

消費者委員会 消費者安全専門調査会  
(第31回)  
議事録

内閣府消費者委員会事務局

# 消費者委員会 消費者安全専門調査会（第31回） 議事次第

1. 日時 平成29年6月12日（月） 14:00～17:59

2. 場所 消費者委員会会議室

3. 出席者

（委員）

山本座長、藤田座長代理、相澤専門委員、市瀬専門委員、西田専門委員、  
村田専門委員

（消費者委員会委員）

大森委員、中原委員、樋口委員

（事務局）

黒木事務局長、福島審議官、丸山参事官、友行企画官

4. 議事

（1）開 会

（2）第30回消費者安全専門調査会における議論の整理

（3）事故情報の分析の試行の結果について

（4）暮らしの事故に関するアンケートの結果について

（5）閉 会

## 《 1 . 開会 》

○山本座長 定刻になりましたので、皆様、本日はお忙しいところをお集まりいただきまして、ありがとうございます。ただ今から第31回「消費者安全専門調査会」を開催いたします。

それでは、配付資料の確認を事務局からお願いいたします。

○友行企画官 お手元に配付資料一覧がございまして、今日は資料1から資料7まで配付させていただいております。それから、机上配付資料として一番後ろに1と2と2種類つけてございます。

以上です。

---

## 《 2 . 第30回消費者安全専門調査会における議論の整理 》

○山本座長 それでは、最初の議題は「第30回消費者安全専門調査会における議論の整理」です。

事務局から説明を5分程度でお願いします。

○友行企画官 それでは、資料1を御覧いただけますでしょうか。前回の消費者安全専門調査会における議論の整理でございます。

4月27日に開催されたものでございまして、少し時間が空いておりますけれども、まず「ヒートマップについて」ということございまして、プラスアルファ・コンサルティングの発表について議論の振り返りがございまして、まず1のところでございますが、ヒートマップによる事故情報の分析では、「こうした単語が頻出する」とか「こうした書き込みが多い」といったことが分かるけれども、実際にどのような事故が起きたのかということについては、一つ一つ中を見て内容を確認する必要があるという議論がございました。

それから、自由記述の全ての部分を人間が読むことはできないが、ヒートマップを活用することによって、当たりをつけることができるという議論がございました。

また、同じくヒートマップについて、マクロな分析とミクロな分析の両方が必要であるけれども、ヒートマップはマクロな分析のためのものという理解でよいのではないかという御意見もいただきました。

また、これに関連いたしまして、ヒートマップあるいはSNSの情報については、人の価値判断部分と事実に関することが同時に発信されているので、事実に関して把握するのであれば、そうした観点で情報を見ないと誤った判断につながるというような意見がございました。

次に、「病院に対する調査について」ということをごさいます、これは海外の事故情報データバンクの話をごさいます、アメリカの例を取り上げた部分をごさいます、事故情報の全国推計を行うためには、調査対象の病院のサンプリングをしっかりと考えておかないといけないといった議論ですとか、世界全体で見ると、病院をベースとした事故情報システムが大半となっている。消費者庁も病院から事故情報を収集していると思うが、病院の選定に当たっては、全国推計などが行えるような設計にすればよいのではないかといった意見もごさいました。

それから、その病院のサンプリングについて、アメリカにおいては、救急医療を行っている病院を対象としているため、軽度の傷害などの緊急性の低い疾患等については別の調査で扱うという棲み分けがなされているのではないかという指摘もごさいました。

4、アメリカの病院が調査に協力する理由についてということで、アメリカは100ぐらいの病院が調査に協力しているのですけれども、それに対してNEISSの方は金銭的な対価を払っておりまして、金銭的なメリットだけではなくて、病院側に調査に協力する意思がないとなかなか協力してもらえないと思うという議論もごさいました。

これに関連いたしまして、日本の医療機関ネットワークは病院に調査に協力してもらっているという状況だけでも、アメリカのように予算をつけて情報収集するといったことの必要性を感じたという意見もごさいました。

それと、医療機関ネットワークは現在公表されておりませんので、集計情報しか外部からはアクセスできない、見るができないが、どういう情報が登録されているのか見たいと思うといった意見もごさいました。

一方で、日本では、病院からだけではなくて、メーカーや消費者から直接事故についての報告があるので、詳しい事故のプロセスが分かるといったメリットもあるといった意見もごさいました。

次に、情報のコード化のごさいます、情報のコード化の重要性と自由記述の活用についてのところで、NEISSの利用拡大の重要なポイントは情報のコード化にあると思われるといった意見もごさいました。

それから、その情報の入力のところでは、入力に慣れると、訓練をすれば入力に問題はないと思うということで、入力に際してある程度訓練をすることが重要だという意見もごさいました。

その他のところをごさいます、アメリカの政府部内における事故情報の共有について、政府部内に対しては、一般には公開していない、例えば民族とか人種などの情報も提供しているのではないかといった意見もごさいました。

それと、事故情報の分析に関する国際的な連携については、事故情報の分析について、各国の機関が連携した取組は、今のところないと思うといった意見もごさいました。

以上ごさいます。

○山本座長 それでは、ただ今の説明の内容につきまして、御質問、御意見のある方はお

願いたします。よろしいでしょうか。

### 《 3. 事故情報の分析の試行の結果について 》

○山本座長 それでは、今日の本題に入りたいと思います。「事故情報の分析の試行の結果について」です。本専門調査会では、専門委員の皆様には事故情報データベースに登録されているデータを御覧いただき、各委員の専門とする分析技術を活用した分析をお願いしております。

委員からは、それぞれ20分程度で御説明をいただきたいと思います。本日は5名の委員から御説明をいただきますので、前半と後半に分けて、まず前半は藤田座長代理、村田専門委員に御説明をいただき、その後、30分程度質疑を行い、その後、10分間休憩を取りまして、相澤専門委員、市瀬専門委員、西田専門委員から御説明をいただき、その後、50分程度質疑、それから全体を通した質疑を30分ほどといったスケジュールで進めてまいります。

それでは、早速ですけれども、藤田座長代理から御発表をお願いしたいと思います。

○藤田座長代理 それでは、藤田の方から御報告したいと思います。お手元の資料では資料2-1になります。よろしく申し上げます。

私の方は「事故情報の流れから見た効果的な仕組みの構築」ということで、まず1ページ目、この検討が始まったときにもお出しした情報なのですが、消費者のところで発生する事故の情報を分析・活用して、事業者の方にフィードバックするまでをここにモデル化していて、この流れ全体を見渡しながら、どのようなことをやれば事故情報がうまく伝わって、未然防止ないしは拡大防止につながっていくのかということを中心に検討いたしました。

次のページは検討方法になります。畑村洋太郎先生が提案されている思考展開法というものがございます。それを使ってワークショップの形式で課題解決策を導出いたしました。ワークショップには、消費者委員会の事務局、それから消費者庁からも御参加いただいて、議論をしてきました。都合5回、ワークショップを開きまして、今日の発表に至っております。

まず「あるべき姿・全体像の把握」として、思考平面図というものがありますけれども、これは議論の種になるような項目を、小さいものから大きいものまで思いつくものを全部出すというものです。その後、それを上位概念でくくっていきまして、全体的に構造化していくという作業になります。これは第1回で行っております。規模的には、そこがございますように、抽出項目は約700項目、10人で抽出しております。上位概念にまとめたものが130ぐらい、最上位の概念2個ということで、かなり大規模なものになりました。

今日はここには持ってきていないのですが、大きさは、模造紙大の大きさのものが2枚ぐらいになっています。これを課題に落としまして、思考展開図というものに

しております。それがこの2列目にあります「課題設定・解決策立案」ということです。

この思考展開図は、お手元にあります資料2-2ということで添付しております。これをざっと見ていただきますと、大きくは、活用できる事故情報をどうやって作るのかという情報そのものに関するものです。これは簡単に申しますと、分かりやすく記述したり、分析を考慮した記述にするとか、伝達を考えて作らなければいけないとかということが挙がっています。

その資料の2枚目がもう一個の観点です。この事故情報をシステムとしてどのように使っていけばいいのかという視点で課題と解決方法がまとめられているのが、その次の資料になります。大きくは、情報の質・量をどうやって高めていくのかという話、それから、分析の質とスピードをどうやって高めるのかということです。もう一個は、これは非常に重要なのですけれども、情報伝達の確実性をどうやって向上させるのかということが挙がりました。詳細は割愛しますが、そこに課題要素、解決方法というのを作りまして、解決案としてまとめるようにしております。きょうは、この先でこれをサマライズしたもので御説明していきたいと思います。

もう一度正面の資料に戻っていただいて、これをやった後に第3回で「解決案の整理・共有」をいたしまして、まんだらとか仮想カタログを使った方法で一つの解が見えるような形にしてみました。今日は仮想カタログのものを御紹介したいと思います。

それから「解決案の具体化」ということで討議をしまして、全体をまとめて、さらに「ブラッシュアップ」ということで、第5回なのですけれども、簡潔表現モデルとここでは呼んでいますが、このモデルを使って情報を記述すると、非常にエッセンスが分かりやすく記述できるのではないかという試行をやってみました。これも今日後ほど御紹介したいと思います。

次へ行きますと、先ほどの試行展開図を表の形にサマリーを描いたものがこのスライドになります。ここで、今日は全てお話しすることは時間の都合でできませんので、この色が着いているところです。簡潔表現モデル、フォーラムの話、安全情報マイスターという考え方、そういったこと、画面で言うと黄色のところをお話ししたいと思います。

それから、用語の統一とコード体系というのももちろん出てきまして、これは先ほどの資料にありましたように非常に重要なもので、データバンクの中核になるものとなっていくのではないかと考えています。

次のページですけれども、システム化の観点で今日御紹介するのは、一つは消費者事故フォーラム、それから体験型テーマパークということを考えてみてはどうかということ、黄色のところですけれども、御紹介したいと思います。

あと、最新の分析手法の話も議論の中で出たのですが、これにつきましては、ほかの専門委員の方からお話しされることと同じことを言っていますので、ここでは割愛したいと思います。

次のページですけれども、解決策の具体化をどのように図ったのかということ、少しお

話したいと思います。事故情報を構造化して見るということを目指してやってみたのが、簡潔表現モデルと言っているものです。こういう表現モデルを作ると事故の特性が非常に分かりやすく表現できます。今から3件について事例を御紹介したいと思います。それから、フォーラムとポイント制等、幾つかのものについては仮想カタログという手法でイメージを作ってみましたので、それも御紹介したいと思います。

まず、次のページになりますけれども、事故の簡潔表現モデルとはどんなものなのかということをお説明します。これは失敗学の観点から、その出来を4つの要素で簡潔に表現したものでございます。この図にありますように、1つは要因です。事故の引き金となる事柄を要因として抽出します。それから、特性として、事故の出来を特徴づけるような性質です。からくりと呼んでもいいのですけれども、そういうものを抽出すること。それから、制約とありますけれども、事故を特徴づけるような背景を抽出します。結果、どういふことが発生したのかという結果になります。

実は、事故が起きたときというのは、この絵にもありますように、結果の方からこの中身は見えなわけですが、この見えない部分をどのように記述するかで伝わり方は随分異なってきます。これをうまくやることによって、その下にありますように、事故の特性を知識化することができて、これで自分で危険を予測・回避できるような学習をすることもできます。それから、特性を変えることで再発を防止できると考えられます。

では、1枚めくっていただきまして、1つ事例を御紹介します。ここにありますように、子供が自転車に乗っていて、丁字路の坂道を下っていたら、前方に壁があるのでブレーキをかけたのですけれども、十分に効かず壁に激突したという事故事例、これは過去もありますし、今でも起きていると思います。

これを右にある図のように分解しまして、制約のところ、これは背景です。子供が自転車に乗っているという状況を書きまして、要因として、坂道を下る、それからもう一つ、前方に壁があるということです。特性として、ブレーキをかけたのですが十分に効かないというのがあって、それで結果的に、壁に激突して負傷したと、このように事故事例を分解、構造化します。これを事故の簡潔表現とここでは呼んでいます。

このように分解しますと、特性の深掘りによって真の原因追求と対策がやりやすくなります。なぜ、ブレーキをかけても十分に効かないのかという問いを立てることができずし、これはブレーキの問題なのか、運転の問題なのかというふうに分かってきて、これの再現実験をやってみようかという話になって、大きくは、停止距離の予測が甘いということが出てくるとすると、これは安全教育の方にフィードバックが必要になってきますし、もしブレーキの性能が悪いということになりますと、製品の改良ということになってきます。

ここまでですと、この事故を構造化して表現しただけになるのですけれども、これをさらに知識化ということで、右の方です。ブレーキのあるものは条件次第で停止距離の予測を誤ることがあるというふうにも、もし、知識化したとすると、これは自動車、車椅子等、

ブレーキの付いているものは全て当てはまります。このように考えると、ブレーキの付いているものは停止予測距離を見誤ることがあるということを知っていれば、そういうものに乗ったときにはおのずと気を付ける、あるいはそういう教育ができる可能性があると言えるのではないかとということです。

また、その下に枠で囲ってありますけれども、こういう再現のビデオ、例えば15秒ぐらいで教育ビデオを作ってみせるというようなシナリオも、この要素、制約、要因、特性、結果、これが含まれているビデオを作成すれば、非常に分かりやすいビデオができるのではないかとということです。

同様にして、次のページに行きますと、これは現在言われているもので、結構ホットな事例だと思えますけれども、ライターで何かに火を着けた後、胸のポケットに入れたら残り火があってシャツが焦げてやけどしたということです。これも同じように簡潔表現に展開することができます。

そうすると、やはり同じように特性の深掘りということをやってみますと、何で残り火が生じるのか。これはライターの問題なのか、利用者の問題なのかと例えば自問したとすると、再現実験をやってみるということもありますし、ライターの構造上の問題でしたら製品改良に、利用者の問題だと注意喚起にということになってくるかと思えます。

さらに、知識化というのは、残り火があり得るものは不測の事態を招くということで、知識化すると、例えばライターだけではなくて、ガスコンロも当然こういうことが起き得ますし、たき火、花火、火を扱っているものについてはほぼ当てはまるのではないかとということになります。そうすると、こういうことを連想するような動画を作ったりして、啓蒙活動を図っていくこともできるのではないかとということです。

もう一つだけ御説明します。これも最近の事例ですけれども、10代、20代の女性がネットで見つけた芸能事務所のオーディションに応募したら、合格・採用通知が来たので、事務所と契約したところ、高額な入学金を要求されたということです。

これも制約と要因2つ、それから内容を吟味せずに契約するという特性です。それに結果が書いてあります。

これも同じように特性のところを見ていきますと、内容を吟味せずに契約したと、これは契約の内容の問題なのか、それとも消費者側の問題なのかという問いを立てることができます。これもいろいろなパターンがあると思えますけれども、業者側の問題であれば、注意・指導が必要ですし、消費者側の問題であれば、契約について知っておくべきことがあるのではないかとということです。

ここでは、知識化として、契約行為のあるものは十分な内容吟味がどうしても不可欠だということを知っておくことが重要なのではないかとということで、他に展開できるものとしては、高額品の購入、あるいは保険、それから就業も含めて、そういうものは契約行為がありますので、同じではないかとということです。こういうものも同じように啓蒙あるいは再発防止のためのビデオなどを作ったりすることができるのではないかとということです。

ざいます。

このように、簡潔表現に落としていくと、伝達に非常にメリットがあるのではないかと  
いうことです。

次に行きますと、今度は消費者事故フォーラムということで、情報をここにありますよ  
うに事故関係者がそれぞれシェアしていくことが非常に重要だろうという結果になりまし  
て、これも事故発生後に被害を最小化したり、未然防止に非常に役に立つのではないかと  
いうことで、例えば消費者庁や国民生活センターが、事故関係者が互いに話し合えるよ  
うな場としてフォーラムを開催してはどうかということなのです。

効果としては、このフォーラムが情報ハブになれることで、情報共有が進んで、製品改  
良とか事故回避につながっていくのではないかとということなのです。

これは仮想カタログとして作ったものが、お手元の資料2-3にございます。時間の関  
係で詳しく読んで説明はしないのですけれども、基本的にはこのような、これをカタログ  
だと思って見ていただくといいと思うのですけれども、暮らしの安全をみんなで考えよう  
ということやってみるとどうかなということなのです。

例えば裏へ行ってもらいますと、議事次第で、これも例ですので、データバンクの利用  
状況とか活用状況、在り方などについて、消費者の側、事業者の側、事故を分析する機関、  
行政機関などがディスカッションを行うようなイベントをやってはどうかということ  
です。このような感じで、これはあくまで例なのですけれども、現実には情報をシェアする場を設  
けたらどうかというのが、この一つでございます。

もう一つは、体験型テーマパークというのも同じようにイメージを作ってみました。こ  
こでは、SNSとかウェブの情報をもっと活用したらどうかということがデジタルの世界で  
あったと思うのですけれども、一方で、実体験を通して事故を体感・実感することで、消  
費者自らが考えて判断できるようにするという取組もあり得るのではないかと  
いうこと  
です。

