The University of Osaka
Institutional Knowledge Archive

Title	質的データ分析支援ソフトウエアの機能と背景にある考え方 : Atlas.ti7 とNVivo11 の比較から
Author(s)	樋口, 麻里
Citation	年報人間科学. 2017, 38, p. 193-210
Version Type	VoR
URL	https://doi.org/10.18910/60474
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

〈研究ノート〉

質的データ分析支援ソフトウエアの機能と背景にある考え方
――Atlas.ti7 と NVivo11 の比較から

樋口 麻里

要旨

質的データ分析支援ソフトウエア(CAQDAS: Computer-Assisted Qualitative Data Analysis Software)を、有効に利用するには、各ソフトウエアがどのような分析を行うために開発されたのかを理解する必要がある。そこで本稿では、主要なソフトウエアであるAtlas.ti7とNVivo11を取り上げ、それぞれの機能の特徴から、各ソフトウエアの背景にある分析方法に対する考え方を探った。

その結果、次のことが示唆された。Atlas.tiは、データに即した抽象度の低いコードを作成し、それらを関連づけながら概念を構築する機能を特徴としていた。これは、データの文脈に根差した分析を重視する、グラウンデッド・セオリー・アプローチの方法に適合的であった。一方、NVivoは、ミックスド・メソッド・アプローチといった、より多様な方法への対応を志向していた。これは、コーディングの機能に、概念構築とデータの分類という二つの異なる機能が備わっていることから考えられる。この機能は、量的データを含む大量のデータを分類して、データ全体の特徴を把握することに適していた。

キーワード

CAQDAS、QDAソフトウエア、Atlas.ti、NVivo、質的データ分析

1．CAQDASの機能の背景にある考え方

質的データ分析支援ソフトウエアは、CAQDAS(Computer-Assisted Qualitative Data Analysis Software)と呼ばれ[1)]、グラウンデッド・セオリー・アプローチに基づいて開発されてきた（佐藤 2008）。多量かつ多様な形態のデータを扱うためには、データの体系的な整理と効率的な検索が欠かせない。電子化されたデータを読み込み、様々な検索機能をもつCAQDASを利用することで、「質的データのほぼ完全な脱文脈化、そしてまたきわめて効率的なデータベース化」（佐藤 2008:34）が可能になる。質的データの分析から理論を形成するには、フィールドでの気づきやデータの解釈、分析結果のまとめ方などについて、アイディアのメモを分析プロセス全体を通して作成することが重要である(Strauss and Corbin 2008=2012)。CAQDASでは、データだけではなく、こうしたメモについても、データと組み合わせて整理し、検索することができる。また、分析のための図表を作成し、それをデータやメモと合わせてCAQDASのファイル内で保存することができる(Corbin and Strauss 2015)。これは、分析者自身だけではなく、分析者以外の人々がデータの分析プロセスを確認し、その妥当性について判断する助けにもなる(Strauss and Corbin 2008=2012)。さらに、こうした機能を組み合わせることで、仮説のエビデンスあるいは反証例を示すデ

ータを容易に見つけ出すことができ(Kelle et al. 1995)、データ分析の試行錯誤が容易になる(Denzin and Lincoln 1998; Strauss and Corbin 2008=2012:429)。

CAQDASを効果的に利用するには、それぞれのソフトウエアが、どのような分析を行うために開発されたのかということと、その分析方法を実現するための主な機能の両方を理解する必要がある。質的データの分析では、「コーディング」のように方法の名前が同じでも、分析者の考え方によって、それを具体的にどのような手順で行うのかが異なる。CAQDASの背景にある、こうした分析方法に対する考え方を理解していなければ、ソフトウエアの各機能を使って意味のある結果を得ることは難しいだろう。その場合、どのような機能があるのかや、分析者が頭に描いている分析方法に対応する機能があるのかを、手探りで確かめていくことになる。また、複数の機能をどのように組み合わせて、最終的な結果に至るのかという道筋が見えにくくなる。それに対して、ソフトウエアの分析方法に対する考え方と、機能との関連を知ることで、ソフトウエアの全体像を理解しやすくなる[2)]。また、自分の目的に合わせて、ソフトウエアの機能を選択できるようになる。

そこで、本稿ではCAQDASのうち、世界で広く利用されているAtlas.ti 7とNVivo11を取り上げ、それぞれの主要な機能を比較する。そして機能の特徴をもとに、その背景にある分析方法に対する考え方を推論する。具体的には、それぞれのソフトウエアの考え方が、グラウンデッド・セオリー・アプローチにおけるどのような立場に親和的なのか[3)]、また、グラウンデッド・セオリー・アプローチの枠を超えて、どのような使い方ができるのかを探る。

2. CAQDASの中心的機能：コーディング

グラウンデッド・セオリー・アプローチは、量的調査を専門とするBarney G.Glaserと、質的調査の専門家であるAnselm L. Straussが、共同で開発した質的データの分析方法で、社会調査を通じて体系的に獲得されたデータから理論を発見することを目的としている(Glaser and Strauss 1967=1996)。この方法の特徴として、データ収集とデータ分析の同時進行、データのコーディング・継続的比較・理論的メモによる分析、理論的サンプリングによるデータ収集がある（佐藤 2008:106）。

グラウンデッド・セオリー・アプローチは、その後に登場した質的データ分析の方法に影響を与えている（佐藤 2008:106）。とりわけ、分析プロセスの中で中心的な作業であるコーディングは、質的データを分析するための手順として広く採用されている。コーディングでは、データを内容毎に区切り、そのデータの文脈や言葉の意味を表すラベルを貼っていく作業を行う（戈木・クレイグヒル 2013）。このラベルをコードと呼ぶ[4)]。これらのコードをまとめていくことで、最終的には、調査対象に関する現象を説明する理論をつくり上げていく。したがって、データに即した(グラウンデッドな)理論を導き出すには、データの文脈を反映したコードが作成できるかが重要になる。

現在流通している代表的なCAQDASとしては、NVivo、Atlas.ti、MAXQDAがあり、1980年代にこれらのソフトウエアの初期バージョンが開発された(Paulus, Lester and Dempster 2014)。CAQDASは、グラウンデッド・セオリー・アプローチの方法を基調としてデザインされているが、これはCAQDASの登場時期と、グラウンデッド・セオリー・アプローチが、質的調査の方法論の領域で普及した時期とが重なっ

たためである(Di Gregorio and Davidson 2008)。コーディングという手順が、質的データの分析方法で広く用いられるようになったこともあり、コーディング機能は、CAQDASの中心的な分析機能となっている。そのため、NVivoおよびAtlas.ti、そしてMAXQDAといった、主要なCAQDASのいずれにもコーディング機能が備わっており、その操作も似通っている。しかし、各ソフトウエア開発者の方法に対する考えによって、コーディング機能の操作性や用語法、画面表示が異なる。こうした傾向は、コーディング以降のプロセスにおいて使用する機能でも同様である。したがって、コーディング機能の特徴を把握することで、他の発展的な機能についても理解しやすくなる。

