

事故情報データの 品質向上に向けて

平成29年6月12日 第31回消費者安全専門調査会

(公財)統計情報研究開発センター 村田 磨理子

事故情報データベースシステムの利用

- 事故情報トピックス
- 事故情報を閲覧する
 - フリーワードで検索する
 - 詳しい条件で検索する
- 関係機関からの注目情報・お知らせ
- 事故情報データベースからのお願い
- 事故情報データベースからのお知らせ



事故情報を閲覧する フリーワード検索

- 例えば「充電器」で検索すると、1,142件がヒットする
- 「検索結果を集計」してみる
- 集計に指定できる項目
 - 被害者年代
 - 発生年月
 - 発生場所(施設用途)、
〃(項目)、〃(都道府県)
 - 商品など分類(大分類)、
〃(中分類)
 - 事故内容
 - 傷病内容
 - 傷病程度



検索結果を集計



- 被害者年代で集計すると、合計件数は1,088だが、年代別の内訳の和は110件(ヒットした件数1,142件と合計件数の違いもある)
- 「なお、集計結果の合計値には、無記入の件数も含まれます。」「複数選択が可能な項目の集計においては、各項目の件数の合算値と合計(件数)とが一致しない場合があります。」との注記あり
- つまり、件数の9割は、被害者年代が無記入なのだろうか？

検索結果を集計 つづき

集計軸	内訳の和 (件数)
被害者年代	110
発生年月	757
発生場所(施設用途)	522
発生場所(項目)	183
発生場所(都道府県)	194
商品など分類(大分類)	1,512
商品など分類(中分類)	1,529
事故内容	1,084
傷病内容	131
傷病の程度	124

- 内訳の和の件数が大きく異なる、つまり、無記入の状況は集計軸(項目)によって大きく異なると予想される
- 商品など分類は、複数選択が可能なため、内訳の和が件数を超えていると予想されるが、大分類と中分類の件数が一致しないのは疑問

事故情報詳細

事故情報詳細

リチウム電池内蔵充電器 事故情報ID:0000299452

印刷する

このページを開じる

種別	事故情報
発生年月日	2017年05月02日
発生場所	都道府県:東京都
商品など分類	住居品 - 他の住居品
商品など名称	リチウム電池内蔵充電器
型式・ロット	LP-MBV50WH
事故内容	火災事故
事故の概要	宿泊施設で当該製品を充電中、当該製品及び周辺を焼損する火災が発生し、1名が軽傷を負った。現在、原因を調査中。
傷病内容	
傷病の程度	詳細:軽傷1名
被害者人数	1人
被害者年代、性別	
事業者	事業者名:株式会社MSソリューションズ 事業区分:輸入事業者
原因調査状況	
事故原因	
措置状況	
備考	A201700074
情報提供元	関係機関 消費者庁・経済産業省・農林水産省 製品安全データベース

同一事故情報

同一事故情報はありません。

類似の情報

0000253852 - リチウム電池内蔵充電器

種別:事故情報

発生場所:長崎県

傷病内容:

事故内容:火災事故

0000265559 - リチウム電池内蔵充電器

種別:事故情報

発生場所:鹿児島県

傷病内容:

事故内容:火災事故

0000231144 - リチウム電池内蔵充電器

種別:事故情報

発生場所:埼玉県

傷病内容:

事故内容:火災事故

0000231845 - リチウム電池内蔵充電器

種別:事故情報

発生場所:大阪府

傷病内容:

事故内容:火災事故

0000279207 - リチウム電池内蔵充電器

種別:事故情報

発生場所:東京都

傷病内容:

事故内容:火災事故

0000260477 - リチウム電池内蔵充電器

種別:事故情報

発生場所:

傷病内容:

事故内容:火災事故

- 事故情報詳細について、事故の概要の自由記述文の内容と区分分けされている項目、例えば「商品など分類」、「商品など名称」との対応を個別にみると、事故の内容が類似であっても異なった区分に分類されている場合があるといった、分類の揺らぎが散見される

事故情報詳細 つづき

- 前述の分類の揺らぎは、商品の全体をみるか部品をみるか、本体と付属品があるときに、本体をみるか付属品をみるかといった、定義の違いがあると思われる
- そのほか、利用者の視点では、次のような要望も考えられる
 - 「事故原因」、「措置状況」をまとめて知りたい
 - 「被害者年代、性別」の項目があるので、検索結果の集計に「被害者性別」も追加できないか

フリーワード指定と項目指定

事故情報検索・集計_事案 × 事故情報詳細(リフォーム電池内) +

jikojoho.go.jp/ai_national/search/aiccc09-accidentsearchcondition.do

項目指定

事故情報ID:

種別 (複数可) ☐ 事故情報 ☐ 危険情報

商品など分類: 大分類 中分類

商品など名称:

