.0
54	事業者-事業者区分3	6	0.00	16.1	0	0.00	0.0	0	0.00	0.0
55	事業担当部署-所在地3	0	0.00	16.1	0	0.00	0.0	0	0.00	0.0
56	事業担当部署-電話番号3	0	0.00	16.1	0	0.00	0.0	0	0.00	0.0
57	事業担当部署-担当部署3	0	0.00	16.1	0	0.00	0.0	0	0.00	0.0
58	事業担当部署-担当者役職3	0	0.00	16.1	0	0.00	0.0	0	0.00	0.0
59	事業担当部署-氏名3	0	0.00	16.1	0	0.00	0.0	0	0.00	0.0
60	製造・輸入時期3	0	0.00	16.1	0	0.00	0.0	0	0.00	0.0
61	購入・販売時期3	0	0.00	16.1	0	0.00	0.0	0	0.00	0.0
62	設置期間3	0	0.00	16.1	0	0.00	0.0	0	0.00	0.0
63	使用期間3	0	0.00	16.1	0	0.00	0.0	0	0.00	0.0
64	措置状況詳細	6,961	0.15	231.2	0	0.00	0.0	0	0.00	0.0
65	原因究明等-原因調査状況	5	0.83	2.2	1	1.00	1.0	1	1.00	1.0
66	原因究明等-事故原因	31	0.13	27.5	0	0.00	0.0	0	0.00	0.0
67	原因究明等-事故原因詳細	24,992	0.21	348.0	0	0.00	0.0	0	0.00	0.0
68	原因調査機関-機関名	6	0.00	32.2	0	0.00	0.0	0	0.00	0.0
69	原因調査機関-連絡先	11	0.00	25.3	0	0.00	0.0	0	0.00	0.0
70	通報者情報-氏名	12	0.00	20.5	0	0.00	0.0	0	0.00	0.0
71	通報者情報-属性	753	0.43	17.0	1	0.55	15.0	0	0.00	0.0
72	通報者情報-職業	19	0.61	13.7	12	1.00	13.7	0	0.00	0.0
73	通報者情報-住所	78	0.67	9.3	50	1.00	9.3	0	0.00	0.0
74	通報者情報-連絡先-TEL	5	0.00	16.6	0	0.00	0.0	0	0.00	0.0
75	通報者情報-連絡先-FAX	3	0.00	19.3	0	0.00	0.0	0	0.00	0.0
76	通報者情報-連絡先-E-MAIL	2	0.00	31.5	0	0.00	0.0	0	0.00	0.0
77	最終更新年月日	2,487	1.00	16.7	2,318	1.00	16.7	60	1.00	16.8
78	受付年月日	21,972	0.90	19.9	18,783	1.00	21.4	0	0.00	0.0
79	情報元システムID	198,101	1.00	13.6	121,363	1.00	16.0	8,308	1.00	11.0
80	情報提供機関名、担当課名-所属	1,665	0.88	30.6	1,505	1.00	38.1	120	1.00	9.8
81	情報提供機関名、担当課名-氏名	25	0.00	15.0	0	0.00	0.0	0	0.00	0.0
82	情報提供機関名、担当課名-TEL	23	0.00	9.6	0	0.00	0.0	0	0.00	0.0

