

短報

計量テキスト分析を用いた井戸水等水質検査相談内容（2009-2022年度）のデータ分析

古賀智子

当所では、井戸管理者を対象として主に井戸水の検査業務を実施しており、業務の中で水に関する様々な相談を受ける。本研究では、相談内容やそれに対する回答を対象としてデータ分析を行い、当所に持ち込まれる相談内容及びその回答の特徴を把握することを目的とした。データ分析の手法の一つである計量テキスト分析（テキストマイニング）を用いテキスト型データを分析した結果、相談者から具体的に提示される分析項目は主にマンガン、農薬、細菌であることや、過去の職員の対応状況などが明らかとなった。

[キーワード：井戸水、相談、計量テキスト分析、テキストマイニング]

1 はじめに

福岡県では、福岡県飲用井戸等衛生対策実施要領¹⁾に基づき、飲用井戸水について毎年一回以上定期的に検査を実施することを井戸管理者に推奨している。当所では、県民に向けて井戸水の検査窓口を設置しており、理化学検査を水質課（当課）、細菌検査を病理細菌課で実施している。この業務において、井戸水をはじめとした水に関する相談を受けることがあり、その内容を記録している。持ち込まれる相談内容は多岐に渡り、個々の相談に対して都度職員で話し合い、どのように回答するかを決めている。そのため、相談内容にどう対応するかは職員の経験によるところが大きく、誰もが画一的に対応できるようにマニュアル化することは難しい。

そのため本研究では、過去に受けた相談内容及びその回答を分析することによって、当所に持ち込まれる相談内容とその回答の特徴を見出し、知見の継承に資することを目的とした。対象データの一部が自由記述であったため、計量テキスト分析（テキストマイニング）を用いてデータ分析を行った。

2 方法

2・1 対象データと使用ソフトウェア

対象としたデータは、2009年度から2022年度までに職員が受けた井戸水などに関する相談を聞き取った記録とそれに対する回答とした。記録した項目は、“相談に用いられた手段（電話・来所など）”、“相談内容のカテゴリ”、“相談内容”、“相談内容に対する回答”であった。相談内容のカテゴリは“臭気”、“着色”、“異物”、“農薬”、“検査項目”、“検査受付”、“検査結果”、“井戸水以外”、“その他”とした。“相談内容”及び“相談内

容に対する回答”は自由記述式であった。

計量テキスト分析にはフリーソフトウェアのKH Coder3 (Windows版)²⁾を用いた。計量テキスト分析とは、議事録や新聞記事など数値化できないテキスト型データを計量的に分析するための手法で、テキストマイニングとも呼ばれている。先行事例として、白井及び高橋³⁾が水道利用者からの異臭味苦情について計量テキスト分析を活用している事例がある。今回は、“相談内容”及び“相談内容に対する回答”が自由記述式であるためKH Coder3を用いることとした。本研究で利用したKH Coder3の機能を以下に説明する。

1) KWICコンコーダンス

抽出した語が文章中でどのように用いられているかを探ることができる。テキスト型データの解析では、抽出語そのものだけでなく、それが使用された前後の文脈を確認することが重要であるため利用した。

2) コロケーション統計

KWICコンコーダンスを行った語について、その語の前後にどんな語が出現していたかを容易に読み取ることができるため、関連語の検索に有用である。

3) 共起ネットワーク

データ中で出現パターンが似通った語、つまり強い共起関係にある語を線で結んでネットワークを描く手法である。可視化することでデータを理解しやすくなり、データ中にあらわれたテーマやトピックを読み取ることができる。

2・2 対象データの前処理

対象データを用いて計量テキスト分析を行う前に、データの前処理を行った。表1及び表2に前処理内容を示す。前処理内容は、例えば記録者によって“Cu”や“銅”な

ど元素記号で記述したり和名で記述したりして表記が統一されていないため、元素名は全て和名とした。また、“pH”は一つの単語として認識されないため、強制的に一つの単語として抽出されるようにした。そのほか、前処理を実行すると“前”や“月”などが一つの単語として抽出されたが、本研究においては特に意味をもたないと考えられたため使用しないこととした。

これらの前処理を行ったデータを解析対象とした。

表1 単語の統一

元の単語	修正後
Cu, Mnなどの元素記号	銅、マンガンなどの和名
当所、保環研、弊所	研究所
試験	検査
飲料	飲用

表2 強制抽出語及び使用しない語

強制抽出語	使用しない語
pH	前 月 気 場合

3 結果及び考察

2009年度から2022年度の間に当課が受けた相談は182件で、年平均13件であった。その内、電話109件、来所71件、メール2件であった。

相談内容のカテゴリごとの件数を図1に示す。一つの相談が複数のカテゴリにまたがることもあるため、相談内容のカテゴリごとの件数の総和は全相談件数よりも多くなっている。“その他”が55件と最も多く、カテゴリに分類できない様々な相談が持ち込まれていることが分かった。次いで多かったのが、“検査受付”に関するもので45件、“着色”に関するものが26件であった。最も少なかったのは“臭気”に関するもので5件であり、当所は異臭に関する相談は少ないことが分かった。