機能としては、消費者事故の展示であったり、専門家による再現実験です。科学実験の  
ようなものであったり、最近はやりの仮想現実などを使った体験コースを設置することも  
可能かなと。インターネットだけではなくて、媒体としては漫画であったり、子供向けの  
絵本であったり、ゲームなどを使って情報媒体が自由にそこで利用できるようにしてはど  
うかということ  
です。

それによって、幼稚園、小学校を含め、校外学習に使ってもらったり、家族での訪問先  
で定着するというのも考えられるのかなということ  
です。消費者事故の防止については、消費者の関心を高めるというのは非常に効果があると議論の中  
でもありまして、体感・実感を伝えることが、こういうテーマパークを作ること  
のできるのではないかと  
いうのがこの例  
でございます。

これも仮想カタログで一応作りまして、2枚目になりますけれども、作ってみました。  
下の方に「事故を身近に」ということで、事故の展示室だったり実験室、体験コースでVR、

バーチャルリアリティの技術でやってみようという話とか、裏にありますように、事故をテーマごとに知ってみようということで、場所別に展示をしたり、道具別、原因別みたいな形で示すことでもきるかなということです。あるいは小学生以上には漫画で示したり、それから絵本、ゲームということも考えられるかなということでございます。

もう一つ、くらしの安全掲示板と安全情報マイスターということで、これもあくまで仮説としての解決策なのですけれども、家庭とか学校でヒヤリハットの情報を、現在の収集ルートだけでは集め切れないような情報を消費者自ら積極的に収集に協力していただけるような仕組みを何か考えてみたらどうかというのが上がりまして、機能的には情報が蓄積していくような掲示板型の情報サイトです。SNSでTwitterのようなものとどんどん流れていってしまうのですけれども、掲示板のような仕組みですと蓄積されていきます。そこへのインセンティブとして、ポイント制による安全情報マイスターみたいな称号を作って、貢献者を表彰するようなこともあるといいのではないかなと。あくまで一案ですけれども、そういうことが議論の中で出ました。

効果としては、もちろん消費者事故と消費者自らが関わるということがポイントになっていまして、網羅性を向上させる、あるいは関心を持って参画できるようなものにできるのではないかなということです。これも同じようにお手元の仮想カタログの3枚目に示してみました。

これを作る仕掛けは、今、既存のいろいろなウェブの仕掛けとかそういったもので十分できるものでして、後はこれをどのようにうまく運営していくかが課題になってきますけれども、まずはこういう取組もあるのではないかなということを示してみました。

今、3つぐらいお示ししましたけれども、資料2-2の展開図には、もう少しいろいろな課題について解決方法を結構たくさん網羅的に考えてみたのが今回の取組だったのですけれども、まとめとしましては、消費者事故情報の発生から活用まで、ワークショップでやってみて、解決案の検討を行いました。都合5回ぐらい、1回2時間ぐらいの議論になりました。

一つは、事故の簡潔表現モデルをやると、分かりやすく表現できて、危険回避とか再発防止のコンテンツ作成がうまくできるのではないかなということが分かりました。

もう一つは、いろいろ解決案を具体化するという意味合いで、実際に作ってみることはできませんので、仮想的なカタログということで作って、検討メンバーでイメージを共有したというのが3つ目でございます。

消費者にできるだけリアルな体験を促すような仕組みということで3つ取り上げて、イメージを作ってみたというのが今回の御報告内容になります。

最後に「今後の課題」としまして、後から分析の中でも議論があるかと思うのですけれども、事故情報の自由記述というのはかなり有用だと考えています。これをもっと立体的に記述できるように、例えば事象、経過、原因、対処、総括、知識化のような6項目による記述みたいなことをやったり、あるいは第三者的に記述するのではなくて、むしろ事故

に遭遇した人が一人称でどのように感じたのかというのを書くことが考えられます。

もう一つは、事故情報の正確性・均質性を向上させ、利用者が見たくなる、欲しくなる情報を作る仕組みを検討する必要があるのではないかとことです。

それから、伝達媒体の多様化は、ぜひ考えていかないといけないことかなと思います。

事故情報の入力負担の軽減というのは、これから出てくる技術を使って軽減していく可能性はかなりあるのではないかとことです。特にAIを使った対話入力とか、自由記述から自動的に項目を抽出して入力するような仕組みはあり得るのではないかと考えました。

駆け足になりましたけれども、以上になります。ありがとうございました。

○山本座長 いろいろ伺ってみたいこともあるかと思いますが、まず、村田専門委員から御説明をいただきまして、お二人の報告についてまとめて質疑あるいは議論をしたいと思います。

それでは、村田専門委員からお願いします。

○村田専門委員 私の方からは、事故情報データバンクシステムに蓄積されているデータですね。こちらの「事故情報データの品質向上に向けて」というテーマでお話をさせていただきます。

この事故情報データバンクシステムですけれども、実は私はこちらの専門調査会のお話をいただいた去年の秋頃に初めてこのシステムの存在を知って、使わせていただいたところなのですけれども、この事故情報データバンクシステムを実際に私の方で使っているのは一般利用者向けのものだけなのですが、こちらのシステムのトップページを見て、フリーワード検索ができるとか、いろいろと今現在の注目情報などが載っているということで、割と見やすいし、面白そうなホームページだなということで、いろいろと思いつくようなフリーワードを入れて検索してみて、その結果を見て、なるほど、こういう情報とかが蓄積されているのだということで、いろいろと試していました。

その中で、初めはフリーワードで検索していたのですが、その結果を集計できるということに途中で気が付いて、集計となりますと、私はふだんも統計を作ったり、結果の分析をしたり、データの分析をしたり、統計自体の品質について分析・確認をしたりとやっているもので、集計と見るとつい集計をしたくなるのですが、その集計をするときに、ちょっと、あれと思ったことがありました。

最初に思ったのは、集計できる項目が最初から決まっているのですが、それ自体はそれほど、決まっているということ自体は普通によくあることなのですが、この項目がちょっとこれで足りるのかなというのは、まず一つ思いました。

次に、実際に検索結果を集計してみると、さらにあれと思うことがありまして、例えば被害者の年代で結果を集計できるのです。この例を作ったのはつい最近なのですけれども、そのときも同じように、スマホ関係のキーワードでやっていたのですが、「充電器」とやったところ、ここでは1,142件ヒットしたのですが、それを年代別で出すと、年代別のところに出ている数字がやけに小さいことに気が付きました。この例では、実際に年代別の内訳

を足し上げると110件しかないのです。そうすると、調査でもよくあることなのですからけれども、では、残りの1,000件以上は年代が分からなかったのだろうか。普通の統計の感覚だと、それにしても1,000件は多過ぎだということがありまして、被害者の年齢はかなり重要な項目なのではないだろうかと思いつつ、いろいろこのシステム自体を利用させていただいていました。

細かいことですが、このとき被害者の年代が一番入っていないで、全部で110しか入っていないのですけれども、その他にも集計できる項目を総当たりで見ても、1,100件ヒットするのにも関わらず、その内訳の数が大分少ないものもあるといった状況に気が付きました。

フリーワード検索の他にも、このデータバンクシステムでは、項目というか、区分を指定して集計するという機能もあります。

すみません。その前に、一件一件の事故情報を更に見たときなのですからけれども、先ほどの内訳を足し上げると件数に合わないというのを見て、では、事故情報そのものにはどういった項目が入っているのかなということ、1,100件も全部見るのは大変なので、幾つかピックアップして見ていくと、いろいろと項目が並んでいるのですが、実際に入っていない項目といったものがある他に、よくよく数十件ぐらい見ていくと、自由記述で事故の概要とかが書いてあるのです。ただ、ここのシステムに表示される自由記述文が比較的短い文章なので何十件も見ることができたのですけれども、それを見ていると、そこに出てくる内容の例えば商品とかの分類の区分が、事故の内容が同じなのに違う分類になっているということもあって、その分類の揺らぎがあるのではないかという点に、普通の利用者として見たときに気が付きました。

もう一つは、ここに表示される項目なのですからけれども、自分がこの事故情報を知りたいと思ったときに、これだけだとよく分からないなと思ったのは、例えば事故の原因が何だったかとか、その事故が起こった結果どうしたのかというようなことが、一部分は入っているのですけれども、全部の詳細の情報には入っていないといったものがあつたりとか、あと、表示される項目には被害者の年代、性別というのがあるのですが、検索結果の集計のときに性別が無かったなというのが、最初に引っ掛かったのはそこなのですからけれども、その辺りで、余りそれは消費者事故では関係ない項目なのだろうかとか、実際のデータを使う前にいろいろな予測を立てて見ていました。

ちなみに、性別の問題なのですからけれども、フリーワード指定の他に項目指定ができるようになっていまして、そちらには性別を選ぶというところがあるのですが、実はこの1,142件のうち、性別を入れてヒットしたのは7件のみということで、これもちょっと不思議な結果だと思いつつ、データバンクシステムの中身を見ていました。

その後、このデータバンクシステムに蓄積されているデータをファイルの形でまとめていただいて、その分析を行っているわけですが、自分自身のこういった疑問点を含めて、データの品質向上をどういう観点で見ていくかということをお初めに考えたのですが、デー

タの品質向上に期待されることとして考えられるのは、例えばこれまで検索・集計に利用されていなかった項目を適切に分類して検索・集計を実行することができたら、利用者にとってもっとニーズの適合性が向上するのではないかとといったことや、利用の拡大が期待されるのではないかとといったことを考えました。

他にも項目の揺らぎが、数十件見た範囲でもちょっと揺らぎがあるのかなといったこともありましたので、この辺の基準を明確にすることによって、事故情報の比較とかそういうものが向上できるのではないかと。

あとは、これまで統計といった形では公表されていないと思うのですが、そういった統計を作るとか、統計的な分析を行うといったときにもデータの説明性というものが上がるのではないかと。その点については、さらに一定の標準的な項目とか分類基準を示せば、より多くの消費者から、いろいろなルートで情報提供が容易になるのではないかとことや、情報入力を実際に行う人たちの負担などの軽減も期待されるのではないかと。そういった観点でデータ分析を試行しています。

ただ、こういったデータの品質向上に期待されることについて、こういったことを目標にして、こう改善すればこう良くなるというところまでばっちり結論が出ればよかったですけれども、実際に今回の試行では、なかなかこのデータが複雑といったことや、非常に手間がかかるといったこともありまして、課題があるということの洗い出しぐらいにとどまっているのですけれども、そういった範囲で報告をさせていただきたいと思います。

データの品質というのは、今回いただいた蓄積されたデータで、しかも、そのデータファイルに収録された内容のみを分析するという形で行っているのですが、本来はデータの品質というのは、設計とかデータの収集、データのコード化、編集、加工とか訂正を含みますけれども、それとどういう形で公表するかといった処理過程の全てが影響して品質というものが決まります。今回は蓄積されたデータファイルの分析の範囲にとどまるので、主なものはデータのコード化とか編集に直接関係するものばかりなのですけれども、一部については、その設計やデータの収集などに関する示唆も得られています。

まず、分析対象にした事故情報データです。こちらの分析対象のデータは、これまで蓄積されてきたデータということで、項目としては83項目あり、その中には一般向けウェブサイトでは表示されない項目も含まれています。また、期間としては登録年月日が2009年10月1日から2016年12月31日ということで7年分ほどのデータがあり、全部で20万レコードほどがあります。

最初に、データバンクシステムを検索したときに、もしかして入っていないとか、無記入の項目が多いのかなという点が一番気になっていましたので、一番最初に項目ごとにどのくらい値が入っているかという状況を調べてみました。その結果、項目によっては空白が非常に多くて、例えば95%以上空白がある項目が83項目のうち半数近くあります。40のうち13は全て空白だったので、この13項目はいただいたのですけれども、何も情報がなくて使うことができないといったものでした。

また、この空白の多さ、少なさなのですけれども、これは情報提供元、参画機関ということですね。それによって、その多さが違うということがあります。

その空白の原因なのですけれども、これは無記入とか、該当なしとか、恐らく参画機関ごとによって採る項目が違いますので、そのもともとのオリジナルの情報に対応する項目がないといったような理由があるかと思います。ただ、とにかく思ったよりも入っていない情報が多いなという印象をここでは受けていました。

具体的な数字ですけれども、小さくて見にくいかと思うのですが、お手元の資料を見ていただければと思いますが、青が入っていない割合で、赤が入っている割合です。最初の種別はほぼ赤なので、ほとんど入っています。この種別というのは、事故情報か危険情報かという区別です。その区別に関しては、ほぼ全て入っています。

ところが、そのすぐ下が発生日時という項目で、これは事故の発生の時期のことなのですけれども、これは3割ちょっとが青になっていて、発生日時が入っていないデータが20万件のうち3割ぐらいあるといったことを示しています。

発生場所については、半分ちょっとぐらいは入っていない状況になっています。

私が検索システムの結果を見たときに、結局、結果がどうなったのかを知りたいといったのが、多分、措置状況とかが一番近い項目なのですが、それに至ってはほとんど入ってなくて、データとして空白でなかったのは1割にも満たないような部分的なものでした。

右側は被害者に関わるものなのですけれども、最初の被害者の人数が余り入っていないのは、被害者1人とかそういうことが多いのかなとは想像がつきますが、その後、問題のところで年代なのですけれども、6割ぐらいは空白になっています。

その下の性別なのですが、集計とかには使えなかったのですけれども、性別は年代よりは少し多目に入っていることが確認できて、これによって性別の分析もデータを使えば可能なのかなと考えて、この後、進めています。

これなのですが、情報提供元による違いがあるということをお話ししました。こちらが実際の具体的な例で、左側が発生日時の入力状況なのですけれども、トータルでは3割強が空白となっています。これを情報提供元の参画機関ごとに見ると、例えば参画機関A、Bに関しては、全て発生日時は入っている。参画機関Cになると、2割ぐらいは空白がある。その下の参画機関Dになると、半分ぐらい空白があるけれども、残り半分は入っているといった形で、一番下の参画機関Hは、発生日に関する情報が全く入っていないといった状況で、かなり情報提供元によってデータの入れ方について、必ずしもデータの採り方、収集時のものではないと思うのですけれども、データを作り上げたときの入り方によっては違いがあるということが明らかになっています。

同様に、右側は被害者の性別のものなのですけれども、この性別も機関によっては全く入っていないところと、あとは7割近く入っているところ、全て入っているところということで、参画機関によってかなり扱いが違うといった状況になっています。

次に、発生日時についてももう少し見ますが、発生日時全体での空白は3割強です。とこ

ろが、問題はもう一つありまして、入っているといても、発生日時という項目からは想像がつかないような言葉も入っています。左側に実際に入っている言葉を入れてあるのですけれども、年だけが入っているとか、年月が入っているとか、年月日が入っている、時間とか分まで入っているとといったものもありますし、あとは複数の時点が並列的に入っているとか、大体の時期が入っているとか、期間が入っているとか、あとは頻度が入っているものとか発生条件ですね。夏、気温30度以上とかというのが発生日時という項目に入っている状況で、他にも購入から1年と。購入の時点が分からないと、そこがいつだか分からないのですけれども、そういった基準となる時点からの経過・遡及時間といった発生日時が入っています。

ここで発生日時を一つ取り上げた理由としては、例えばSNSとかである話題についての件数が多くなったときに、それが発生日時とどういう関係にあるのかを見ようとしたときに、この発生日時を直接集計できれば、そういったものとの分析もできますし、あとはこの情報が参画機関に通報されてくる受付日時ですね。それと発生日時が実際にどのくらい近いのか、離れているのかというのを見ると、それによって対応の仕方も変わってくると思いますし、そういう点でも発生日時を取り上げて、もう少し見てみたいと思います。

この発生日時なのですけれども、今、入っている内容からフォーマットを統一したとしたらどうかという問題なのですが、とにかく今の状況では非常に使いづらい。私もつい先日まで、なぜこういうことになっているのかははっきりとは分かっていなかったのですが、お手元の机上配付1を御覧いただきますと、2ページの上から順番に、情報提供元、種別、発生日時とあるのですけれども、この発生日時の入力の方法が、年という項目と、月日というものと、時間という項目に分かれています。年の方は記述と書いてあるので、多分西暦で入れることを期待しているのですけれども、実際には平成も入っていたので、もしかしたら年号でも可としているのかもしれませんが、ここは実際に数字なりを入れる項目になっています。月日はプルダウンで選べるということなので、1月から12月を選んで、1日から31日をプルダウンで選ぶという形式です。時間のところは記述なので、何時何分を入れるということを期待していると思うのですが、それ以外の言葉も恐らく何でも許容するという形になっているのだと思います。

