

Maciej Brosz

Zastosowanie pakietu NVivo w analizie materiałów nieustrukturyzowanych

Przegląd Socjologii Jakościowej 8/1, 98-125

2012

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Maciej Brosz
Uniwersytet Gdański

KOMPUTEROWE WSPOMAGANIE BADAŃ JAKOŚCIOWYCH. ZASTOSOWANIE PAKIETU NVIVO W ANALIZIE MATERIAŁÓW NIEUSTRUKTURYZOWANYCH

Abstrakt Artykuł poświęcony jest wykorzystaniu oprogramowania NVivo w analizie danych jakościowych. Główny wątek opracowania odnosi się do problemu redukcji danych jako procesu towarzyszącego analizie jakościowej. Wykorzystanie oprogramowania nie musi oznaczać niekontrolowanego procesu zmiany postaci danych i tym samym utraty istotnych wymiarów doświadczeń zgromadzonych w materiale badawczym. Najnowsze narzędzia z rodziny CAQDAS (np. NVivo 8, 9) pozwalają zminimalizować ów problem poprzez umożliwienie pracy na materiałach źródłowych w formie niezmodyfikowanej. Artykuł przedstawia przykłady analizy danych w postaci tekstów, obrazów, nagrań audiowizualnych. Dodatkowo przedstawione zostały wybrane narzędzia wspomagające proces kodowania.

Słowa kluczowe komputerowe wspomaganie danych jakościowych, utrata informacji, redukcja danych, kodowanie, NVivo, metodologia teorii ugruntowanej

Maciej Brosz, asystent w Zakładzie Socjologii Stosowanej Instytutu Filozofii, Socjologii i Dziennikarstwa Uniwersytetu Gdańskiego, analityk w Pracowni Realizacji Badań Socjologicznych UG. Główne zainteresowania to socjologia mieszkalnictwa, rewitalizacja zdegradowanych obszarów miejskich, problematyka bezrobocia oraz metodologia nauk społecznych.

Dane adresowe autora:

Uniwersytet Gdański, Instytut Filozofii,
Socjologii i Dziennikarstwa
ul. Bażyńskiego 4, 80-952 Gdańsk
e-mail: maciej.brosz@post.pl

Niniejszy artykuł poświęcony jest prezentacji doświadczeń zdobytych w toku pracy badawczej, w której wykorzystano oprogramowanie przeznaczone do analiz danych jakościowych. W momencie podjęcia prac badawczych z wykorzystaniem narzędzi komputerowego wspomaganie badań jakościowych, czyli narzędzi typu CAQDAS (Computer Assisted/Aided Qualitative Data Analysis), spośród dostępnego na rynku w 2004 roku oprogramowania wybrano program NVivo¹.

¹ Obok NVivo na rynku istnieje wiele innych programów przeznaczonych do wspierania analiz jakościowych. Są

Historia² narzędzia sięga roku 1981, kiedy opracowano pierwszą wersję programu NUDIST (*Non-Numerical Unstructured Data Indexing Searching and Theorizing*), w roku 1997 pojawiła się jego zmodyfikowana wersja pod nazwą N4. Ko-

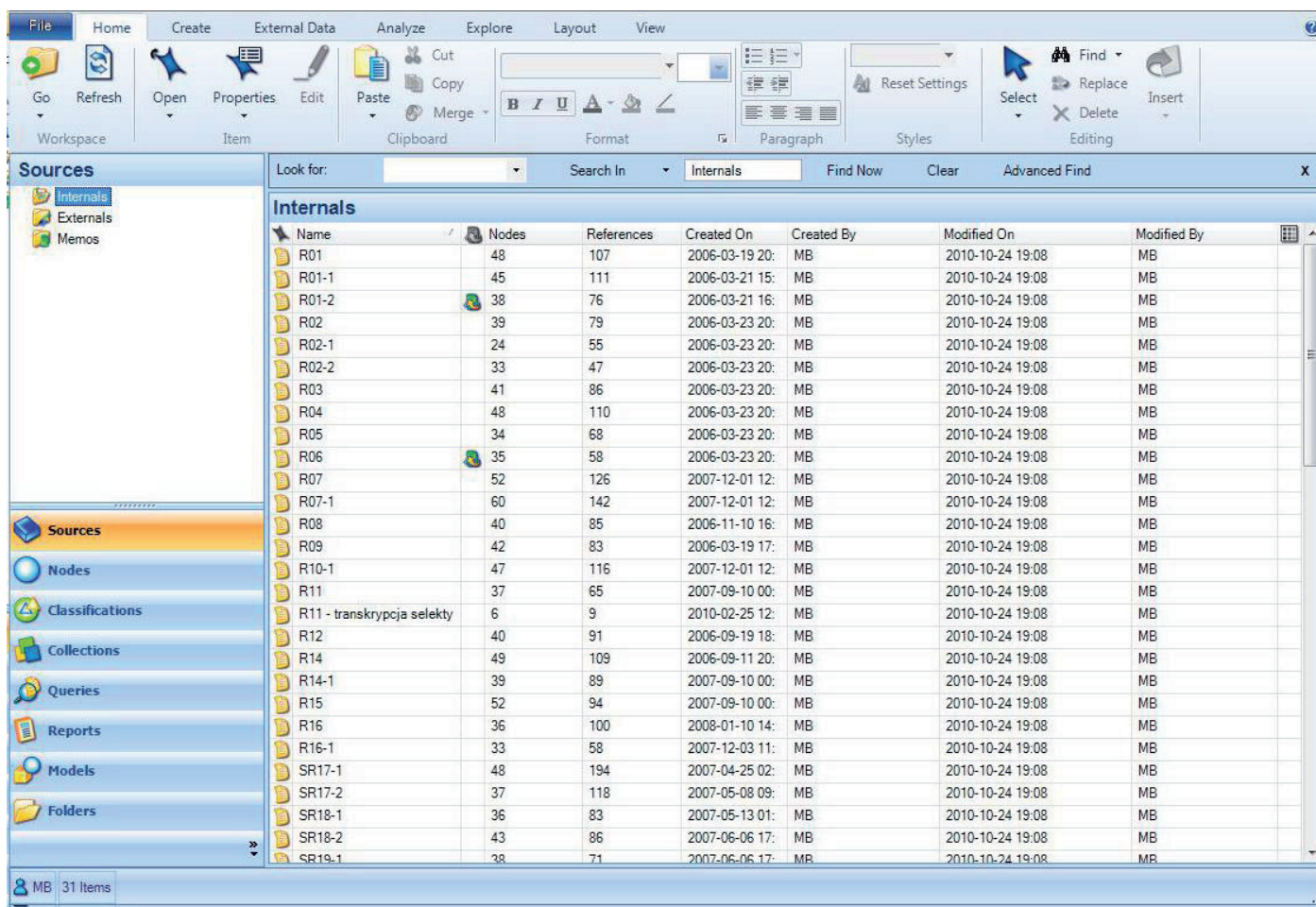
one zróżnicowane pod względem specyficznych funkcji, modelu działania, poziomu komplikacji oraz przede wszystkim przeznaczenia. Programy, które były testowane pod kątem użyteczności w pracy autora to, obok Nvivo, także: Atlas.ti, Ethnograph, N6, Nudist, XSight, MaxQDA, HyperResearch, C-I-Said, QDA Miner, AnSWR, WeftQDA, HyperQual, QualRus. Bogata i na bieżąco aktualizowana lista oprogramowania wraz z opisem funkcji poszczególnych narzędzi jest dostępna w Internecie na stronie <http://caqdas.soc.surrey.ac.uk/>.

² Na potrzeby niniejszego opracowania pominięto historię zastosowań oprogramowania w badaniach jakościowych w szerokim rozumieniu. Przykładowe pozycje bibliograficzne podejmujące to zagadnienie to: Seale (2008), Bieliński, Iwińska, Rosińska-Kordasiewicz (2007).

lejne wersje to N5 oraz wprowadzona w 2002 roku wersja N6. W tym samym roku wydano wersję programu pod nazwą NVivo w wersji 2.