次に、相談内容とその回答についてKH Coder3を用いて解析を行った。表3に相談内容とその回答の頻出語上位22位までを示す。相談内容と回答どちらも1位は“検査”、2位は“水”であった。そのほか、上位10位までを井戸水の水質検査に関する一般的な用語が占めた。相談内容の頻出語をみると、14位に“結果”という単語が出現した。KWIC コンコーダンスをみると、検査結果の解釈に関する相談がいくつかみられた。相談者が検査結

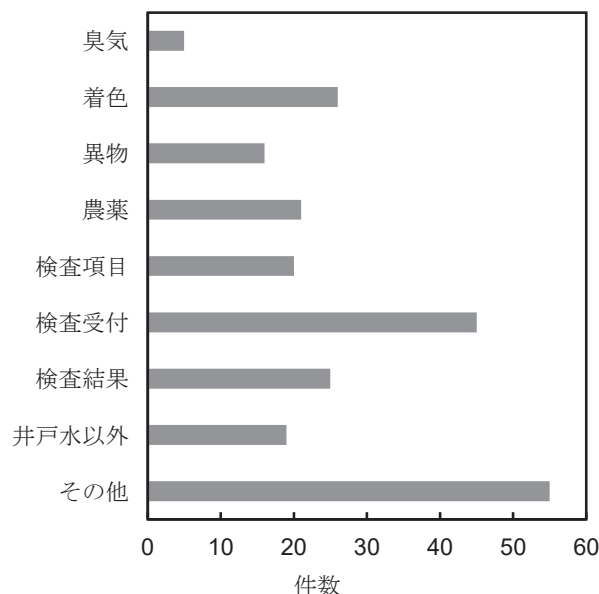


図1 相談内容のカテゴリごとの件数

果を提示するだけでなく、そこから何が言えるのかを分かりやすく説明して欲しいという要望があることが分かった。また、具体的な水質分析項目の名称である“マンガン”、“農薬”、“細菌”という単語が出現しており、当所ではこの3項目に対する相談が多いことが分かった。

“マンガン”のKWIC コンコーダンスをみると、洗濯物やタイルが黒くなるため検査をして欲しい、という相談がいくつかみられた。“農薬”については、自宅や近隣で散布された農薬が井戸水に混入していないか気になるので検査可能か、との相談が複数みられた。“細菌”については、一般細菌が基準を超えていたが煮沸すれば飲用しても問題ないか、という相談が多かった。

次に回答の頻出語をみると、相談内容には出現しなかった“銅”が上位に出現した。“銅”という単語は、風呂場のタイルが青くなるなど、青い着色についての相談のときに出現していた。“銅”が相談内容の頻出語上位に出現せず回答のみにも出現していることから、前述した黒い着色の原因となるマンガンの場合は相談者から具体的にマンガンの検査をして欲しいとの発言があるのに対して、青い着色については相談者は原因が分からないことが多いことが示唆された。そのため、青い着色に関する相談があった場合には、銅が原因である可能性に言及し検査を提案することを考慮に入れる必要があることが分かった。そのほか、22位に“容器”という単語が出現した。コロケーション統計の結果をみると、“容器”の単語の前後に“料金”や“項目”などの単語が出現していた。このことから、当所の水質検査内容全般についての問合せが多くあることが分かった。検査内容全般の問合せは、

表3 相談内容（左）とその回答（右）における
出現回数上位22位の単語

相談内容の頻出語			回答の頻出語		
順位	単語	出現回数	順位	単語	出現回数
1	検査	151	1	検査	277
2	水	114	2	水	120
3	井戸水	89	3	項目	101
4	使用	71	4	分析	99
5	依頼	45	5	説明	86
6	分析	44	6	飲用	60
7	飲用	43	8	水道	35
8	水道	35	8	水質	59
9	水質	33		伝える	
10	井戸	31	10	可能	55
11	研究所	27	11	対応	54
12	問い合わせ	26	12	基準	52
13	可能	25	13	細菌	51
14	結果	24	14	銅	50
16	マンガン	23	15	水道	48
17	基準	21	16	依頼	47
19	原因	18		研究所	
	出る		18	一般	44
	浄水		19	確認	40
	細菌		20	結果	38
			21	理化学	35
			22	容器	34

総務課が担当することが多いが、当課でも電話や窓口において分析項目、料金、採水方法などをスムーズに説明できるようにしておく必要性があることを認識した。最後に、相談内容及び回答に関する共起ネットワークをそれぞれ図2及び図3に示す。図2の中で“言う”という単語に着目すると、“業者”や“水道”と共起関係があった。KWICコンコーダンスをみると、“言う”は全て“言われた”という活用形で使用されていた。このことから、業者や水道局に問い合わせ、アドバイスを受けた上で当所に持ち込むという傾向があることが分かった。また、図3では“紹介”という単語と“民間”や“機関”の共起関係が強い結果となった。KWICコンコーダンスをみると、当所で検査が困難な検体の場合に民間の検査機関を紹介していることに起因していた。当所で検査困難な検体とは、例えば①水道水や飲用井戸水以外の検査依頼、②当所の検体搬入希望日時に持ち込めない依頼、③井戸水に含まれる固形物の分析依頼、などであった。この他に、農薬の分析依頼については農薬の種類が不明な状態で検査を相談される場合が多く、当所では農薬の一斉分析の検査を受けていないため、民間の検査機関を紹介することがあった。表3において農薬に関する相談が多いことが分かったが、当所で分析できない場合は民間の検査機関