番号	フィールド名	参画機関③			参画機関④			参画機関⑤		
		辞書サイズ	網羅率	平均 バイト数	辞書サイズ	網羅率	平均 バイト数	辞書サイズ	網羅率	平均 バイト数
0	No	5,955	1.00	4.0	7,009	1.00	3.8	4,911	1.00	3.8
1	事故情報ID	20,060	1.00	9.5	17,653	1.00	8.8	7,901	1.00	9.0
2	事故概要	1,992	1.00	42.8	4,761	1.00	27.5	1,274	1.00	25.5
3	種別	2	1.00	12.0	2	0.99	12.0	1	1.00	12.0
4	発生日時	4,366	0.98	15.6	3,165	0.86	16.8	2,906	1.00	16.7
5	発生場所	0	0.00	0.0	65	0.97	9.4	47	1.00	9.3
6	商品など分類	25	0.84	33.9	56	1.00	31.8	32	1.00	30.5
7	商品など名称	2,016	1.00	48.7	4,783	1.00	28.2	1,280	1.00	25.7
8	事故内容	2	1.00	10.9	20	1.00	14.3	12	0.90	12.0
9	事故内容詳細	11,800	1.00	107.4	7,780	0.79	96.5	4,858	1.00	113.8
10	傷病内容	0	0.00	0.0	23	0.37	11.3	14	0.10	11.0
11	傷病の程度	2	0.01	10.4	7	0.08	13.7	2	0.14	12.6
12	措置状況	0	0.00	0.0	2	0.31	9.1	3	0.72	9.1
13	情報提供元	1	1.00	40.0	1	1.00	12.0	1	1.00	48.0
14	情報提供元URL	1	1.00	27.0	1	1.00	38.0	1	1.00	47.0
15	登録年月日	28	1.00	16.9	583	1.00	16.7	710	1.00	16.7
16	備考	0	0.00	0.0	1,109	1.00	49.5	7,901	1.00	15.1
17	通知方法	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
18	被害者の情報-人数	0	0.00	0.0	154	0.41	4.3	10	0.29	4.0
19	被害者の情報-氏名	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
20	被害者の情報-所属-職業	0	0.00	0.0	16	0.00	15.3	0	0.00	0.0
21	被害者の情報-所属-住所	0	0.00	0.0	30	0.01	9.5	0	0.00	0.0
22	被害者の情報-年代等-年代	0	0.00	0.0	41	0.13	8.7	11	0.02	8.8
23	被害者の情報-年代等-属性	0	0.00	0.0	1	0.00	24.0	0	0.00	0.0
24	被害者の情報-性別	0	0.00	0.0	3	0.13	6.0	3	0.02	6.0
25	発生場所-施設用途	0	0.00	0.0	10	0.79	12.2	0	0.00	0.0
26	発生場所-施設名	0	0.00	0.0	50	0.01	13.6	0	0.00	0.0
27	発生場所-発生場所	0	0.00	0.0	24	0.01	10.3	0	0.00	0.0
28	型式・ロット・生産国	3,995	0.65	37.2	9,201	0.66	27.2	4,342	0.68	15.0
29	傷病内容詳細	0	0.00	0.0	171	0.01	18.8	0	0.00	0.0
30	傷病の程度詳細	1	0.20	6.0	186	0.11	12.0	66	0.28	10.5
31	事業者-事業者名1	1,357	0.65	32.0	6,414	0.87	30.5	1,464	0.67	42.2
32	事業者-事業区分1	0	0.00	0.0	8	0.86	15.3	3	0.66	15.0
33	事業担当部署-所在地1	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
34	事業担当部署-電話番号1	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
35	事業担当部署-担当部署1	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
36	事業担当部署-担当者役職1	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
37	事業担当部署-氏名1	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
38	製造・輸入時期1	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
39	購入・販売時期1	0	0.00	0.0	2	0.00	17.0	0	0.00	0.0
40	設置期間1	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
41	使用期間1	385	1.00	7.4	0	0.00	0.0	0	0.00	0.0
42	事業者-事業者名2	0	0.00	0.0	468	0.04	27.9	2	0.00	33.0
43	事業者-事業区分2	0	0.00	0.0	8	0.04	15.5	1	0.00	15.0
44	事業担当部署-所在地2	0	0.00	0.0	1	0.00	27.0	0	0.00	0.0
45	事業担当部署-電話番号2	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
46	事業担当部署-担当部署2	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
47	事業担当部署-担当者役職2	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
48	事業担当部署-氏名2	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
49	製造・輸入時期2	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
50	購入・販売時期2	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
51	設置期間2	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
52	使用期間2	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
53	事業者-事業者名3	0	0.00	0.0	16	0.00	20.8	0	0.00	0.0
54	事業者-事業区分3	0	0.00	0.0	6	0.00	16.1	0	0.00	0.0
55	事業担当部署-所在地3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
56	事業担当部署-電話番号3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
57	事業担当部署-担当部署3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
58	事業担当部署-担当者役職3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
59	事業担当部署-氏名3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
60	製造・輸入時期3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
61	購入・販売時期3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
62	設置期間3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
63	使用期間3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
64	措置状況詳細	4,679	1.00	253.8	1,282	0.23	185.2	1,014	0.74	186.0
65	原因究明等-原因調査状況	4	1.00	6.8	4	0.34	6.3	3	0.88	6.1
66	原因究明等-事故原因	16	1.00	25.2	29	0.27	36.3	9	0.06	45.9
67	原因究明等-事故原因詳細	9,816	1.00	349.2	1,462	0.32	130.6	5,616	0.90	476.3
68	原因調査機関-機関名	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
69	原因調査機関-連絡先	0	0.00	0.0	0	0.00	0.0	9	0.00	28.4
70	通報者情報-氏名	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
71	通報者情報-属性	0	0.00	0.0	733	1.00	24.7	0	0.00	0.0
72	通報者情報-職業	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
73	通報者情報-住所	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
74	通報者情報-連絡先-TEL	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
75	通報者情報-連絡先-FAX	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
76	通報者情報-連絡先-E-MAIL	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
77	最終更新年月日	31	1.00	16.9	667	1.00	16.7	217	1.00	16.8
78	受付年月日	2,135	1.00	16.9	1,844	1.00	16.7	1,917	1.00	16.7
79	情報元システムID	20,060	1.00	9.9	17,653	1.00	10.9	7,901	1.00	10.0
80	情報提供機関名、担当課名-所属	18	1.00	16.2	3	0.99	12.2	1	1.00	15.0
81	情報提供機関名、担当課名-氏名	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0
82	情報提供機関名、担当課名-TEL	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0