事故内容 (複数可)

☐ 火災事故	☐ 発煙・発火・過熱	☐ 点火・燃焼・消火不良	☐ 破裂
☐ ガス漏洩	☐ ガス漏れ	☐ 燃料・液漏れ等	☐ 化学物質による危険
☐ 漏電・電流等の障害	☐ 製品破壊	☐ 部品脱落	☐ 機能故障
☐ 転落・転倒・不安定	☐ 操作・使用性の欠陥	☐ 交通事故	☐ 誤飲
☐ 中毒事故	☐ 異物の混入・侵入	☐ 腐敗・変質	☐ その他

傷病内容 (複数可)

☐ 骨折	☐ 脱臼・捻挫	☐ 切断	☐ 擦傷・挫傷・打撲傷
☐ 刺傷・切傷	☐ 頭蓋(内)損傷	☐ 内臓損傷	☐ 頸部・背髄の損傷
☐ 筋・腱の損傷	☐ 窒息	☐ 熱傷	☐ 凍傷
☐ 皮膚障害	☐ 感電障害	☐ 一酸化炭素中毒	☐ 食中毒
☐ その他中毒	☐ 感覚機能の低下	☐ 呼吸器障害	☐ 消化器障害
☐ その他の傷病及び諸症状	☐ 不明	☐ 呼吸器障害	☐ 消化器障害

傷病の程度 (複数可)

☐ 不明	☐ 医師にかからず	☐ 治療1週間未満	☐ 1～2週間
☐ 9週間～1カ月	☐ 1カ月以上	☐ 死亡	

登録年月日 (西暦): 記入例 2010年02月01日
 年 月 日 から 年 月 日まで

発生日 (西暦): 記入例 2010年02月01日
 年 月 日 から 年 月 日まで

発生場所(住所): 都道府県

発生場所(施設用途) (複数可)

☐ 住宅	☐ 店舗・商業施設	☐ 学校	☐ 病院・福祉施設
☐ 公園	☐ 道路	☐ 公共施設	☐ 河川・山・川等自然環境
☐ 車内・船内・船内	☐ その他		

発生場所 (複数可)

☐ 階段	☐ 浴槽・風呂場	☐ 台所	☐ 玄関
☐ 居室	☐ 洗面所	☐ ベランダ	☐ 庭
☐ 廊下	☐ 昇降機(エレベーター)	☐ エスカレーター	☐ 動歩道
☐ 自動ドア	☐ 回廊	☐ その他	

発生場所(施設名):

被害者の人数: 人から 人まで

被害者年代 (複数可)

☐ 0歳以下	☐ 1～4歳	☐ 5～9歳	☐ 10歳代
☐ 20歳代	☐ 30歳代	☐ 40歳代	☐ 50歳代
☐ 60歳代	☐ 70歳代	☐ 80歳以上	

被害者性別 (複数可)

☐ 男性	☐ 女性	☐ 不明
-----------------------------	-----------------------------	-----------------------------

このページの先頭へ ▲

- 事故原因、措置状況は項目指定に対象にはなっていない
- 項目指定で被害者性別を指定すると、1,142件のうち7件のみがヒット

データの品質向上に期待されること

- 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に寄与する

本報告について

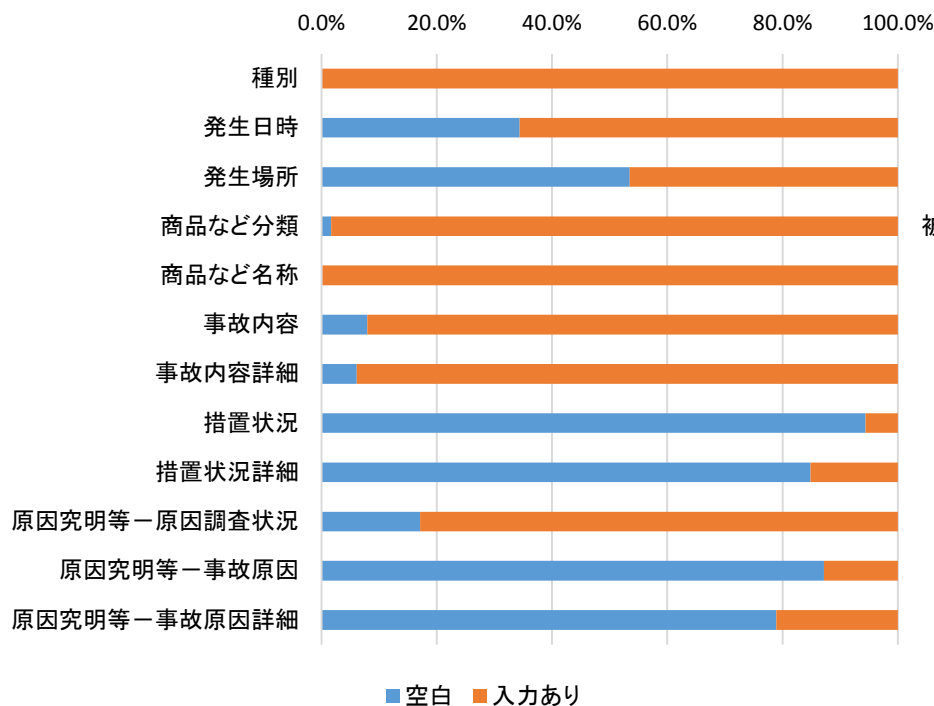
- データの品質は、設計、データの収集、データのコード化及び編集、公表といった処理過程のすべてが影響する
- 本報告は、既存の完成されたデータファイルの内容を利用した分析の範囲にとどまる
- 分析結果からは、データのコード化及び編集に直接関係するもののほかに、設計やデータの収集などに関する示唆も得られる