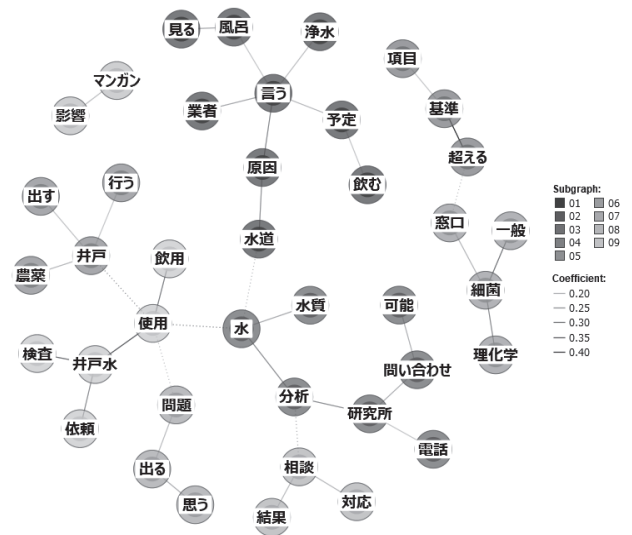


図2 相談内容の共起ネットワーク（Subgraph：比較的強く共起関係にあるグループ、Coefficient：共起関係の強さを表す係数）

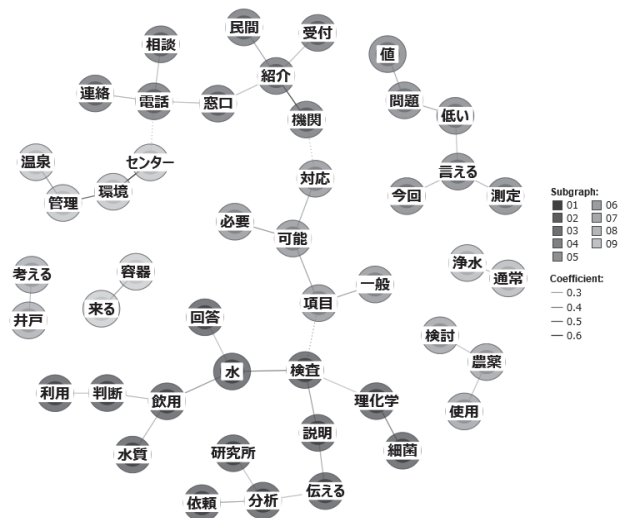


図3 回答の共起ネットワーク（Subgraph：比較的強く共起関係にあるグループ、Coefficient：共起関係の強さを表す係数）

の紹介をスムーズに行うことができるように体制を整えておく必要があることが分かった。

4 まとめ

本研究で示された当所に持ち込まれる相談内容やその回答の主な特徴を以下に示す。

- ・相談者は検査結果から何が分かるのかを知りたいことが多いため、水質検査結果を提示するだけでなく検査結果の解釈を分かりやすく説明する。
- ・黒い着色の原因の一つにマンガンがあるという知識を

有している相談者は多いが、青い着色の原因について不明である人が多い。相談内容をよく聴いた上で、適宜銅の検査を提案する。・農薬については、農薬の種類が不明の状態に相談にくる場合が多く、当所で一斉分析を実施していないため民間の検査機関を紹介する。

・水道水や飲用井戸以外の検査依頼、当所の搬入希望日に持込み困難な依頼、井戸水に含まれる固形物の分析依頼などについては当所で検査は困難であるため、民間の検査機関を紹介する。

・細菌については、一般細菌が基準を超えていたが煮沸すれば飲用しても問題ないか、という相談が多いため飲用は推奨しない旨を説明する。

文献

- 1) 福岡県飲用井戸等衛生対策実施要領，令和元年12月9日。
- 2) 樋口耕一：社会調査のための計量テキスト分析—内容分析の継承と発展を目指して—，第2版，2020（ナカニシヤ出版）。
- 3) 白井隆太，高橋英司：計量テキスト分析を活用した異臭味苦情内容の解析，水道協会雑誌，90，12-18，2021。

(英文要旨)

Data analysis of consultations for water examination by quantitative content analysis

Tomoko KOGA

*Fukuoka Institute of Health and Environmental Sciences,
Mukaizano 39, Dazaifu, Fukuoka 818-0135, Japan*

We provide well water examination services and receive various requests for consultations about water. The purpose of this study was to understand the characteristics of these consultations by analyzing the consultations requested and their answers. We used quantitative content analysis (text mining) to quantitatively analyze text-type data. This analysis revealed that the analytes requested by the clients were mainly manganese, pesticides, and bacteria; and how the staff had dealt with the clients in the past.

[Key words; well water, consultations, quantitative content analysis, text mining]