次に、上記で得られた情報に基づき、人手により分類項目をタイプ別に分類した。図表2-10に、この分類作業のために設定したタイプ種別をまとめる。タイプの分類にあたっては、辞書サイズおよび平均バイト数を参照しながら、具体的な登録情報内容を目視で確認した。ただし、個々の情報提供元における登録環境の詳細（たとえばレコード登録時に選択リストが登録者に提示されるか等）は考慮していないため、現在の運用環境の実態と本調査の分類は必ずしも対応しないことに注意が必要である。

【図表2-10】分析に用いた登録項目の種別

種別	説明	分類項目の例
I	識別番号	《事故情報ID》など
P	個人情報で分析対象としないもの	《通報者情報-氏名》など
L	場所	《発生場所》など
D	時間・時期	《発生日時》など
TO	リスト選択で辞書が固定されているもの（データベース運用期間を通して語彙が変化しないもの）	《事故種別》など
T	リスト選択で対象が固有名であるもの（データベース運用期間中に語彙を追加する可能性があるもの）	《商品など名称》など
TS	リスト選択に近いが一般名詞であるもの	《傷病内容詳細》など
S	自由記述文	《事故内容詳細》など

#### (イ) 言語解析の適用可能性に関する調査

(ア) の調査結果に基づき、《情報提供元》が《国民生活センター》で、《事故種別》が《事故情報》である85,778レコードを抽出して言語解析の適用可能性を調査した。対象となるレコード群には、自由記述項目の平均バイト数が約890バイトと、言語解析の適用に十分な分量のテキストが含まれている。また、《事故種別》を《事故情報》に限定したことから、《傷病内容》や《傷病の程度》の網羅率が1.0となり、本調査のテキストマイニングの目的に合致したサンプルとなっている。調査では、さらに、網羅率等を踏まえて、元データに含まれる83個の分類項目から合計17個を解析対象として選んだ。ここで、《発生年》と《発生月》は元の事故情報データでは1つの分類項目として《発生日時》にまとめられていたが、季節やイベントと関連がある事故を分析するため、異なる分類項目として再定義した。また、商品に関する情報、事業者に関する情報はそれぞれ、《商品情報》および《事業者情報》としてまとめた。以上により、合計で14個の分析用分類項目を設定した。

解析にあたっては、これら14個の分類項目をさらに、「絞り込み項目」と「表示項目」の2つに振り分けた。「絞り込み項目」とは、図表2-10のL（場所）、D（時間・時期）、T0（リスト選択）であり、値を指定することで条件に合致するレコードを絞り込むことができる。「表示項目」とは、T（固有名）、TS（一般名詞）、S（自由記述文）であり、辞書サイズが大きいことから絞り込みには適さないと判断した。分析に用いたデータの概要を図表2-11に示す。