事故情報データについて

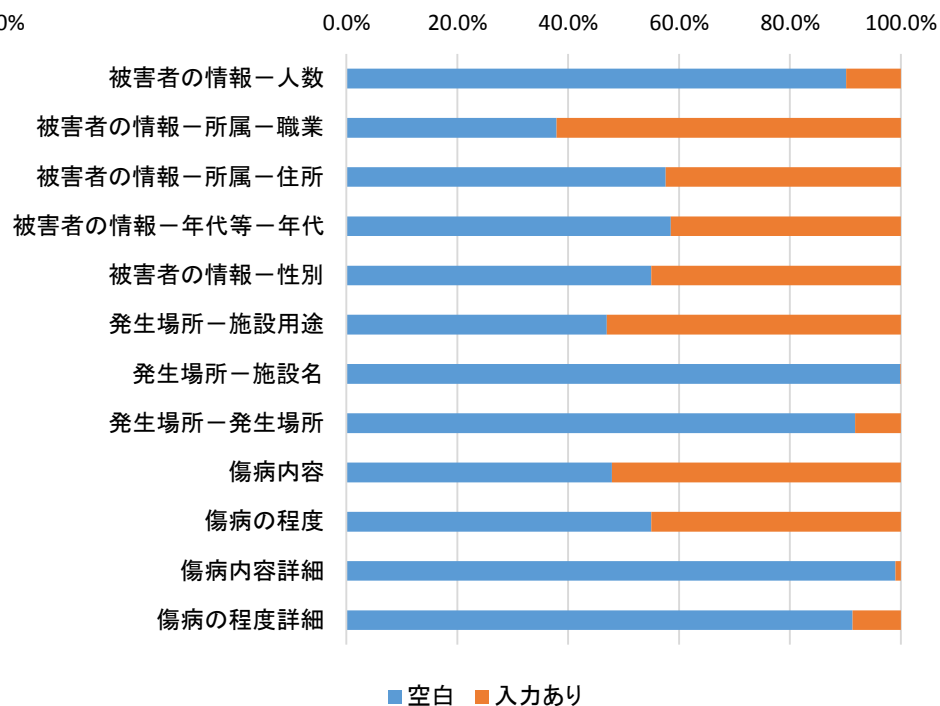
- 分析対象のデータ
 - 83項目（一般向けウェブサイトでは非表示の項目を含む）
 - 198,135レコード（登録年月日が2009年10月1日から2016年12月31日まで）
- 項目の収録状況
 - 項目によっては、空白が非常に多い（95%以上が空白である項目は40項目、そのうち13項目はすべて空白）
 - 情報提供元によって、空白の多さは異なる
 - 空白の原因は、無記入、該当なし、元の情報に対応する項目がないなどの理由が考えられる

