

論文

ソーシャルワーク研究における テキストデータ分析に関する一考察

日和恭世[†]

要約：本稿の目的は、研究法における信頼性や妥当性の意味を問い直し、ソーシャルワーク研究における質的研究のあり方について考察することである。質的研究は信頼性や妥当性が担保されにくいと指摘されることが多いが、この課題を解決するために大きな役割を果たすのがテキストデータ分析ソフトである。膨大なデータを扱う場合、これまで手作業で行っていた部分を自動化することによって、効率的に分析することができる。しかし、テキストマイニングだけではテキストデータの文脈を捉えることができないことから、ソーシャルワーク研究においてテキストデータを分析する際には、量的にも質的にも分析することが必要である。

キーワード：テキストデータ分析, 信頼性, 妥当性, QDA, テキストマイニング

目次

1. 問題意識と研究目的
2. 質的研究における課題と新たな方向性
 - 2-1. 質的研究における信頼性と妥当性
 - 2-2. 質的研究におけるデータの加工と分析
3. 先行研究の検討
 - 3-1. テキストマイニングを用いた先行研究
 - 3-2. テキストマイニングの課題
4. ソーシャルワーク研究におけるテキストデータ分析ソフトの活用
 - 4-1. テキストデータ分析ソフトの可能性
 - 4-2. 今後の課題

1. 問題意識と研究目的

近年、ソーシャルワークにおいては、EBP (Evidence Based Practice) の重要性が指摘され、EBP との関連で研究方法のあり方がくり返し議論されている⁽¹⁾。そのなかでたびたび論じられているのが、何をエビデンスとするかということであり、アメリカ医療

[†]同志社大学大学院社会学研究科博士後期課程

*2013年6月27日受付, 2013年7月1日掲載決定

政策研究局（Agency for Health Care Policy and Research：AHCPR）をはじめ、Gambrill（2006）、McNeece & Thyer（2004）など様々な研究者がエビデンスの強さをもとにしてエビデンスのレベルを作成している。例えば、アメリカ医療政策研究局（AHCPR）の分類によると、エビデンスレベルの高い順に、Ia：ランダム化比較試験をメタアナリシスしたもの、Ib：少なくとも一つのランダム化比較試験によるもの、IIa：少なくとも一つによくデザインされた非ランダム化比較試験によるもの、IIb：少なくとも一つによくデザインされた準実験的研究によるもの、III：よくデザインされた非実験的記述的研究によるもの、IV：専門家委員会の報告や意見によるもの、とされている（正木ら2006）。これらのエビデンスレベルによれば、データに基づき、客観的に結果を示すことができる量的研究に比べて、質的研究はエビデンスレベルが低いとされる。その理由としては、質的研究で扱うテキストデータは、研究者の視点や判断などによって分析、解釈されるため、量的研究のように同じ手続きを踏めば誰が行っても同じ結果が出るわけではなく、信頼性や妥当性が担保できにくい点があげられることが多い。

しかしながら、ソーシャルワークが対象とする事象は、複雑・多様なものであり、数字だけでは表すことのできないものである。特に、クライアントの置かれている状況やニーズは個別的で多岐にわたり、それぞれの社会的文脈において捉えなければならないものである。そのため、ソーシャルワークにおいては、量的研究では捉えることのできない知見を得るため、グラウンデッド・セオリー・アプローチやKJ法、事例研究などの質的研究が積極的に実施されてきた。その意味で、社会における質的研究の意義は相当地に大きなものであると考えられる。

そこで、本稿では、研究法における信頼性や妥当性の意味を問い直し、ソーシャルワークにおける質的研究のあり方について考察することを目的とする。

2. 質的研究における課題と新たな方向性

2-1. 質的研究における信頼性と妥当性

質的研究とは、既存の研究方法にとらわれることなく、我々が生きる日常生活の場で起こる現象について、「現象の新たな側面を発見したり、実証的なデータに基づいて新たな理論を生み出したりすること」（Flick＝2002：9）を目的とした研究方法である。

質的研究の歴史は古く、社会学や心理学の領域では19世紀末から20世紀初頭から用いられていたと言われる。しかし、次第に数量化や標準化を目指す量的研究が主流となり、量的研究を用いることが科学的であると認識されるようになった（Flick＝2002）。だが、今日、質的研究の意義が改めて問い直され、人文・社会科学に限らず、医学や看護学などにおいても質的研究を用いた研究が積極的に取り入れられるようになってい

る。

Flick によれば、質的研究の基本的な特徴は、①研究対象に適した方法と理論を選ぶこと、②異なった様々な視点を考慮に入れ分析すること、③研究者の自分の研究に関する反省をも考慮に入れること、④様々なアプローチと方法を用いること、の4つであるという (Flick = 2002: 7-11)。