W chwili podejmowania decyzji o wyborze programu dostępna była wersja 2. Dopiero na przestrzeni kolejnych lat ukazały się wersje 7, 8 oraz najnowsza – wersja 9 (data udostępnienia – 21 października 2010 roku)³. Do najistotniejszych zmian, które w nich wprowadzono należą: uproszczenie interfejsu (od wersji 7); rozszerzenie możliwości stosowania procedur kodowania na materiały nieedytowalne, takie jak zdjęcia, nagrania dźwiękowe oraz filmowe (od wersji 8); wprowadzenie różnorodnych metod i form

³ Od października 2011 roku dostępna jest wersja 9.2.



Rys. 1. Główne okno NVivo 9

przeglądania powiązań pomiędzy kategoriami kodowymi (*Nodes*) oraz możliwość posługiwania się arkuszami danych, charakteryzującymi badaną zbiorowość lub zbiór kategorii (od wersji 9). Ostatnia wersja programu obejmuje również wsparcie pracy zespołowej w sieci (NVivo 9 Server). Doświadczenia przedstawione w artykule są oparte na pracy z wersjami 2, 7, 8 oraz przejrzeniem możliwości „dziewiątki”.

Główny wątek opracowania odnosi się do problemu redukcji danych jako procesu towarzyszącego analizie jakościowej. Wykorzystanie oprogramowania nie musi oznaczać niekontrolowanego procesu zmiany postaci danych i tym samym utraty istotnych wymiarów doświadczeń zgromadzonych w materiale badawczym. Najnowsze narzędzia z rodziny CAQDAS pozwalają zminimalizować ów problem poprzez umożliwienie pracy na materiałach źródłowych w formie niezmodyfikowanej (na żywym niemalże materiale, tj. nagraniach audio i filmach) oraz oferując zróżnicowane formalnie procedury porządkowania, klasyfikowania i kodowania danych.

CAQDAS, czyli komputerowe wspomaganie analizy danych

W niniejszym artykule podtrzymuje się ton narracji metodologicznych odnoszących się do wykorzystywania oprogramowania w badaniach jakościowych. W literaturze przedmiotu akcentuje się z jednej strony możliwości tego typu narzędzi, z drugiej zaś ich ograniczenia.

Literatura poświęcona oprogramowaniu przeznaczonemu do wspomaganie lub prowadzenia analiz jakościowych dzieli się na dwa główne

rodzaje. Do pierwszej grupy przypisać można opracowania charakteryzujące rolę tego typu narzędzi w procesie badawczym, wpływ wykorzystania narzędzia na osiągnięte rezultaty i wynikające z tego ewentualne zagrożenia (zob. Lonkila 1995; Coffey, Holbrook, Atkinson 1999; Bieliński, Iwińska, Rosińska-Kordasiewicz 2007). W pracach tych podkreśla się, iż nawet najlepsze oprogramowanie nie zastąpi intuicji badacza oraz jego wyobraźni socjologicznej, a przeniesienie na narzędzia komputerowe ciężaru interpretowania danych skutkować będzie jedynie generowaniem bezsensownych lub co najwyżej jałowych konstatacji. Artykułowane wątpliwości odnoszone są również do stylu pracy, jaki narzuca korzystanie z komputerowego wsparcia. Jednym ze wskazań jest utożsamienie obróbki danych jakościowych z ich kodowaniem, „chcąc prowadzić badania jakościowe z wykorzystaniem komputera należy zakodować posiadane dane” (Lonkila 1995: 48–49). Programy wspomagające badania jakościowe nie ograniczają się w różnorodności oferowanych funkcji tylko do rozmaitych form kodowania. Współczesne wersje oferują oprócz możliwości kodowania wiele innych funkcji, które wykraczają poza jego wspomaganie.

Drugą grupę opracowań stanowią instrukcje obsługi, *tutoriale*, fragmenty podręczników, które mają za zadanie przybliżyć czytelnikowi, czyli potencjalnemu użytkownikowi, zasady posługiwania się danym narzędziem (zob. Trutkowski 1999; Babbie 2003; Seale 2008; Gibbs 2011). O ile teksty pierwszej grupy spełniają swoją funkcję znakomicie – pozwalają wypracować stanowisko względem modelu użytkownika i wkomponowywania oprogramowania w pracę badawczą

– to opracowania z drugiej grupy stosunkowo szybko się dezaktualizują ze względu na rozwój technologii (pojawianie się nowych funkcji oraz wzrost mocy obliczeniowej dostępnych urządzeń). Nie oznacza to jednak, że opracowania te stają się bezużyteczne, ponieważ rozwiązania w nich opisywane są obecne w kolejnych wydaniach poszczególnych programów.

Wątpliwości i zarzuty formułowane w literaturze przedmiotu powinny być traktowane jako istotne wskazania uwrażliwiające i uświadamiające badaczy, z czym wiąże się wkomponowanie w proces badawczy elementu wsparcia informatycznego. Obok metod tradycyjnych – angażujących papier i ołówki – istnieją również te angażujące moc obliczeniową procesora, równoległość dokonujących się procesów czy dynamiczne sortowania tysięcy jednostek wyodrębnionych w toku analizy danych. Potencjał ten należy w sposób kontrolowany i świadomy wykorzystywać z jednej strony w celu podniesienia jakości opracowań, a z drugiej – aby rozsądnie dysponować czasem. Niewątpliwą zaletą korzystania z oprogramowania jest oszczędność czasu, co w warunkach prowadzenia komercyjnej działalności badawczej ma ogromne znaczenie⁴. Zaznaczyć też należy, iż „komputerowo wspomagana analiza danych jakościowych (CAQDAS) [...] nie jest oczywiście odpowiednia do wszelkich działań⁵ podejmowanych przez bada-

4 O ile proces kodowania danych metodami tradycyjnymi i wspomaganymi komputerowo zajmuje w przybliżeniu tyle samo czasu – dane trzeba dokładnie przeczytać i uporządkować – to przeszukiwanie danych, ich wizualizacja, możliwości sortowania i tworzenia kwerend przemawiają na korzyść analiz obejmujących korzystanie z CAQDAS.

5 Jeśli oprogramowanie tego typu nie znajdzie bezpośredniego zastosowania w ramach prac badawczych, może być z powodzeniem użyte jako narzędzie służące do za-

czy wobec danych jakościowych, ale dostarcza znaczących korzyści” (Seale 2008: 234).

Na potrzeby niniejszego opracowania zostały przyjęte następujące założenia. Po pierwsze, komputerowe wspomaganie danych nie jest jedyną i najlepszą metodą opracowywania materiału w badaniach jakościowych. Proponuje się traktować ten model analiz jako rodzaj propozycji wzbogacającej metodologię badań jakościowych. Różnorodność w tej sferze skutkować może cenną możliwością dokonywania wyboru i dopasowywaniem stosowanych rozwiązań względem prowadzonej działalności. Pożądany obraz kondycji metodologicznego oprzyrządowania socjologii przedstawiałby sytuację dynamicznego współistnienia różnych, a przez to wartościowych sposobów dokonywania przygotowania, obróbki i interpretacji danych pochodzących z badań.

Po drugie, stosowanie oprogramowania komputerowego w toku badania nie może eliminować aktywności intelektualnej badacza. Komputer i program niczego samodzielnie nie stworzą. Ta rola pozostaje przypisana analitykowi. W mniejszym stopniu struktura oprogramowania ingeruje w proces opracowywania materiału i interpretacji, tym doskonalszym jest narzędzie, gdyż w mniejszym stopniu współtworzy świat badania niezależnie od woli analityka. Przezroczystość interfejsu oraz algorytmu programu

zarządzania gromadzonym przez każdego naukowca zbiorem materiałów bibliograficznych, cytatami, fiszkami czy zeskanowanymi odręcznymi notatkami. Wykorzystanie mechanizmu kodowania do wielowymiarowego indeksowania gromadzonych informacji pozwala dynamicznie i efektywnie selekcjonować informacje na potrzeby rozmaitych opracowań.

komputerowego stanowi jego atut, który należy uwzględnić, dokonując wyboru narzędzia.

Po trzecie, im większa różnorodność możliwości ujmowania informacji i danych przewijających się w toku procedury badawczej, tym większa szansa na zrealizowanie postulowanej przezroczystości narzędzia analitycznego, jakim jest określony program. Stosowanie narzędzia, które traktuje dane jednorodnie, skutkować będzie ograniczonym modelem generowanej analizy. Formalna różnorodność uzyskanych danych pociąga konieczność dopasowywania stosowanych procedur, aby uwydatnić tkwiący w określonych źródłach potencjał informacyjny.