項目の収録状況

項目別収録の状況

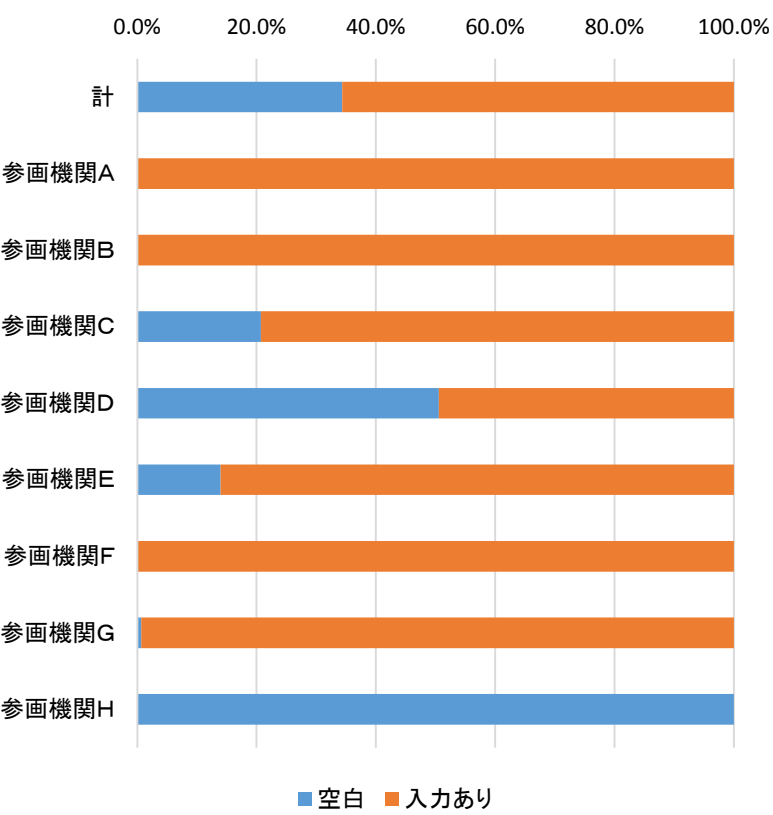


項目別収録の状況

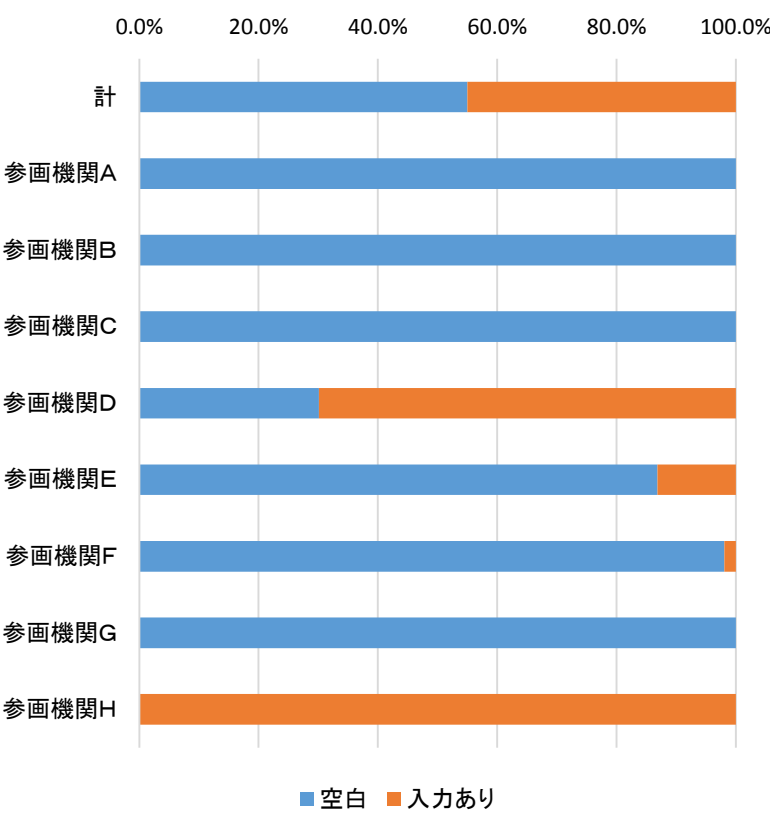


項目の収録状況(情報提供元による違い)

発生日時の入力状況 情報提供元別



被害者性別の入力状況 情報提供元別



計は、参画機関A～Hと外5機関を含む

収録形式の設定：発生日時

実際に収録されている内容は、
次のように様々な形式になっている

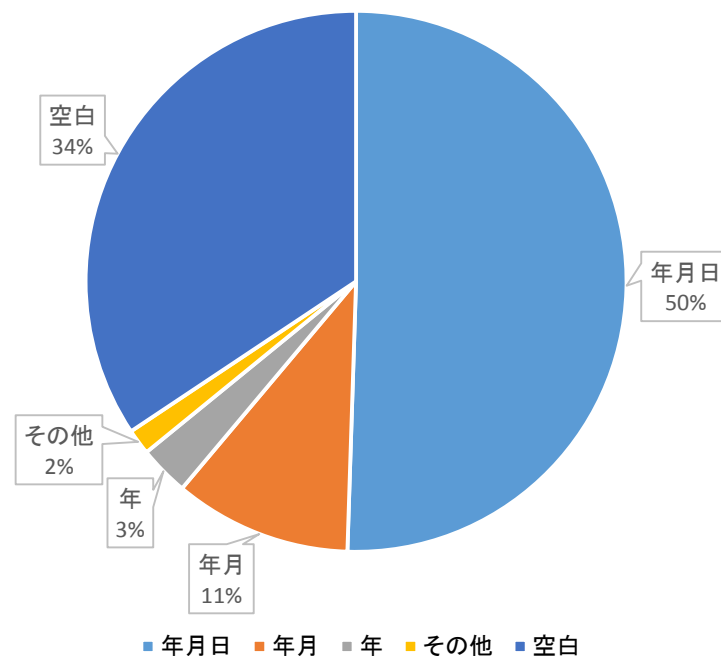
2001年
2001年4月
2001年1月16日
2001年02月22日8時
2009年06月04日15時30分頃
2007年09月23時
07年8月、08年2月、09年5月、11年2月
2006年夏頃
2007～2016年(継続中)
3～4年前から半年ごとに発生
夏 気温30度以上
購入から1年
約1カ月前
別件でアイドルリングストップ不具合の修理後から