その結果、戻りますが、これはもともと3項目に分かれて入力をされているものが、こちらで分析する用のデータは全てこれがつながった形でいただいているのだと思います。実際にこういったデータをExcelで分析する場面を具体的に御説明すると、例えばもともとの3項目が3つのセルに並んで順番に入っていたら、どこで区切られるのかがはっきり分かるのですけれども、それが1つのセルにつながった状態で全部入っていると、そもそもどこで区切っているのかを再現するのは余り簡単ではないのです。こちらでも分析するに当たっては、そこを分解するのはかなり手間で、途中で諦めて、ざっくりした集計で諦めたのですけれども、そういった状況になっています。

まず1つ言えるのは、そもそも入力が3つに分かれていたのであれば、そこは分けてお

いてほしいということなのですが、年月日も西暦なら西暦と、そういったものを決めた方がいいのではないかとすることは分析をされていて思いました。あと、これはもちろん入力必須ではないので、無記入というのも当然あり得るのですが、例えば購入から1年とかいったときに、それは一体何年のことなのか。例えば少なくとも西暦で何年のことなのかというのが分かるような情報まで相手からいただけたらなというのは、正直なところ、分析に必要なのではないかなという観点で思っています。

様々な形で入っているものなのですが、無理やりに何とか年月日とかを所定の形式に直すと、年月日までとれるものが半分ぐらいありました。約20万件のうち半分ぐらいは年月日でとれます。6割ぐらいは年月でとれます。

年月日でとれる半分ぐらい、これについて受付までの日数を実際に集計してみました。ちょっと細かいことを言えば、受付の年月日、こちらも空白のものが多くはないのですが、空白の場合もありますので、実際に集計に使えたのは8万件ぐらいなのですが、それを事故の発生時点から何日後に受付したかというのをグラフ化したものです。赤い方が事故情報で、青い方が危険情報なのですが、今、グラフでは30日のところで区切っています。これを見ると、事故情報の赤い線は30日のところで80%まで行っているのです。つまり、発生時点から見ると30日後までには8割が受付されているということになります。それに対して危険情報の方は、30日では6割程度。これはこの後、だらだらとすごく長い期間続いていって、私が確認した中では最長20年というのがありまして、ちょっとそれは信憑性を疑いたくなるような感じなのですが、かなり長い期間たってから通報することもあり得るということです。

一応、中央値で比較すると、事故情報は9日目までに5割が受付されていて、危険情報は14日までに5割が受付されています。

ここに注意を書かせていただいて、同じような言葉がこの後、繰り返し出てくるのですが、20万件のうち、10万件くらいは実際には日時が入っていないのです。入っていない10万件もこれと同じような形になっているかは全然分かりませんので、これが本当に全体の姿を現しているかどうかは保証できないのですが、データが入っている範囲ではこういった傾向が見られるということが分かります。

次が、情報提供元による違いもやはりありまして、こちらは事故情報だけに限定して見えています。そうしますと、参画機関ごとに見ますと、参画機関1と2は割と早い時期に受付されることが多くて、3は途中までゆっくりなのですが、その後は割と短期間でかなり受付数が増えるという特徴があり、参画機関4については、受付までの日数は比較的長くかかるといったものがあります。これに関しては、通報者に関する属性とクロスして見れば分かるのではないかと思ったのですが、通報者に関する属性がいただいたデータでは空白が多くて、余りはっきりとしたものが出ませんでした。ただ、考えられるのは、早い時点で受付されているのは消費者から直接通報されているものだろうと思います。受付までに長い時間がかかるのは、むしろ事業者からの情報提供があったものだと見ていま

すが、この辺りも、このデータの利用者にとっては、こういった情報はあらかじめ分かっておいた方がいいのかなということで、それによって他の事故の概要とかも違いが出てきていると思いますので、そういった辺り、情報をあわせて提供していただけると使い勝手がよくなると思います。

この日数に関する考察なのですが、今のグラフや集計等から単純に分かることなのですが、例えば1つは、受付までにある程度日数がかかるケースがありますので、今、受付の件数が増えているからといって、必ずしも今、発生が増えているということではないという点は注意が必要だと思います。特に危険情報の場合は、4割近くは1か月以上前に発生した案件ということが、この日時のデータから分かります。発生から時間がたっている案件については、例えば類似の案件の報道発表等に触発されて通報するといった行動もあると思いますので、そういった報道発表とかメディアの動きもあわせて分析しないと難しいところがあるのではないかと思います。

最後、繰り返しになりますけれども、データの利用者については、こういった参画機関ごとの特徴ももう少し説明があると、データの使い勝手がよいかと思っています。

この日時については、もう一つ、これも試みてみたのですが、時刻の情報が入っているものが1割弱ぐらいあったので、それを取り出して、見てみました。事故情報と危険情報別に棒グラフになっているのですが、12時が突出しているのは、多分、入力の上の問題だと思います。実際に事故内容の自由記述文を確認すると、12時に起きているとは思えないようなものが12時に入っているもので、例えば時間が分からないものはその日の真ん中の正午に入るとかいうルールがもしかしたらどこかであるのかと思うのですが、この時間帯の分析をあえてやってみた結果、そこがちょっと異常な値になっています。

ただ、この12時を除けば、事故は午前よりも午後の方が多くなっていて、18時から20時に多く発生している傾向が見られるかと思います。ただ、このデータを使えるのが余りにも一部分なので、これが本当に全体を示しているのかは保証ができないという結果になっています。

それから、今まで単一の項目の集計などはデータバンクシステムとかでも結構見てきたのですが、クロス集計をもう少しできないかということを見てみました。被害者の年齢、性別といった属性とか、商品と事故内容のクロス集計で事故の特徴とかが大分分かるのではないかということで、クロス集計を行ってみたのですが、実際は、空白が余りにも多いとか、収録形式の問題があって、単純にクロス集計をやっても結果の解釈は非常に難しいのですが、ここでは一例として、被害者の性別と商品の分類のクロス集計を行ってみました。

なるべく解釈を容易にするために、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行っています。その結果、集計表が右上に出ていて、性別の構成割合を左側のグラフに出しています。トータルでは、男性が2万2,293人、女性が6万750人、不明・空白が3万7,620人で、女性が男性の2.6倍になっているという傾向が見られます。余りにも女性が

多いのですが、よく見ていくと、商品の分類ごとには男女比の偏りが少ないものとか、異常に不明・空白が多いものもあれば、男女比の偏りが非常に大きいものとかもあります。実数で見えていくと、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特に目立っていることが分かると思います。クロス集計によって、こういった被害者の属性とその商品との関係を容易に明らかにすることができるのですが、今の状態のデータだと空白が多くて、この辺りの情報を取り出すのも難しいといったことは一つ挙げられます。

その次なのですが、事故情報と危険情報という形でデータが分かれています。傷病の程度が被害に関する項目であるのですが、危険情報はそもそも傷病の程度というか、けがとかが起きていないというのが原則的なものなので、データとしても当然、傷病の程度は全然入っていないという形になっています。これを単純に合算してしまうと、不明・空白がやけに多いといった結果になるのですが、これは仮に危険情報を「人的被害なし・物損事故」のみといった区分にして、それと合算しますと、商品の分類ごとに、傷病の程度が重いものはどういったものに多いのか、そういった単純な処理でも分かるようになるといった例を示しています。

こちらは危険情報は全て「人的被害なし・物損事故」とみなして、その上が「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」といった区分で並べ直してみますと、こちらの構成割合のグラフからは、保健・福祉サービスとか保健衛生品、被服、食料品は人的被害が出ている割合が比較的高いということが分かります。また、その中でも保健・福祉サービスや保健衛生品は、治療1か月以上とか死亡の割合も比較的高いといったことがはっきりしてきます。それに対して、光熱水品とか家電製品、車両・乗り物は「人的被害なし・物損事故」の割合が多いといった傾向が見られます。

ここまでの項目の集計なのですが、もう一つ、最初に思っていた自由記述で分類の揺らぎが出ているのではないかとといった問題なのですけれども、ここは実は情報提供元で自由記述の部分から商品とか事故内容の分類ルールをどうしているのか、具体的に見せていただく機会がありまして、それを確認すると、そのルールに沿っていればそんなに揺らぎが出るとは思えないといったことが確認できました。では、そうすると、検索したときに揺らぎを感じたのはなぜかということなのですが、データを見ますと、恐らく情報提供元の違いです。それによって分類の揺らぎが表面化しているのだと思います。それぞれの情報提供元の参画機関では、自分の機関での分析に必要な分類とかを使っていると思うのですが、それが多分、機関ごとの揺らぎを生み出して、一括してまとめているということから、そういった揺らぎを利用者が感じてしまうのではないかと思います。

ただ、それを定量的に示すような分析結果を作るところまで私の方ではできなかったのですが、そこはまだ分析の余地があるのですけれども、一つはそういうことがあります。

それと、項目別に見ると、事故原因とか措置状況については空白が非常に多いというお話をしたのですけれども、別の項目の自由記述文にその事故原因とか措置状況が推察され

るような内容が結構含まれていまして、そこから事故原因、措置状況といった新たな項目を起こすというか、その内容を埋めることができるのではないかという場合にも気が付きましたが、すみません、それに関しては私の方で具体的にこうすればいいのではないのかというところまでは行き着けていないので、今後まだ分析が必要かなと思っています。

また、その自由記述の扱いなのですが、情報提供元の分類の揺らぎというか、分類の違いというものがありますので、例えば事故情報データベースシステムとして統一的にアフターコーディングを行うといったことも一つの方法としては考えられます。公的統計の分野でも、自由記述を統計として集計している例は幾つもあります。例えば経済センサスの産業分類とか社会生活基本調査の生活行動分類については、独立行政法人統計センターで長いこと研究されていて、今は既に実用化されているといった事例もあります。

ちょっと長くなってきたので、この辺り、あとははしょっていきますが、自由記述に関してなのですけれども、分類区分はやはりもう少し見直していった方がいいのではないかと、この辺りが幾つかありまして、例えば事故内容の分類区分なのですけれども、集計すると、その他が4割を占めています。その他が4割となると、4割の中身をもう少し細分化したいなという要望もあるのではないかと思うので、この辺りは情報提供元の参画機関によつての分類の違いも影響しているのかもしれないですけれども、4割の内容は精査して、細分化ということも考えられるのではないかと。

あとは自由記述文のままの項目と、それを分類した項目の間で不整合な場合がたまにあるのです。例えば傷病の程度には、先ほどの「医者にかからず」から「治療1か月以上」「死亡」といったようなものと、あと、傷病の程度詳細という自由記述の項目の2つがあるのですが、この辺りが整合していないものがありました。例えば傷病の程度詳細に死亡と読み取れるにも関わらず、項目にコード化されている方には「死亡」が入っていないで空白になっているという例があります。

あとは、先ほどの危険情報に関して、この傷病の程度が何も入っていないということをお話ししたのですが、やはり「人的被害なし・物損事故」と、何もないというものも区分としては作っておいた方が、本当に分からなかったのか、それとも被害はなかったのかといった区別は非常に重要だと思うので、そういった細かいところも見直した方がというか、使い勝手をよくするためには考えた方がいいかなと思っています。

それから、データベースへの入力なのですけれども、必須の項目がほぼないのですね。最初に、事故情報か危険情報かの区別は全て入っているというお話をしたのですが、それ以外は必須項目ではないのです。ただ、どうしても分析に入れてほしい項目というのはあると思います。それは研究者とか、行政機関等の利用者とか、一般利用者とか、それぞれニーズは違うと思うのですけれども、その辺りのニーズを明らかにして、入力を推奨する項目を明確にした方が情報を入力してくれる参画機関の方々もやりやすいのではないかと感じています。これは具体的にどういう項目なのかというところまでは私の方でニーズの把握ができていなくて、この先の分析ということになりますけれども、そういった点が気

になっている点です。

この後ですが、技術的な話なのですけれども、ふだん使いなれているデータと比べて、この辺りがもう少し使い勝手がよくなればなという、単純に技術的な問題を2枚にわたって載せていますが、ここは割愛させていただきます。

最後にまとめ。まとめと書いてありますが、これは全て今後の課題というふうに考えていただければと思います。このデータ品質の向上のために、自由記述文とコード化された項目を明確に分けた方が非常に使い勝手はよくて、利用者拡大という意味で、やはりコード化は非常に重要だと思いました。諸外国の例で、アメリカのNEISSのデータの利用者が拡大されたのもコード化されたことが大きな理由だという御発表がありました。今回使ってみて、やはりコード化するのは非常に重要だと思いました。

自由記述文からの分類は非常に難しい問題ですが、場合によっては、この事故情報データベースシステム独自のアフターコード化を考えてみてもいいのではないかと。

あとは単純な話ですが、項目間の関連性です。この辺りも明確に定義するといったことです。

あと、今、空白が多いためですけれども、利用者のニーズに応じて埋めるべき、提供すべき項目をもう少し明確にした方が、利用者にとって非常に使い勝手がよくなると思いますし、あとは、データの説明文書があると非常に分析する側は助かります。今回、基本的にはデータだけでいただいたので、かなりその辺り、内容を理解するのに時間がかかってしまったという点もありました。

報告は以上です。

○山本座長 ありがとうございます。

それでは、ただいまの藤田座長代理、村田専門委員からのそれぞれの御説明につきまして、質疑を行いたいと思います。いかがでしょうか。事務局の方からも何かありましたら、聞いていただいて結構です。

それでは、相澤専門委員からお願いします。

○相澤専門委員 藤田座長代理に一点ご質問です。御発表の資料の7ページに示して頂いた、事故の簡潔表現モデルは大変分かりやすいと思います。この簡潔表現も出ると、15ページ目の「今後の課題」にある「6項目による記述」の関係について少し御解説をいただいてもよろしいでしょうか。

○藤田座長代理 まず、今、データベースの中に入っている情報を一般の人に伝えるときにどのように加工すれば分かりやすいのかという観点で作ったのが、簡潔表現モデルになります。まとめのところに書いてある6項目の表現は、自由記述のところはすごく有用だろうなど。今後なのですけれども、今後やっていくときに本当に自由というよりは、幾つかの項目を自由記述の中に入れていくような枠組みを作った方が、より深い分析に行けるのではないかと考えて、それで今後の課題のところへ上げました。ですので、今後の課題でもし例えば6項目の記述ができれば、これは知識化まで実は入っていますので、先ほど

の簡潔表現モデルが更に簡単に抽出できるのではないかと思います。

実は、要因と特性というのは原因の中に記述されることが非常に多いとかという関係がありますので、更に分解して、うまく、多分、制約のところは事象とか経過のところに書かれることが多いですので、そういう関係になるかと思います。

○山本座長 更にいかがですか。

○相澤専門委員 ご説明、ありがとうございます。

テキストマイニングという観点からみると、7ページの簡潔モデルの方がより深い理解に基づくものであるということですね。

○山本座長 その他にいかがでしょうか。

それでは、お願いします。

○友行企画官 すみません。事務局からお尋ねさせていただいて申しわけありませんが、村田専門委員からアフターコード化という話があったのですけれども、それをもう少し具体的に、どういう形でやるのかというのを教えていただければと思うのです。

○村田専門委員 今、具体例を持っていないので、なかなか。机上配付資料1に事故情報登録編集項目一覧というのを御作成いただいたかと思うのですけれども、これで、例えば分かりやすいのは「商品など分類」ですね。こちらは大分類、中分類という形で作っていますが、これがまさにアフターコードの一例なのですね。恐らくこれは自由記述とか聞き取ったものの中から、商品が何だったかというのを改めてコード化し、コード化ということか分類・区分化しているといったものです。だから、一つは、これが長い自由記述文の中から商品という点に着目して作ったアフターコードになります。

同じように「事故内容」についても、火災事故とか発煙・発火・過熱といった分類を作っているのですけれども、これも長い自由記述の文章というか聞き取りの中から、アフターコードによって、火災事故なのか、燃料・液漏れなのかというのを決めていると思うのですけれども、そのことを一般的にアフターコードと言っていて、アフターと付くのは、アンケートなどをイメージしてもらえばいいのですけれども、そのときは選択肢が事前に用意されていますね。その選択肢のどれかというのを回答者に選んでもらうのですけれども、それに対してアフターコードの方は、回答の内容を分析する側、集計する側でコード化するという意味でアフターと付いていますが、基本的にはそういうことです。

今回、データバンクシステムの独自のアフターコード化をと言ったのは、1つには、分類が決まっているのですが、その詳細の対応付けは恐らく情報提供元によっても違いがあるのではないかとか、そのあたり、多分、違いを吸収するような形でこれまでシステム設計をされていると思うのですけれども、やはり7年分蓄積したデータを見ると、その他が多いという問題とか、文章を見て一人の人間が判断すると、ちょっと違う分類かなと感じてしまうとかいうのがあるので、それを踏まえた上で、データバンクの発信したい情報として、どういう分類が必要なのかという観点で、独自のアフターコード化をということで書かせていただきました。