本節では、NvivoとAtlas.tiのコーディング機能から、それぞれのソフトウエアがコードをどのように捉えているかを検討する。

2.1 Atlas.ti――データの文脈を重視したコード

Atlas.tiは、コーディングをデータに即して行うことと、コーディングを文脈単位で行うことを重視しているように見受けられる。

まず、コーディングの作業画面について検討する。データ中の一か所に複数のコードを与えることは、しばしば起こる。1文や1段落といった、文章の構成単位でコーディングを行うのではなく、データの文脈単位でコーディングする場合に、こうしたことが起こりやすい。これは、1文の中に複数の文脈が含まれることがあるからだ。コーディングの作業画面で、一か所に多数のコードが集中すると、NVivoやMAXQDAではコードが重なってしまい、各コード名の最初の数文字しか表示されなかったり、コーディングしたデータの近辺に、コードが表示されなくなる。そのため、どのコードがどの部分に与えられているのかが、確認しにくい。それに対してAtlas.tiでは、コードの表示領域の幅を任意に拡げて、コード名をすべて表示することができるため、複数のコードが同一の文章や段落に集中する場合でも、コードとデータの対応が把握しやすい。

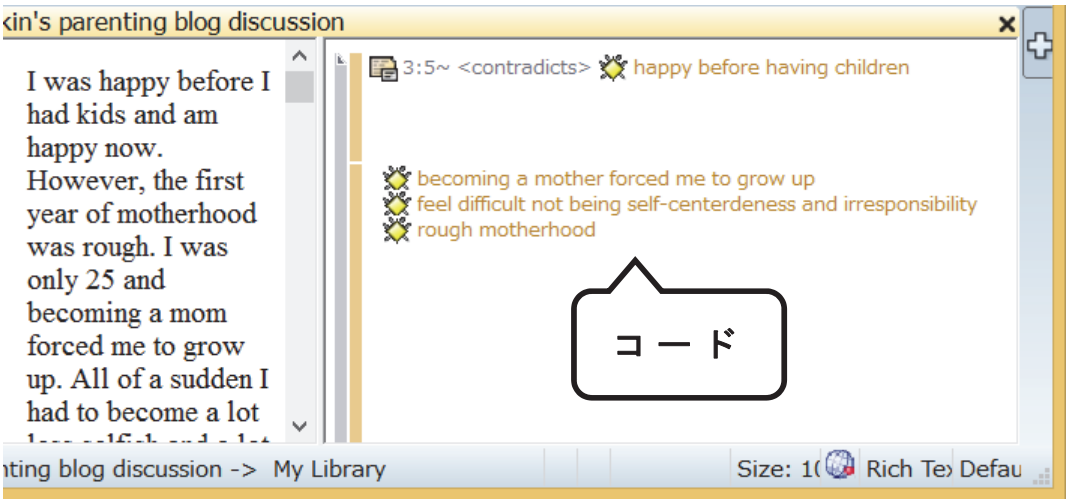

図1 Atlas.tiによるコーディング作業画面

Atlas.tiのマニュアルでは、Atlas.tiのコーディング機能は、グラウンデッド・セオリー・アプローチにおけるコーディング法を全て含んでいるわけではないこと、また、その使い方はグラウンデッド・セオリー・アプローチに限定していないことが言及されている[5)]。例えばAtlas.tiは、NVivoを含む他のCAQDASと同様に、コードをデータのどこからどこまでに与えるのかという範囲を任意に指定することができる。したがって、複数の文脈を含むデータの広い範囲に、コードを与えることもできる。その際には、言及されているトピックを表すような、抽象度の高いコードを割り当てて、データをまとまった単位で取り出すといったこともできる。こうしたコーディングの仕方は、グラウンデッド・セオリー・アプローチというよりも、ライフ・ストーリー研究や、ナラティヴ・アプローチなどに利用できるかもしれない。

しかしAtlas.tiでは、コーディング単位を細かくしても、コードとデータのコーディング部分の対応関係を確認しやすい。この特徴からは、データに根差した分析を重視する、グラウンデッド・セオリー・アプローチの考え方に忠実であろうとする、Atlas.tiの開発姿勢がうかがえる。

また、Atlas.tiには、コード・マネージャー(Code Manager)と呼ばれる機能があり、作成したコードを一覧表示し、管理することができる。コード・マネージャーを開くと、それぞれの列に名前がついている中で、「grounded」という列が目に入る。これは、各コードがいくつのデータに与えられているかを表す数値である。例えば、あるコードをデータ中の3か所に付与した場合、このコードがデータに根差している(groundedな)度合いは「3」ということになる。こうした用語法は、Atlas.tiに特有のものである[6)]。こうした名前の付け方からも、Atlas.tiは、グラウンデッド・セオリー・アプローチを重視して開発されたと見て取れるだろう。

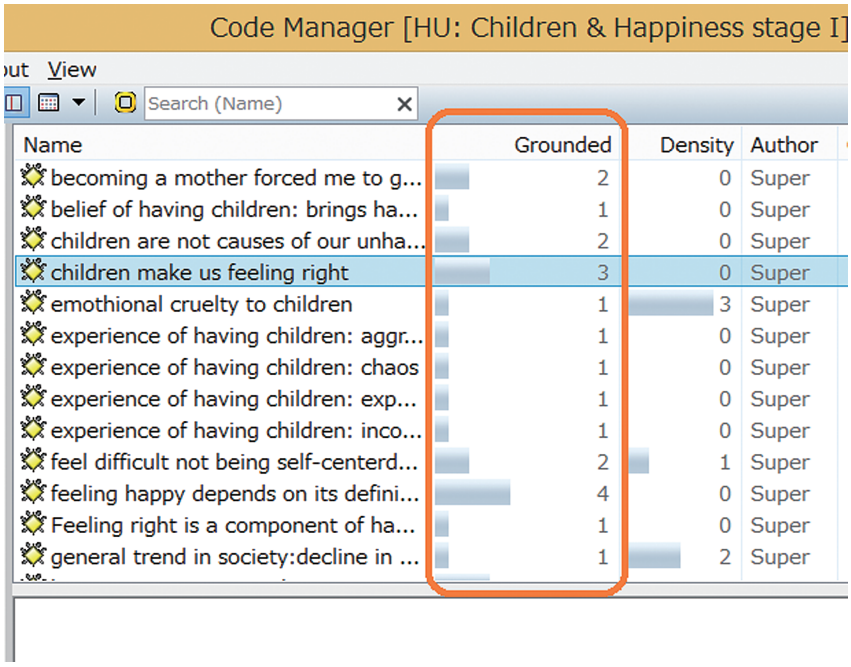

図2 Atlas.tiのコード・マネージャー

2.2 NVivo――概念と分類の2種類の意味を持つノード

NVivoは、グラウンデッド・セオリー・アプローチにも対応しているが、データ全体を俯瞰する機能の開発にも力をいれている。これは、「ノード」の機能に反映されている。

NVivoが、Atlas.tiやMAXQDAと大きく異なる点は、「コード(code)」という言葉の代わりに「ノード(node)」という用語を使用していることである[7)]。ノードは、グラウンデッド・セオリー・アプローチのオープン・コーディング機能だけではなく、データ整理機能も追加したものと考えると理解しやすい[8)]。オープン・コーディングに利用する「概念ノード(theme node)」と、データ整理に用いる「ケースノード(case node)」とがある[9)]。

先に述べたように、データの任意の範囲に、コードを割り当てる機能は、NVivoにも基本的機能として組み込まれており、概念ノードが用いられる。ただし、NVivoの場合は、抽象度の低いコードを手作業でデータに細かく割り当てていくという、データの細かな文脈を重視するコーディングを行うのは容易ではない。これは、データは横書きの縦スクロールで表示されるのに対して、コードは縦書きの横スクロール表示となっているため、データに振られたコード数が多くなった時に、コーディングの作業画面で、どのコードがデータのどの部分に対応しているのかを、視覚的に確認しにくくなるためである。その背景として、「特定の単語があれば、このコードを与える」といった、ルールに基づく自動コーディング機能を使うことが想定されているのかもしれない。抽象度の低いコードについては、目で見ながら手作業をする必要はないという考え方がうかがえる。