- 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
- 現在または基準となる時点からの経過・遡及時間

形式の統一について

- 発生時点や受付時点を基準とした時系列分析を行うに当たって、現状のデータは使い勝手がよくない
- 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
- 最初の情報提供を受ける際に、少なくとも「年」は必須とするなどの対応が必要ではないか
- 現状のデータに対して形式統一処理をしたとすると、右のような構成割合になり、半数は「年月日」が利用でき、約6割は「年月」の情報を利用できる

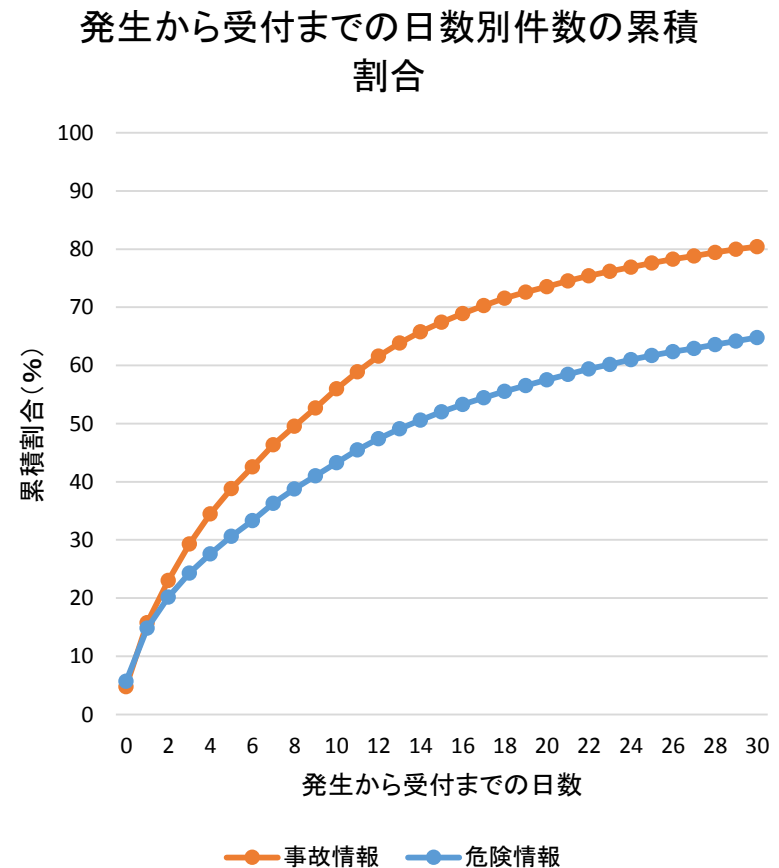
発生日時の収録形式の構成割合



事故発生から受付までの日数

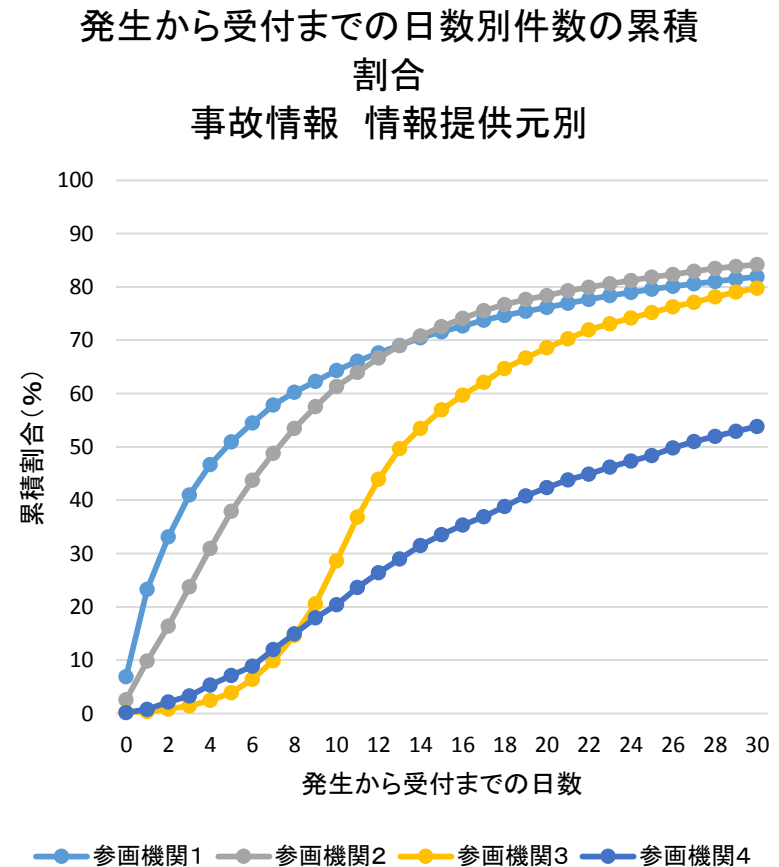
- 年月日を取り出せた82,698件を集計した
- 事故情報は発生から30日までに8割が受付され、危険情報は30日で6割程度が受付される
- 中央値で比較すると、事故情報は9日、危険情報は14日である