Prezentowane poniżej doświadczenia nie powinny być odczytywane w tonie wpisywania się lub nie w obszar metodologii teorii ugruntowanej lub analizy jakościowej w szerszym rozumieniu (QDA). Opisywane procedury mogą z powodzeniem zostać zaadoptowane na potrzeby przedsięwzięcia badawczego reprezentującego radykalne nurty teorii ugruntowanej, ale też dobrze komponować się z jakościowymi opracowaniami innego typu.

Status formalny danych a procedury kodowania

Problem stratnej kompresji informacji

Procedurom analizy danych jakościowych towarzyszy zjawisko ich redukcji. Proces ten, o ile jest kontrolowany, stanowi kluczowy element opracowywania materiału i w takim rozumieniu stanowi specyfikę procesu interpretacji prowadzącej do wypracowania modelu teoretycznego. Jak piszą Miles i Huberman, „redukcja danych od-

nosi się do procesu selekcji, koncentracji, upraszczania, odrywania i przekształcania danych” (2000: 11). Każdej fazie obróbki danych towarzyszy zmiana ich postaci, formuły i znaczenia. Szczególnie istotne jest w tym miejscu uwzględnienie, na ile badacz jest świadomy owych przekształceń skutkujących formalnym i treściowym oddaleniem się danych od źródłowej postaci, czyli empirycznie wypreparowanych elementów świata społecznego. Przyjmując, że proces ten przebiega względnie świadomie (por. Glaser, Holton 2004), będzie można określić go mianem analizy lub interpretacji, jeśli natomiast zmiany zachodzą w sposób niekontrolowany, to rezultatem będzie bliżej nieokreślona impresjonistyczna wizja określonego wymiaru rzeczywistości.

Proces redukcji stanowi jeden z elementów obróbki danych jakościowych i jest immanentną cechą jakościowego procesu badawczego. Proces ten należy odróżniać od skutku, jakim jest wspomniane wyżej oddalenie się danych od ich wyjściowej postaci. Użyteczną ilustracją tego zjawiska jest porównanie go do procesu kompresji danych w ramach komputerowej obróbki materiału dźwiękowego. Konwersja dźwięku z postaci analogowej do cyfrowej, a w tym ostatnim przypadku na przykład z formatu surowego WAV do popularnego formatu MP3, skutkuje utratą jakości dźwięku. Kompresja polega w tym przypadku – w przybliżeniu – na eliminowaniu składowych dźwięku niedostępnych z osobna ludzkiemu zmysłowi słuchu. Okazuje się jednakże, że proces ten nie pozostaje bez znaczenia dla jakości odbioru dźwięku. Miłośnicy muzyki (i nie chodzi tu o tak zwanych audiofilów) zwracają uwagę na fakt, iż dźwięk pozbawiony pełnego spektrum alikwotów nie jest odzwier-

ciędlaniem dźwięku rzeczywistego (oznacza to utrudnione rozpoznanie barwy głosu znanej osoby lub brzmienia charakterystycznego dla danego instrumentu)⁶.

Gromadzenie i rejestrowanie informacji skutkuje generowaniem danych. Przeistoczenie informacji w dane to proces polegający na utrwaleniu i związaniu informacji z określonym sposobem zapisu. Oczywistym jest, że już na tym etapie dochodzi do zjawiska redukcji i towarzyszącej mu utraty informacji. Znakomitym przykładem jest tu posługiwanie się transkrypcjami nagrań wywiadów przez osoby, które nie przeprowadzały osobiście wywiadów. W takim przypadku możliwość odniesienia się do istotnych czynników sytuacyjnych, klimatu rozmowy czy nastroju badanego jest wysoce ograniczona, skutkiem czego może być nietrafna interpretacja znaczenia słów czy nawet sensu całego wywiadu (por. Riessman 1993).

Zróźnicowanie danych wykorzystywanych w analizie jakościowej a możliwości współczesnego oprogramowania

Wskazany wyżej problem utraty pokładów znaczeń, towarzyszący procesowi gromadzenia czy obróbki danych, nie jest traktowany tylko jako bariera, która skutecznie blokuje możliwości interpretacyjne. Posługiwanie się metodami tradycyjnymi, nazywanymi czasem analogowymi, może sprawiać wrażenie, iż badacz pozostaje

⁶ Zjawisko to można sprawdzić w warunkach domowych, przeprowadzając prosty eksperyment. Należy nagrać przy pomocy komputera kilka wypowiedzianych zdań, a następnie poddać nagranie kilkukrotnej kompresji na przykład do popularnego formatu MP3. Porównanie brzmienia oryginalnego nagrania i jego wersji skompresowanej będzie więcej niż zaskakujące.

w pełnym kontakcie z badanym fragmentem rzeczywistości społecznej dzięki utrwalonym w swej pamięci doświadczeniom z terenu. Współcześnie rozwijane oprogramowanie skutecznie wspomaga (i tylko wspomaga, nie zastępuje) ów proces utrwalania wrażeń. Dysponując filmem lub nagraniem audio, oprócz sporządzania, a następnie lektury transkrypcji, badacz może równolegle je oglądać lub przesłuchiwać. W trakcie odsłuchiwania nagrań lub oglądania filmów poszczególnym fragmentom można przypisywać stosowne komentarze i uwagi. Zauważalny jest zatem ścisły związek tekstu przepisanego z tym utrwalonym na nagraniu czy filmie. Ten sam mechanizm pozwala na wiązanie treści zakodowanych bezpośrednio na nagraniu, fotografii czy też powiązanej ze źródłami – transkrypcji. Jeśli przyjąć, że rezultatem opracowania jest nota teoretyczna (w dowolnej fazie jej opracowywania), analityk ma możliwość natychmiastowego skonfrontowania swoich spostrzeżeń z danymi bazowymi – zredukowanymi w sposób kontrolowany.

Materiały poddawane analizie jakościowej można określić jako nieustrukturyzowane. O ile metody ich pozyskiwania, porządkowania przyjmują formułę pewnych procedur czy nawet algorytmów, to uzyskiwanych rezultatów w postaci gromadzonych źródeł nie da się przyrównać do danych, jakimi dysponują analitycy materiałów o charakterze ilościowym. Próba sprowadzenia bogactwa i formalnej różnorodności treści nagrań wypowiedzi lub uchwyconego na fotografii kadru życia społecznego do standaryzowanego oznaczenia rekordu bazy danych stanowiłaby znaczącą utratę ważnych informacji. Z drugiej strony udana próba zbudowania bazy danych

obejmującej wszystkie istotne wymiary wydaje się pozostawać w sferze marzeń.

Formalne zróżnicowanie materiałów w kontekście ich użytkowania w komputerowo wspomaganą analizę danych wynika z ograniczeń, jakie nakłada posługiwanie się tego typu narzędziami. Korzystanie z nawet najlepszego oprogramowania nie wyeliminuje konieczności przyjęcia pewnych konwencji komunikowania się z potencjałem obliczeniowym komputera.

W kontekście prezentowanych tu doświadczeń omówienia wymaga zróżnicowanie postaci materiałów źródłowych wykorzystywanych na potrzeby analizy. Można przyjąć, iż forma utrwalenia informacji zdobytych w trakcie badań wpływa na sposób ich późniejszej obróbki komputerowej.