○山本座長 よろしいですか。

その他にいかがでしょうか。

それでは、西田専門委員からお願いします。

○西田専門委員 コメントに近いのですけれども、藤田さんの、先ほど相澤先生からもお話があった簡潔表現というか、こういう整理がすごく役に立ちそうだなと思ってお聞きしていたのですが、ちょうど子供の衝突というか、自転車の事例は少し私も過去に分析したことがあって、そういう意味もあって非常に分かりやすかったのです。例えば、坂道を下ってぶつかるというもので、現象としては、ブレーキをかけようとしたけれども十分に効かなかったというのが統計データから出てくる。これの原因を探ってみると、実はブレーキのレバーの幅が子供の手に合っていないくて、かけようと思ってもかけにくい。つまり、自転車の整備不良、椅子の高さとか、ハンドルの高さとか、ブレーキの幅を調整しなければいけないのですけれども、どうもそれがうまくいっていないというのが出てきて、そうするとやはりそれが一つ介入ポイントというか、製品改善で言えば、ちょっと調整しにくいのですね。ドライバーを差して回さなければいけなかったりするところを、ぐりぐりと手で回せるような製品改善につなげることがいいというようなメーカーに対するヒントにもなったし、そもそもそういうところを調整しなければいけないという認識というか理解がないのですね。ほとんどないので、要するにそういう教育もしていかなければいけないというヒントにもなったということです。

人工知能とかそういうものから出てきたものプラス、やはり人の知恵なり観点を入れて原因を探っていく。人と人工知能の協業というか、ハイブリッドで使っていくみたいなのが非常に大事なのかなと思って、そういう意味でも非常に面白い事例かなと思って、感心して見ていました。ちょうどそういう事例があったので御紹介させていただきました。コメントです。

○山本座長 その他に。

それでは、お願いします。

○大森委員 私も、藤田座長代理の発表ですごくうれしいなと思っています。というのは、やはり事故情報がなかなか一般の消費者にとって難しく、分かりにくいもので、消費者庁のデータも、一般消費者は見にくいというところがあって、今回の委員会も持たれているのだと思うのですけれども、この御提案は一步踏み出していける内容かなと、とてもうれしく思っています。

やはりフォーラム的なものをするとうごく認知度が上がるし、いろいろ人がそこでディスカッションすることによって、そのときの大切なテーマが深められるということもあります。もちろん、この簡潔表現というのはすごくわかりやすいなと思って、同じ印象なのですけれども、例えば体験パークみたいなことを書かれておりましたが、これは京都の京あんしん館でしたか。ああいう子供のところが既にありますね。そういうところに高齢者のものも含めて拡大するとかいう形で、もっと身近なものになって、いろいろな方が行か

れるようになるというなと思いました。

先ほどの村田先生の御発表の方で、空白が多くてクロス集計がしにくいというお話がありまして、クロス集計できないと余り面白いデータが出てこないと思うので、関係機関にいろいろ御協力をお願いしないといけないところで、どのように使われるのかというのが見えてこないとなかなか協力する方も、しんどいことを言われるけれども、何に利用されるのか分かりにくいので、その辺は藤田座長代理のアイデアをばっと打ち上げていただいて、こういう形で進んでいくのだというのが見えると、協力もしていただけるのではないかなと、あわせて感じました。

○山本座長 ありがとうございます。

その他にいかがでしょうか。

それでは、市瀬専門委員からお願いします。

○市瀬専門委員 藤田座長代理の発表に関してですが、2ページの全体の検討対象を見ると、ソーシャルメディアから情報を収集・蓄積する部分と、提供する部分と2通り矢印があります。収集と蓄積に関しては、提言として、くらしの安全掲示板と安全情報マイスターという形で最終的に固まったのかと思うのですが、提供の方に関して何かありましたら、提言いただけるといいかと思いますが、いかがでしょうか。

○藤田座長代理 提供の方は余り検討ができていないのですけれども、1つは、今回の話でいくと、掲示板そのものはリードとライトと両方の機能を多分果たすと思っていて、掲示板で収集・蓄積されたものをさらに掲示板とかで結果を提供するだとか、あるいは現在もSNSやメール経由でデータが開示されていますけれども、そういうルートに乗っかっていくのかなと思っています。

大きく今回気が付きがあったのは、メールとかSNSというのはどんどん流れていってしまう情報なので、どこかに蓄積型の参照、それから書き込みができるようなものがあった方がいいのではないかというのが、今回の提案の中で出たものです。

○山本座長 市瀬専門委員、よろしいですか。

その他にいかがでしょうか。よろしいでしょうか。

お二人の専門委員からは、とりわけ情報のインプットの部分とアウトプットの部分に関して、非常に具体的に現在の問題点を示していただいた上で、今後どのように改善をしていったらよいかという御提言もいただきまして、私もなるほどなといろいろ思いました。

藤田座長代理の御発表の中で簡潔表現モデルというものが提案をされておりましたけれども、私がばっと思ったのは、失敗学からの引用であるということもありまして、『失敗百選』というのがありますね。あれはもともと法律の分野の『判例百選』から転用してつくられたようなのですけれども、『判例百選』を考えてみますと、まず事例があって、それについてどういう判断が行われて、今後これをきっかけにしてどのように判例の判断が行われていくことになるのか、あるいは今後どのようなことを検討しなければいけないのかということコンパクトなページの中にとにかくエッセンスをまとめるという作業で、そう

いうものを示していくと、いろいろ事故の防止等に役立てられるのではないか。ただ、百選というのは冊子体ですので、なかなかアクセスが難しいというところもあるのですけれども、そのところでテーマパークとか非常に具体的に、特に現代の若い人、あるいは子供の情報の取り方に合わせた形で情報を提供していくということを提言されていて、非常になるほどなと思いました。

村田専門委員の御報告に関しましても、確かに、まず情報の出どころごとにいろいろなことをやっていて、それが最終的に集計されて、活用されるということを余り想定されないうで、それぞれのところがそれぞれのやり方でやっているというのが現状であろうと思いますので、そのところをこれからどのように使っていくのかという絵を示した上で、それではどのようにその情報を入れるときに気をつけていったらいいのかということをもっと更に検討する必要があります。更に言えば、もっと単純に、その他項目が非常に広過ぎて、それだけではなかなか情報として使いづらいといった非常に単純なところに関しても、現在のやり方にはいろいろ問題があるということ指摘いただきまして、これからは情報が統一的に使われる。利用者のことも考えてシステムを作っていくかなくてはならない。恐らく入力をする際にも、今のいろいろな技術を使えば、あるいは項目をもう少し工夫していけば、入力する側の手間等も省ける部分があるのではないかとということかと思っておりますので、その辺は確かにトータルでいろいろ考えていく余地があるのではないかと、私もお話を伺いまして思いました。

それでは、ここで10分間休憩いたしまして、3時40分に再開をいたします。それでは、休憩いたします。

(休 憩)

○山本座長 それでは、そろそろ後半の御発表と討論に移りたいと思いますが、後半は相澤専門委員、市瀬専門委員、西田専門委員から御説明をいただき、3人の方の発表内容につきましてまとめて質疑をし、更にその後、本日の5名の委員から御説明をいただきました内容全体につきまして質疑をしたいと思っております。

それでは、相澤委員からお願いします。

○相澤委員 それでは、相澤から報告をさせていただきます。よろしく申し上げます。

お手元の資料4-1です。

私は言語解析技術が担当ということで仰せつかりました。本発表では、目的、分析する事故情報の分野、分析に用いられた技術の概要、明らかになった点、課題と感じられた点の順番に御紹介をいたします。

まず目的です。ここでは2つの目的を設定しております、第1点目は分類項目の網羅性や整合性を調査すること。第2点目は自由記述項目に対する言語解析の適用可能性を調査することです。

次に分野です。分析項目ごとの俯瞰的な調査では、先ほども御紹介がありました全レコード19万8,135件を対象に、簡単な統計を調べました。

言語解析の適用可能性に関する調査では、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という条件を満たす8万5,778件をサンプルデータとして抽出して、そのデータに対して調査を行いました。先ほど村田専門委員から御指摘がありましたとおり、大変欠損値が多いデータで、情報発信元による違いも大きいため、なるべく均質になるように絞り込んだものです。さらに分析の例では、その中でも特に子供の重大事故に関わる4,459件を選択して例を示す形にしております。

さて、3ページ目は、分類項目ごとのデータ登録状況に関する俯瞰的な調査についてのご報告です。先ほど御説明にありましたとおり分類項目は83個ありますが、ここでは、それぞれの分類項目について、辞書サイズ、網羅率、平均バイト数の3種類の数値を調べました。

1番目の辞書サイズは、項目ごとの登録文字列の異なり数です。つまり、データベースのレコード中に登録されている文字の種類を数えました。例えば事故種別でしたら事故情報と危険情報の2つがあるので、その場合は異なり数は2となります。自由記述の場合は項目ごとに登録されている文字列はほとんどの場合互いに異なるので、辞書サイズは登録レコード数に近くなります。

2番目の網羅率は、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です。つまり値が登録されているデータの割合です。

3番目の平均バイト数は、項目ごとに登録されている値の文字列長さで、欠損値を除くデータについてのバイト数平均です。ここで分析のためのデータはUnicodeのUTF-8と呼ばれる文字コーディングを使っているので、日本語の場合は1文字3バイトとなっています。英語は1文字1バイトです。

以上3種類の数値は、情報発信元のURLが明記されていた発信元別に集計するようになっています。具体的な集計結果については、お手元の机上配付資料にあります。文字も小さく情報量も多いので見づらいことと思いますが、先ほど村田専門委員から御指摘のありましたとおり、情報提供元ごとにデータの欠損値や、平均バイト数などに偏りがあること、未登録でゼロとなっている分類項目が多くある点につきまして、実感していただけることと思います。

次は、人手による分類項目の分類です。先ほどご説明がりましたが、データベース中の分類項目は、登録するときにプルダウン方式で限られた語彙の中から1つを登録する場合、あるいは複数の候補を選択する場合、それから、自由記述の場合とあります。ただ、ここではむしろ内容に注目して、資料の表にあるような種別を設定して、各目を人手で分類しました。たとえば、プルダウンメニューについても、その中身を考えてデータベース運用期間中に語彙が変化するものとししないものという形で区別してあります。タイプについて簡単にご紹介しますと、IDは識別番号で分析の対象にはならない。Pとあるのは個人

情報、例えば通報者情報・氏名などで明らかに分析の対象にはならないものです。それから、場所、時間・時期と続き、プルダウンで辞書が固定されているもの、プルダウンで恐らく今後の状況によっては語彙を増やす必要があるもの、それから、リスト選択に近いが一般名詞であるもの。たとえば、傷病内容詳細などです。最後が自由記述文で、これがいわゆるテキストマイニングの対象になる項目です。

これらの項目についてさらに検討をした結果を踏まえて、最初に述べました通り、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グループに焦点を当てて次の分析を行いました。このデータは、自由記述項目の平均バイト数が平均で890バイトで、かなりリッチな内容となっています。また、傷病内容あるいは傷病の程度の網羅率が1.0で、比較的欠損値が少ない均質なレコード群となっています。

その概要をお手元の資料4-2の表1にまとめてあります。例えば発生日時を発生年と発生月に分ける、商品など名称、型式、ロット、生産国は商品情報にまとめる等をして、全部で14個の分析用の分類項目を設定しています。

次に、これを更に2つのグループに振り分けました。1番目のグループは絞り込み項目で、値を指定して条件に合致するレコードを絞り込むための項目です。2番目は表示項目で、辞書サイズが大きいことから絞り込みには適さない項目です。先ほどの表1にも書いてありますが、絞り込み項目というのは、たとえば、発生日時が1月で、被害者の年代が10代など、そのような組み合わせで、ある特定の条件を持つレコードをグループ化するために使う項目です。表示項目は自由記述である事故内容詳細などです。

ここまでが前処理で、次にいわゆる言語解析を適用します。ここでは2つのステップを踏んで行っています。第1番目のステップAは、絞り込み項目を使ったレコードのグループ化。第2番目のステップは、グループごとの項目要約表示です。絞り込み項目を使ったレコードのグループ化は、すでにご説明した通り、指定した10個の絞り込み項目の任意の組合せを使って、レコードどうしをグループ化する操作です。例えば4月に発生した治療1週間未満の事故といった具合です。計算機を使えば自由に組合せを作ることができますので、サンプルデータを対象にした処理では、絞り項目数を2個から3個に指定して、160万ぐらいのグループを自動的に生成しています。ここで、160万個ものグループを人手で確認することは不可能なので、このグループの中から特徴的で人手でチェックしたいグループのスコアづけが必要になります。本調査では試行的なスコアを用いて、特徴スコアが上位のグループを選べるような形にしてあります。こういったスコアの計算方法は研究対象でもありまして、今後、検討する必要のあるところだと思います。

2番目のステップBでは、得られたグループを目視で確認するため各項目の要約表示を作ります。これはグループの項目ごとに、特徴的なキーワードの上位N件、例えば10件を表示します。このキーワードのランキングにも統計的な手法を用いています。

表示する際に用いるキーワードは、事故情報詳細以外の項目については文字列そのもの、事故情報詳細については自由記述文に対して言語解析を適用した結果から得られる名詞句

及び名詞句と動詞の係り受けペアを抽出しています。その結果がお手元の資料4-2の表3です。最も頻度が高かった子供の事故の例になっています。

今回、特に子供の重大事故に注目したいと伺いましたので、国民生活センター事故情報データの中から、特に子供の重大事故に関するものを抽出して、同じ手法で分析しています。具体的には重大事故というのは死亡事故、または1か月以上の傷病に関わる事故で、事故件数は9件+487件です。「子供」の定義は先ほど御指摘のあったとおり非常に揺れていますので、表2に示したように子供とは何かというのを定義しています。具体的には、被害者の情報、所属、職業あるいは被害者の情報、年代等を手がかりに子供を決めました。

その結果として一番大きな事故グループは何だったかという、発生場所・施設用途が店舗、商業施設で事故概要はその他、傷病内容が骨折というものでした。お手元の資料では、このグループについて、各項目ごとに特徴的なキーワードをあげてあります。たとえば一体何年が多かったか、何月が多いのかといったことです。この事故内容詳細の名詞句のところを見ると、特徴的であった名詞、係り受けを見るとけがをするとか、左肘を骨折するとか、そういった情報が表示されていて、グループの特徴を確認できるようになっています。

以上は頻度が多いグループでしたが、特徴的なグループとして上位3件を挙げると、中学生の化粧石けんによる皮膚障害、住宅におけるウォータークーラーによる火傷、住宅における化学物質による危険による呼吸器障害、がありました。こういったグループが子供の重大事故の特徴的なパターンとして挙げられたということになります。

以上で概要の御説明を終わります。次にまとめです。まず前提として、重大事故の分析のようにクリティカルな情報は、最終的には人手で詳細に行う必要があると考えております。ただ、その前段階として、分析対象をいろいろな角度から絞り込んで整理することには、今回の調査で用いたような手法はある程度有効ではないかと思えます。

最後に分析の施行を通じて課題と感じられた点です。今までの御発表でのご指摘と共通する点も多いのですが、情報提供元による情報のばらつきが多いので、入力インターフェースの共通化、あるいは、ガイドラインの共通化などの工夫は必要であろうと感じました。

また、これも繰り返しとなりますが、分析の中で重要な役割を果たす「概念」に関する定義が揺れている、あるいは明示的に定義されていない点を、整備していく必要があると思えます。さらに、絞り込み項目で値が入力されていない欠損値を、自由記述項目のテキストを利用して補完する可能性はあるのではないかと感じました。

これに関して、言語解析でどれくらいのことのできるのかというイメージを、資料の末尾に示しました。同じような名詞をグループ化した例です。20個のグループになっています。大体似たようなカテゴリの名詞がクラスタリングでされていて、こういったものからうまく辞書を作成たり情報を補完したりすることが、テキストマイニングの1つの使い道になると考えられます。

以上です。

○山本座長 ありがとうございます。

それでは、市瀬専門委員からお願いします。

○市瀬専門委員 国立情報学研究所の市瀬と申します。

私の方は事故情報データ分析ということで、私のやったことについて少しお話をさせていただきたいと思います。

まずは私の方でやった分析なのですけれども、分析の目的としましては、現在、事故情報データバンクで非常に大量の事故情報を保持しているということで、そういった事故情報の中から、こういったパターンがあるのかということを見てみたいということで、こちらの方では取り組みました。その際に、事故には一定のパターンがあるだろうということで、パターンを取り出そうと思ったのですけれども、その事故の特徴としてはパターンでもそれぞれ年齢などによって違いがあるのではないだろうかということで、年齢によってどういうパターンがあるのか。そして、事故のときに事故の重大さというのもいろいろと違いがあるので、それぞれの事故の重大さによってどういったようなパターンの違いがあるのかということに関して、こちらでは分析を行いました。

具体的な分析手法なのですが、基本的には頻出パターンマイニングと呼ばれる分析手法を利用しています。この頻出パターンマイニングというのはどういう手法なのかというと、事例同士の中で頻繁に共起する事象を抽出するような技術になります。