質的研究では、人々の経験や認識等、量的研究では明らかにすることができない実態を説明できるという点で大きな意味を持つ。しかし、その一方で、研究者の解釈が研究の本質的な部分になるため、その解釈や結論に至るまでのプロセスが見えにくい。そのため、量的研究で用いられている評価基準からすると、信頼性や妥当性が担保できにくいように捉えられてしまうのであろう。しかし、ここである疑問が生じる。それは、質的研究に量的研究に用いられている信頼性や妥当性などの評価基準をそのまま当てはめることができるのであろうか、というものである。そこで次に、信頼性と妥当性の意味について検討することにした。

量的研究において、信頼性とは、「尺度の安定性や精度の高さ」のことであり、「安定性とは、測定者が異なっても、何回測定しても、同一の事象については常に同一の結果が得られることを意味する。」とされる (和気 2010: 93)。これは、再現可能性とも言い換えることができるものである。一方、妥当性とは、「尺度が意図する概念を正確に反映して測定しているかどうかを問うもの」であるとされる (和気 2010: 93)。簡単に言えば、「その研究で見ようとしているものをしっかり見ることができているか」ということである。

質的研究において、量的研究のような信頼性や妥当性などの基準を応用するのか、それとも、質的研究に適した評価基準を新たに作るのか等、質的研究の評価基準をめぐっては様々な議論がなされている (Flick = 2002)。例えば、質的研究では、データ記録を形式化すること、インタビュー法の実習をすること、データ解釈の際に解釈者間で手続き等の共通理解を得ること、などの方法で信頼性を高める工夫がなされている。Flick によれば、質的研究において、信頼性という基準は「データやさまざまな研究手続きの『確実性 (dependability = 当てにできるかどうか)』の検証」という意味で捉えられ (Flick = 2002: 275)、結果自体の確実性は問題にされていない。

つまり、量的研究では、分析結果を不変的なものとして捉えているのに対して、質的研究では、分析結果は可変的なものであると捉えられており、分析結果よりも結果を導き出す手続きの確実性に焦点が当てられているのである。

では、妥当性についてはどうであろうか。妥当性ということば自体の意味するところは、量的研究の場合と違いはないが、質的研究における妥当性は、ある現象に対して研究者が導き出した解釈がどの程度データに根拠をもっているかという点で問われるとい

う (Flick=2002)。質的研究では分析によって事実そのものを表すのではなく、フィールドや語りのなかにある根拠に基づいて、研究者がその研究対象に対するひとつの見方を表現している。そのため、質的研究において妥当性を高めるためには、ひとつの研究において複数の手法を用いたり、異なる理論的立場をとったりすることで、より多面的に捉えることや、インタビュー内容を文字に変換した後に調査対象者に発言内容が正しいかどうかを確認するメンバー・チェック等、トライアングレーションが重要になると考えられる。

このように、質的研究と量的研究における信頼性、妥当性の意味するところは全く同じではないものの、質的研究においても研究の質を担保するために、信頼性や妥当性を高めるための努力は絶えず行なわれている。

2-2. 質的研究におけるデータの加工と分析

2-2-(a). テキストデータの質的分析

前述したように、質的研究は、ある現象を深く理解するためには非常に重要な研究方法である。しかし、扱うデータが大量になると、データを分類、整理するにはかなりの時間と労力が必要となる。この悩みを解消するために、近年、テキストデータ分析のためのソフトウェアが開発されている。このようなソフトウェアは、「文字テキスト情報を文書型データベースとして体系的に整理し、また分析していくために開発されたコンピュータ・プログラム」であり (佐藤 2008 b: i), 総称して「QDA (Qualitative Data Analysis) ソフトウェア」または「CAQDAS (Computer Assisted Qualitative Data Analysis Software)」と呼ばれる。佐藤によれば、これらのソフトウェアは欧米ではすでに 20 年以上も前から使用されているが、日本語に対応できるものがなかったことから、これまで日本ではあまりなじみがなかったという (佐藤 2008 b)。しかし、近年、日本語対応のソフトが開発され、日本においても質的研究に取り入れられるようになってきている⁽²⁾。

QDA ソフトの具体的な機能としては、①文書データセットの管理、②文書データに対する編集およびコーディング、③コード同士の関係に対応した分析モデルの構築、④特定のコードに対応するカードの検索・抽出、の 4 つがある (佐藤 2008 a: 135-136)。具体的には、QDA ソフトを活用することによって、紙媒体では管理が大変であったテキストデータを効率的に分類・整理することやテキストデータを見ながらコーディング作業を行うことができるようになる。また、図解機能もあることから、コーディングしたものを体系的に整理し、図式化することも可能である。

このように、QDA ソフトは「基本的には、紙媒体でおこなわれてきたデータ分析の作業の手順を電子化したもの」であり (佐藤 2008 b: 54), 「時間の節約」, 「研究の質

の向上」, 「データ管理の簡素化」などのメリットがある (Flick=2002: 311-312)。しかし, これらのソフトを導入したからと言って自動的に分析が行われるわけではない。その意味で, 量的研究で用いられるようなデータを投入すれば結果が抽出される統計分析ソフトとは大きく異なっていると言える。