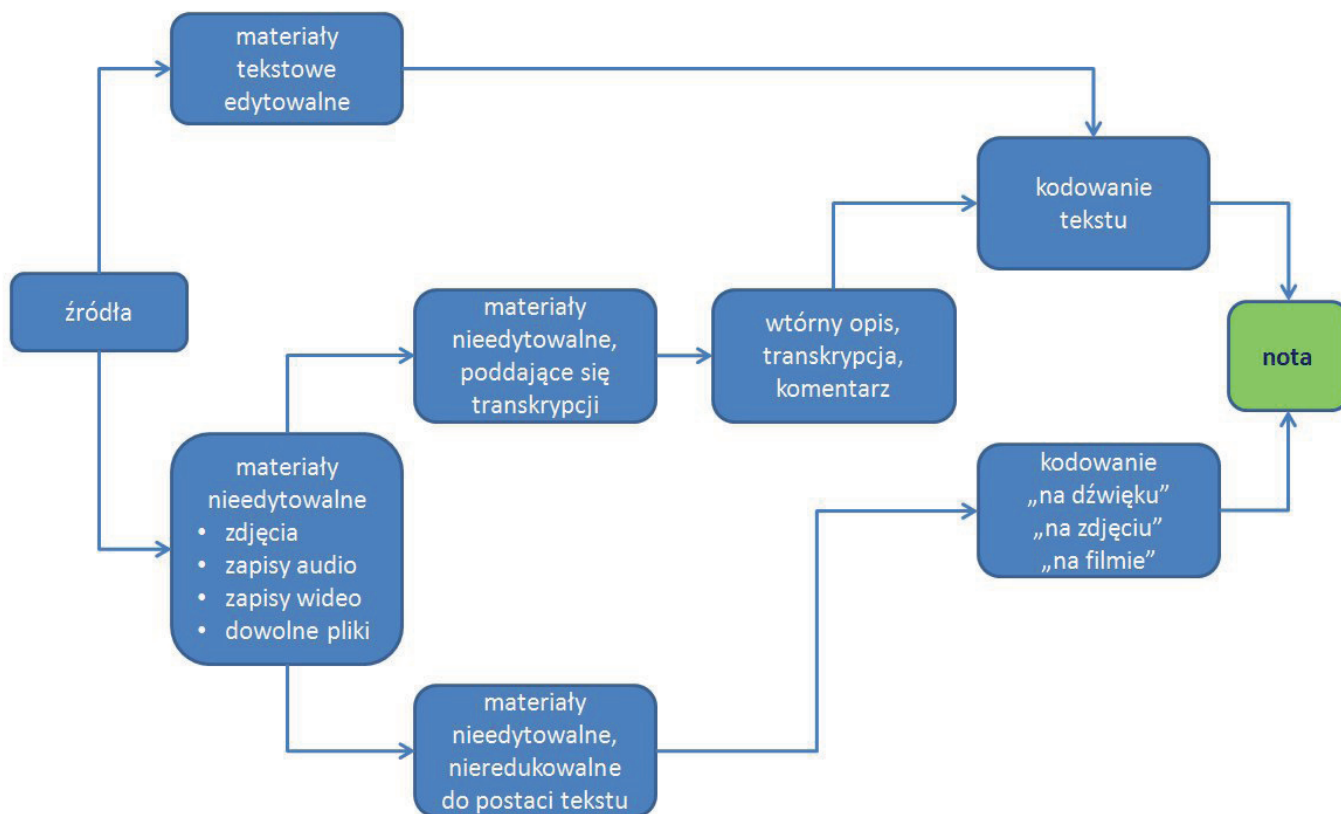
Podstawową i z reguły najczęściej wykorzystywaną formą utrwalania danych będą pliki tekstowe stworzone przy użyciu edytorów tekstu. Pliki tego rodzaju reprezentują kategorię materiałów edytowalnych, po ich zaimportowaniu w obręb projektu NVivo badacz może z miejsca przystąpić do analizy, kodowania czy przeszukiwania. Wyróżnikiem tego typu danych jest możliwość przeprowadzania operacji typu kopiuj/wklej bez zmiany sensu i postaci informacji. Dla przykładu operacji takiej nie można wykonać bez konsekwencji w postaci utraty ważnych informacji na materiałach wizualnych. Nawet najlepsze oprogramowanie nie zaklasyfikuje gestu czy ciekawej sytuacji utrwalonej na zdjęciu.

Zdjęcia, nagrania audiowizualne czy zeskanowane dokumenty pisane odręcznie pozostają dla oprogramowania i komputera jedynie nic

nieznaczącą sekwencją bitów. Takie znaczenie musi zostać nadane materiałom przez badacza. Ten typ dokumentów można określić mianem nieedytowalnych. Niech ilustracją w tym miejscu stanie się brak możliwości wykorzystania polecenia „szukaj” (Ctrl+F) do odnalezienia znajdującej się w naszym zbiorze fotografii, na której widnieje na przykład osoba i towarzyszący jej pies przewodnik, za pomocą wpisania odpowiednich słów. Nie należy się spodziewać, iż po wpisaniu słowa „pies” program odczyta, czym jest pies, jak wygląda i w rezultacie wskaże określony obiekt.

Wymienione wyżej dane nieedytowalne wymagają opracowania, czyli przekonwertowania ich do postaci, która będzie czytelna dla algorytmu programu i jednocześnie zadowalająca dla analityka. Proces ten, polegający na konwersji jakości wizualnej, dźwiękowej, opiera się na sprawdzeniu postaci informacji do formy tekstu. Można w tym miejscu odnieść model konstruowania danych do procesu transkrypcji czy tekstualizacji. Procesowi temu towarzyszy oczywiście pewna utrata informacji. Jednakże w zależności od precyzji, zdolności językowych badacza strata informacji może być kontrolowana lub przynajmniej ukierunkowana.

Współczesne oprogramowanie pozwala dokonywać pewnych operacji bezpośrednio na zgromadzonym materiale, nawet w postaci nieedytowalnej. Przykładem może być kodowanie przeprowadzone bezpośrednio na nagraniu wywiadu, a nie na jego transkrypcji, co zostanie docenione przez badaczy, którzy oprócz lektury zapisów wywiadu preferują odsłuchiwanie nagrań (Kaufmann 2010).



Rys. 2. Procesy kodowania na tle formalnego zróżnicowania danych

Status formalny danych: teksty, obrazy, dźwięki i filmy

Możliwości jakie oferuje program stają się dostępne w momencie, gdy w pracy analityka wykorzystywane są wszystkie możliwości klasyfikowania i porządkowania zgromadzonych danych. Dane wykorzystywane w badaniach jakościowych charakteryzuje wysoki poziom zróżnicowania formalnego. Ograniczenie formy opracowywanych danych tylko do materiałów tekstowych stanowiłoby jednocześnie ograniczenie możliwości dokonywania w nie pogłębionego wglądu badawczego. Program NVivo umożliwia wprowadzanie w obręb projektu badawczego danych o wysokim stopniu zróżnicowania. Mogą to być materiały tekstowe, zdjęcia, nagra-

nia dźwiękowe i filmowe. Możliwości programu są pod tym względem rozległe – badacz ma możliwość wprowadzania danych w różnych formatach plików. W sytuacji, w której zgromadzony materiał ma postać plików nieczytelnych dla programu można posłużyć się mechanizmem tworzenia wirtualnych reprezentacji materiałów (*Externals*) i pracować z danymi w postaci umożliwiającej ich komentowanie, tekstualizację czy transkrypcję. Oznacza to, iż na potrzeby analizy badacz może wykorzystać każdy rodzaj danych, który można sprowadzić do postaci pliku komputerowego (*File link*), odnośnika internetowego (*web link*) lub dowolnego obiektu, który badacz przy pomocy dostępnych sobie form artykulacji będzie w stanie opisać, a następnie wykorzystać na potrzeby prowadzonych badań.

Przy założeniu, że badacz posługuje się standardowymi formatami tekstów (rozszerzenia: .txt, .rtf, .doc, .docx), fotografii (.bmp, .jpg, .gif, .tiff), materiałów dźwiękowych (.mp3, .wav, .wma), filmowych (.mpq, .mpeg, .avi, .mov, .mp4), import danych nie będzie stanowić bariery.

