

テキストマイニングによる FGI データの分析

Analysis of Focus Group Interview Data by Text Mining

1. 緒言

今回の研究では、原子力施設の安全確保についてリスクコミュニケーションのあり方を検討することを目的として実施されたフォーカスグループインタビュー (FGI) のデータ (議事録) を、テキストマイニングの手法を用いて分析した。本稿では分析の結果、および得られた知見について述べる。

2. FGI データの分析

2.1 FGI の概要と、その対象者

FGI は、司会者、専門家、記録者および、インタビュー対象者という構成で実施されたものである。対象者は、都市部および立地地域 (福島県立地市町村) の男女である。主となる質問の内容は以下の4つである。

- (1) リスク情報を活用した原子力発電所の検査見直しに対する印象
- (2) 関係主体に対する信用
- (3) 見直しにより節減されたコストの適切な用途
- (4) 非原子力分野へのリスク情報を用いた検査の適用の妥当性

2.2 分析の実施手順

上記 FGI の議事録の内容を、インタビューの内容ごと、対象者ごとの2通りの分け方により分析した。

2.2.1 インタビューの内容ごとの分析手順

- (1) 各議事録を、対象者のカテゴリ (都市部、立地地域、男、女) ごとに主となる4つの質問に切り分ける。
- (2) 切り分けたものを、形態素解析ソフト「茶筌」を用いて形態素に切り分ける。
- (3) 分析に有用と考えられる形態素以外を除去する。
- (4) 得られた形態素をキーワードとしてヒストグラムを作成し、各カテゴリを比較する。

2.2.2 対象者ごとの分析手順

- (1) 各議事録を、対象者の発言ごとに切り分ける。
- (2) 前項と同様に、「茶筌」を用いて有用と思われる形態素を抽出する。
- (3) 得られた対象者ごとのキーワードの組をベクトルとして、式(1)のようにして対象者2人の組み合わせごとに余弦を計算する。
- (4) 得られた余弦を親近度として評価し、比較する。

3. 数式

$$\mu_{ij} = \frac{\vec{v}_i \cdot \vec{v}_j}{\|\vec{v}_i\| \|\vec{v}_j\|} \quad (1)$$

対象者 i と j のキーワードのベクトルをそれぞれ \vec{v}_i 、 \vec{v}_j とし、親近度 (余弦) を μ_{ij} とする。

4. 結果・考察

分析を通じて得られた主な知見は、下記に示す通りである。

- (1) 都市部、立地周辺とも男性では、自身の職業経験に基づくと考えられる専門用語がキーワードとして多く見られた。
- (2) 節減されるコストの適切な用途については、立地女性は教育や地元への還元についてのキーワードが多かった。
- (3) 非原子力分野について、女性は都市部、立地とも食べ物に関するキーワードが多かった。
- (4) リスク情報活用の是非に関する質問全般に関して、対象者のどのカテゴリにおいても肯定的なキーワードが多かった。
- (5) 親近度として余弦を用いた評価については、値が高いからといってその二人の発言者が同じような発言をしていると認められるものではなく、今回の場合では余弦が必ずしも親近度として適当であったとは言えなかった。

5. 結論

今回のテキストマイニングの手法は、有用と思われるキーワードの抽出に関して若干の曖昧性を含んではいるものの、原文をほとんど読まずにそれに近い結果が得られたという点である程度有効なものであったと考えられる。

参考文献

- 1) 林俊克, Excel で学ぶテキストマイニング入門, オーム社(2002).