つまり, テキストデータを質的に分析する際, これまで手作業で行っていた作業を自動化することによって, 質的研究におけるいくつかの課題は解消されるものの, 結局のところ, テキストの解釈の部分は研究者一人ひとりに委ねられているのである。

2-2-(b). テキストデータの量的分析

QDA ソフトがテキストデータを質的に分析するのを補助するものであるのに対して, テキストデータを量的に分析すること助けるソフトウェアも開発されている⁽³⁾。質的データを量的に分析する方法としては計量テキスト分析や計量内容分析などがあるが, それらの分析手法をソフトウェア化したものが「テキストマイニング」と呼ばれるものである。

テキストマイニングとは「テキストデータをさまざまな計量的方法によって分析し, 形式化されていない膨大なテキストデータという鉱脈のなかから言葉 (キーワード) どうしに見られるパターンや規則性を見つけ, 役に立ちそうな知識・情報を取り出そうとする手法・技術」であり (藤井 2005: 10), 「自然言語処理, 統計解析, データマイニングなどの基盤技術の上に成り立っている。」ものであるとされる (松村ら 2009: 1)。言い換えれば, テキストデータという大きな山を掘り進め, 集めたものを統計的に分析することによって新たな知見を発見する, という宝探しのようなものである。

テキストマイニングは, これまで商品に関する感想や要望等, 消費者の生の声を把握することが重要視されるビジネスやマーケティングの領域において頻繁に用いられてきた。しかし, 近年, 学術領域においても積極的に取り入れられている。

藤井によれば, 「テキストマイニングのプロセスは, テキストデータの収集, 集められたテキストデータの分析, 分析結果の解釈という3段階をとる。」という (藤井 2005: 26)。分析のプロセスだけを見ると一般的な調査と違いはないが, 大きく異なるのは, まず集められたテキストデータを, 意味をもつ最小の言語単位である形態素に区切っていく作業を行うことである。これを「形態素解析」という⁽⁴⁾。さらに, 助詞などの分析にあたって不要な言葉を削除したり, 同じ意味を表す言葉を規定したりすること等によりキーワードを抽出し, 数量的に処理できる状態にする⁽⁵⁾。次に, 形態素に分解した言葉を対応分析, 主成分分析, クラスタ分析などの手法を用いて分析し, その結果を布置図や樹形図などであらわすというプロセスをたどる。

このように, テキストマイニングではテキストデータを統計的に処理できる状態に変

換することができるため、統計解析が可能になり、その結果、「数量化」、「可視化」することができるのである。つまり、テキストマイニングは質的研究のなかに量的研究の要素を取り入れた新しい方法であると言える。

これまで見てきたように、QDA ソフトもテキストマイニングもテキストデータを分析するという点では同じであるが、コンピュータを通して行う作業は全く質の異なるものである。つまり、テキストデータの分析には、テキストデータを質的に分析する方法とテキストデータを量的に分析する方法の2つがある。テキストデータを質的に分析することはこれまでに数多く行われてきたが、質的に分析するだけでは見出すことのできないものも存在するのではないかと考えられる。その意味で、テキストデータを量的に分析する方法への期待は大きいのではないだろうか。

そこで、次にテキストデータを量的に分析するテキストマイニングについてさらに詳しく見ていくことにしたい。

3. 先行研究の検討

3-1. テキストマイニングを用いた先行研究

では、ソーシャルワークにおいてテキストマイニングを用いた研究はどの程度行われているのであろうか。

前述したように、テキストマイニングの手法を用いてテキストデータを分析する際には「Word Miner」や「KH Coder」などの分析ソフトを使用することが多い。Word Minerとは、形態素解析、テキストデータの量的データへの変換、多変量解析などを行うソフトウェアのことである。一方、KH Coderは、樋口が開発した計量テキスト分析を行うためのソフトである。計量テキスト分析とは、「計量的分析手法を用いてテキスト型データを整理または分析し、内容分析を行う方法である」とされ、内容分析の一種または一部として位置づけられている(樋口 2006: 18)。Word Minerは高額であるが、KH Coderはフリー・ソフトウェアであり、手軽に利用できるものである。ここでは、これらのテキストマイニング専用の分析ソフトを用いてテキストデータを量的に分析しているものを「テキストマイニングを用いた研究」と捉え、先行研究をレビューすることとする。

論文検索サイト CiNii で「テキストマイニング」「ソーシャルワーク」と検索したところ、2件ヒットした。また、「テキストマイニング」「福祉」では26件がヒットした⁶⁾。これらの論文のテーマを概観したところ、「テキストマイニング」「福祉」の検索結果26件中には、「テキストマイニング」「ソーシャルワーク」で検索してヒットした2件中1件が含まれていた。この26件のなかには、心理学や建築学などの他分野の論文も含まれており、純粋に福祉の領域における研究であると判断できるものは26件中