Organizacja danych

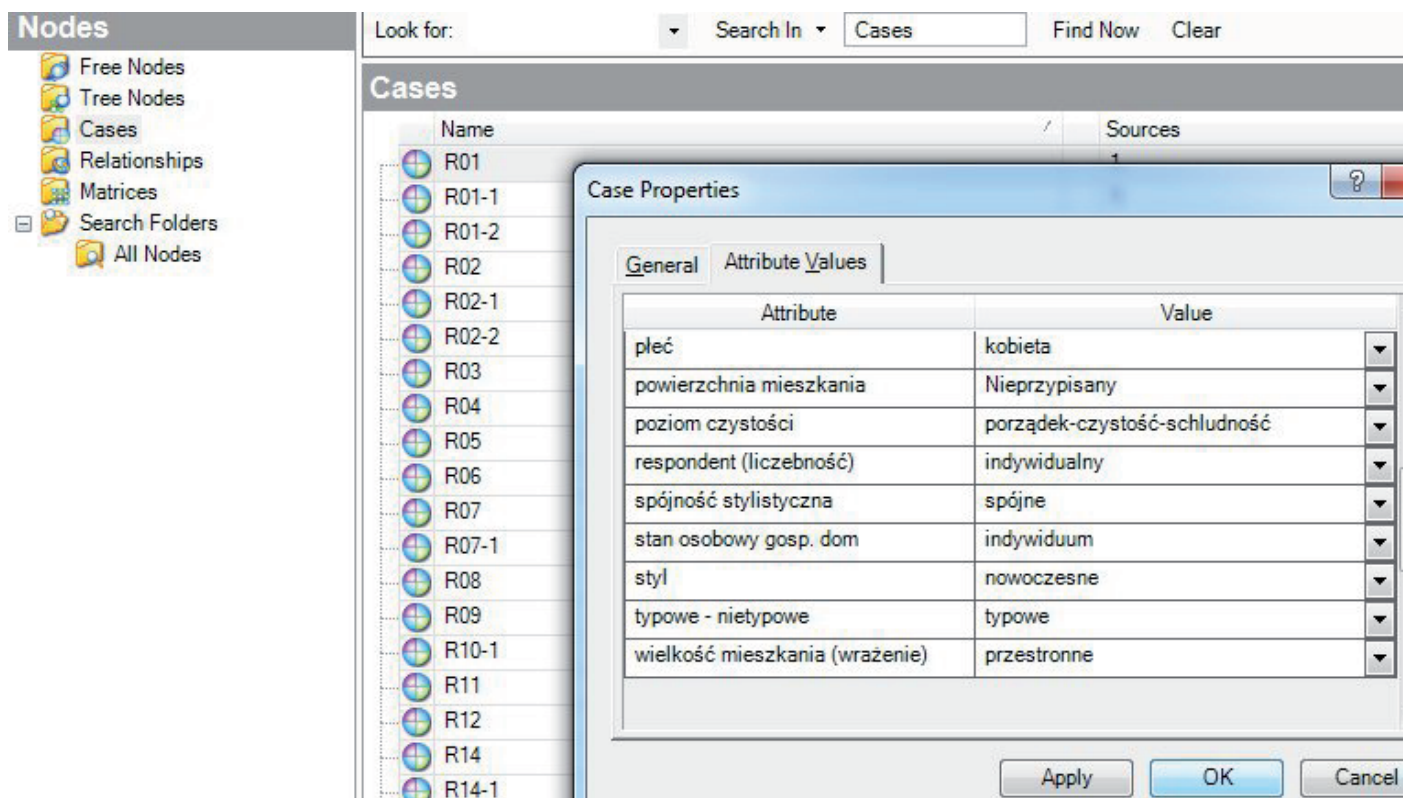
Sposób uporządkowania danych w NVivo tworzy rodzaj hierarchii. Ze względu na specyfikę jakościowego podejścia wszelkie hierarchiczne uporządkowania wydają się sztuczne, biorąc jednakże pod uwagę na poły instruktażowy charakter artykułu, przyjęta została perspektywa pozwalająca dokonać przeglądu możliwości narzędzia. Zbiory informacji oraz kategorie klasyfikacyjne w obrębie NVivo tworzą następujący szereg:

- zbiory dokumentów (*Sets, Collections*);
- pojedyncze dokumenty źródłowe (teksty, obrazy, materiały audio, filmy);
- wirtualne przypadki (*Cases*) reprezentujące poszczególne osoby lub sytuacje będące przedmiotem badania oraz ich atrybuty;
- noty (*Memos*);
- kategorie kodowe (*Nodes*) wyodrębnione na podstawie podstawowej analizy tematycznej (zob. Charmaz 2009);
- zdarzenie, wątek (np. opisywane w kilku akapitach);
- akapit, jednorodna jednostka wypowiedzi, jednorodna scena filmowa;
- kategorie kodowe wyodrębnione w ramach kodowania wiersz po wierszu;
- zdanie, wiersz;
- wyrażenia, związek frazeologiczny, zbitka słowna;
- słowo.

Do powyższej klasyfikacji należy wskazać narzędzia pozwalające ustalać relacje pomiędzy poszczególnymi, wymienionymi wyżej elementami. Wymienić w tym miejscu warto narzędzia linkujące ze sobą dowolne elementy treści oraz graficzne oznaczenia relacji, które również mogą zostać potraktowane jako dokument i jako takie być interpretowane.

Konstruowanie danych w NVivo

Punktem wyjścia do pracy w programie jest import zgromadzonych materiałów do projektu (NVivo wykorzystuje ujednolicony typ plików o rozszerzeniu .nvp). Operacja ta sprowadza się do wybrania opcji importowania plików i wskazania ich lokalizacji – reszta procesu przebiega automatycznie i nie wykracza poza zwykłe uruchomienie operacji otwierania plików. Nie jest to jednak koniec możliwości wstępnej obróbki danych na poziomie importu. Konstruowanie danych w obrębie narzędzia komputerowego to oprócz ich kompletowania i porządkowania również opatrywanie ich atrybutami. Mechanizm przypisywania atrybutów poszczególnym materiałom źródłowym odpowiada w NVivo procedurze gromadzenia informacji o respondentach przy pomocy pytań metryczkowych. Jeśli każdy dokument odpowiada określonej badanej sytuacji, konkretnej osobie czy też dowolnej, istotnej dla badacza charakterystyce, to można przypisać danemu dokumentowi taki właśnie atrybut, tworząc tym samym kryterium klasyfikacyjne, które pozwoli przyporządkować ustalone przez analityka wartości każdemu przypadkowi. Zabieg ten pozwala operować kontrolowanym obszarem zmienności i tym samym rozszerzyć możliwości przeszukiwania zbioru



Rys. 3. Atrybuty są przypisywane poszczególnym przypadkom objętym badaniami

danych, a w momencie zakodowania materiału – porównywać różnice w obszarach poszczególnych kategorii na tle wyróżnionych uprzednio zmiennych.

Przypisywanie atrybutów powinno być traktowane tak samo skrupulatnie jak samo kodowanie. Im więcej atrybutów zostanie wprowadzonych oraz przypisanych wszystkim zebranym i sukcesywnie dołączanym do zbioru materiałom, tym większa szansa na możliwość przeprowadzania w kolejnych fazach badania interesujących kwerend. Lista atrybutów oraz poszczególnych wartości nie jest nigdy gotowa – badacz samodzielnie konstruuje poszczególne wymiary i w każdej chwili może poddać je pożądanym modyfikacjom.

Procedury kodowania danych w NVivo

Podstawową zaletą korzystania z programów do obróbki danych jakościowych, reprezentowanych przez NVivo, jest porządek, jaki rozwiązania tego rodzaju wprowadzają w proces pracy z i nad materiałem. Wszystkie informacje są ulokowane w jednym pliku, który można utożsamiać z realizowanym projektem. Wszelkie dane, takie jak transkrypcje, różnego rodzaju noty, wszelkiego typu i funkcji zapiski, znajdują się w jednym miejscu, co w przypadku projektów realizowanych metodami tradycyjnymi może być trudne do osiągnięcia.