(注意) 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推測することは困難である



事故発生から受付までの日数

- 事故情報に限定して、情報提供元ごとにみると、日数の分布に違いがみられる
- 違いは、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、分析が深まると考えられる



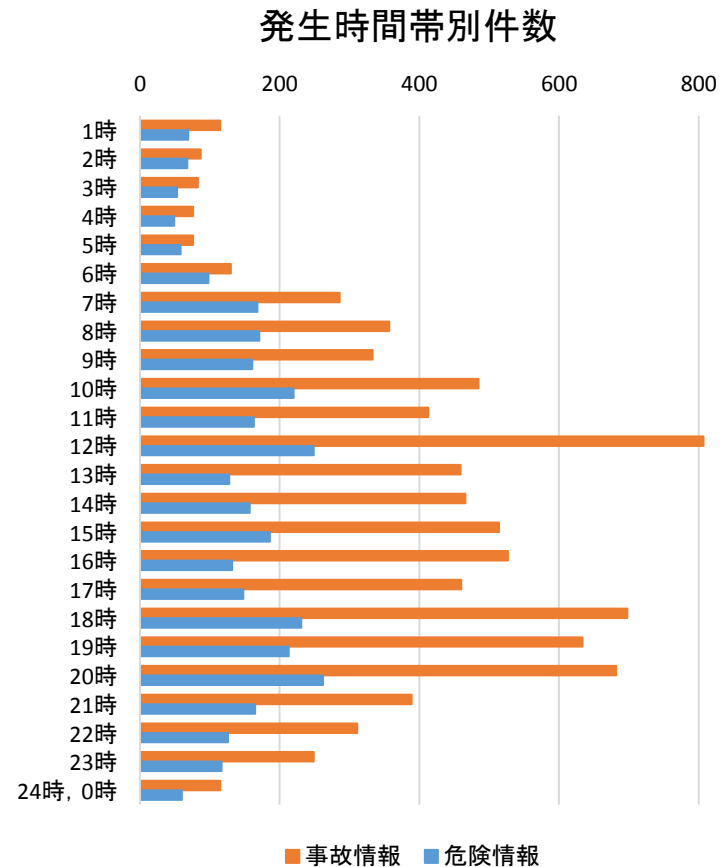
注意：日数利用可・不可の比率が情報提供元ごとに一律ではないため、本当の経過日数の分布とは異なる可能性がある

日数に関する考察

- 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。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
- 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ことが重要と思われる
- 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（日数に限らず、全般にわたって）

事故発生の時間帯

- 発生日時から時刻の情報を取り出せた12,204件について、時間帯別に件数を集計した
- 12時の件数が多いが、事故内容を確認すると、実際に12時台に発生したと考えられないものが散見される。発生日時を入力する際のルールを確認する必要があるだろう
- 12時を除いてみると、事故は午前よりも午後、特に18～20時台に多く発生している傾向がみられる