12 件であった。

次に、「計量テキスト分析」をキーワードとし、「計量テキスト分析」「ソーシャルワーク」、「計量テキスト分析」「福祉」と検索したところ、該当したのはいずれも同一論文 1 件のみであった。

計量テキスト分析は、計量的手法を用いて内容分析を行う手法であることから、さらに「内容分析」をキーワードとして検索を試みた。その結果、「内容分析」「ソーシャルワーク」では 2 件、「内容分析」「福祉」では 70 件がヒットした。ただし、この中には看護学や心理学など他分野の論文もあり、福祉の領域における研究であると判断できるものは 70 件中 19 件であった。そのうち、計量的に内容分析を行っているものは 1 件のみであった。

テキストデータの分析手法であるグラウンデッド・セオリー・アプローチや KJ 法についても同様に CiNii にて「ソーシャルワーク」、「福祉」をキーワードに先行研究を検索した。その結果、福祉の領域における研究と判断できるものは、グラウンデッド・セオリー・アプローチが 46 件、KJ 法が 30 件であった⁽⁷⁾。これらの検索結果から、ソーシャルワークにおいてテキストマイニングの手法をとっている研究は現段階ではそれほど多くはないといえる。テキストマイニング専用の分析ソフトを使用している先行研究の詳細は表 1 の通りである。

まず、先行研究の分析手法に着目すると、大半が Word Miner を用いて分析を行っていることがわかる。また、Word Miner では、比較的簡単に多変量解析を行うことができるため、頻出語の関連性の分析だけでなく、対応分析やクラスター分析を行っているものも多い。

次に、分析対象について見てみると、①アンケート調査の自由記述をテキストマイニングを用いて分析したもの、②インタビュー調査のテキストデータをテキストマイニングを用いて分析したもの、③先行研究の記述内容をテキストマイニングを用いて分析したもの、の 3 つに分類することができる。

また、内容に関しては、①次世代育成行動計画や市町村障害者基本計画等、福祉計画策定のために住民のニーズを把握することを目的とするもの、②クライアントのニーズ把握を目的とするもの（福祉計画策定以外）、③ある事柄についての対象者の意識や認識の抽出を目的とするもの、④ある概念の概念整理を行うことを目的としたもの、に整理できる。

これらの先行研究から、テキストマイニングを用いた研究のメリットとしては、大量のテキストデータの処理が可能であること（鳩間ら 2004；伊藤ら 2008；田垣 2009）、データの恣意的な解釈を回避できること（鳩間ら 2004；南 2005；小野 2006；山頭ら 2011）、バラバラに見えるデータから共通性を見出すことが可能であること（小野

表1 テキストマイニングを用いた研究

研究者	対象とするデータ	方法	概要
鳩間亜紀子 児玉桂子 田村静子 (2004)	質問紙による 自由回答	Word Miner 対応分析	介護福祉士とホームヘルパーを対象に高齢者向け住宅改造の効果に関するアンケート調査を行い、自由記述データ252件を分析することで、住宅改造の評価指標と効果測定にかかわる要因を導き出している。
南彩子 (2005)	質問紙による 自由回答	Word Miner 対応分析 クラスター分析	社会福祉専攻の大学生99名を対象に「高齢者観」についてのアンケート調査を行い、自由記述データを分析することで、学習や実習などの経験によって学生の高齢者観にどのような変化が見られるかを検討している。
峰島里奈 (2005)	インタビュー調査	Word Miner 対応分析 クラスター分析	児童期・青年期に親を亡くした青年6名を対象にインタビュー調査を行い、その結果を分析することによって、遺児のニーズを明らかにしている。
小野セレスト摩耶 (2006)	質問紙による 自由回答	Word Miner クラスター分析	次世代育成行動計画策定のための調査の一環として行った記述式アンケート調査のデータ553名分を対象に、就学前の子どもをもつ母親の子育ての不安・負担のそれぞれ内容や両者の関係性について考察している。
志村健一 (2007)	先行研究(論文)	KH Coder	畠山龍郎氏が執筆した『ソーシャルワーク研究』の「巻頭言」20編を分析し、畠山氏のソーシャルワーク論について考察している。
坪井真 (2007)	先行研究(文献)	Word Miner	1960年代移行に発表された方面委員に関する研究と代表的な社会事業の通史研究について、「方面委員」の概念に関する記述部分を分析することで「方面委員」の概念を整理している。
伊藤利恵 渡辺俊之 (2008)	質問紙による 自由回答	Word Miner	認可保育所に勤務する保育士498名を対象に家族支援についてのアンケート調査を行い、自由記述データを分析することにより、家族支援に対する保育士の意識と保育所におけるソーシャルワーク機能について考察している。
田垣正晋 (2009)	質問紙による 自由回答	Tiny Text Mining 主成分分析 クラスター分析	市町村障害者基本計画のニーズ調査のうち、自由記述データ1,809名分をテキストマイニングとKJ法を用いて分析し、住民のニーズを把握している。また、これらの分析を通して、テキストマイニングと質的研究の分析手法の併用について検討している。
高橋一公 (2010)	質問紙による 自由回答	KH Coder 対応分析	介護福祉士を目指す短大生68名を対象に「老人観」についてのアンケート調査を行い、自由記述データを分析することで、一般的な老人イメージと自己の老人イメージを明らかにしている。
山頭照美 飛永高秀 (2011)	インタビュー調査	KH Coder	軽費老人ホーム、養護老人ホーム、ケアハウスの入居者41名を対象に「すまい」についてのインタビュー調査の結果を分析し、頻出語の関連性を考察することにより、「すまい」の捉え方を明らかにしている。
京 俊輔 (2012)	インタビュー調査	Word Miner 対応分析	知的障害者の性に対する教育や支援を効果的に展開する必要があるとの問題意識から、施設職員6名と保護者9名にインタビュー調査を行い、その結果を分析し、両者の知的障害者の性に対する意識を比較している。
峯山 太 (2012)	インタビュー調査	KH Coder 多重応答分析	地域包括支援センターのソーシャルワーカー3名を対象にグループインタビューを実施し、その結果を計量テキスト分析にて分析することで、ソーシャルワーカーがソーシャルワークの焦点をどのように捉えているかを考察している。