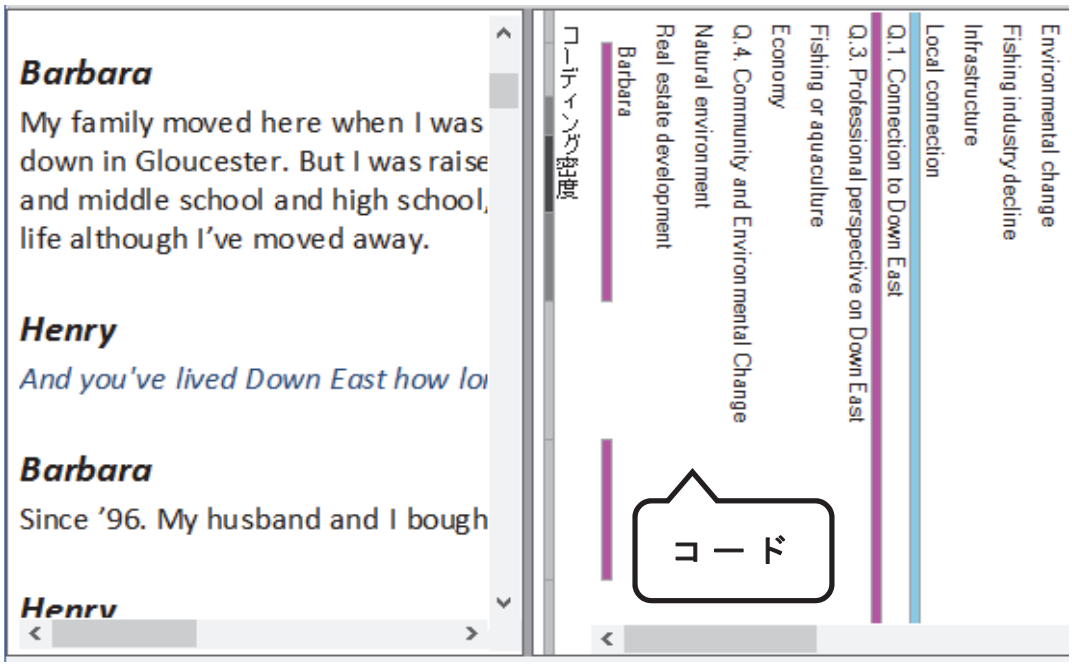

図3 NVivoによるコーディング作業画面

NVivoでは、コードをツリー形式でまとめることができる。類似するいくつかのコードの上位に、それらのコードの特徴を表す抽象度の高いコードを置く。この上位のコードには、下位のコードが付与された

データを関連づけることができる。例えば、「湯呑」「茶碗」「汁椀」のコードを、それぞれデータの1カ所に与えたとする。そして、これらをまとめる上位のコード「食器」を作成する。すると「食器」コードは、「湯呑」「茶碗」「汁椀」コードが与えられた3カ所と自動的に紐づけされる。

この機能は、コードの高次化によって概念を作成する方法に適していると言える。また、コーディング作業画面の特徴と合わせて考えると、NVivoではデータの細かな文脈に沿って、帰納的にコーディングを行うというよりは、自動コードやあらかじめ理論的枠組みを設定し、その枠組みに従ってコードをデータに割り当てる方法に特に適しているようである。

加えて、NVivoには、データファイル自体をコーディングできるという特徴がある。この機能は、データファイルの分類に使う。例えば、インタビューのトランスクリプションをWordファイルで作成したとする。Nvivoでは、このWordファイルに対して、コードを付与することができる。この場合、ソフトウエアの操作上は、概念ノードとケースノードのどちらでも、データファイルのコーディングができる。但し、ケースノードのみに、属性を指定する機能がついていることから、データファイルのコーディングは、ケースノードを用いることが想定されていると考えられる。

ケースノードの属性機能は、量的調査データに対応させて考えると理解しやすい。ケースノードは対象者ごとの回答票のID番号で、属性機能は対象者の年齢や居住地といった属性変数である。量的データでは、ID番号ごとに各変数の値をエクセルファイル等に入力して整理する。そうすることで、特定の変数の値を指定し、それがどの対象者に該当するのか、あるいはその値をとる対象者が何人いるのかということを集計できる。NVivoでは、そうした集計を、属性機能を使って行うことができる。トランスクリプションのファイルに「Aさん」というケースノードと、「年齢＝40代」という属性値を割り当て、他のファイルにも、同様のコーディングをする。そして、4.1節で取り上げる検索機能と組み合わせることで、インタビューの人数が増えたときに、40代の対象者は誰かといったことや、どの年齢層が多いのかを調べることができる。

このように、NVivoのコーディング機能は、使い方の異なる概念ノードとケースノードとを含んでおり、概念ノードは上位のコードを作成して概念を作る方法に適していた。このことから、NVivoでは、多様なレベルでデータを分類する方法として、コーディングを位置づけていると言えるだろう。

3. Atlas.tiのファミリーとネットワーク・ビュー・マネージャー

本節では、Atlas.tiの応用的機能として、ファミリー(family)とネットワーク・ビュー・マネージャー(Network View Manager)を取り上げる。そして、それらの機能が、データの文脈に即したコーディングを重視する考え方と、どのように関連しているのかを検討する。

3.1 ファミリーを用いたグルーピング

ファミリーは、データ(primary document)、コード、メモをグルーピングする機能である[10)]。本節では、Atlas.tiに特徴的な、データをまとめるファミリーと、コードをまとめるコード・ファミリーについて取り上げる。

まず、コード・ファミリー(code family)は、類似するコードをまとめる機能である。オープン・コーディングを進めていくと、コードの一覧機能で表示されるコードが多くなり、どのようなコードがあるのかが分かりにくくなる。そこで、コードの特徴を表す名前をつけたファミリーで、いくつかのコードをグルーピングする。Atlas.tiでは、全ての操作画面で、注目したいファミリーに含まれるコードだけを表示することができる。データの文脈ごとに

具体的なコードを付与していくと、コードの種類が数百を超えることもある。こうした際、コード・ファミリーを使って、作業画面に表示されるコードを絞り込めば、オープン・コーディングの作業が進んだ段階でも、新規のコードを作成するかどうかを、既存のコードを見ながら吟味することができる。分析者の解釈に基づく手作業のコーディング方法では、解釈の基準やコード名のつけ方が時間の経過とともに変化しうる。コード・ファミリーを使って、既存のコードが容易に参照できれば、類似する現象に異なる名前のコードをつけたり、名前は異なるが概念的意味は同じコードを作ってしまったりという、コーディングのゆれを抑えることができるだろう。

コード・ファミリーは、query tool の機能と併せて使うことで、特定のコードを検索することができる。この機能は、オープン・コーディングが収束に向かい、概念を構築していく段階で有用になるだろう。query tool では、複数のコード・ファミリーに共通して含まれているコードや、特定のコード・ファミリーのいずれかに含まれる（あるいはいずれにも含まれない）コードなど、ファミリーを使って、様々な条件を指定しながらコードを検索することができる。これは、概念の中心的テーマを考えるヒントとなるだろう。例えば、「having children: unhappy」と「view for parenthood」の2つのファミリーに共通して含まれるコードが、いくつかあったとする。この場合、これら共通のコードは、having children と parenthood に関する概念を関連づける役割を果たしているか、あるいは having children と parenthood は、一つの概念としてまとめられるかもしれない、といったことが予想できる。