具体的なイメージを持っていただくために、どういったようなところで使われる技術なのかということを紹介させていただきますと、例えばスーパーの買物かごに入っている商品があったときに、どういった組合せで商品が買われているのかといったときに使われる分析手法になります。そういった状況でこういう技術を使いますと、例えばソーセージを買う人の多くがロールパンを購入していくなんていったことが買物かごの中に入っている情報から分かって、この人はきっとパンにソーセージを挟んで食べるんだなということが想像できたりします。あと、このような理解可能なパターンだけではなくて、例えばおむつを買う人の多くはビールを購入するなんていうようなパターンが出てくると、子供が別にビールを飲むわけではないのですけれども、これはお父さんがおむつを買いにきたときに、ついでに一緒に買っていくなんていう、そういったような人間がなかなか想像しがたいようなものも出てくる分析手法になります。

これを適用するに当たりまして、事故情報データで頻繁に共起する語を利用するということで、どのようなパターンがあるのかということをごちらの方で試してみました。

次に、データの全体的な処理について説明させていただきます。全体的な処理としましては、各カテゴリの事故概要、商品など分類の2つに着目しました。この2つを取り出した主な理由なのですが、事故概要、商品など分類というのは比較的欠損値が少なく、ほとんどのデータに入っているということから、この2つをまず選び出しております。あと、実際に事故内容詳細も後ほど用いているのですが、事故内容詳細に比べると事故概要、商品など分類は割といろいろな事故に関して同じ言葉が使われており、それぞれ余り内容が

細かくないため、グループ化に利用しやすいということで、これらの情報をまず利用しています。そこから名詞を抽出してきまして、その抽出された名詞が頻繁に共起する事例をグループとして抽出しております。その後、ここで抽出されたグループというものが、実際に内容としてどういったような事例が含まれるのかというものを詳細に見るために、各グループ内の事例の事故内容詳細からこれも名詞を抽出してきまして、そして抽出された名詞で頻繁に共起する名詞を列挙するというような形の処理を行っております。

次に、対象としたデータについて説明をさせていただきます。対象となったデータは全てで19万8,135件あるのですが、それらのデータをまず重傷度と年齢によって分けます。重傷度は死亡した事例と重傷、ここでは治療1か月以上のもを重傷と考えてグループ化しております。あと、その他のものというような形で3段階に分けております。また、年齢の方は9歳以下、10～59歳、60歳代、70歳代、80歳以上ということで、年齢を分けてデータを取り出しております。

このデータの中からまず死亡のデータだけを取り出しますと、全体として1,517件あります。また、重傷・死亡、治療に1か月以上かかった事例と死亡した事例と合わせると、全体で1万3,621件あります。このデータと、あと年齢の区分としましては、ここでは10～59歳を除きまして4つのグループを分析の対象としております。9歳以下と60歳代、70歳代、80歳以上ということで、9歳以下が3,889件、60歳代が1万3,191件、70歳代が8,972件、80歳以上が4,366件というような形になっております。

先ほどの重傷度と年齢の組合せを両方見ますと、それぞれの組合せで14個の分析カテゴリが出てきます。重傷・死亡の場合と死亡の場合、そして9歳以下で全てのデータを対象にした場合と、9歳以下で重傷・死亡だった場合、9歳以下で死亡だった場合、60歳代の全てのデータ、60歳代で重傷・死亡の場合、60歳代で死亡の場合、70歳代の全ての場合、70歳代で重傷・死亡の場合、70歳代で死亡の場合、80歳以上全ての場合、80歳以上で重傷・死亡の場合、80歳以上で死亡の場合というような形で、分析のカテゴリを14個設定しております。それぞれのデータの件数に関しましては、ここに記載したとおりになりまして、少ないものと53件のものから、一番多いもので1万3,621件というような形で、非常にばらばらの数のデータに関してこちらで分析を行っております。

次に、分析の結果になりますが、一番最初、カテゴリ1ということで重傷・死亡のデータを対象にした場合です。重傷・死亡のデータは全部で1万3,621件ありますが、その中で先ほどお話したような方式で、グループはどういうものがあるのかというのを取り出すと、122個のグループが取り出されてきております。その中で5個以上の語が取り出されたグループは、全体として30個ありました。例えばどういったようなものがあるのかといいますと、上のところにあるように洗顔、衛生、保健、品、石けん、類、化粧とか、このような形のグループが取り出されています。この語は先ほどお話をしました事故概要、商品など分類から名詞を取り出してきたものになります。

この結果の見方なのですが、ここのグループの例として、ID24というもので、IDはそれ

ぞれのグループに振られています。この場合ですと例えばID24は洗顔とか衛生、保健、品、石けん、類、化粧という語が出てきたものになります。そこの出現数が全部で492件ありまして、その中でどういったものが出てくるかというのを更に分析した結果が、ここのところに書かれている顔、皮膚、受診といったようなところになります。2段目に書かれているのが事故内容詳細の頻出語の組合せになります。ここのところでは全て列挙するわけにはいかなかったもので、代表的と思われるものを私の方で選択をして、ここのところでお見せしております。

こちらの表なのですが、基本的には語に関しては語1から語7まで書いてありますが、順番は特に指定されておりません。

他にどういったようなものが見られたのかというと、例えばID28を見ると住宅、設備、建物、構成、材というようなグループがありまして、そこを見ると中には救急とか転倒、骨折といったようなものがあるので、住宅の中で転倒して骨折してしまったことが重傷の事故に結び付いていることが読み取れるかと思えます。あとID68などを見ますと、住居、生活、用品というような形のグループがありますが、そういったところを見ると製品、建物、火災といった頻出語が見られるので、何らかの製品によって火災が引き起こされて、それが重大な事故に結びついているのであろうといったことが、こういったところから読み取ることができます。

カテゴリ1、重傷・死亡のところは医療とか保健といった関連の事例が割と多いような印象を私の方では受けました。

カテゴリ2は死亡の案件になりますが、こちらはグループ数が全部で102個出てきました。その中で6語以上のグループということで、割とたくさん語が出てくるようなものを調べてみると大体24個ぐらいありました。どういったパターンがあるのかというのを表示したのがここの表になります。見方としては同じです。どういったものがあるのかというのを少し見てみますと、例えばID12にベッドとか介護、用品、生活、住居といったグループがありますが、こういうところで何が起きているのかというのを見てみますと、すき間、ボード、首という形で書かれているので、恐らくベッドの隙間か何かに首を挟んでしまって、介護中に亡くなってしまったなんていう事例が割と多く見られるのかなということとか、あとID19を見ますと課外、活動、体育、指導という形で出てきます。このグループの中にも複数のパターンを詳細のところで見ることができまして、その中で見えてきたものとしては体操、練習なんていうようなグループとか、野球、バッティングとか、サッカー、部、活動というようなことで、サッカーとか野球とか体操とか、そういったような部活の最中に何か問題があって、死亡してしまう事例が結構な数が見られるとか、医薬品は投与、中毒、過剰というようなことがありますので、多分、薬の投与に関してトラブルがあったのかと思われそうです。石油ストーブのところには、火災、全焼というようなことがあります。石油製品が火災を引き起こすというののもかなりの死亡に結びつく、重大な事故に結びつく可能性のあるものなのかと思えます。このカテゴリを見ると火災とか介護、

部活動、医薬品といったような事例が割と見られるように思われました。

次がカテゴリ3ということで9歳以下になりますが、同じように取り出しますと、これは9歳以下全体になりますので、軽傷から重傷まで全てを含むのですが、そういった中でトラブルとして見られるものとして建物、設備のグループを見ますと、例えばカビ、アパート、部屋なんていうようなことがあって、アパートの中でカビが生えてしまって、それによって何か病みみたいな形でトラブルになってしまった事例とか、スイミングスクールで何か問題になった事例とか、あと、娯楽、遊具のグループの中を見てみますと、滑り台、骨折とか、ショッピングセンターとか、そういったものが見られて、滑り台で滑っている間に何かトラブルが起こるなんていう事例が非常に多く見られるというところが分かってきます。

少し別のものとしては、外食とか食品なんていうものがありまして、そのところで何が起きているのかを見ますと、メニューとかアナフィラキシーとか、そういったものが見られまして、これは恐らく外食のときにアレルギー表示がしっかりとしていなかったために、アレルギーのトラブルが起こるなんていうのが9歳以下の事例では結構見られるのではないだろうかということで、このカテゴリではアレルギーとか、遊んでいる最中とか、そういったような事例が特徴的に見られるように感じました。

次ですが、カテゴリ4ということで9歳以下の重傷・死亡した例になりますが、建物、設備、自動ドア、けがなんていうようなものとか、乗り物、車両というところで車とか指というものができて、恐らくこれは車で指を挟んでしまって重傷になってしまったなんていう事例がそれなりに出てくるのだろうとか、あと先ほども出てきましたが、滑り台によるトラブルが非常に多いというのが重傷・死亡のところで見られるということで、割と重傷や死亡に結び付くものとしましては、乗り物とか遊んでいる最中のもので問題が出てくるのかなと思われまます。

9歳以下の死亡になりますと、死亡の事例自体がかなり少なくなるのですが、グループ数としては24個ぐらいのグループが出てきております。4語以上のグループを見ますと、例えば認可外、福祉、保健、保育、施設というものとか、体育、教科、保健その他みたいなものがあつたりしまして、そういったところの中で何が起きているのかというのを見ると、体育の教科のときには運動場で何かトラブルが起きたりとか、プールで水泳の最中にトラブルが起きたりする事例とか、保育施設で就寝中に幼児が何かトラブルがあつて死亡するというような事例がそれなりの数、見られました。このカテゴリのパターンとしては、学校活動や保育中の事例が非常に多いのかなと思われまます。

次が60歳代になりますが、60歳代になりますとグループ数が99個出てきます。どういったものがあるのかというと、設備、建物のところでは骨折とか転倒とか治療ということで、60歳代になると割と転倒する事例が多くなってきます。こちらの9歳以下と比べると、9歳以下だと転倒してというよりは、滑り台のようなもので遊んでいることによってトラブルが起きているのですが、60歳代だと建物とかで転倒というようなことで骨折を起こした

りとかします。別の事例としては食料、調理というので、これは冷凍ピザに農薬が入っていた事例が出てきているのだと思いますが、割と60歳代の人がこういったもので事故に遭っているとか、美容、サービスでは毛を染めて皮膚がトラブルになったものがパターンとして出てきています。60歳代になると割と医療とか保健、サービスといったところにトラブルが多く見られました。

カテゴリ7が60歳代の重傷・死亡で、そこではシックハウスとか頭痛、転倒、骨折とか、あと医療のところでは歯医者、抜歯、視力、矯正といったような医療サービスによって非常に重大な事故・死亡の深刻な問題が起きやすいということで、先ほどの60歳代全体と比べると、より深刻なトラブルがこちらでは出てきております。

死亡した例のところを見ますと、先ほどのところでも出てきましたが、ベッド、すき間とか、食料のところでは喉、窒息というようなものが出てくるので、喉に詰まらせてしまって死亡する事例も割と多くなっているのかと思います。

次が70歳代です。70歳代も医療に関して結構多く、そういったところを見るとインプラントとか、これは小麦の石けんのものだと思いますが、そういったものが出てきたりとかしております。

重傷・死亡の例ですと、インプラントとか注射とか金属アレルギーとか、こういったようなところが医療・福祉分野のトラブルとして出てきており、重傷や死亡に結び付いたりすることが多いのです。70歳代から見られるものとしてバス、サービスに関して重傷や死亡の事例が見られるようになってきます。バス停、停車、転倒、骨折ということで、バスが止まろうとしたとき、発車しようとしたときに、転倒して骨折してしまうというのが非常に大きな事故に結び付いているというようところが、70歳代の特徴になってくるのかなと思います。

その次、死亡の事例を見ていきますと、先ほどバスで転倒というところがあるのですが、バスで転倒すると骨折して非常に重傷になる可能性はあるのですが、死亡というところで見ますとそういった事例はほとんど見られなくなります。70歳代でも食品のグループで団子・喉とか、パン・喉という形で喉に詰まらせて亡くなってしまう事例がみられます。あと70歳ぐらいから車椅子というものが出てきます。車椅子でどのような形の事故になるのかというのを見ますと、水田に転落していたりとか、踏切、列車なんていうのが出てきているので、踏切で渡ろうと思って間に合わなかったとか、そういったようなことが死亡事故につながっているのかなと思います。

80歳以上になりますと介護の話が割と出てきて、介護とか入れ歯、歯医者などがみられます。また、同じようにバス停のところで転倒とか、そういった事例が80歳代以上では見られまして、重傷のところも割とバスに関するトラブルみたいなものも多く見られます。死亡の方では、老人ホームでご飯を喉に詰まらせてしまったとかが見られ、老人ホームといったようなところが80歳代の特徴かなと思います、介護、保健などに関する事例が割と多いような印象を受けました。

こういった形で私の方では事故のパターンにこういったようなものがあるのかについて、全体を網羅して見るような形のアプローチをとったのですが、このデータを分析するに当たって課題としては、一番大きかったのは未入力ของデータが非常に多いということです。今回は入力が多いデータを対象として分析を行いました、実際に先ほどから何度も話題になっておりますが、例えば性別などのデータが入ってくると、更に細かい分析が可能になってくるだろうということが考えられます。欠損が多いと統一的な解析が困難になるので、その部分をどうするのかというのは今後考えていく必要があります。例えば発生場所を入れると県別の傾向調査とか、そういったようなことができるようになってきますので、分析の目的に応じて、こういったデータをどうやって整備するのかということを考えていく必要があると思います。

あと記入方法、入力用語が割と統一されていないという問題もあります。先ほど年齢というところを使いましたけれども、年齢の書き方も実は代表的なものとしましては80代とか70代とか年齢が明確に分かる書き方もあれば、中にはそのような書き方ではなくて小学生とか書かれおり、入力が統一されていないので、似た事例を機械的に抽出するというのが非常に困難な問題があります。入力の際に標準化ができるものは標準化するなど、少し解決する必要があるだろうと考えられます。

留意点ですが、私の取り組んだ方法というのは語の共起頻度に基づく分類のため、グループ内の事故が全く同じ形態とは限りません。そこのところは注意して見る必要があります。先ほど少し紹介しましたが、同じキーワードのグループ内でも幾つかパターンが存在する場合があります、それが異なるパターンと見なされることもあります。あと、リスト中の似たグループの多さが同様の事例の多さを意味するものではないです。先ほど出現数というものも表の中で書きましたけれども、この出現数というのは複数のところに分類されている可能性もありますので、リスト中の似たグループの多さが事例の多さを意味するものではないという点は注意が必要です。このように分析手法により、固有の特性があるため、データ分析の際には、特性を理解したうえで、目的に応じた分析をする必要があります。

あと、データ分析は全体の傾向をつかんで、更に深い分析を行うための糸口としてやるものなので、これが全体の傾向全てではないという点も留意が必要で、詳細な分析には人の目による個別事故の確認が必要であるかと思ひます。

以上になります。ありがとうございます。

○山本座長 ありがとうございます。

それでは、西田専門委員、お願いします。

○西田専門委員 それでは、産総研の西田から報告させていただきます。

私のは比較的ピンポイントの分析になっていまして、大規模なデータがあるときに、人海戦術で読み込めばいいのですけれども、重要な案件を見つけてほしいというニーズが結構ありますので、そういうことができるかなということ試してみたものについてお話し

たいなと思っています。特にこういう心身機能とか認知機能が変わりやすいお年寄り、子供が製品事故、致死性が高いということがあるので、それを見てみたというもので、今の市瀬先生の発表に非常に近いところもあるかなと思いますけれども、御説明したいと思います。

今回お借りしたデータは20万弱ということで、これを使って、やったことは優先付けができないか、重要領域を自動抽出できないかということで、具体的には製品と傷病の程度が書いてありますので、その中から死亡とか重傷1名とかいう情報が書いてあるものの数をカウントして、全体の数、いろいろな製品事故の中から軽傷を含んだものの中で、重傷になっている率が高いものを出してみたということです。製品事故になったときに重傷率が高い、この製品で事故すると重傷になりやすいですよというものをピックアップしてみようというものをやりました。

10歳以上、20歳以上という感じのカテゴリになっているので、10歳未満のところだけを抽出して、今の重傷率が高いものを製品ごとに並べてみたものなのですけれども、こういう形でパワーウインド、コック、シーソー、滑り台、湯たんぼ、バス、遊具、自転車、プールという形で、この辺が事故というか届出が出た場合に重傷、死亡が占める割合が非常に高いというもので、全体の数がすごく少ないものもあるのですけれども、そういうものになっています。

次が表になっていまして、これが全体の数だと重傷の数が書いてありまして、基本的にはこの割り算で出しています。すごくパワーウインドみたいに、これは1位なのですけれども、全体の数がそもそも少ないものもあれば、自転車のように多いものもあるということになっていますが、そこは見ないで今回、重傷率を出してみたということで、比較的こんな単純な集計をするだけでも重要なところが出てくるかなと思っています。