2006), などが考えられる。

しかし、先行研究が少ないことや研究者が十分にテキストマイニングの活用方法を身につけていないことから、いくつかの課題も有している。例えば、質的データを量的データに変換するための作業や質的データを量的データに変換した後の統計解析の手続きが不十分であることなどがあげられる。特に、質的データの量的データへの変換に関しては、「分かち書き処理手続き、置換語や削除語抽出の手続きなど検討すべき課題は多い。」とされる(小野 2006: 46)。また、テキストマイニングを活用したからといって、必ずしも新しい知見が得られるわけではない(高橋 2010)。データ数によっては、元のテキストデータを丁寧に読むことによってより深く理解することができるかもしれない。

したがって、今後テキストマイニングを活用していくには、質的データを量的データへと変換する作業の透明性の確保や統計解析の技術の習得、また、テキストマイニングの対象とすべきデータ量であるかの吟味などが求められるのではないかと考えられる。

3-2. テキストマイニングの課題

前述したように、テキストマイニングはこれまで質的研究の大きな課題とされてきた信頼性や妥当性を一定程度担保することができるという点では、これまでにない画期的な分析方法のように思われる。しかし、果たして本当にテキストマイニングを用いることによって、分析者の主観に偏ることなく客観性を確保することが出来るのであろうか。

藤井は、テキストマイニングのメリットは大きいものの、限界があることも肝に銘じておく必要があると述べている。具体的には、①テキストデータ自体がデータとして曖昧であること、②形態素解析において文章自体のもつ意味合いを考慮しなければならないこと、③テキストデータだけでは回答者の意図や真意は把握できないこと、の3点があげられている(藤井 2005: 26-27)。言い換えれば、統計的な分析を行うことによって「数量化」や「視覚化」をすることは可能であるが、データの収集やデータの分析に至るまでの部分に曖昧さが存在するという点である。この点は先行研究レビューからも明らかになった部分であり、テキストマイニングを活用する際の課題のひとつであると考えられる。

情報処理においてよく用いられる格言のひとつに「garbage in, garbage out」というものがある。これは、コンピュータにゴミのようなデータを入力すれば、結果として出てきたものも価値のないものになる、という意味のことばである。つまり、コンピュータに入力するデータそのものが価値あるものでなければ、いくら客観性を担保するためにテキストマイニングの手法を用いても、得られた結果は信頼性や妥当性のないものになってしまうということである。前述した藤井の指摘にもあったように、テキストマイニングを用いるに当たっては、テキストデータを分析する前段階である形態素解析やそれ

に付随する不要語の削除や類義語の規定、意味のある言葉の抽出等が分析の質を左右する非常に重要な作業になる。言い換えれば、テキストマイニングでは、統計解析を行う前に何が不要であり、何が同じであり、何が意味のある言葉なのか、といった研究者の解釈や判断が不可欠であり、これらの解釈や判断を抜きにしては成り立たない分析手法であると言える。したがって、テキストマイニングにおいては、テキストデータを分析するまでの段階でいかに信頼性や妥当性を担保するか、ということが大きな課題になるであろう。