Zanim zostaną zaprezentowane procedury kodowania konieczne jest ustalenie statusu kate-

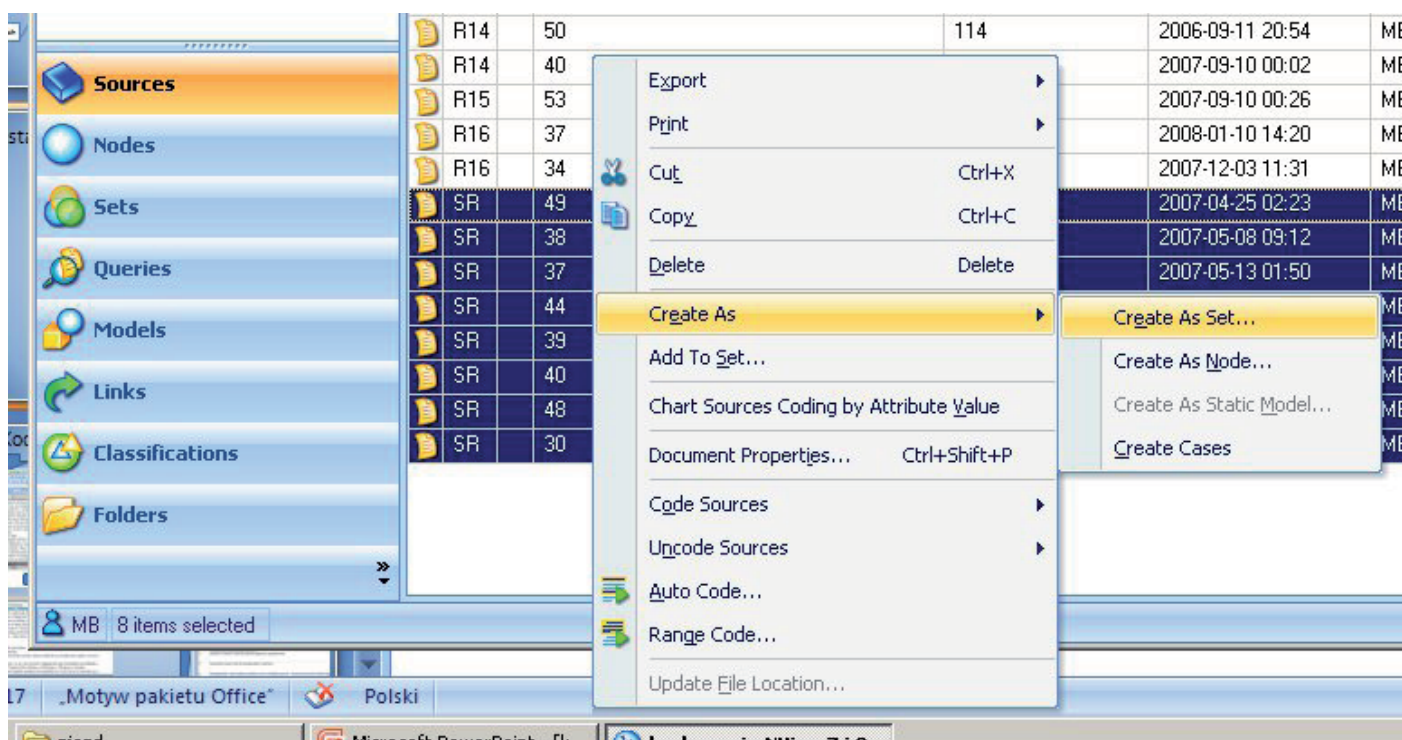
gorii kodowych. Pojęcie kategorii w kontekście prezentacji możliwości narzędzi komputerowych nabiera przede wszystkim znaczenia proceduralnego. Oznacza to, że pojęcie to odnosi się w mniejszym stopniu do opisu gęstego rozumianego jako strategia interpretacyjna, a w większym do zabiegów klasyfikacyjnych dokonywanych na zebranych materiałach. I tak, wykorzystywane w programie NVivo określenie *Nodes* dotyczy raczej określonego agregatu, aniżeli zidentyfikowanego obszaru działania społecznego. Kategorię kodową należy w tym miejscu rozumieć jako zbiór informacji jednorodnych pod jakimś, znanym dla badacza, względem.

Grupowanie danych – *Sets, Collections*

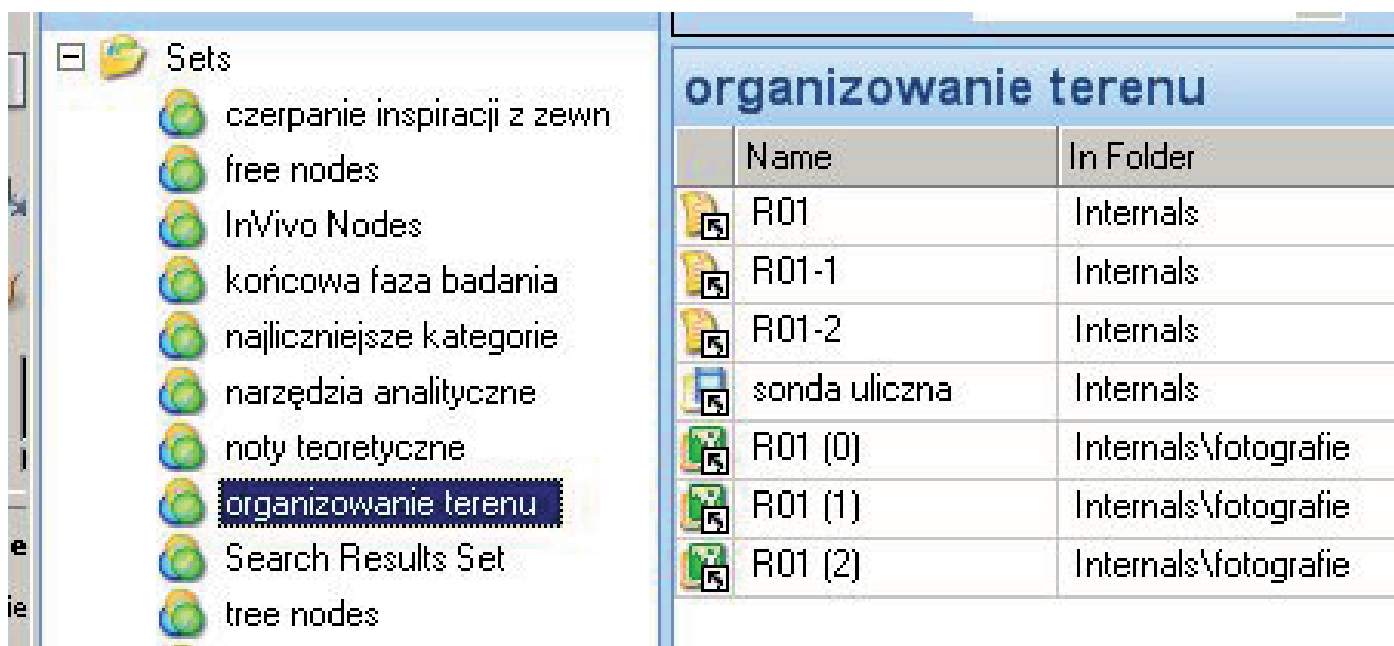
Operacją pozwalającą utrzymać porządek w opracowywanych i wytwarzanych materiałach (dokumenty źródłowe, kody i kategorie,

noty, atrybuty, odnośniki i in.) jest ich grupowanie. Posługiwanie się zestawami danych (*Collections, Sets*) pozwala na szybki dostęp do aktualnie opracowywanych materiałów lub wyselekcjonowanie tych charakteryzujących się przypisanym im walorem. Idea zbiorów jest przykładem realizacji postulatu obecnego w metodologii teorii ugruntowanej, wyrażanego stwierdzeniem *all is data*. W istocie wszystko, co uda się umieścić w ramach projektu, może zostać pogrupowane i uporządkowane.

Wskazanie i zaznaczenie przy pomocy myszki porcji dowolnych dokumentów umożliwia utworzenie lub dołączenie ich do zbioru danych. Kryterium agregacji może być dowolne i to, w jaki sposób uporządkuje się swoje materiały (wywiady utworzone w początkowej fazie badań, grupa kategorii kodowych obejmujących określony typ zjawisk, specyficzna grupa respondencka itp.)



Rys. 4. Grupowanie danych



Rys. 5. Porządkowanie różnorodnych materiałów przy pomocy zbiorów danych (Sets)

pozostaje kwestią wyobraźni analityka. Praktyka pokazuje, iż bardzo użytecznym wykorzystaniem zbiorów jest porządkowanie danych reprezentujących różne jakości. Utworzenie zbioru, który obejmuje materiały tekstowe, zdjęcia, nagrania audio oraz filmy, jak i specyficzna grupa not służąca opisaniu analizowanego przypadku lub reprezentująca określoną fazę interpretacji, sprzyja sprawnemu poruszaniu się w zebranych i obrabianym w danym momencie materiale.