この製品名に関してはテキストマイニングを使って、このカテゴリ名がもともといただいたExcelに書いてあるのですけれども、遊具と書いてあったり、非常に粒度が大きく書いてあるのです。それよりはもう少し絞り込んだ方がいいのではないかとということで、これはその項目を使わずに自由記述から読み取って製品情報を出したものです。

高齢者は、先ほど市瀬専門委員からもお話がありましたけれども、バス、はしご、手すり、ドア、自動車、湯たんぼ。湯たんぼはお湯が漏れて火傷したという事故です。湯たんぼ、踏み台、ベッド、自転車、車椅子、石けんという形で、意外というか、私にとって意外ですけれども、バスが非常に重傷率が高いという結果が出てきて、今そういう年齢で考えた場合に、非常にターゲットにすべき製品が出てくるというのは大きいのかなと。製品まで絞り込めるとメーカーの方なんか話したときに、アイデアが出やすいというか、出てくる可能性があるので、先ほど藤田座長代理のああいうモデルがありましたけれども、一個一個そういう事象を考えていくと予防につながりやすいのか、対策につながりやすいのかなと思っています。

これも次に表が示されています。すごく多いですね。バスは1位なのですけれども、数

としてもすごく多くて、これは大問題なのかなと思います。非常に個人的なことですが、関西に出張したときに神戸辺りによく出張に行くのですけれども、その地区はすごくお年寄りが多いのです。なのでバスの運転手がお年寄りが立っている間は絶対に発進しないです。それに比較すると都バスはがんがん、乗ったらすぐ発車して、そういうものも相当差があるのではないかと思って、これは公式の見解でも何でもありませんけれども、少しバスに関しては対策が必要なのかなというのも出てきて、こういうものやっていますけれども、こういうものの自動抽出、注意喚起を2013年にやっているみたいですが、どういふところを重点化しなければいけないか。そういうところの効果が本当に出ているのかどうかというところなんかを見ていけるのかなと思いました。

今回、トレンドと言っているいろいろな行政の施策をしたときに、その効果はどれくらい出ているのかというのを使えるといいなと思ったのですけれども、データが少なかったこともあったので、以前ここに日本スポーツ振興センターの方に来ていただいたことがあったと思うのですけれども、あのデータをお借りして、それでトレンド分析を試みたものです。2008年から2015年なので、今回のデータも期間としては近いかなと思うのですが、もう少しデータ数があるので、こちらで分析したもので、これは何でもよかったですけれども、柔道は比較的問題視されているところがあるので、その障害が年々どうなっているかという変化を見たものです。

柔道はすごく人気落ちてきているので、絶対数が下がってきているのもあって、骨折はすごく減っている。ちょっと見づらいのですけれども、この辺は骨折とか挫傷とかそういうものが入っています。それから、なぜか分からないけれども、ある年から急に靭帯損傷というけがが急浮上して、一定になっているというのも出てきて、これは原因が分からないのですけれども、恐らく急にこの年から上がるというのはおかしいので、診断基準が何か変わったのではないかと思います。落ちていたものと、脱臼のようにほとんど変わらないものがある、何が対策をすべきか、何が効果が表れているのかというのは、そういうものが見られるのかなと思います。

実は傷病名に関しては項目立てがされていて、Excelでやれば誰でもすぐ出せるようなものの例です。これはテキストマイニング不要ですぐ出るものなのですが、次はもう少しテキストマイニングが必要なもので、柔道は技が非常に重傷に関連すると言われていまして、それでこれはある先生が調べられて、大外刈りという技と背負い投げという技が非常に死亡が高い、それから、重傷になる率が高いということが分かったもので、それで2012年に相当キャンペーンを張ったのです。物すごい介入をした。そうするとその後、背負い投げに関しては激減をしまして、一方、大外刈りは余り減っていないということで、これは効果が出てきている。これは教育効果とかそういうものですが、そういうことでやったことがちゃんと表れているのかどうかというものの分析には役に立ちそうなので、これから先ほどのバスなんかはどういふのではないですかね。かなり今、数が多いので、何かやったときに本当に出てきているかどうかを見ることがいいのかなという

ことで、これはテキストマイニングが役に立っています。こういう技は一切書いていませんので、項目としては出てきていません。自由記述文に埋め込まれているのです。でもそういう情報が結構入っている。ここに抜き出すと項目立てされていないものに関して変化が見られる。逆に言うと次の年からこれを項目立てして、ちゃんと明示的に入力するような項目にしていきましょうねという項目のデザインのヒントにもなるのかなと思います。ということで、これは分かりやすい事例かなと思って紹介させていただきました。

そのほかにも、今回の重傷と軽傷という表記が書いてあるのですけれども、このスポーツ振興センターのデータにも給付金、保険のお金がありまして、お金が高い方が重傷、安いと軽傷という、これは仮説です。大体そんなものではないかということで並べてみると、大体こういうロングテールというか、べき乗則というか、重傷は起きないです。非常に少なく重傷のものは起きて、大半のものはずっと軽傷。統計をとるとどれも大体そのようになります。その比較をして何が重傷の要因になるのか、何が重傷にしてしまう要因なのかというのを分析するソフトを今作っているのですけれども、例えばそれで分析してみると、走って転んだという似たような状況をずっと集めてみて、その中である物体が加わると非常に重傷になりやすい。それは何なのかとって見つかったのがハードルです。走って転んだという、縄跳びをして転んだりいろいろするのですけれども、ハードルが加わると非常に押し上げることが多いということで、例えばこういうものが出てくると、これはメーカーの人と協力して、普通、足を掛けるところをこのまま前へ突っ込んで骨折したりするので、このように開いて、足が引っ掛かると観音開きで開くので倒れない。これを競技で使うと怒られますけれども、練習中はこういうものを使って少し安全に練習をして、本当の競技のときはここが開かないようなものでやるみたいなことをすると減るのではないかということで、非常にそういう予防につながるという事例が出てくるのかなと思います。なのでこのようなものに関して今トライしているところです。

今回、限界と有効性ということなのですけれども、基本的には単語の数え上げなので自由記述に出てきたら数えてしまうのです。それが本当に事故の原因だった製品かどうかはさておき、書いてあったら一応数えてしまうというところがあるので、要因でない可能性はあるということで、絶対数が重要な応用には適切ではないかと思います。以前、国民生活センターにヒアリングしたときにも、裁判の資料か、そういう行政の資料だったか、何かのときに結構数が大事という資料があるみたいで、それは向かないかなと。嘘が入ってしまうところがある。ただし、大きなトレンドを見たいとか、効果が出てきているとか、何割か違っている可能性があるけれども、大きくどうなのかと見るようなものに関しては非常に有効ではないかということで、重要課題のピックアップとかトレンド分析なんかには使える可能性がある程度あるのかなと思います。絶対数というよりは相対的にまずいのか。だんだん効果が出てきているのかというのを見るのに向いているのかなと思っています。

あと、これは先ほど藤田さんのを見て追加したのですけれども、我々はデザインワーク

ワークショップといっているいろいろなメーカーの方なんかと集まって、対策を考えるワークショップというか、グループワークみたいなことをするのですけれども、あれはいきなりやるとすごく発散するのです。こんな問題本当に世の中にあるのかとか、どうだこうだという話になるのですけれども、人工知能を使って課題の焦点化、こういう課題は確かにある、こういう重傷事故は確かに起こっているというところまで絞り込んだ後に変人の人に入ってもらって、解決策を出していくという、これは上はAI、下が人間の変人とか知恵者が入る。そういう連携ができると非常にいいのかなと。この上がないとすごく発散するのです。なのでこういう応用はすごく向くのかなと思って、先ほど面白いなと思ってお聞きしていました。

自転車は先ほどもお話したのですけれども、これです。まさにテキストマイニングをすることでこういうものが出てくるのです。坂を下ってぶつかってしまったよと。これもそういうものをどうコントロールできるのか調べてみると、これはその後、我々がやったものですが、手が1年生から6年生まで20%も伸びるのです。ブレーキの幅は伸びないので、それで幅が合っている場合と合っていない場合で調べてみると、実は静止するのにかかる、ブレーキが効き始めるのにかかる時間がかかなり違う。1メートルぐらい差が出てしまうということで、それでここをちゃんときちんとブレーキの調整をしましょうと。

ちなみにブレーキが一番かかる、力が入るところはどこですか。ブレーキのレバーが1番にあるとき、2番にあるとき、3番にあるとき、知っていますかと。これもほとんど知られていないのです。ちょっと馬鹿みたいですが、我々は研究して、1番が一番握力が出るということで、1番の位置にレバーを調整するといいいですよという教育コンテンツにして、現場にフィードバックをしたというものです。ちなみに自転車調整をちゃんとしていますかという統計をとりますと、サドルの高さは8割の人はやっている。みんなやっている。でもブレーキの幅とかハンドルの高さはほとんどやっていないということなので、これは介入すべきというか、ちゃんとやらなければいけないというのが出てきて、そういうコンテンツを作ってやったというものです。

先ほどどなたかからの発表で村田先生だったかな、何か数値データがすごく偏るという話があったと思うのですけれども、これはCPSCのデータでたんすに上ったときにテレビが落ちてきて、事故になったという統計データがあるのですけれども、そのときに何歳のときにその事故が起きたかというヒアリングをするのですが、そのときに2歳、3歳、4歳でピークが立つのです。これは明らかにいいかげんに答えていて、あのおとき2歳ぐらいだったとか、ピークが本当にそこに立つわけがないと私は思うのですが、立つのです。だからこういう現象がヒアリングするときに出てくるので、その限界はあるのかなということで紹介させていただきました。

そういうことで、少し周辺の話もさせていただいたのですけれども、基本的にはそういう重要トピックを抽出したり、トレンドを出すのに非常に有効ではないか。絶対数の議論をするようなものには向かないのかなという印象を持っています。

御清聴ありがとうございました。

○山本座長 ありがとうございました。

それでは、ただいま3人の委員から御説明のありました内容につきまして、質問、意見のある方は御発言をお願いいたします。いかがでしょうか。

それでは、お願いいたします。

○大森委員 直接関係ないかも分からないのですが、私は兵庫県の宝塚市なので、いつも阪急バスに乗って、止まらないうちに立ち上がって叱られています。なので兵庫県のバスと都バスでデータを調べると、いい結果が出ると思います。

先ほどの柔道のキャンペーンを張って事故が減ったということなのですが、ハードルのように機械で調整するとか、ヘルメットをかぶりましょうとか、そういうことは分かるのですが、柔道の場合はどのようなキャンペーンだったのですか。

○西田専門委員 柔道の場合は、大外刈りと背負い投げというのは、起こる学年が決まっているのです。これは高1と高3で起きるというのが分かっていて、背負い投げというのは背負って投げに行くのですが、そのときに投げに行った人が床に頭を突っ込んで骨を折るといふ、プールの飛び込みと同じ現象が起きる。だからベテランで起きる。オリンピックみたいなああいうすごくうまい人で起きることが分かっていて、そういうものに関して低いところで回さないようにすごく強い指導をする。コーチングです。大外刈りは初心者で起きると分かっていますので、初心者にはそれはしない。慣れてからやるというような、基本的には教育システムを変えていったということの効果みたいです。

○大森委員 分かりました。ありがとうございました。

○山本座長 その他にいかがでしょうか。事務局の方からでも結構です。

○西田専門委員 先ほどの市瀬専門委員のものと私の分析に近い傾向のものがあるな、いろいろ入っているなと思っていて、バスとかプールとか、なのでやはりああいう分析をすると、いろいろな手法があるのですが、出るものは出てくるということがあって、非常にそういう意味では有効性があるのかなと思いました。

例えば入れ歯とか出てきていましたけれども、最近、私が消防庁のデータを分析したところ、入れ歯とか義歯での窒息事故、あれを飲み込んだ事故が意外と多いというものが出てきていて、同じようなものが非常に出てきているなど。入れ歯とかああいうものを飲み込んでしまっというものが起きているということなので、そういうものをピックアップするのに非常に有効性があるのではないかという印象を持ちました。

○山本座長 私も2人の報告で、特に60歳、70歳ぐらいの方のバスの事故というものが非常に浮き出てきて、そうかと思ったのですが、これはかなり注目されているのですか。現に消費者委員会とか消費者庁で。どうなのでしょう。

○友行企画官 注意喚起されたことはあります。

○山本座長 2013年に注意喚起をしていて、しかし、今でもそういう傾向が依然としてあるのですか。

○友行企画官 その関係で先生方にお伺いしたいのですけれども、今、先生方は20万弱のデータを分析していただいていますけれども、一時に機械にかけて経年とか、年を区切ってはやっていないと思うのですが、年を区切ってやることについて傾向が見えてくるとか、その辺りはどのように考えたらいいのでしょうか。

○西田専門委員 その傾向はある程度の量があるものに関しては出てくると思います。だからバスなんかはやっていませんけれども、出てくるのではないか。そもそも少ないものは難しいのですが、多いものはかなり出てくるのではないかと思います。

○市瀬専門委員 バスのように既に注目している事故というものがあつた場合は、例えばカテゴリにしっかりとバスの項目を作るなりして、その後、バスの事故がどうやって推移するのかをちゃんと簡単に追えるような形でデータを作成すると、その後の状況の把握が容易になり、施策にいかせるのではないかと思います。

○友行企画官 あと、先生方がデータを見ていただいて、予想外の結果が出たとか、これは意外だったなというところがもしあれば、教えていただきたいのですけれども。

○山本座長 いかがでしょうか。

○村田専門委員 結論としては最後に西田専門委員から報告があつたように、全体の発生している量を見るには向かないのではないかとというような、トレンドなどを見るのも全体に発生している量が分からないのではないかとということが理由だと思うのですけれども、私は性別とか傷病の程度別に商品分類とクロスとして見たときに、何か実感と違うかなという分布傾向が見られたのです。それは全体で発生しているものの中から通報してくる割合というものがそもそもかなり偏っているから、そういうことが起きているのではないだろうかというのは見て感じました。なのでデータとして取り込まれているようなものの発生のモデルとかメカニズムが、物すごく結果の分布に効いているようなところが印象としては受けています。

○西田専門委員 それはそのとおりだと私も思っています、要するにバイアスがかかっているデータかという意味では、かかっているのだと思います。国民生活センターは割と通報に近くないですか。届出とか、何か言いたいことがある人が言うパターンが多いのではないかと思っています、そういうものと、もう一個よりバイアスが低いものとしては、病院で集めている医療機関ネットワークのデータがあるのです。あれはけがをした人をとるので非常にそこでバイアスがかからない。病院に行ったというバイアスがかかるのですけれども、むしろそれは重傷度が高いものなのでピックアップという意味では悪くないフィルターだと思うのです。だから国民生活センターのものと、そういう医療機関みたいなものを比べると、バイアスがどうなのかというのは出てくる可能性があるかなと思います。

○山本座長 ありがとうございます。

何か消費者委員会でこんなことは余り意識していなかったけれども、気が付いたということはあるですか。先ほどのデータを見て。お願いします。

○樋口委員 今後の議論につながるかもしれない点ですが、例えば市瀬専門委員と西田専門委員が19万のデータを分析されたということなのですが、将来、こういう分析をルーチン化して、素人とは言いませんけれども、普通の職員が、分析手法のアドバイスを受けて、実際にデータバンクからいろいろな分析をしていくという体制を作った場合に、どのくらいの時間と人数を必要とするかをお伺いしたいと思います。つまり今回、19万の分析をされて、その結論としてある種のヒントがいろいろ出てきたり、有益な結果が出てきたりしたと思うのですが、これは従来の消費者庁がいろいろ取り組んでいる業務の中でも、極々、断片的には出てきた話であったように思います。ですから、こういう分析を是非、ルーチンに導入してみたいなと思うのですが、その労力が物すごく大きいとか、人数が必要であるとか、ソフトがたくさん必要であるということになると、あるいは、専門の方しかアクセスできないという話になると、なかなか難しいなと思います。もちろん、実際の分析に当たっては、データが統一されていないとかいろいろな問題はあると思うのですが、もしそういうところがある程度整備されるとしたら、お二人が今回行った分析がどのくらいの仕事量なのか、どのくらいの体制を敷けばいいのか、今回、言われたようなことをある程度ルーチン化してやっていくということがあるとしたら、どのような感じでしょうかというその辺のイメージをお伺いしたいと思います。全くラフな感想で結構ですが・・・。

○山本座長 なかなか難しい御質問かもしれませんが、今はなかなかデータがそもそもそろっていないとか、いろいろな問題があって非常に御苦労されたわけですがけれども、ある程度その辺りの問題が解決されたと仮定した場合にどうなのでしょう。