そこで、次に、このようなテキストマイニングの課題を解決する方法について考えていくことにしたい。

先行研究のなかには、テキストマイニングとその他の質的研究の手法を組み合わせることで、この課題の解決を目指すものもある。例えば、田垣は、障害者基本計画のニーズ調査の自由記述をテキストマイニングと KJ 法を併用して分析し、両者の併用のあり方を検討している。田垣によれば、テキストマイニングでは頻度の低い言葉は分析することができないのに対し、KJ 法では言葉の頻度に関わらず意味の連関を検討できることから、「テキストマイニングだけでは、新しい知見を発見したり、従来漠然と知っていることを言語化したりすることは難しく、KJ 法の併用が必要である」という（田垣 2009: 79）。また、稲葉らは、がん告知の可否に関するフォーカスグループインタビューについて、テキストマイニングとグラウンデッド・セオリー・アプローチとを組み合わせた「グラウンデッドなテキストマイニング・アプローチ (GTMA)」という新たな手法を用いて分析している。これは、テキストマイニングの統計的分析による可視化（テキストデータの定量化）とグラウンデッド・セオリー・アプローチによる研究者自身の解釈（テキストデータの定性化）という両者のメリットを活かした手法であり、片方の手法だけでは見出すことのできない新たな知識を構築できる可能性があるとする（稲葉ら 2011）。

そもそもテキストマイニングとは、テキストデータを形態素に分解し、単語の出現頻度や相関関係などを導き出すものである。「数量化」や「視覚化」をすることによって、質的研究が抱えるデータの分析や解釈における研究者の客観性の問題を解消しようとするものであることは前述したとおりである。これに対して、グラウンデッド・セオリー・アプローチは、実証主義的研究への批判から 1960 年代にグレーザーとストラウスによって提唱された、データに密着した分析から理論の生成を目指すものである（木下 2003）。佐藤によれば、QDA ソフトの多くはグラウンデッド・セオリー・アプローチの発想をもとにデザインされており、テキストデータの解釈に重きが置かれているという（佐藤 2008）。一方、テキストマイニングのソフトには統計解析の手法が取り入れられているため、分析結果を客観的に示すことは可能であるが、テキストデータのもつ意味

の連関を検討することはできない。したがって、テキストマイニングは、テキストデータを分析するものではあるが、あくまでも数量化や統計解析などを最終的な目的とする量的研究に近いものであるといえる。

4. ソーシャルワーク研究におけるテキストデータ分析ソフトの活用

4-1. テキストデータ分析ソフトの可能性

ソーシャルワークにおいて、クライアント集団のニーズ把握やソーシャルワーカーの業務実態を明らかにする研究等の場合には、量的研究が実施される。量的研究によって、個々の対象の特徴ではなく、対象の全体像を客観的に把握することが可能となる。しかし、量的研究の場合、その多くが選択肢法を用いたアンケート調査によるものであるため、得られるデータは断片的なものになってしまう。なぜならば、選択肢法を用いたアンケート調査の場合、回答者の回答は質問項目に規定された範囲にとどまり、質問項目として設定されていない事柄については、回答を得ることができないからである。この課題は、アンケート調査に自由記述欄を設けることによって解消することができ、自由記述によって研究者が想定していないような回答を得ることも期待できる。

しかし、量的研究の場合、回答数が膨大になることが多いことから、並行して得られる自由記述の量も多くなり、その一つひとつを丁寧に読み、そこから何らかのルールを発見するにはかなりの時間と労力を要すると考えられる。そのため、自由記述のテキストデータは、これまで活用されることなく、眠ったままになっていることも多かったという（藤井 2005）。

このような課題を解決するのがテキストデータ分析ソフトである。テキストデータ分析ソフトを用いることによって、自由記述欄に記入された大量のテキストデータを短時間で効率的に分析することができるようになる。また、インタビューなどにおいてもテキストデータ分析ソフトは当然有意義であると考えられる。

質的研究は、量的研究では捉えることができない実態を把握する場合に用いられる。例えば、クライアント個々のニーズや状況の理解、一人ひとりのソーシャルワーカーのものの見方や考え方などを明らかにする場合などが考えられる。質的研究の場合、解釈の手續きの透明性が求められることは前述したとおりであるが、もし、テキストデータ分析ソフトを活用することができれば、手作業の場合とは比べものにならないほど、テキストデータの分類・整理が容易になり、データ保管のスペースを削減することもできる。

このように、扱うテキストデータの量が膨大である場合にテキストデータ分析ソフトを活用することは、研究の質やスピードの向上に貢献すると考えられる。

また、前述したように、テキストデータ分析ソフトを活用することによって、テキストデータを量的に分析することも可能であることから、質的研究の場合でも量的研究の場合でも、テキストデータを量的に分析することで、これまで得られなかった新たな知見を見出すことができるのではないだろうか。

例えば、クライアントを対象としたインタビュー調査を実施した場合、語と語の結びつきから、分析者の解釈では明らかにすることができなかつた事柄を導き出すことができると考えられる。