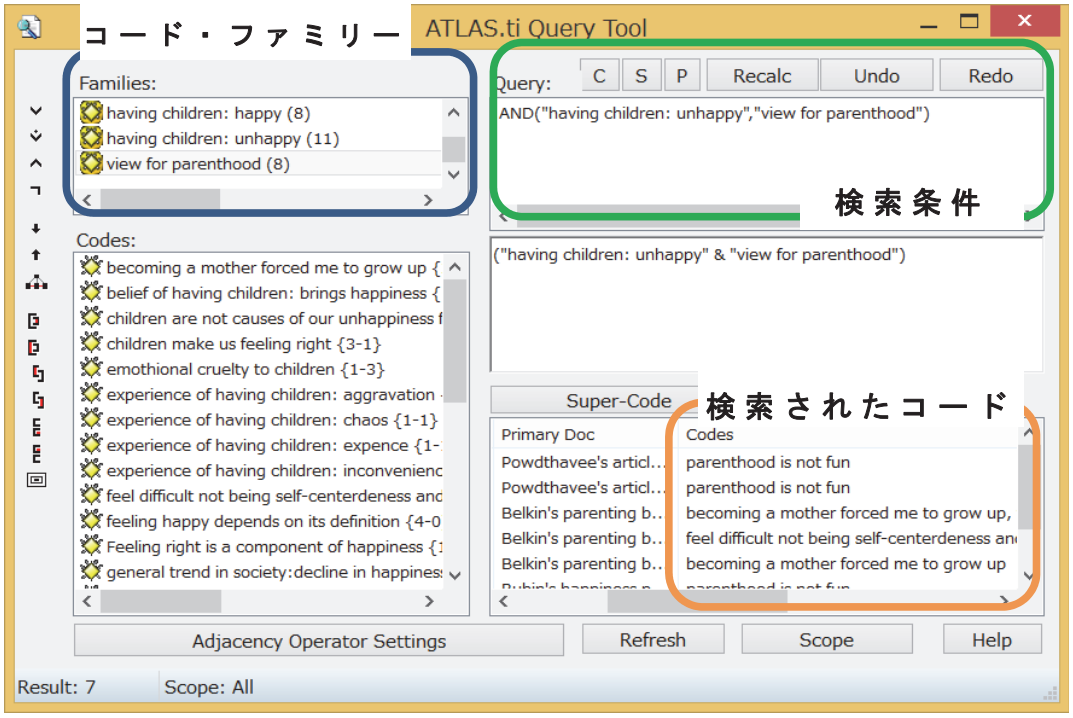

図4　コード・ファミリーと query tool によるコードの検索

グラウンデッド・セオリー・アプローチでは、全体像の分からない概念を明らかにしていく（戈木・クレイグヒル 2013）。その際に、手がかりになるのがコードである。コード・ファミリーの機能は、コー

ディングのゆれを抑制すると共に、概念の特性を表すコードの探索を容易にする。それによって、コードから概念を構築する分析プロセスが促される。

次に、データをまとめるファミリーの特徴は、データの属性として使用できることである。例えば、インタビュー調査では、トランスクリプションを対象者ごとのWordファイルにして作成することが多い。ファミリーの機能を使うと、対象者の性別や学歴、居住地といった属性ごとに、Wordファイルをグルーピングできる。この機能とコードを組み合わせて使うことで、コードの分布を観察できる。表1は、ブログに投稿されたコメント(blog)と研究報告に関する記事(research report and comments)を使った例である[11)]。ブログと記事のデータファイルは、いずれも複数ある。これらのデータを、それぞれブログと記事の種類別にファミリーとしてまとめ、3種類のコード(having children: happy、having children: unhappy、view for parenthood)が[12)]、どちらの種類のデータに多いかを検討した。表1から、ブログのデータには「having children: happy」のコードが多く、記事のデータには、「view for parenthood」が多いという傾向が分かる。

表1 プライマリードキュメント・ファミリーを用いたデータの種類別コード数

		Code Families			TOTALS
		having children: happy	having children: unhappy	view for parenthood	
Primary Document Families	*blogs	12	9	5	26
	*research report and comments	4	10	10	24
	TOTALS	16	19	15	50

こうしたコードの分布を属性別に確認する機能は、他のCAQDASにも備わっており、Atlas.tiに限ったものではない[13)]。Atlas.tiでは、データファイルを単位としてグルーピングして、属性を設定する。コードの分布をファイル群ごとで比較することで、ファイルに共通する特性を探索することができるだろう。

3.2 コード間の関連に基づいた概念構築を支援するネットワーク図

概念の構築プロセスで使われるAtlas.tiの中心的な機能には、コード、メモ、データ間の関連を描画するネットワーク・ビュー・マネージャー(Network View Manager: 以下、ネットワーク図とする)がある。この機能では、コードとコードがどのような関係にあるかを、矢印と関係性を表す言葉で関連づけることができる。

概念化の方法には、大きく分けて2種類ある。それは、修正版グラウンデッド・セオリー・アプローチ（木下 2003）が指摘する、コード間の関連から概念を作り上げる方法と、一次コード、二次コードと、複数のコードをまとめて階層化することで抽象度を上げる方法である。Atlas.tiは、どちらの方法にも利用できるが、このうち前者の方法により適しているようである[14)]。

階層化の方法では、類似するコードをグルーピングし、さらにそれらのグループをまとめるという作業が繰

り返される。Atlas.ti の場合、コードのグルーピングでは、コード・ファミリーを使うことになるが、ネットワーク図では、コード・ファミリー間に任意の線(リンク)を引いて関連づけることができない。これは、グルーピングによって概念を構築することで、概念がデータから乖離してしまうことを防ぐためと考えられる。

Atlas.ti のマニュアルでは、Glaser と Strauss(Glaser and Strauss 1967=1996)が、「is property of」や「brings」といった関係性を示すリンクを使用して、概念を関連づけたことに言及しており[15]、ネットワーク図が、グラウンデッド・セオリー・アプローチにおける概念の構築作業を想定して設計されていることが分かる。ネットワーク図で、関係性を指定したリンクづけができるのは、データに直接根差しているアイテムに限定されており、それはコード間および任意に切り出したデータの一部(quote)間のみである。したがって、コード・ファミリーとそこに含まれるコードは、自動的に破線の双方向矢印で結ばれるのみで、この矢印に関係性を記入することはできない。