【図表2-11】言語解析を適用した85,778レコード集合に関する統計データ

項目種別	項目分類	分類項目 (オリジナル)	分類項目 (分析用)	辞書サイズ	網羅率	平均バイト数
D	絞り込み項目	発生日時	発生前	46	0.57	4.0
D			発生月	12	0.53	2.0
L		発生場所-施設用途	発生場所-施設用途	9	0.70	13.4
L		発生場所-発生場所	発生場所-発生場所	10	0.10	7.7
T0		被害者の情報-所属-職業	被害者の情報-所属-職業	10	1.00	13.0
T0		被害者の情報-年代等-年代	被害者の情報-年代等-年代	11	0.91	8.2
T0		被害者の情報-性別	被害者の情報-性別	3	0.99	6.0
T0		事故内容	事故内容	17	1.00	11.1
T0		傷病内容	傷病内容	20	1.00	19.6
T0		傷病の程度	傷病の程度	7	1.00	15.4
TS	表示項目	事故概要	事故概要	23,564	1.00	17.1
T		商品など名称	商品情報	32,485	1.00	24.6
T				型式・ロット・生産国	19,051	0.31
T		事業者-事業者名1	事業者情報	31,033	0.82	18.3
T		事業者-事業区分1		1,034	0.02	44.1
T		事業者-事業者名2		6,196	0.16	17.3
T		事業者-事業区分2		58	0.00	39.1
S		事故内容詳細		事故内容詳細	85,761	1.00

本調査における分析は、(A)絞り込み項目を使ったレコードのグループ化、(B)グループごとの項目要約表示、の2つのステップから構成される。

#### (ステップA) 絞り込み項目を使ったレコードのグループ化

図表2-11で絞り込み項目として指定した10個の項目の任意の組み合わせを指定して、これらの項目について値が同じレコードどうしを

グループ化する。たとえば、分類項目《発生月》には《1月》から《12月》の値が入力されていて辞書サイズは12、分類項目《傷病の程度》には《医者にかからず》、《治療1週間未満》、《1～2週間》、《3週間～1か月》、《1か月以上》、《死亡》、《不明》のいずれかが入力されていて辞書サイズは7である。たとえば、絞り込み項目として、《発生月》と《傷病の程度》の2つを指定する場合には、「《4月》に発生した《治療1週間未満の事故》」など、組み合わせ数合計 $12 \times 7 = 84$ 個のグループが得られることになる。この方法では、任意個の絞り込み項目を自由に組み合わせて、機械的に多数のグループを生成することができる。一方で、生成されるグループの数が膨大になり、人手で確認すべきグループを見つけることが難しくなるという問題がある。そこで、得られたグループの中から注目すべきグループを選別するため、当該グループに含まれる事故レコード数と意外性を表す統計尺度（相互情報量）を掛け合わせた尺度を使って、グループの順位づけを行った。

#### （ステップB）グループごとの項目要約表示

（ステップA）で得られた各グループの項目ごとに、指定した数の特徴的な「キーワード」を表示する。これによって、各グループの特徴を項目別に概観することが可能になる。ここでキーワードとは、《事故内容詳細》以外の項目については入力されている文字列そのもの、《事故内容詳細》については自由記述文に対して言語解析（形態素解析と係り受け解析）を適用した結果から得られる名詞句、および名詞句と動詞の係り受けペアとする。さらに、抽出した名詞句や係り受けペアが各グループにおいてどれくらい特徴的であるかを統計的な尺度（頻度×相互情報量）を用いて計算し、そのスコアに基づき上位のものを出力した。

#### エ 分析の試行を通して明らかになった点

子どもの重大事故情報に、上記の分析技術を適用した例を以下に示す。

まず、上記で作成したサンプルデータで《傷病の程度》が《1か月以上》または《死亡》であるレコードの中から、《被害者の情報-所属-職業》および《被害者の情報-年代等-年代》を手がかりに、中学生以下と判断できるものを抽出して、《0歳以下》、《1～4歳》、《5歳以上未就学児》、《小学生》、《中学生》のラベルを付与した。登録レコードの中には、「40歳代の小学生」なども存在するため、不適切と思われるレコードの

一部は人手で除いた。

全体で、《死亡》事故9件、《1か月以上》の傷病にかかわる事故487件が抽出された。判断の目安とした値の組み合わせ、および、抽出されたレコード数を図表2-12にまとめる。