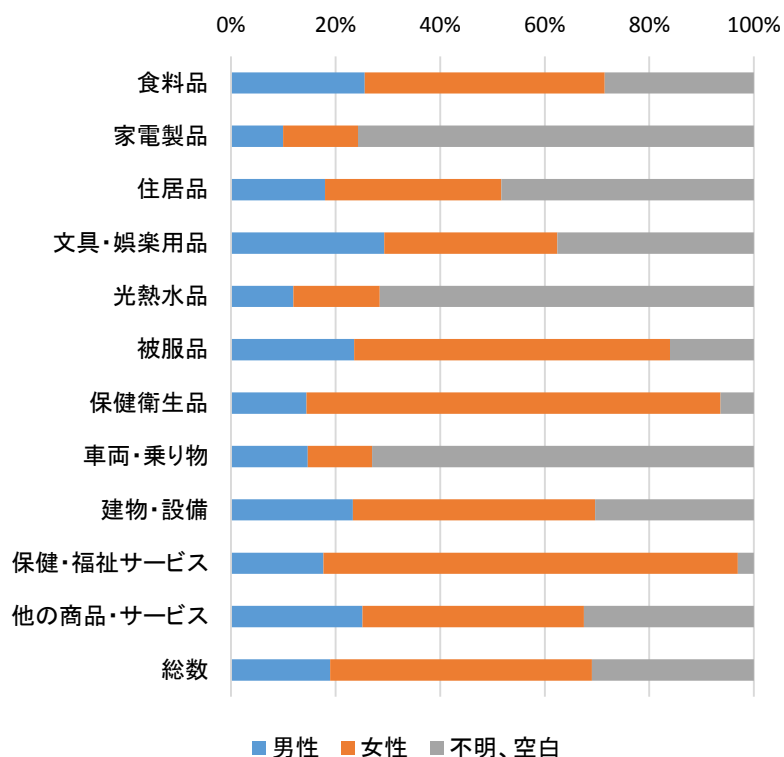
(注意)時刻を利用できないものが大部分であるため、本当の時間帯別の分布とは異なる可能性がある

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
- 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、被害者性別と商品の分類とのクロス集計を試みる
 - 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集計
 - 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点に注意が必要

クロス集計の例

商品など分類別被害者性別の構成割合
—情報提供元に限定



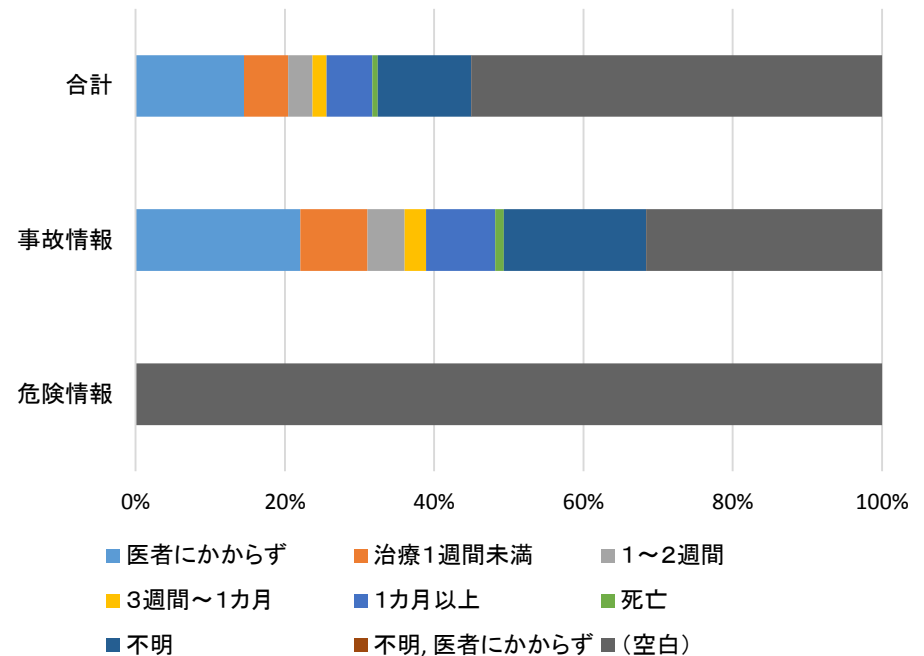
	男性	女性	不明、空白
食料品	6,975	12,535	7,793
家電製品	1,133	1,628	8,616
住居品	3,082	5,789	8,287
文具・娯楽用品	1,339	1,517	1,718
光熱水品	187	260	1,125
被服品	875	2,247	595
保健衛生品	4,831	26,468	2,136
車両・乗り物	1,618	1,355	8,051
建物・設備	2,194	4,362	2,860
保健・福祉サービス	3,881	17,445	667
他の商品・サービス	3,959	6,659	5,117
総数	22,993	60,750	37,620

- 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
- 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

事故情報と危険情報

- 情報提供元によって、事故情報のみの提供と、事故情報と危険情報の両方を提供している場合がある
- 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
- 「人的被害なし・物損事故」の区分を追加するか、「傷病の程度は事故情報にのみ収録する」旨の説明が必要