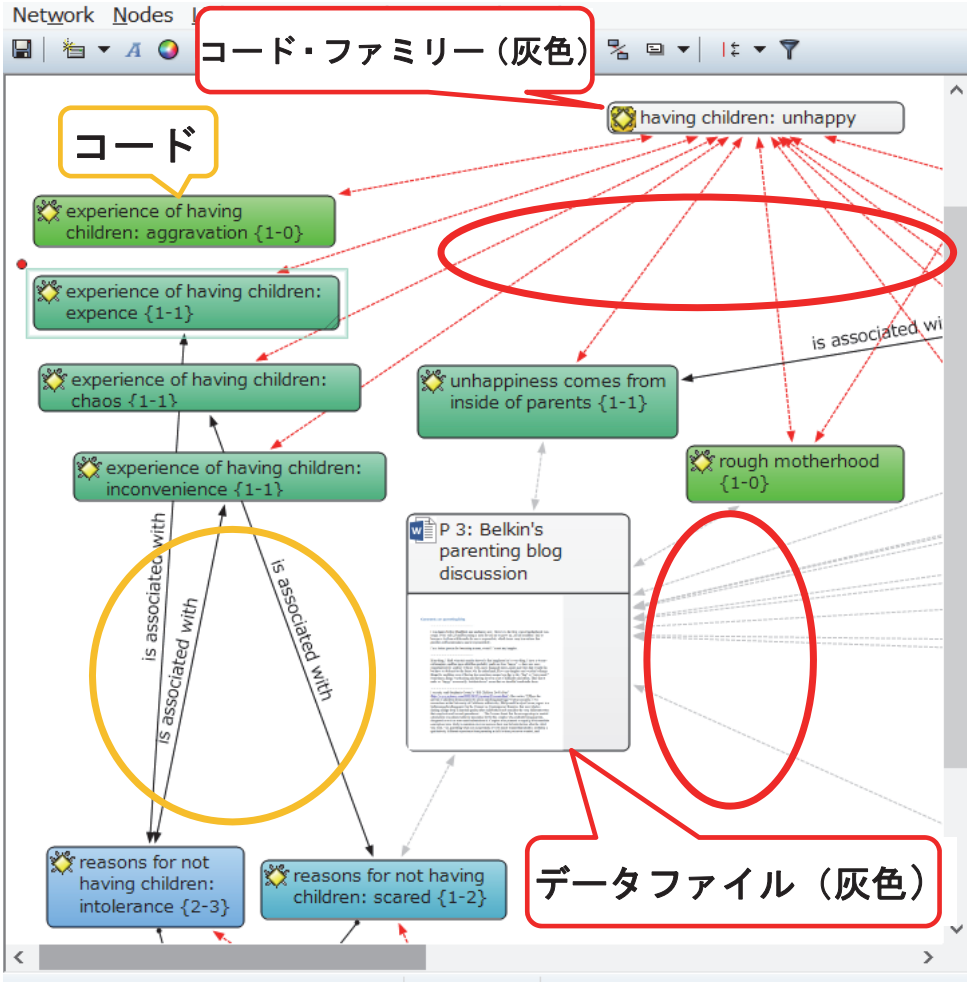

図5　ネットワーク図を用いた概念の構築（一部のみ抜粋）

（赤色楕円：コードとコード・ファミリー間の矢印、データファイルとコード間の矢印。

オレンジ色楕円：コード間の関係性を指定した矢印。）

こうしたネットワーク図の機能は、一見すると不便かもしれない。しかし、関係性を指定したリンクづけをコード間に限定することで、データに密着しているのは、コードであって、ファミリー(グルーピングされたコード群)ではないことが視覚的に示される。それによって、概念の構築がコード間の関連の検討を通して行われるようになり、概念がデータに根差していることが担保される。こうしたコード・ファミリーの機能的制限は、階層化してつくられた高次のコードは、データから乖離する可能性があることを指摘している、修正版グラウンデッド・セオリー・アプローチの立場（木下 2003）に親和的である[16)]。

現在は、他のCAQDASにもネットワーク図と類似する描画機能が付加されているため、Atlas.tiでなければコード間の関連を示す図が描けないということはない。NVivoではネットワーク図の画面に、任意の図形や文字を書き込めるといった、Atlas.tiにはない機能もある。この書き込み機能を使えば、1つのネットワーク図内で、コードをリンクづけしながら概念を描くと同時に、概念単位で複数のコードを枠で囲い、概念間の関連づけも描くことができる。

しかし、データに根差したコーディングのための機能(2.1節)と、概念化とグルーピングを区別するためのファミリー機能、コード間の関係性を描くネットワーク図の機能を総合して考えると、Atlas.tiは、ネットワーク図の機能に敢えて制限を加えているのだろう。この制限を設けることで、データに直接根差したコード間の関連に注目しながら、概念を構築することを促しているものと考えられる。

4. NVivoの行列コーディングと自動コード

NVivoでは、概念の構築だけではなく、データを分類する方法としても、コーディングが捉えられていた。本節では、NVivoの応用的機能である、行列コーディングと呼ばれる機能と自動コードを取り上げて、それらの機能がこうしたコーディングへの考え方と、どのように関連しているのかを検討する。

4.1 行列コーディングによる多角的なデータの検索とミックスド・メソッド・アプローチへの対応

行列コーディングは、データ検索機能の一つである。クロス表を作成し、特定のコードが、どの対象者、あるいは属性に振られているのかを確認できる。この機能によって、データを様々な角度から検討し、概念を精緻化できるだろう。

例えば、図6は、「Natural Environment」と「Economy」の2つのコードの数を、対象者名のコード別に示した結果である[17)]。

「Natural Environment」コードが最も多いのは、Barbaraである。また、「Economy」コードは、BarbaraとCharlesで同頻度である。これらの結果から、「Economy」コードの数は、BarbaraとCharlesで似通っているが、Barbaraの方が「Natural Environment」コードが多いのはなぜだろうか、という新たな疑問が得られるかもしれない。焦点化した疑問があると、注目するデータを絞ることができる。この例の場合、Barbaraのインタビューデータに戻り、「Natural Environment」のノードが振られたデータのうち、Barbara特有の表現や経験は何かといったことを再検討することで、それまで気がついていなかった「Natural Environment」の概念の構成要素を発見できるかもしれない。行列コーディングを積み重ねることで、注目するデータを絞り込み、概念の構成要素を豊富にし、概念が精緻化できる。概念についての「分厚い記

述」(Geertz 1973=1987)をするには、これらの要素をデータから抽出できているかということと、不足しているデータが何かを分析者が理解していることが重要である。

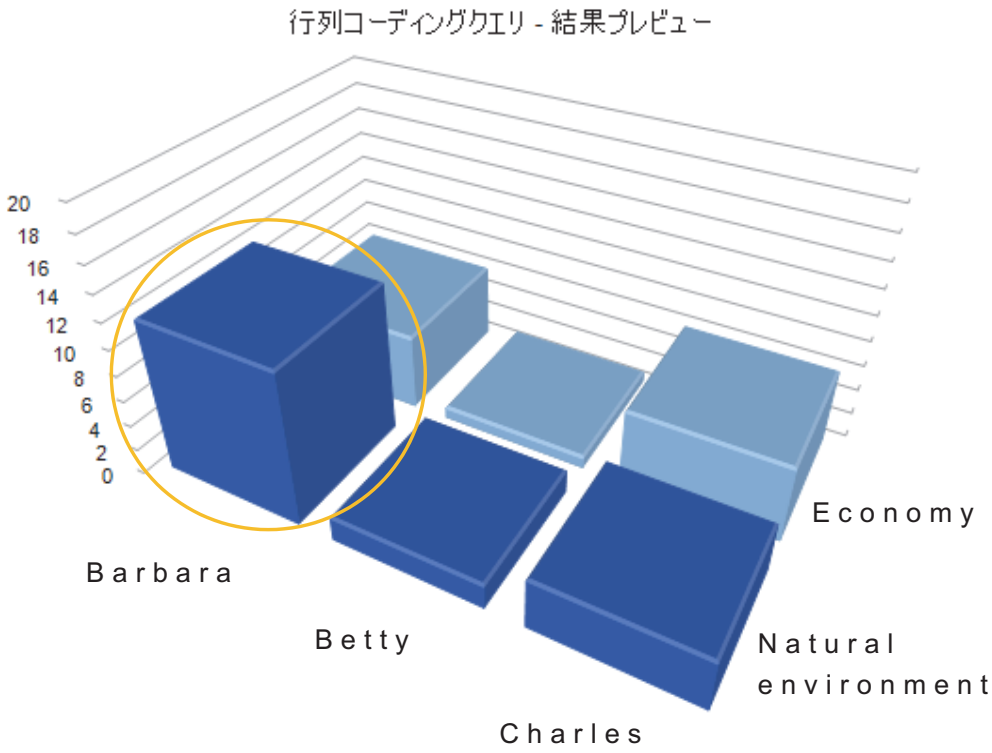

図6 行列コーディングを用いた対象者別のノード分布

図6の例について、さらに年代との行列コーディングを行ったとする。その結果、Barbaraは30代で、30代の対象者は「Natural Environment」のノードが最も少ないことが分かったとする(図7)。