○西田専門委員 印象で言う話ですが、ある程度のルーチン化は可能ではないかと思っ  
ていまして、以前、ソフトウェア会社の方がここでプレゼンされて、非常にインターフェ  
ースが使いやすいから、我々が出したのはそこはいけていないです。ビジュアルライズをそ  
んなに凝ったことをしていないので、でもそこをきちんと見せるような機能をつけると、い  
わゆるそういうものの、それほどの専門家でなくても使えるようなものになっていくの  
ではないかと思えます。もちろんそういう講習会をすとか、こういうふうに見たらいい  
ですよということあると思うのですが、それは十分可能性があるのではないかと。す  
ごくそういうものの専門家でないと使いこなせないようなものではないのではないかと  
思います。

○山本座長 何か市瀬専門委員からありますか。

○市瀬専門委員 私も同じように感じておりまして、前回お話をいただいた人に聞  
きましたら、その分野のデータ分析の専門家が使うよりかは、実務でデータを見て  
いる人が使っていますみたいな話がありましたので、そういった商用のソフト  
みたいなものを導入するような形でも対応できるのかなと思います。

あと、今回は、割と全体的な傾向を見るためにそれぞれが持っている技術や何かを使  
って分析しましたがけれども、データを入れる際にもう少しカテゴリ分けとかし  
っかりしていただくだけでも、全体の傾向というのはかなり見やすくなると思  
います。そこら辺のカテゴリ

をどうやって設計して、どのようなデータの集め方をするかというのを工夫するだけでも、Excelとかでも分析できるような形にはできるのかなと思います。

○山本座長 テキストマイニングという点で、相澤専門委員から何かございますか。

○相澤専門委員 単純なマイニングはツールを使えば大丈夫ではないかと、私もそういう印象を持っています。ただ、今回のデータは死亡や重大事故のように非常にクリティカルな情報を含みますので、簡単に統計処理だけでは済ませられないと思います。総件数を考えると実はそれほど数は多くないので、そこは是非人手で分析されるのがよいと感じています。

○山本座長 それでは、藤田座長代理、お願いします。

○藤田座長代理 市販のソフトとか、今だとクラウドのサービスなんかで、少し研修を受けたぐらいのレベルでデータ分析ができるようなものが次々出始めているのですが、一番専門家の方と違うところは、西田専門委員の資料にもありましたけれども、限界とか有効性を見るところです。そこが専門の方とは経験的にも理論的な理解の話も含めてハードルがあるのかなと思います。

だけれども、それを踏まえた上でトレンドのところを見ていくというのは、それはそれで有効な部分があると思いますので、だんだんハードルは下がってくると思われれます。時間は多少かかるとは思いますけれども。

○山本座長 ありがとうございます。

それでは、お願いします。

○樋口委員 この先の議論に若干関係してくるのですが、今非常に有益な分析結果を皆様から聞かせていただいたので、データバンクの実務においてもこうした分析ができればいいのではないかと考えているのですが、一方でそのためには例えば入力を各機関がある程度そろえることが必要になってくるわけです。その際にこの消費者庁のデータバンクの置かれた立場というものを考えますと、各機関に対して、それぞれデータの統一により分析上のメリットが各機関側にも生じることを理解してもらう必要があるのではないかと思います。例えば、こういう分析ができるわけだから、年齢と男女の入力をこのように統一的にやってくれとか、今まではそこまでやっていなくても、あなたの機関はよかったのかもしれないけれども、全体としてこういう分析をかけることによって、こういうメリットが皆さんにも還元することができるので、したがって、入力方法を統一してくださいとか、そういうことを言わないと、これまで分析をしてきた各機関の人たちは、今のスタイルを急に変えるのはなかなか難しいのではないかと。そんなことを一方で少し考えた次第です。

したがって、うまくデータベースの分析のメリットというものをそれぞれの機関の人たちに理解していただいて、そして入力に空白がないとか、そういう定型的事を全国津々浦々に徹底しなければいけないという話に繋がる訳です。例えば、年齢1つとっても多分それぞれの機関の入力の仕方の問題、それぞれの現場におられる方の入力の仕方の問題と

かいろいろなことに関わってくるので、メリットが非常に大きければそれぞれの機関が重い腰を上げるのかなと思います。その辺の狙いを明らかにしていくこともぜひ御検討いただければと感じました。

○山本座長 今の点は重要なことで、こういうメリットがそちらの機関にとってもあります。それから、入力をする場合にも、こういう技術を使えばもっと便利にできますということを書いていかないと。もっと多くの労力をかけてくれという言い方ではだめなのでしょうね。こういうメリットがありますということを示していただいて、協力を求めていくという、そこが実際は非常に重要なところだと思いますけれども、何かございますか。それでは、お願いします。

○西田専門委員 今、本当に大切な御指摘だと思います。消費者相談センターというか、ああいうところにおられる人はかなり意識が高いというか、そういう問題に対して何とかしなければいけないということを考えている人が非常に多い気がして、ただ、データがどう分析されて、どう役に立つのかというのはよく分かっていないということがあると思うので、先ほど交流、コミュニケーションという話があったのですけれども、実はこういうふうに関に立ったとか、そういうものを知らせる場であるとか、消費者の方にフィードバックするいろいろな教材であるとかの資料を現場にきちんと返すであるとか、そういうコミュニケーションがもう少しあるといいのかなと、印象ですけれども、そういうものができるといいのかなと思います。

○山本座長 ありがとうございます。

その他にいかがでしょうか。お願いします。

○大森委員 資料7のアンケートを見せていただいたのですけれども、商品のメーカー名を知りたいという消費者の意見がすごく多いのです。一般消費者にとってメーカー名、商品名が出ないとなかなか分かりにくい。どうしても行政は営業妨害になってはいけないとか、産業の保護に立つわけですけれども、消費者は実際分かりにくいです。行政のものは具体的な商品名とかメーカー名というのがなかなか出ないので、ここはもう少し消費者視点にシフトしていただくと、こういう事故情報なんかも興味を持って見てもらえるようになるのではないかとというのが1つ。

あと、藤田座長代理がいろいろSNSとの情報交換みたいなことを提案されていますけれども、このアンケートでも若い世代はSNSを使う、高齢者は使わないとはっきり色分けされておりまして、これはインターネットでアンケートを集めていますので、インターネットを使う人ばかりなのですけれども、SNSに関して事故情報に関しては使っていないところがあって、みんなが今、情報を集めるのは、若い人はどんどん行政の情報よりSNSの情報に行っているんで、行政の空回りにならないように、ここをちゃんと連携していかないといけない大切な部分だなと、藤田座長代理からの提案も安全の掲示板だとか、ポイント制などの御提案をいただいているのですけれども、この辺は真剣に考えていかないといけないのではないかと。若い人はポイントをもらえるとか、そういうものに慣れているので、ポイ

ント制とかいうのはすごくいいのではないか。アンケートを集計したものがどう役に立っているかというのも表彰制度みたいなものがある、この団体はすごく貢献してくれたとか、そういう目に見える形で交流できるともっと進んでいくのかなという気がします。

○山本座長 ありがとうございます。

今、少し全体的な話も出てきておりますけれども、先ほどの3人の委員の方の報告に関して、さらに質問等がなければ全体的な話に移っていこうと思いますが、いかがでしょうか。3人の委員の報告に関して、さらにここを聞いておきたいということがございましたら。

全体といたしまして、私も西田専門委員が最後にまとめていただいたのですが、重要課題をピックアップする。こういうところに気を付けなくてはいけないという全体的な状況の把握あるいはこういう年齢層、こういう性別あるいは属性の人に対して、特にこういう点に気をつけるように情報を発信しなくてはいけないとか、こういった全体の把握であるとか、あるいは経年変化。対策を打った、その対策が効果をどれくらい上げているのかとか、そういうトレンドを把握するとか、あるいは何らかの大きな事故が起きたときに似たようなことで、過去どういったものがあったのかということやぱっと調べるといったような場合には、大変有効なツールではないか。まさに人手でやらないといけないとすると大変なので、そのところでいろいろな情報技術が役に立つのではないか。もちろんそこから先は、更に人の目で見なくてはいけないということが当然にあるわけなので、それで人間が見なくても全てが解決するというわけではもちろんないのですが、そのところの全体を把握する、あるいは検索を迅速に行うといった点では、非常に有効なツールとして使えるのではないかと私も思いました。

#### 《 4. 暮らしの事故に関するアンケートの結果について 》

○山本座長 そろそろ全体的な話に入っていきたいと思っておりますけれども、今、暮らしの事故に関するアンケート調査結果について少し言及していただきましたので、あるいはこれについても報告いただいて、これも踏まえて全体的な討論をしようと思っておりますけれども、よろしいですか。

では、資料7につきまして事務局から説明をお願いします。

○友行企画官 それでは、お手元の資料7でございますけれども、事務局でインターネット、ウェブ調査の方法でアンケートを5月末から6月頭にかけて行いました。対象は男女ともそれぞれ10代、20代、30代、40代、50代、60代、70代以上をターゲットに入れて、結果としてはそれぞれの年代ごとに100人ずつがちゃんとデータとしてとれるような形で、合計1,200名という形になっております。回答者の中で70代以上の人はおりませんので、60代までとなっております。回答者の居住地とか、回答者の職業などは5番と6番にあるようになっております。

中身に入りますけれども、1枚おめくりいただきまして、まず1番のところですが、暮らしの事故に関する情報の入手先としてはどこが多いですかということですが、やはりテレビのニュースが圧倒的に多くなっております。次いでインターネットのニュースサイトという形になっております。

3ページを見ていただきますと、テレビのニュースは圧倒的にどの年代も多いのですが、例えば女性10代、20代とか男性10代、20代につきましては、Twitterから取りますという人たちの割合も高くなっております。

2番の消費者庁が発信している事故情報に関する公表の認知度というところをございまして、これは事前アンケートには重大事故とか非重大事故、重大製品事故、注意喚起のURLを添付してございまして、そこで確認していただくような、その上で知っているか知らないかということ聞いております。知らないと答えた人がどの項目についても8割以上という形になっております。

おめくりいただきまして4ページですが、次に消費者庁ウェブサイトとかデータバンクシステムとかTwitterとかFacebook、子ども安全メールについて、どの程度知っていますかということ聞いたものが3番でございまして、ウェブサイトは知っているという人も2割超おりましたが、その他については知らないという方が多くなっておりました。

4番のところですが、知っているというふうに答えた方について、どのぐらいのペースで見えていますかというのを聞いたものが4番でございまして、事故情報データバンクや消費者庁Twitter、消費者庁Facebookは「よく見る」と「時々見る」を合わせると50%ぐらいになってございまして、知っている人は、これらの3つについては半分ぐらいの人はよく見えていますという形になっております。

5ページでございまして、消費者庁が運営している各種情報サイト、データバンクシステムですとか、子ども安全メールについて、何で知ったのですかということ聞いてみますと、特徴的なのは子ども安全メールについては地区の広報で知りましたという結果が出てございまして、親といいますか、そういう人たちが登録していることが恐らく多くて、そういう方たちはこういう広報なんかも見ているのかなということ想像させるような結果になっております。

6番は事故情報の有用性ということで、こうした事故情報データバンクですとか、そういったものについて「役に立った」と「やや役に立った」を合わせると60%になってございまして、それをちゃんと知っている人は、中身的には役に立っているという方は6割もいらっしゃるという結果になっております。

6ページでございまして、今度は直近、消費者庁が公表しました注意喚起について、これもアンケートの際にはURLを張り付けてございまして、実際に見ていただいて、それを知っていましたかということ聞いたものが7番でございまして、「事故の内容や注意すべき点まで良く知っていた」という方と「事故が起こっていることは知っているが、その詳細や、注意すべき点は知らなかった」「知らなかった」というのは、このような割合になっており

ます。

性別・年代別に見ますと、女性50代、60代の方は割と事故が起こっていることは知っているがという形の方が46%とか42%と数字が高くなっているのが特徴的となっています。

7ページに行きまして、食品による子供の窒息事故というのは何で知りましたかということを探ねたところ、やはり最初の回答と似通っていますが、テレビのニュースという方が圧倒的に多くなっています。

8ページでございますが、食品による子供の窒息事故という公表資料のレイアウトについて見ていただきまして、見やすいレイアウトか、やや見やすいかとかいろいろ聞いておられますと、見やすいという回答と、どちらかというで見にくいという回答が半々になっております。この公表レイアウトは9ページぐらいの公表資料になっておりまして、絵があったりとかいろいろしていますけれども、全般的には字が多目で、ページ数も9ページなのでちょっと多いと思った方もいるのかもしれない。ただ、見やすい、やや見やすいという方も半分ぐらいという結果になっております。

その公表資料のボリュームにつきましては、「適当だと思う」という方が55.3%となっております。「多いと思う」という方も4割ぐらいいたということでございます。

9ページでございますが、今後どんな媒体で事故情報を公表してもらえるといいですかということ聞いたところですが、テレビのニュースですとかテレビの情報番組というものが、やはり想像どおり棒グラフが長く伸びておりまして、年代別の表のところを見ていただきますと、若い方、男性10代とか20代ですとか女性10代、20代については、Twitterとかでも言うてもらおうといいという回答の割合も高くなってきております。

10ページ、先ほど大森委員からも少し言及がございましたけれども、消費者庁から発信される事故情報で、どのようなことが発信されるといいですかということで、具体的な商品名ですとか、事故の内容や状況というところが上位に来ておりまして、それと同時に事故につながるように注意すべき点とか対応策というのも併せて出してほしいというところも割合として高くなっております。

11ページのところは暮らしの事故に遭った場合の相談先ということでございまして、事故に遭ったときにまず真っ先にどこに言うかということでございますが、メーカーが一番多くなっておりまして、その次は購入した販売店とかサイトとなっております。消費生活センター、国民生活センターに言うというのは大体2割ぐらいになっています。

それから、特に相談しないという回答も選択肢の中にありますが、性別・年代別のところを見ていただきますと、こちらは若い人たちが特に相談しないと回答する割合が高くなっております。

12ページに行きますと、暮らしの事故に遭った際のSNSとの関係でございますけれども、事故に実際に遭った場合にSNSにアップしますかということ聞いたところなのですが、SNSを利用していないという方がそもそも多いわけなのですけれども、46.5%は利用していない。それから、緑色のところで「暮らしの事故に関してはSNSに情報をアップしない」

という方が3割方いらっしゃいます。残りの方は、まずは相談先に相談した上でアップする、それから、すぐにSNSにアップするという方が21%ぐらいいらっしゃるという結果になっております。

このところにつきましては、性別・年代別の表を見ていただきますと、SNSを利用していないという一番右端のところですが、これが30代でんこんなに段差がつくものかなというぐらいははっきり分かれています、30代以上になるとSNSは利用していないという人の数字が急に上がるという結果になっております。

13ページのところは、暮らしの事故に実際に遭ったのではなくて、遭いそうになったときのSNSへの情報発信の状況でございますが、こちらでもSNSは利用していないという方が多くて、アップしていないという方も3～4割ぐらいいるという結果になっています。

14ページのところは、このアンケートに自由記述の部分を設定しております、アンケートの回答者からの意見・要望としては、以下のようなことがあったということをもとめております。

公表の在り方については、事故情報を公表していることを初めて知りましたとか、こういうものがあるのだったらもっと目に付く形で公表してほしいとか、そのような意見がございました。

最後に15ページのまとめのところでございます。現時点ですぐ言えることとしては、このぐらいかなと整理しておりますが、消費者はテレビのニュースなどといった報道を通して事故情報を入手しているため、事故情報を公表する際には、そういったテレビニュースなどに上げられやすい形になっているといいのではないかとということがまず1つ。それと、10代、20代に限ると事故情報の入手手段としてはTwitterも割と使われていると言えるのではないかとということがわかりました。

事故情報などを発表するについては、併せて事故の予防策とか対応策なども公表することが有用なのではないかということも言えるのではないかと出てきております。事故に遭った場合に製造メーカーや販売店にまずは連絡するということが回答内容としては多くて、SNSにより事故情報に関する情報を載せるという意欲は乏しいということもあるでしょうし、そもそも利用していないという方が数字的には出ています。ただ、10代、20代の若い人はSNSの利用率は高いので、こういった世代はそういう手段を使うことも有効だと言えるのではないかとこの辺りがわかりました。

以上でございます。

○山本座長　ありがとうございました。

このアンケートの趣旨について何も申し上げませんでしたけれども、以前、最初の方ですかね。この会議で消費者が実際、どのように情報の媒体を使っていて、消費者庁等の事故情報の発信についてどのように把握しているのかという実態を少し見ておいた方がいいのではないかとこの御意見がありましたので、消費者委員会の方で急遽アンケート調査をしていただきまして、その結果が今日このような形で出ているということでございます。

今のアンケート調査の結果自体に関して、何か御質問はございますでしょうか。特にございませんか。

それでは、今日の5人の委員の御報告と、今のアンケート調査の結果を踏まえまして、全体的な質疑、討論を行いたいと思いますけれども、いかがでしょうか。全体を通しまして、あるいはピンポイントに絞っていただいても結構ですが、御意見あるいは更に質問し忘れたのだけでも、こういったことを聞いておきたいという質問でも結構ですし、ありましたらお願いしたいと思いますが、いかがでしょうか。