また、先行研究のように、概念整理にテキストマイニングを用いることも大変意義のあることであると考えられる。日本のソーシャルワーク研究は、欧米から大きな影響を受けている。加えて、ソーシャルワーク自体、他の学問領域の多様な知見をもとに構築されていることもあり、ソーシャルワークの領域では似たような概念があまり整理されずに用いられている感がある。例えば、ソーシャルワークには生活モデル、ストレングスモデル、心理社会的アプローチ、問題解決アプローチ、エコロジカルパースペクティブなど、多様なモデルやアプローチ、パースペクティブが存在する。もともとはモデル、アプローチ、パースペクティブはそれぞれ異なる概念であるが、研究者によって各概念の定義づけが異なるため、それぞれの違いを理解するには困難を極める。このような場合、様々な文献で用いられている各概念の定義をテキストマイニング用のソフトに入力することによって、用いられている言葉の重なりなどの特徴をつかむことができる。このように、概念整理を行うような基礎的な研究を行う場合にテキストマイニングを用いることによって、バラバラに使用されている概念を多少なりとも整理することができるのではないだろうか。

しかしながら、インタビュー調査によって得られたテキストデータを分析する研究の場合、テキストマイニングの手法だけで分析すると、インタビューイーの語りに込められた真意がこぼれ落ちてしまう可能性がある。その意味で、テキストマイニングは、「数量化」「可視化」を行うことで、あくまでも質的研究で課題であるとされる客観性や再現性などの部分を担保するための手法であり、テキストマイニングだけでテキストデータの本質を分析できるものではないと考えられる。したがって、テキストデータの分析を行う研究の場合には、テキストマイニングのみを用いるのではなく、グラウンデッド・セオリー・アプローチやKJ法などの手法を併用し、テキストデータを量的にも質的にも分析することによって、より信頼性のある分析結果を得ることができるのではないだろうか。

4-2. 今後の課題

近代科学においては、普遍性、論理性、客観性が求められるとされる（中村 1992）。

ソーシャルワークもこれらの3つの原理をもとに科学化を目指してきた歴史がある。その意味で、ソーシャルワーク研究においてテキストマイニングを用いることは、3つの原理のうち普遍性や客観性を担保することにつながり、実証的な研究を行う場合には意義があると考えられる。しかしながら、従来のこのような科学のあり方では、データを操作的に対象化することにより、分析において見落とされてしまう部分があることも推察される。このような点から、心理学や社会学などの分野においては従来の実証主義的な科学の捉え方に対する批判も多く、ソーシャルワークにおいても「科学化」とは何かを改めて問い直す必要があるとの指摘もある（狭間 2000；平塚 2007）。

テキストマイニングは再現性という点では確かに有用な分析手法である。しかし、量的分析はテキストデータの一部は明確化することができても、そこに含まれる内容を多面的に総合的に分析、解釈することはできない。そのため、テキストマイニングを単独で用いるのではなく、従来の質的研究方法を併用することによって、より多角的な分析ができるであろう。そうはいつても、ソーシャルワークにおいてテキストマイニングを用いた研究は緒についたばかりである。したがって、テキストマイニングの可能性については、今後の研究の動向をみながらさらに検討していくことが必要であろう。

注

- (1) 『ソーシャルワーク研究（相川書房）』では、「ソーシャルワークの研究手法」（2004）、「エビデンス・ベースド・ソーシャルワーク」（2008）、「ソーシャルワークの研究手法」（2009）、「ソーシャルワーク・リサーチの技法」（2010）などの特集が組まれている。また、ソーシャルワーク学会では、シンポジウムにて「社会福祉実践（ソーシャルワーク）における研究方法を問う」（2005）、「ソーシャルワークの研究手法論」（2006）、「エビデンスに基づくソーシャルワーク実践の科学化－実践事例の分析とその理論化」（2007）などのテーマが取り上げられている。
- (2) 代表的なものとしては、「MAXqda」、「ATLAS.ti」、「NVivo」などがあり、これらのソフトウェアの基本的な操作方法について解説している文献もある（佐藤 2008 a；佐藤 2008 b）。
- (3) テキストマイニング専用のソフトには、「Word Miner」や「SPSS Text Analytics for Surveys」といった高額なものから、「茶筌」や「KH Coder」など無料でダウンロードできるものまでいくつかの種類がある。また、社会福祉や心理学などの人文・社会科学の領域においてテキストマイニングを使うための入門書も出版されており、当初に比べると初心者にも親しみやすくなっている（藤井ら 2005；松村ら 2009）。
- (4) これらの作業は「分かち書き」とも呼ばれる。
- (5) 分析に不要な語を削除する作業は「削除辞書」の作成、同義語や類義語を置き換える作業は「置換辞書」の作成と呼ばれることもある。
- (6) CiNii では、「社会福祉」をキーワードとして検索した結果は「福祉」をキーワードとした場合の検索結果に含まれることから、ここでは「福祉」をキーワードとして検索している。
- (7) 「グラウンデッド・セオリー・アプローチ」「ソーシャルワーク」は 13 件、「グラウンデッド・セオリー・アプローチ」「福祉」は 46 件ヒットした。そのうち、福祉の領域における研究と判断できるものは、46 件であった。KJ 法については、「KJ 法」「ソーシャルワーク」は 2 件、「KJ 法」「福祉」は 54 件ヒットした。