Poszczególne elementy projektu mogą być w NVivo grupowane wielokrotnie. Każde źródło, nota, relacja czy dowolny dokument roboczo utworzony na potrzeby obróbki materiału może być przypisywany do dowolnej liczby zbiorów. Analitycy pracujący z dużą liczbą dokumentów docenią możliwość wyodrębniania porcji materiału poddawanej bieżącej obróbce. Skład danego zbioru może być dynamicznie kształtowany i tym samym poddawany częstej aktualizacji.

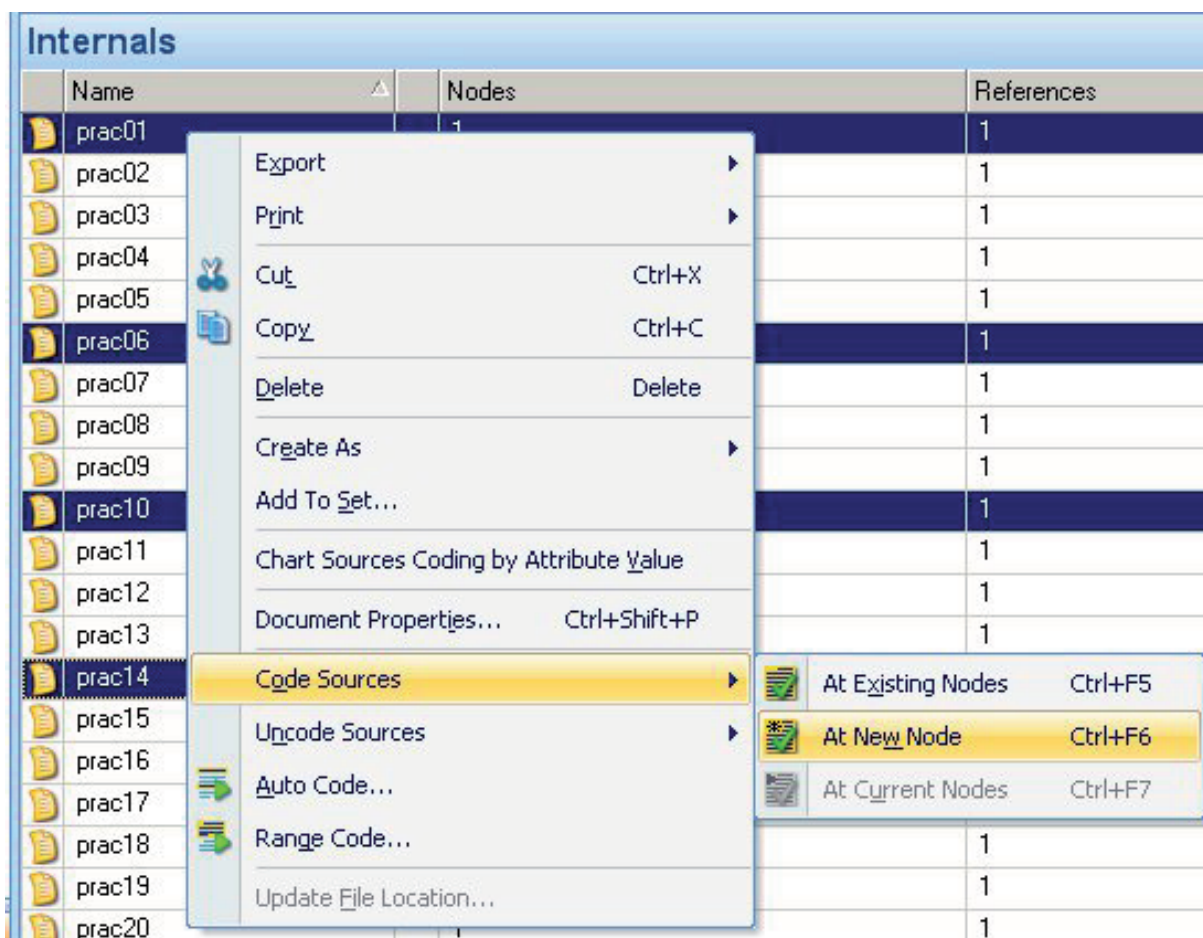
Kodowanie całych źródeł

Do poszczególnych kategorii kodowych mogą być kwalifikowane całe źródła (*Sources, Internals*) bez uwzględniania wewnętrznych zróżnicowań tematycznych. Zabieg ten stanowi rozwinięcie możliwości, jakie oferuje procedura grupowania danych. Kompleksowe kodowanie źródeł pozwala na umieszczenie w obrębie określonej kategorii kodowej, obok szczegółowo zakodowanych fragmentów tekstu, nagrania czy obrazu, całego dokumentu. Rozwiązanie to jest użyteczne w przypadku, kiedy tematyka danego dokumentu kompleksowo wkomponowuje się w obręb danej kategorii lub stanowi wartą uwzględnienia ilustrację.

Kodowanie dokumentów źródłowych

Kodowanie tekstu edytowalnego

Możliwość wielokrotnego „pokawałkowania” i klasyfikowania (Glaser, Holton 2004) poszcze-



Rys. 6. Kodowanie całych źródeł

gólnych, zupełnie drobnych oraz obejmujących nawet i kilka stron fragmentów tekstu to elementarna funkcjonalność programu typu CAQDAS, stanowiąca kluczowe narzędzie pracy analitycznej. Narzędzie to może być wykorzystywane do przeprowadzenia prostego uporządkowania zgromadzonego materiału, czyli podstawowej analizy tematycznej lub do bardziej zaawansowanych form kodowania, na przykład wiersz po wierszu lub wydarzenie po wydarzeniu (Charmaz 2009).

Czynność kodowania polega na zidentyfikowaniu treściwego fragmentu, zaznaczeniu go na czarno przy pomocy myszki, a następnie przeciągnięciu tak oznaczonego fragmentu na ikon-

kę symbolizującą odpowiednią kategorię. Tak należy postąpić w przypadku, gdy kategorie kodowe zostały uprzednio przygotowane lub wypracowane. Nie jest to jedyny sposób zakodowania fragmentu tekstu. Twórcy programu zadbali o to, aby użytkownik mógł wybrać jeden z trzech sposobów oznaczania fragmentów tekstu. Obok opisanego sposobu istnieje możliwość wywołania menu kontekstowego prawym przyciskiem myszy i wybrania opcji *Code*, a następnie jednego z wariantów *Code Selection At Existing*, *New* lub *Current Node*.