それでは、お願いします。

○藤田座長代理 コメントなのですけれども、今のアンケート調査の結果についてです。事故の情報をインターネットで見るといふ、インターネットのニュースサイトと書いてあるのですけれども、恐らくインターネットのニュースサイトの特徴として、新聞であったり、テレビの情報番組のサイトとか、あるいはテレビ局がやっている番組で発信された、要は従来のメディアから発信された情報を、インターネット上で集めてニュースサイトとして公開している場合が非常に多いのです。ですので情報源をたどっていくと、実は新聞とかテレビの情報番組のものであることが1つ言えると思うのです。

インターネットのニュースサイトでそれを確認するか、テレビのニュースで確認するかというのは、スイッチをつけて、ながらで情報を得られるか、あるいは自分で取りに行くことになれているかということの違いではないかと私はこれを見ていて思いました。ですのでインターネットのネイティブ世代の若い人たちは、自分で情報を取りに行くことにすぐなれていますので、テレビの番組でも例えばYouTubeのような、それを断片的に挙げてあるようなものを自分で取りに行き、見ているのはテレビの番組のある断片だったりということが多いということですね。

年代が上がるに従って、ある時間帯にテレビをつけておく習慣というのがあったりするので、それでラジオとかテレビからニュースを得ているという、大きく2つあるのかなと思います。情報源が実は従来のメディアから出ていることがほとんどだということと、情報の取りに行き方の違いかなと思います。

○山本座長 ありがとうございます。

私も一言申し上げますと、若い世帯でも意外とまだテレビのニュース等から情報を取るといふ人の割合が、私はもっと低いと予想していたのですが、まだこのぐらいの数字なのだなという印象を受けました。ただ、これもテレビ番組についてもオリンピックまでに同時配信をインターネット等で推進するという議論がされていますので、そうするとテレビとかインターネットの利用の仕方も今後さらに変わって行って、この数字も少し動いていく可能性があるかなと思いましたけれども、現段階では割とまだ若者もテレビというふうに言っているんだなという印象を受けました。

SNSにアップするかという話で、そもそも年齢によってSNSは余り使っていないということもあるのですけれども、10代、20代で見て事故、特に重大な事故なんかになるとSNSにす

ぐ情報はアップしないんだなという印象がありましたが、他方で10代、20代と比べると、10代はSNSに事故の情報をアップするという割合が、20代に比べてはつきりと高いという傾向があるので、この辺りも更に今後トレンドが変わっていく可能性もあるのかなと見て思いました。

非常に興味深い検討結果だと思いましたが、その他にいかがでしょうか。全体的なコメント、御意見、御質問、何でも結構です。

私、相澤専門委員の報告で1つ聞き忘れたのですが、11ページにスコアの話が出てきます。スコアの計算方法は今後の検討課題とあるのですが、これはもう少し具体的に言うと、どのようなスコアの付け方を想定されているのか、もう少し説明していただけないでしょうか。

○相澤専門委員 今、使っているのはどれくらいレアなイベントかということと、どれくらい事故件数があるかということを掛け合わせてバランスをとったスコアです。これは、どのような事故に注目したいかによって、件数が多い事故を見たい場合、滅多に起こらないものだけをチェックしたい場合など、ケースバイケースだと思われるので、分析の目的にあわせて検討する必要があると思います。

○山本座長 ありがとうございます。スコアの付け方がいろいろ考えられるので、どういう付け方が有効かということを考えていかなければいけないという御趣旨ですね。分かりました。

その他にいかがでしょうか。それでは、お願いします。

○藤田座長代理 今のアンケートの例えば3ページを見ると、入手先で挙がっているものというのは、実は事故の情報を得る、現状存在するものというのが一覧として挙がっていると思えるのです。逆に言うと事故の情報というのは、このぐらいのメディアというか媒体からしか入手できていないということだったと思うのです。なので私の資料の中にもありましたけれども、もっと多様なメディアを活用して、ここの入手先を広げていくというのは結構重要なのかなと思いました。それは漫画であったり本であったり、テーマパークみたいなリアルな伝達媒体もあると思いますし、インターネットみたいなバーチャルなものだけでなく、この入手先を広げていくことでより狭い隙間にも入っていくような、そういう取組が要るのかなと改めて感じました。

例えばテレビだけにニュースを流せばそれで済むわけではないような気がしますし、ニュース性があるものでないと多分テレビでは取り上げてくれなかったりしますので、それを考えるともう少し多様性を膨らませていく方がいいのかなと思いました。

○山本座長 ありがとうございます。

その他にいかがでしょうか。テーマパークとか具体的な御提案がございましたけれども、テレビで受け身で見るというのではなくて、体で体験をして、事故に関する情報とか、あるいはその対策のとり方とか、そういったものをもっと他の手段を使って感じていただく、知っていただくことは確かに面白いと思いました。



○西田専門委員　すぐできるなと思うのは、子ども安全メールがあって、あれはピックアップがある程度の基準でされていると思うのですけれども、なかなか苦労されているのではないかと思うので、こういうものを今回みたいな手法を使って候補、決めるのは人が決めればいいと思うのですけれども、候補群を挙げるのに使うのはすぐ利用できるかなと思います。

現状は多分、医療機関ネットワークなんかに参加している看護師か何かに、ちょっとこれはという事例を教えてくださいと言って、すごくアナログチックにやっているというようなことを聞いているのですけれども、それは結構当たるのです。人の感覚ってすごい当たるのでいいのですけれども、こういう手法も少し併用していけるといいのかなと思います。

○市瀬専門委員　私の方で感じたのは、それぞれ年代とかによって事故のパターンが違うので、こういう事故があるというのを知らせるのが一番重要なのかなと思います。

そういった中でどうやって事故を知らせるのかというと、先ほど西田専門委員からもありましたけれども、それぞれの年代の人たちがいる場所というのがそもそも違って、子供ですと割と学校とか保育園とかで事故が起こっているというのが出てきた場合には、そういったような保育園とか学校などにつながるようなところに重点的に協力してもらって、情報を出すのが必要でしょう。バスでしたらバスの事業者とかバスの車内に表示するとか、そういったようなそれぞれの情報によっていかし方が変わってくるのかなと思います。そこら辺はどういったパターンがあるのかの分析のもとに専門の方などを交えながら、どのような形で被害に遭いそうな方にリーチしていくのかということを考えていく必要があるのかなと思いました。

○相澤専門委員　ちょっと違う話になってしまうのですが、本日の話の中で、データベースを作る話と、いかにユーザーや消費者に危険情報を周知させていくかという2つのアプローチがあったと思います。前者ではバックエンドでデータベースを作るための非常に地道な努力が必要である。一方で後者は多様な個人にリーチするための多様なパスが必要になる。どちらも、コストのかかる話なので、全部をここでやる必要はないという意識は重要ではないかと感じています。例えばテレビで放送されたものがSNSに流されるパスがあるのであれば、テレビに対して有用な情報を提供することに注力するという考え方もあり得るので、1人で全部をやる必要はないという前提で、コストと効果のバランスを考えることは重要だと思います。

○山本座長　ありがとうございます。

よろしいですか。お願いします。

○黒木事務局長　今回は私どもの方から先生方に事故情報データバンクに入っている情報の分析をお願いしているわけですが、していただいみて御感想で結構なのですけれども、研究対象としてこれはなかなか面白いぞというか、やりがいがあるぞというものなのか、もう少しこのように変えるともっといろいろな人が、研究者の方とかが使えるとか、もっ

といろいろな分析をしてもらえるのではないかとか、お気づきの点がもしありましたら。あと、今回、先生方に非公表の部分も含めて御覧いただいていると思いますけれども、公表の部分だけだとどれぐらいのことができるのかということについて、もしお気づきの点がございましたら教えていただければと思います。

○山本座長 それでは、その点についてお一方ずつ伺いましょうか。

相澤専門委員からお願いできますか。

○相澤専門委員 研究対象としては目的が絞られているという意味で、価値があるデータではないかと思います。ノイズの問題はもちろんありますが、この手の問題、つまりデータの欠損があるとか、多くの不均質な情報源のデータが寄せ集まっているとかは、汎用的でいろいろな領域で今、切実に起きている問題なので、それに耐える頑強な手法が求められているのも事実です。ただ、全体としてデータの品質に大きな課題があったという印象は持っています。

○山本座長 市瀬専門委員、お願いします。

○市瀬専門委員 私の方は情報技術の観点からこのような情報処理をする専門家ということで今回参加させていただいたのですが、こういったようなデータを分析して実際に実務に結び付けていくことを考えたときには、その分野のデータの中身をよく分かっている人が解釈しないとなかなか難しい部分があるなどというのがあります。私のような情報技術者というのは、こういったところで情報処理をすることは幾らでもできるのですが、そこでこの見つけたものに対してどのような価値判断があるのか、何を見付けると施策に生かせるのか、そういったところは専門ではないので、もう少しこのような消費者問題に詳しい方がデータ分析に入っていただくような形で進められると、もう少しいろいろなことが分かるのかなというような気がしました。

○山本座長 では西田専門委員、お願いします。

○西田専門委員 今、お二人の委員がおっしゃられたことに大賛成です。基本的には研究の対象として、これだけ整備されているものが国内ではなかなかないこともあるので、何か申請をすると変な目的に使わない限りはオープンというか、分析、研究に使えるという仕組みがあるとすごくいいのかなと。使いたいという人は結構いるのではないかと思います。

消費者行政ということで、一義的には消費者を守るということに使っていくのですが、メーカーをどういうものか指導するみたいな感じになるのですが、イノベーションのヒントが相当入っていて、もっとよい製品、海外に打って出るような製品をつくるというようなイノベーションの側面でのヒントが詰まったデータベースでもあるので、そういう意味でメーカーなんかにもうまく開放というか、使えるようなものになっていくといいのかなと。そういう経済的なアプローチをしないと、税金だけで解決するというのは限界があるといろいろなところで言われているので、企業もやる気になっていいものを作るみたいな、そういう方向に持っていけるといいのかなと思いました。

○山本座長 先ほど行政機関の協力をいかに得るかというお話がございましたけれども、メーカーにとっても非常に種になる情報なのではないかと思imasるので、企業からの協力を得ながらしていくことがうまくできればいいかなと思imas。

それでは、藤田座長代理、お願imasします。

○藤田座長代理 私はデータの分析というよりは、どちらかというところから情報の流れに着目して分析を試みたのですけれども、まずはデータをどうためていくのかというところから出発したシステムだと思imasるので、そここのところの課題というものはもちろんまだあるのですが、もう少し視野を広げて生の情報が集まってくるころから、こうやって分析して、また消費者に還元していくような流れを意識して、これはプロセスとしてルーチン化できるようなところまで持ち上げると、すごくいいことが起きてくるのかなと思imasました。

やはり情報というのはどんどん伝わって行って、いろいろな人に伝わって行って、形を変えながら還流しているのです。なので還流している流れをつくるというのが重要だなというのを再認識した感じに思imas。

以上です。

○山本座長 村田専門委員、お願imasします。

○村田専門委員 ここまでの各委員の御意見でほとんど出尽くしている感はあるのですけれども、先ほど樋口委員からデータを作っている側にとって、こういうメリットがあるから、なのでこのデータが必要なんだというところをもっと明確にしてというお話もあったのですけれども、データバンクシステムとして一括してデータを蓄積していますが、参画機関それぞれの意図を持ってデータを収集して、それぞれ分析して、それぞれの行政に生かしているところもあると思imasのです。なので関連する人たちのこのデータをどううまく使うかというのは、いろいろな方向性があると思imasっていて、私はデータを見ただけの立場なので、消費者問題の専門家でもありませんし、行政の方にどううまく生かしていくかという専門家でもないのです。その辺りお互いにこういう分析ができたらというような情報交換、それをお伺いした上でこういった分析ができますといったような共同して新しい分析とか、新しい発信といったものを考えていく、検討する場を一時的なものではなくて、もっと長いスパンで考えて広げていったらどうなのかなというところは今回、分析も含めて専門調査会に参加させていただいて、そう思ったところです。

○山本座長 ありがとうございます。

何か事務局長からございますか。

○黒木事務局長 ありがとうございます。参考にさせていただきます。

○山本座長 その他にいかがでしょうか。

それでは、樋口委員、お願imasします。

○樋口委員 今お話がございましたけれども、私も、ぜひ関係機関との情報交換の場を広げていったらいいと思imas。ただ、いろいろ委員の方からお話を伺っていて驚いたことは、意外に簡単なことが実現していないということです。例えば年月日のお話がありまし

たけれども、その辺のミニマムの部分については統一できそうな気がしています、あるいは回答の空白の部分も一部は別に各機関に利益を還元しなくてもすぐにできるのかなと思います。こうした内容についても、消費者庁と各省庁や各機関が調整するとなると、よくある話ですが、各機関がそれぞれうちのやり方でやってほしいと主張して譲らないという話になりかねませんが、この専門調査会が、客観的に御覧になって、そこから話を始めて、せめてミニマムのルールといいますか、そういったところをまず具体化していくということが重要だと思います。1つでもきちんとルールができれば大分結果が違うということを皆様から指摘いただきましたので、せっかくですからその程度のところから始めて、それでさらに突っ込んで情報交流の話に持っていけると流れがいいのかなと、今お話を伺っていて感じました。

○山本座長 ありがとうございます。

すぐできるところから少し時間をかけてといいますか、労力をかけて取り組まなくてはいけないところまで、いろいろあるかと思いますが、確かに非常に単純ですぐできるようなことも先ほどから話を伺っているとありますので、まずはそこからということかもしれません。さらに長期的にどういうふうを考えていくかということかと思いますが、いかがでしょうか。その他にございますでしょうか。

それでは、お願いします。

○友行企画官 何人かの先生もおっしゃっていたのですけれども、分析ツールは1つのツールなので、最後は人の目で確認しないととか、そのようなお話があったと思うのですが、それをどのぐらいまで気にしなくてはいけないかということについての考え方みたいなものがあれば教えていただきたいのですけれども。

○山本座長 その点についていかがでしょうか。

それでは、西田専門委員からお願いします。

○西田専門委員 余りその点についてという感じではないのですけれども、重要なものはピックアップしてきた後には事例を読み込んでいくという作業で、それは全部読むのではないので、今回20万件からピックアップされた重要な領域で出てきたのは恐らく数百件とか、重要なものは数十件とかそういう話になるので、そこも読み込めるところなので、我々の研究所は学生が数千件読みますけれども、その数十件とかそのように絞り込んだものは、ちゃんと読んでいないと意外とおかしなものが混ざっていることもあるのでということではないかと思うのです。

○山本座長 ありがとうございます。

その他に今の点に関していかがですか。

何か事務局の方で更にございますか。

私の分野でも判例を検索すると、それに関連する過去の判例というものが自動的に出てくるというシステムがあるので、大抵は当たっている。しかし、中をよく読んでみると違っているものも混じっていたり、あるいは逆にこういう問題も実はあるんだと

いう新たな問題をそこから発見するということがあつたりいたしまして、例が悪いかもしれませんが、そここのところはどうしても、人の目が入らないと正確にならないし、逆に人の目が入るからこそ、そこから更に新たな発見があるという面もあるのではないかと思います。

その他にいかがでしょうか。是非この場で意見を言っておきたいと思うことがございましたら。

それでは、相澤専門委員、お願いします。

○相澤専門委員 先ほどの年月日の統一に関連して一言だけなのですが、やはり統一される場合にはシステムのサポート、つまり指定されたフォーマットでないと入力を受け付けないなど、そういったサポートを併せて行うことが重要だと思います。

○山本座長 ありがとうございます。

その他に是非御指摘をしたいということがあればお願いしたいと思っておりますけれども、いかがでしょうか。よろしいでしょうか。

本来は何かまとめなくてはいけないのかもしれませんが、もう既に申し上げておりますし、今、各委員でそれぞれ最後にコメントをいただきまして、それでほぼ尽きているかと思っておりますので、改めて繰り返すことはいたしません。

これで事務局の方で報告書のたたき台をつくっていただいて、その上で更に議論をして、最終的には報告書をまとめるという形で進めていきたいと思っております。

今日は5人の委員の皆様には大変お忙しい中を非常に面白い、興味深い報告をしていただきまして、私は何もしていなくて大変申しわけないという気がすると同時に、非常に感銘を受けました。この場で本当にお礼を申し上げたいと思っております。どうもありがとうございました。

これで終わりというわけではなくて、まだ報告書を作らなくてはけませんので、その段階でも今後とも御協力をいただければと思います。

---

## 《 5 . 閉会 》

○山本座長 本日の議題は以上となります。事務局から何かございますか。ございませんか。それでは、お忙しいところお集まりいただきまして、どうもありがとうございました。