この場合、Barbaraは他の対象者に比べて、単に「Natural Environment」のコードが多いだけではなく、同年代の対象者とは異なる傾向をもつことが推測される。こうした一連の結果を総合することで、Barbaraには、同年代の対象者に見られない何らかの特徴(経験や属性)があり、それが「Natural Environment」への言及の多さに影響しているのではないか、ということが予想できる。こうした特徴を明らかにできれば、「Natural Environment」の内容と、それが生起する条件の仮説を導きだせるかもしれない。

コードの分布から、データの特徴的な部分を探し出すには、ある程度多くのデータに、同一のコードが付与されている必要がある。コードをツリー形式でまとめて、上位のコードを作成するようにNVivoが設計されているのも(2.2節)、こうした集計・検索方法を想定しているためだろう。NVivoでは、複数のデータに付与できるような、抽象度の高いコードを振ることで、データを様々な方向から検討し、概念の構成要素を探ることが目指されているようである。

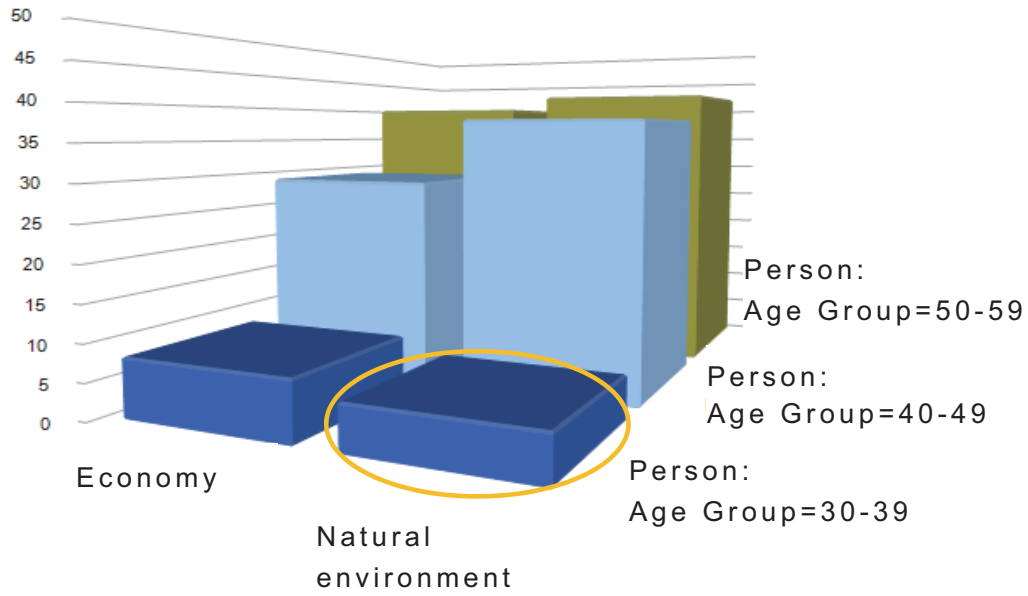

図7 年齢階級の属性値別のノード分布

上記に加えて、行列コーディングは、同一対象者(集団)へのミックスド・メソッド・アプローチを実現する機能としても利用できる。NVivoは、量的データファイルの取り込み機能も備えており、エクセル形式で保存された質問紙調査の質問項目や対象者のIDを、ノードに自動変換して読み込むことができる。例えば、あるプログラムの参加者に、質問紙調査でプログラムの満足度を尋ね、さらに、インタビューでプログラムの参加動機を聞いたとする。ケースノードを質問紙調査の対象者IDに割り当てることで、質問紙(量的)・インタビュー(質的)の2種類の調査データを、同一対象者のデータとして紐づけすることができる[18)]。こうしたデータに行列コーディングの機能を使えば、プログラムへの満足度が高かった人（低かった人）が、インタビューでどのような参加動機を語っていたか、といったことを示せる。

このように、ノードにデータの分類や量的データとの紐づけの役割を持たせて、検索に活用するという機能は、グラウンデッド・セオリー・アプローチの枠組みを超えるものである。NVivoは、ノードの使い方を拡大していくことで、より多様な方法への対応を目指していると言えるだろう。

4.2 自動コードを用いたデータの俯瞰

ノードと検索の機能に加えて、NVivoを開発しているQSR International社が力を入れているのは、言葉の自動抽出機能である。この機能は、現在他のCAQDASにも備わっているが、NVivoは、日本語データにもいち早く対応した。NVivo11では、計量的な形でデータを俯瞰できるように「自動コード(autocode)」の機能が備わっている。自動コードには、いくつか種類がある。①段落に自動的に数値のノードを割り当てるコーディングや、②分析者があらかじめ別のデータに付与したノードを、新たなデータに自動的に

割り当てるコーディング、③NVivoが自動的に言葉を取り出してノードを作成するコーディングである。NVivoのマニュアルでは、大量のデータを素早く整理できることと、データへの新たな洞察を得る手段として、自動コードの機能が紹介されている[19]。本節では、これらのうち手作業のコーディングを介さないことで、データから新たな発見を得ることを目指す、③の機能を取り上げたい。

③の機能は、テキストデータに含まれる熟語、「名詞＋助詞」の組み合わせのうち、一定の出現回数が見られるものをノードとして抽出する[20]。自動コード機能を使うには、あらかじめ頻出語クエリを走らせて、頻出語リストを確認することが重要である。頻出語クエリでは、出現頻度の高い語を一覧として見るだけではく、単語クラウド(word cloud)(図8)やクラスター分析の結果(図9)も表示することができる。それによって、データ全体のおおまかなトピックや、中心的に使用されている語がどのような語と一緒に出現しているのかについて、視覚的に把握できる。

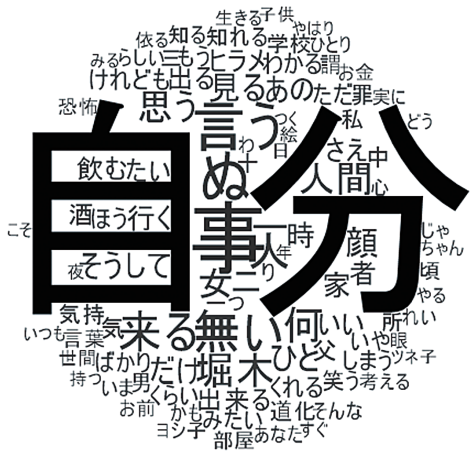

図8『人間失格』の単語クラウド質的調査

では、インタビューや資料が大量に収集できたものの、どこから分析に手を付けてよいか戸惑うことが少なくない。そうした場合に、単語クラウドやクラスター分析を使うと、自分の研究関心に近いデータに当たりをつけることができるかもしれない。データを単語レベルで自動的に視覚化することで、調査者のデータに対する先入観がある程度抑制されることも考えられる。調査者は自身がもつ仮説や研究関心を通して、データを分析する。もし、仮説に対する信念が強いと、事前に想定していなかったデータ間の関連を見落としてしまうかもしれない。また、フィールドに長期間関わる場合は、フィールド内の規範やものの見方を、調査者が無意識のうちに内面化することもある。質的調査では、データの解釈に対する調査者の主観を否定しないが、調査者があらかじめ持っている仮説への固執やフィールドへのコミットメントが、データ解釈のバリエーションを狭めてしまうことは望ましくない21)。単語クラウドやクラスター分析を使うことで、調査者はデータに潜在する、予想外のトピックに気がつくことができるかもしれない。