【図表2-12】子ども事故分析のためのラベル付与

《被害者の情報-所属-職業》	《被害者の情報-年代等-年代》	ラベル	事故発生件数
無職	1～4歳	1～4歳	124
小学生	5～9歳	小学生	95
中学生	10歳代	中学生	80
小学生	10歳代	小学生	45
無職	0歳以下	0歳以下	44
無職	5～9歳	5歳以上未就学児	42
その他	1～4歳	1～4歳	15
その他	5～9歳	5歳以上未就学児	13
その他, 学生	5～9歳	5歳以上未就学児	12
小学生		小学生	8
その他	0歳以下	0歳以下	7
その他, 不明	5～9歳	5歳以上未就学児	3
その他, 不明	1～4歳	1～4歳	3
その他, 学生	1～4歳	1～4歳	3
中学生		中学生	1
小学生	1～4歳	小学生	1

本調査では、絞り込み項目の数を2個から3個に設定して、合計4,459件のグループを出力した。図表2-13に最も事故発生件数が高かったグループの表示例を示す。「|」は区切り記号、括弧内は発生件数である。

また、特徴的なグループの上位3件は、①《中学生》の《化粧石鹸》による《皮膚障害》、②《住宅》における《ウォータークーラー》による《熱傷》、③《住宅》における《化学物質による危険》による《呼吸器障害》であった。これらのグループの詳細は、図表2-13の形式の要約を参照したり、人手でグループ内の事故情報を個別に調べたりすることで、さらに確認することが可能である。

【図表 2-13】本調査における分析手法で得られた事故グループの例

事故発生件数	59
絞り込み項目	《発生場所・施設用途》店舗・商業施設 《事故内容》その他 《傷病内容》骨折
発生年	2013(15), 2016(9), 2012(8), 2014(6), 2008(1), 2010(6), 2015(3), 2009(1), 2011(4)
発生月	01(6), 08(5), 04(5), 12(4), 02(5), 05(8), 03(4), 10(3), 07(4), 09(3)
発生場所-施設用途	(絞り込み項目として設定)
発生場所-発生場所	その他(4), 階段(1), 玄関(1)
被害者の情報-所属-職業	その他, 不明(2), 小学生(19), その他, 学生(3), その他(5), 中学生(7), 無職(23)
被害者の情報-年代等-年代	1~4歳(21), 5~9歳(11), 小学生(19), 中学生(7), 0歳以下(1)
被害者の情報-性別	男性(37), 不明(2), 女性(20)
事故内容	(絞り込み項目として設定)
傷病内容	(絞り込み項目として設定)
傷病の程度	1か月以上(59)
事故概要	遊具(3), 屋内遊戯施設(2), 施設内遊園地のエアースライダー(2), スポーツ教室(2), 遊園地(2), 鍼灸院のマッサージ(1), 他の建物(1), カラオケボックス(1), 商業施設内の遊戯施設の遊具(1), 遊具施設(1), …
商品情報	遊具(3), 施設内遊園地のエアースライダー-スライダー-遊園地・レジャーランド(2), ○○○○(2), 店のキッズコーナーエアマット-遊具-店舗.事務所(1), 有料ボールプール-遊具-遊興施設利用(1), 有料遊具施設-遊具(1), 住居雑品その他(1), 他の建物-他の保育サービス(1), バレエ教室(1), 昇降設備(1), …
事業者情報	○○○○
事故内容詳細(名詞句)	遊具(30), 子供(67), ボールプール(12), 責任(23), 息子(65), 救急車(26), 無料(16), 店側(12), ギブス(10), 腕(15),
事故内容詳細(係り受け)	全治(21), 怪我 を する(18), 骨折(7), 左ひじ を 骨折する(5), 不満(10), ボールプール に 飛び込む(4), 子供 を 遊ばせている(4), 娘 を 遊ばせている(4), 責任 は ない(5), 損害賠償 を 求める(5), 子供 が 骨折する(4), 責任 を 認める(4), 条例 は ある(3), 一切責任 を 負う(3), …

最後にまとめを述べる。重大事故の分析は人手で詳細に行う必要があるが、その前段階として分析対象を絞り込み整理することが有効である。その具体的な方策を検討する際に、本調査で報告した手法が参考になると考えられる。

#### オ 分析の試行を通して課題と感じられた点

事故情報データバンクでは、複数の情報提供元からの事故情報を収集して、横断的に検索する仕組みを提供している。このようなサービスは事故情報の一元的な分析の第一歩となるものであるが、一方で、情報提