文献

- 秋山薊二 (2012) 「第7章 エビデンス情報に基づくソーシャルワークの実践に向けて」『教育研究とエビデンス—国際的動向と日本の現状と課題』明石書店, 205-230.
- Boehm, W. W (1958) *The Nature of Social Work, Social Work*, 3-2. (=1972, 小松源助訳『社会福祉論の展望』ミネルヴァ書房.)
- Flick, U (1995) *Qualitative forschung*, (=2002, 小田博志・山本則子・春日常ほか訳『質的研究入門—「人間科学」のための方法論』春秋社.)
- 藤井美和・小林考司・李政元 (2005) 『福祉・心理・看護のテキストマイニング入門』中央法規.
- Gambrill, E. (2006) *Social Work Practice: A Critical Thinker's Guide*, Oxford University Press.
- 狭間香代子 (2000) 「ソーシャルワークの科学化についての一考察」『東大阪短期大学研究紀要』26, 65-73.
- 樋口耕一 (2004) 「テキスト型データの計量的分析—2つのアプローチの峻別と統合」『理論と方法』19 (1), 101-115.
- 樋口耕一 (2006) 「内容分析から計量テキスト分析へ—警鐘と発展を目指して」『大阪大学大学院人間科学研究科紀要』32, 1-27.
- 平塚良子 (2007) 「ソーシャルワークの科学化—価値の可視化を通して」『社会福祉実践理論研究』16, 19-24.
- 稲葉光行・抱井尚子 (2011) 「質的分析におけるグラウンデッドなテキストマイニング・アプローチの提案—がん告知の可否をめぐるフォーカスグループでの議論の分析から」『政策科学』18 (3), 255-276.
- 木下康仁 (1999) 『グラウンデッド・セオリー・アプローチ—質的実証研究の再生』弘文堂.
- 木下康仁 (2003) 『グラウンデッド・セオリー・アプローチの実践—質的研究への誘い』弘文堂.
- 京俊輔 (2012) 「保護者と施設職員の『知的障害者の性』に対する意識—テキストマイニングを用いた探索的分析」『島根大学社会福祉論集』4, 1-16.
- 正木朋也・津谷喜一郎 (2006) 「エビデンスに基づく医療 (EBM) の系譜と方向性—保健医療評価に果たすコクラン共同計画の役割と未来」『日本評価研究』6 (1), 3-20.
- 松村真宏・三浦麻子 (2009) 『人文・社会科学のためのテキストマイニング』誠信書房.
- McNeece, A. & Thyer, B. (2004) Evidence-Based practice and social work. *Journal of Evidence-Based Social Work*, 1 (1).
- 南彩子 (2005) 「高齢者イメージについて考える (2)—テキストマイニングによる大学生の抱く高齢者イメージの分析」『天理大学社会福祉学研究室紀要』7, 5-19.
- 峰島里奈 (2005) 「児童期・青年期に死別経験をした青年の悲哀過程におけるソーシャルサポートの探索的分析—テキストマイニングによる分析を中心に」『関西学院大学社会学部紀要』99, 221-231.
- 中村雄二郎 (1992) 『臨床の知とは何か』岩波新書.
- 窄山太 (2012) 「人, 環境, 関係および状況を通して考えるソーシャルワークの焦点」『社会福祉学』53 (1), 79-90.
- 佐藤郁哉 (2008 a) 『質的データ分析法』新曜社.
- 佐藤郁哉 (2008 b) 『実践質的データ分析入門』新曜社.
- 志村健一 (2007) 「島山龍郎が記したもの: 『ソーシャルワーク研究』「巻頭言」の計量テキスト分析」『聖隷クリストファー大学社会福祉学部紀要』5, 51-59.
- 田垣正晋 (2008) 『これからはじめる医療・福祉の質的研究入門』中央法規.
- 田垣正晋 (2009) 「市町村障害者基本計画のニーズ調査の自由記述回答に対する KJ 法とテキストマイニングの併用のあり方」『社会問題研究』58 (137), 71-86.
- 坪井真 (2007) 「テキスト・マイニングによる〈方面委員〉概念の分析—方面委員による援助行為の特性分析 (1)」『長野大学紀要』29 (3), 31-42.
- 和気順子 (2010) 「第6章 ソーシャルワークの演繹的研究方法」『ソーシャルワークの研究—実践の科学化と理論化を目指して』相川書房, 89-105.

A Study of Text Analysis in Social Work Research

Yasuyo Hiyori

The purpose of this paper questions a meaning of reliability and validity and considers the way of qualitative study in social work research. It is often pointed out that reliability and validity in qualitative research is hard, but the software for text analysis helps us to solve this problem. When we treat big data, we can analyze the data effectively by using the software. However, it is hard to understand the context of the data by using text mining only. So, when we analyze text data in social work research, it is necessary to analyze quantitatively and qualitatively.

Key words : Text analysis, Reliability, Validity, QDA, Text mining