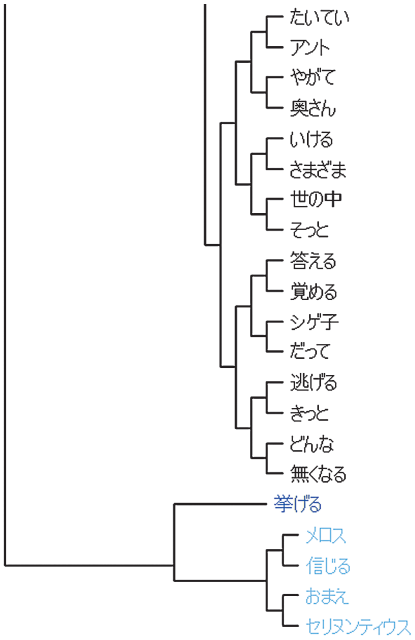

図9 複数の小説データのクラスター分析

NVivoの特徴は、語を自動的に取り出すだけではなく、抽出した語のうち、頻出していた語をそのままコードとして使用できることである。4.1節で見た行列コーディングを使うと、自動コード機能によってコード化された頻度の高い語が、データ全体のどういった箇所で使われているのかを把握することができる。

このように、語の自動抽出とコード化の機能を使うことで、調査者は、データから距離を取って、データをおおよそ俯瞰することができる。ただし、この機能を使って詳細な分析を行い、結果として得られた図を研究発表に掲載することには、ためらいも感じられる。というのも、統計的な分析であれば、内容分析(Content Analysis)のように、体系だった方法が別途に存在するからである。また、そのために開発されたソフトウエアとして、WordStat (Pollach 2011)やKHCoder (樋口 2014)がある。したがって、研究目的への主たるアプローチが、テキストデータの統計的分析である場合は、専用ソフトウエアの利用を検討すべきである。

以上で示したように、グラウンデッド・セオリー・アプローチの分析手法に重心を置いたAtlas.tiに比べると、NVivoはより多様な方法への対応を重視していると言える。

5. 方法論的立場に沿ったソフトウエアの選択

本稿では、Atlas.tiとNVivoという2つのCAQDASについて、その主要な機能を取り上げ、その背景にある分析の考え方について検討した。その結果、Atlas.tiはデータの文脈をより重視しており、その考え方はデータに根差したコードの作成と、概念の構築を支援する機能に反映されていた。こうした特徴からAtlas.tiは、コード間の関連から概念を構築し、特定の対象に適用できる理論の構築を目的とする、修正版グラウンデッド・セオリー・アプローチに適合的であることが示唆された。一方NVivoには、データを多様な角度から俯瞰して、データの全体的特徴を把握することをより重視していた。その考え方は、デー

タの分類を目的とするコーディング機能に表れており、コードを階層化して概念を作成する方法に適合的であった。また、コーディング機能を概念の構築手段としてのみ捉えるのではなく、量的データとの関連づけや自動コードにも使用することで、ミックスド・メソッド・アプローチやテキストデータの計量的な分析など、より多様な方法への対応が目指されていた。

従来、CAQDASは特定の方法に特化したソフトウエアではなく、多様な方法に活用できると指摘されてきた(Pauluset al.,2014; Freise 2012)。そして、分析者自身が各自の方法論的立場に合わせて、CAQDASの機能を使うことが推奨されてきた(Di Gregorio&Davidson 2008：49)。しかし、それぞれの機能の特徴が、どのような分析方法に対応しやすいのか、あるいは対応しにくいのかについては、検討されてこなかった。そこで本稿では、Atlas.tiとNVivoの機能を、グラウンデッド・セオリー・アプローチの分析方法に照らし合わせつつ比較した。これによって、2つのソフトウエアの背景にある分析の考え方と、分析方法への向き不向きとを、一定程度示すことができた。

もちろん、質的データの分析方法には、グラウンデッド・セオリー・アプローチ以外にも多様な立場があり、CAQDASもより多くの顧客を取り込むことを試みている。したがって、CAQDAS間の機能的特徴の相違は、各ソフトウエアがより多くの機能を備えていくことで、今後益々縮小すると予想される(Patton 2002)。しかし、たとえそうだとしても、本稿で明らかにしたようなそれぞれのソフトウエアがもつ、方法に対する考え方の中核は維持されるであろう。

それぞれのソフトウエアには、本稿で取り上げた以外の機能も備わっている。操作方法も含めて、詳細な機能の内容については各マニュアルを参照することが確実だが、質的データを分析するに当たって、これら全ての機能を使う必要はない。というのも、各自の方法論的立場に適合的な、ソフトウエアと機能を選択できることが重要なのであって、ソフトウエアに備わっている機能を全てマスターすることが、分析による発見を導くわけではないからである。

自分が採用したい方法と、ソフトウエアの方法に対する考え方との対応を検討しながらCAQDASを利用すれば、現在使用している機能によって、分析プロセスのどのステージにいるのかを自覚できる。また、次の段階に進むには、どの機能を使うのかも見えてくるだろう。各自の方法論的立場に照らし合わせて、ソフトウエアを選択することが容易になれば、CAQDASはデータの収集や、整理、分析アイディアの獲得を支援するツールとして、質的調査の一層強力な支えとなるだろう。

［付記］

本研究は、JSPS科研費(課題番号「15H06354」)の助成を受けたものです。本研究は、大阪大学研究支援員制度による成果の一部です。研究支援員の寺口季氏(大阪大学大学院人間科学研究科博士前期課程)には、文献資料の収集と文章校正を手伝っていただきました。記して感謝いたします。

参考文献

[1] Corbin, Juliet M. and Anselm L. Strauss, 2015, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory,* 4th edition, Los Angeles: SAGE.

[2] Denzin, Norman K. and Yvonna S. Lincoln, 1998, *The Landscape of Qualitative Research: Theories and Issues,*

Thousand Oaks, Calif.: Sage Publications.

[3] Friese, Susanne, 2012, *Qualitative Data Analysis with ATLAS.ti.,* Los Angeles: Sage Publications.

[4] Geertz, Clifford, 1973, *The Interpretation of Vultures,* New York: Basic Books.（＝ 1987，吉田禎吾訳，『文化の解釈学』岩波書店.）

[5] Glaser, Barney G. and Anselm L. Strauss, 1967, *The discovery of grounded theory: strategies for qualitative research,* Chicago: Aldine Publishing.（=1996，大出春江・水野節夫・後藤隆訳『データ対話型理論の発見――調査からいかに理論をうみだすか』新曜社.）

[6] Di Gregorio, Silvana and Judith Davidson ed., 2008, *Qualitative Research Design for Software Users,* New York: Open University Press.