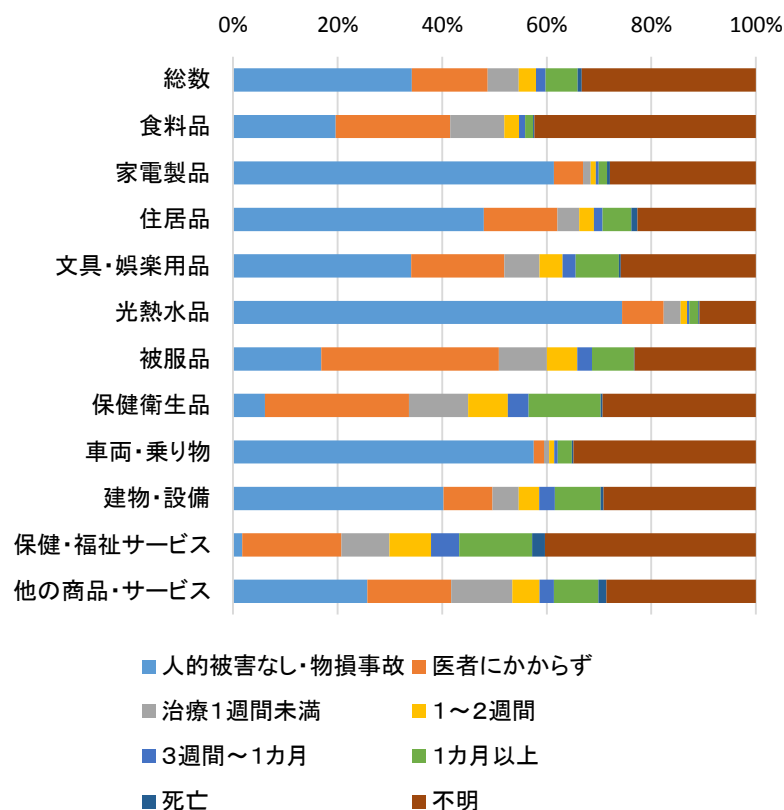
傷病の程度の構成割合



事故情報と危険情報

- 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
- 保健・福祉サービス、保健衛生品、被服、食料品は人的被害が出ている割合が高い
- 保健・福祉サービス、保健衛生品は1カ月以上・死亡の割合も比較的高い
- 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

商品など分類別傷病の程度の構成割合



自由記述の扱いについて

- 情報提供元で商品や事故内容の分類ルールを設定している
- 複数の情報提供元の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すると推察する
- 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が散見される
- 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してはどうか
 - 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

分類区分について: 事故内容

事故内容	度数	パーセント
その他	83,800	42.29
火災事故	21,033	10.62
製品破損	18,848	9.51
(空白)	15,800	7.97
発煙・発火・過熱	14,350	7.24
中毒事故	12,114	6.11
異物の混入・侵入	10,898	5.50
機能故障	5,225	2.64
化学物質による危険	2,754	1.39
破裂	2,406	1.21
部品脱落	2,146	1.08
操作・使用性の欠落	1,936	0.98
転落・転倒・不安定	1,893	0.96
燃料・液漏れ等	1,426	0.72
腐敗・変質	1,039	0.52
点火・燃焼・消火不良	929	0.47
ガス漏れ	497	0.25
漏電・電波等の障害	451	0.23
交通事故	258	0.13
誤飲	174	0.09
ガス爆発	158	0.08

- 「その他」が4割を占める
- 情報提供元によって、「その他」の割合に大きな違いがある
- 情報提供元による分類の違いを考慮に入れて、「その他」の細分化を検討する余地があるのではないか

項目間の整合性について：傷病の程度

- 「傷病の程度」と「傷病の程度詳細」の2つの項目が必ずしも整合していない。例えば、「傷病の程度詳細」に「死亡」の記載があるのに、「傷病の程度」は空白になっているなど
- また、「傷病の程度詳細」には、「人的被害なし・物損事故」の記載が比較的多い。これに相当する区分を「傷病の程度」に追加した方が、登録や利用がしやすいのではないかと
- 比較のため、東京都生活文化局消費生活部の実施した「家庭での電子レンジの安全な使用に関する調査」、「スマートフォンの安全な使用に関する都民へのアンケート調査」をみると、危害及び危害に遭いそうになった経験として、「ケガをしなかった」(危険情報に相当)は、「ケガをした」(事故情報に相当)の2倍以上多く回答されている

データバンクへの入力項目の設定

- 分析対象のデータに収録された83項目のうち、事故情報登録時に入力が必要となっている項目は、情報管理用のIDや情報提供元に関する項目を除けば、種別（事故情報か危険情報かの区別）のみとなっている
- 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ではないか

データ収録内容に関するその他の課題

- 「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、分類の区分ごとに「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－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－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ためかもしれない

データ収録内容に関するその他の課題

- 「原因究明等－原因調査状況」と「原因究明等－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－事故原因」に記載があれば、「原因究明等－原因調査状況」は空白にしないなど
- 「発生場所」、「発生場所－施設用途」、「発生場所－施設名」、「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のではないかと
 - 特に、「発生場所－施設名」は、「発生場所－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

まとめ

- データの品質向上のために
 - 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、事故情報データベース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
 - 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