[7] Kelle, Udo, Ernest Sibert, Anne Shelly, Sharlene Hesse-Biber, and Günter Huber. 1995. "Part Ⅲ Computers and Qualitative Hypothesis Examination, Introduction: Hypothesis Examination in Qualitative Research." pp. 105–112 in *Computer-Aided Qualitative Data Analysis: theory, methods and practice,* edited by U. Kelle, G. Prein, and K. Bird. London: Sage Publications.

[8] 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.

[9] Mandel, Leslie A., 2008, "Doing What It Takes: The Impact of Organizational Issues on Massachusetts School-Based Health Center Success." pp. 151–179 in *Qualitative Research Design for Software Users,* edited by S. di Gregorio and J. Davidson. New York: Open University Press.

[10] Patton, Michael Quinn, 2002, *Qualitative Research & Evaluation Methods,* Thousand Oaks, Calif.: Sage Publications.

[11] Paulus, Trena M., Jessica N. Lester, and Paul G. Dempster, 2014, *Digital Tools for Qualitative Research.* Los Angeles: Sage Publications.

[12] 樋口耕一，2014，『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』ナカニシヤ出版.

[13] 樋口麻里，2013，「統合失調症者の回復過程とその分岐点における居場所――Atlas.ti を利用したグラウンデッド・セオリー・アプローチによる試み」，『社会と調査』(11):85–100.

[14] Pollach, I., 2011, "Software Review: Wordstat 5.0." *Organizational Research Methods,* 14(4):741-744.

[15] 西條剛央，2005，『構造構成主義とは何か――次世代人間科学の原理』北大路書房.

[16] 戈木・クレイグヒル滋子，2013，『質的研究法ゼミナール――グラウンデッド・セオリー・アプローチを学ぶ』医学書院.

[17] 佐藤郁哉 , 2008,『実践質的データ分析入門――QDA ソフトを活用する』新曜社.

[18] Strauss, Anselm L. and Juliet M. Corbin. 2008. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory,* 3rd edition, Thousand Oaks, Calif.: Sage Publications.（=2012，操華子・森岡崇訳『質的研究の基礎――グラウンデッド・セオリー開発の技法と手順』医学書院.）

注

1）CAQDAS は、他に QDAS(Qualitative Data Analysis Software) や、QDA ソフトウエアと呼ばれることがある。本稿では、CAQDAS の表記を用いる。

2）Mandel Leslie A. は、CAQDAS を利用することで、利用者は自身がどのような分析を行うのかについて、考えるようになると指摘している (Mandel 2008:179)。

3）Glaser と Strauss 以降、グラウンデッド・セオリー・アプローチに対する考え方は、様々に枝分かれをしており (Strauss and Corbin 2008=2012)、立場によって採用する方法が異なっている。

4）グラウンデッド・セオリー・アプローチでは、コーディングは、分析者とデータの相互作用を通して、データから概念を引き出すことと説明されている (Strauss and Corbin 2008=2012)。したがって、「データに貼るラベル＝コー

ド」あるいは「データにラベルを貼る作業＝コーディング」と表現することは、コーディングという言葉が持つ意味を単純化しすぎる可能性がある。しかし、本稿の目的は、CAQDAS のコーディング機能と対応させて、コーディングを説明することにあるため、データに貼るラベルをコードとする。

5）Altas.ti7.1 User Guide and Reference (http://atlasti.com/wp-content/uploads/2014/05/atlasti_v7_manual_201312.pdf?q=/uploads/media/atlasti_v7_manual_201312.pdf) を参照のこと。

6）NVivo では、「リファレンス」という名前がついている。

7）Atlas.ti でも、ノードという言葉は登場するが、NVivo とはその意味が異なる。Atlas.ti において、ノードはネットワーク図内に投入されたオブジェクト (コード、メモ等) のことを意味する。

8）NVivo は version11 から、オープン・コーディングに対応するノードを「概念ノード (theme node)」、データ整理については「ケースノード (case node)」と、機能別にノードの呼称が分けられた。また、新たに「感情ノード (sentiment)」が付け加えられたが、本稿では、このうち主要なノードである、概念ノードとケースノードを取り上げる。

9）NVivo の画面上では、概念ノードは nodes、ケースノードは cases と表示される。

10）Atlas.ti8 では、ファミリーの呼称は「グループ」に変更されている。Atlas.ti8 は、本稿の出版直前にリリースされたため、本稿では Atlas.ti7 を扱っている。

11）Atlas.ti7 のサンプルプロジェクトのデータを用いた。コード、コード・ファミリー、データのファミリー (プライマリードキュメント・ファミリー) は筆者が作成した。

12）表 1 では、複数のコードをまとめた、コード・ファミリーを用いている。

13）NVivo では、4.1 節で述べる行列コーディングの機能に対応する。Atlas.ti では、同一ファイル内のデータの特定部分に属性値を振って、収集したデータ全体で、属性ごとのコードの分布を出す、といったようなことはできない。これに対して、NVivo では、同一ファイル内のデータに異なる属性値を割り当てることも可能である。

14）概念構築のための主要な方法として階層化を採用する場合、コードをツリー形式でまとめる機能を利用する CAQDAS ユーザーが多いだろう。この点について、スーザン・フリーズは、Atlas.ti を使うのであれば、概念化の作業は、コードをツリー形式にする code forest 機能ではなく、ネットワーク図を使うように勧めており (Friese 2012)、これは本稿の指摘とも一致する。

15）Altas.ti7.1 User Guide and Reference の p.311 を参照のこと。

16）Atlas.ti を用いて、修正版グラウンデッド・セオリー・アプローチの方法を実際のデータ分析に適用した例として、樋口 (2013) がある。

17）NVivo11 のサンプルプロジェクトのデータを使用した。サンプルプロジェクトでは、あらかじめノードが作成されており、本稿の例でもこれらを用いた。

18）例えば、「Barbara」がインタビュー調査と質問紙調査の両方に回答していたする。「Barbar」というケースノードをインタビューデータと質問紙調査の ID に振ると、Barbara のノードは、両調査での Barbara のすべての回答に割り当てられる。

19）NVivo11 for Windows Help (http://help-nv11.qsrinternational.com/desktop/welcome/welcome.htm) を参照のこと。

20）言葉の抽出のされた方は、データの言語によって異なる。

21）調査者の主観性の問題については、観察や測定、解釈の行為と、それらの対象となる現象をどのように定義するか、といった認識論から議論する必要がある。こうした問題を扱った研究として、西條剛央の「構造構成主義」（西條 2005）による、質的・量的調査法に対するメタ理論がある。

Methodological Philosophy of CAQDAS and Their Functions: Difference between Atlas.ti 7 and Nvivo 11

Mari HIGUCHI

Abstract:

Behind the functionalities of analytical software, there is a methodological philosophy as to how analysis should be done. It is essential to understand such philosophies of each CAQDAS (Computer-Assisted Qualitative Data Analysis Software) in order to use them appropriately for qualitative data analysis. For this reason, this paper explores the functionalities of Atlas.ti7 and NVivo11 to elucidate their philosophical differences.

The Atlas.ti can easily handle codes with low abstraction, which enables users to consider the relationship between codes and to construct concepts or theories grounded on data. The NVivo aims to respond to various methods. It emphasizes extracting the characteristics of large amounts of data by examining the data diversely with applying two functions of coding: constructing concepts and classifying the data.

Key Words : CAQDAS, QDA software, Atlas.ti, NVivo, qualitative data analysis