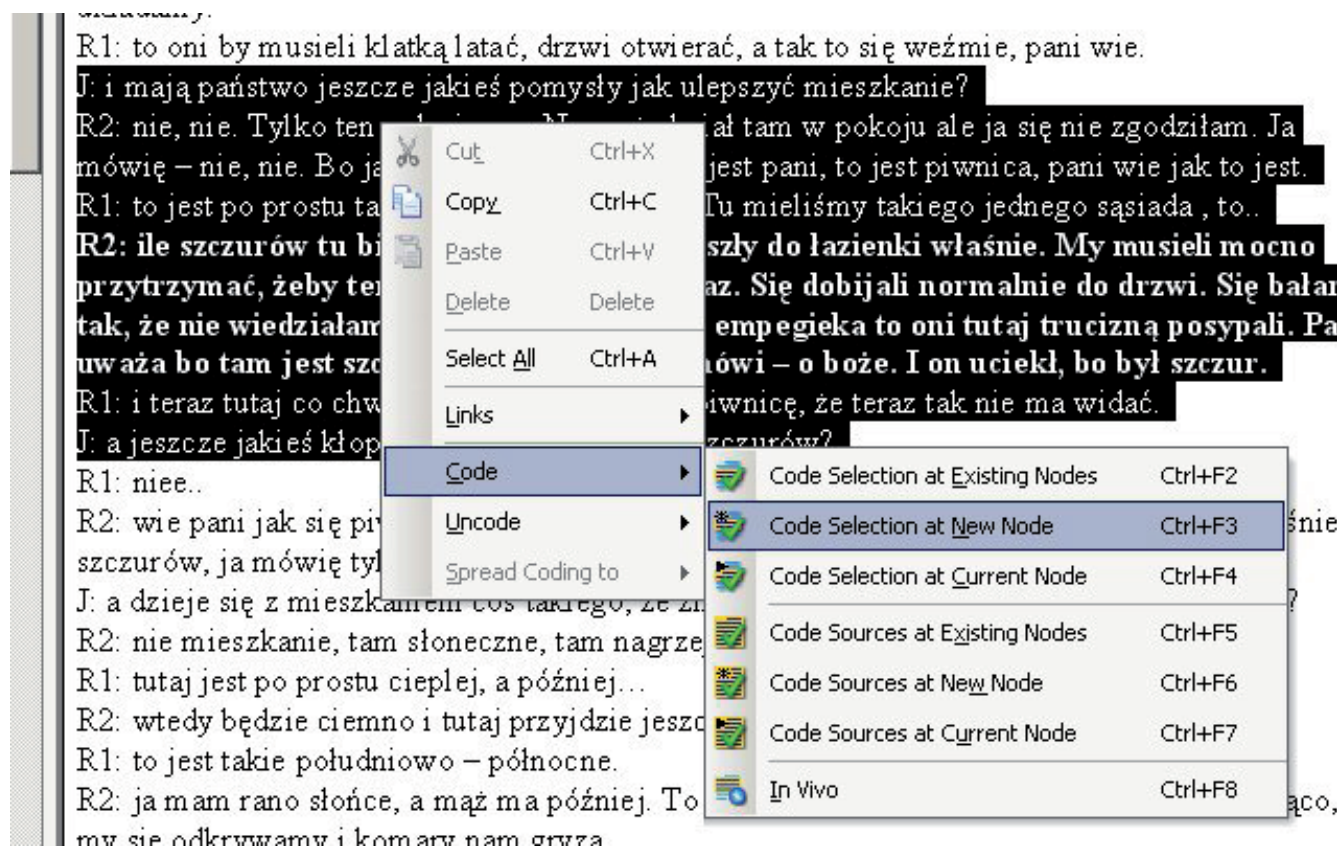
Trzeci sposób kodowania polega na posługiwaniu się skrótami klawiaturowymi. W opinii autora tekstu najszybszą metodą kodowania jest

jednak metoda *drag and drop*, czyli oparta o zasadę przeciągania przy pomocy myszy zaznaczonych fragmentów tekstu.

Tworzenie nowych kategorii kodowych w odniesieniu do oznaczonego fragmentu polega na wywołaniu przy pomocy menu kontekstowego lub skrótu klawiaturowego (Ctrl+F3) okna dialogowego, w którym należy wpisać nazwę kategorii, jej skrótowy opis oraz ewentualnie wskazać lokalizację, w której ma zostać umieszczona kategoria, co ma znaczenie, jeśli projekt jest mocno rozbudowany i ustrukturyzowany przez badacza.

Zaprezentowany wyżej model kodowania docenią badacze, którzy w swoich analizach opierają się przede wszystkim na porządkowaniu

materiału na zasadzie klasyfikacji tematycznej. Ci zaś, którzy realizują założenia metodologii teorii ugruntowanej mogą uznać posługiwanie się mechanizmem kodowania i tworzenia kategorii w ramach procedury wiersz po wierszu za nie do końca wygodne. Ten model kodowania skutkuje powstawaniem dużej liczby kodów, co może sprawiać trudności w pracy z programem. Problem ten powstanie w momencie, w którym badacz zechce przejrzeć swój materiał i podjąć się grupowania kodów. Chcąc porządkować kody oraz jednocześnie podglądać materiał, badacz będzie zmuszony do nieustannego i uporczywego klikania przyciskami myszki, otwierania zakładek, a następnie ich zamykania, co przy dużej ilości materiału może być niewygodne.



Rys. 7. Kodowanie tekstu edytowalnego

Na potrzeby kodowania wiersz po wierszu doskonale nadaje się mechanizm tworzenia przypisów (*Annotations*). Wywołanie tej opcji polega na zaznaczeniu danego wiersza, a następnie kliknięciu przycisku *New Annotation* z podręcznego menu *Links*.

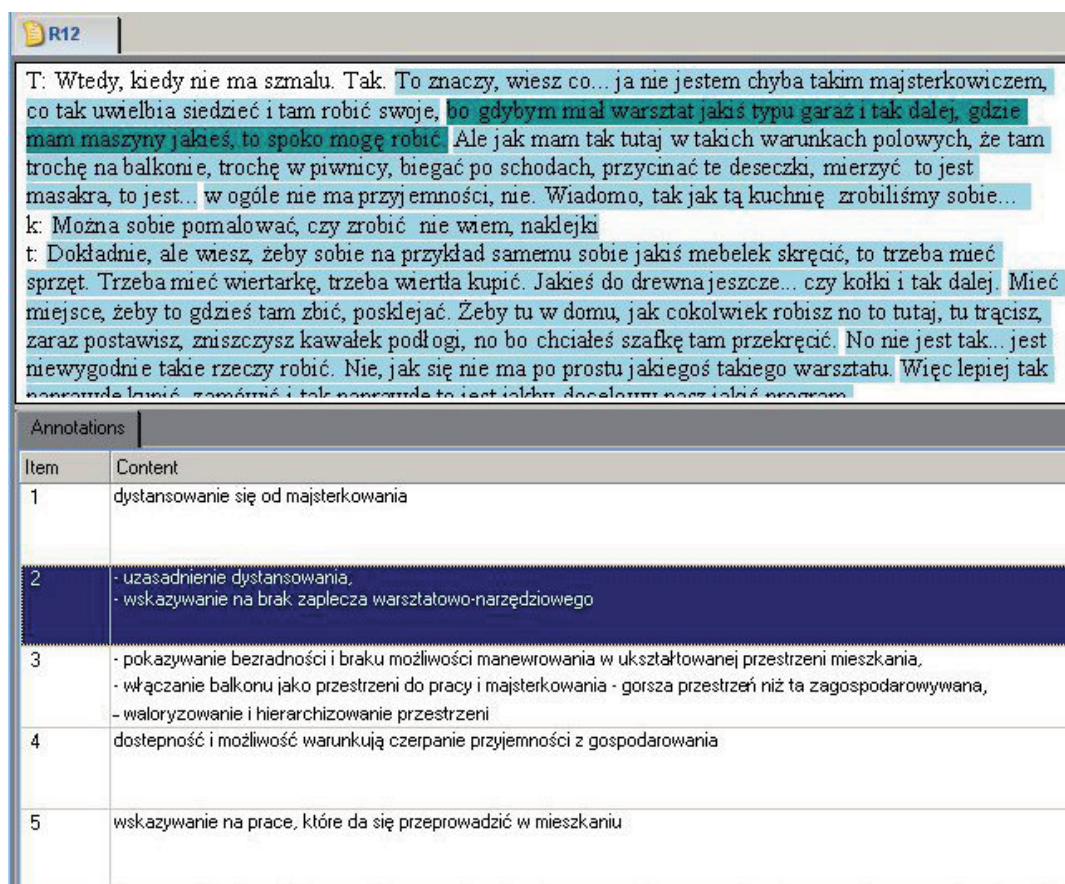
Okno, w którym wyświetlony jest kodowany dokument zostaje podzielone poziomo. Poniżej linii podziału badacz ma możliwość wpisywania treści kodów odpowiadających poszczególnym wierszom. Podczas lektury zakodowanego dokumentu można podświetlać fragmenty tekstu źródłowego poprzez kliknięcie odpowiedniego kodu – przypisu.

Kodowanie wiersz po wierszu należy przeprowadzać na dokumentach źródłowych, o ile to

możliwe, przed uporządkowaniem tematycznym. Zasada ta, zgodna z duchem metodologii teorii ugruntowanej, pozwoli zakodować wszystkie wiersze, nawet te, które podczas analizy tematycznej mogą zostać pominięte. Dodatkowo taki model pracy zdyscyplinuje analityka.

Kodowanie zdjęć

Zainteresowanie badaczy materiałami wizualnymi zaowocowało wprowadzeniem (od wersji 8) możliwości kodowania dokumentów w postaci plików graficznych. W miejsce wykorzystywanych dotychczas transkrypcji – arkuszy kodowych stanowiących reprezentacje danych wizualnych – badacz może potraktować za importowaną do programu fotografię cyfrową niemal jak edytowalny dokument tekstowy.



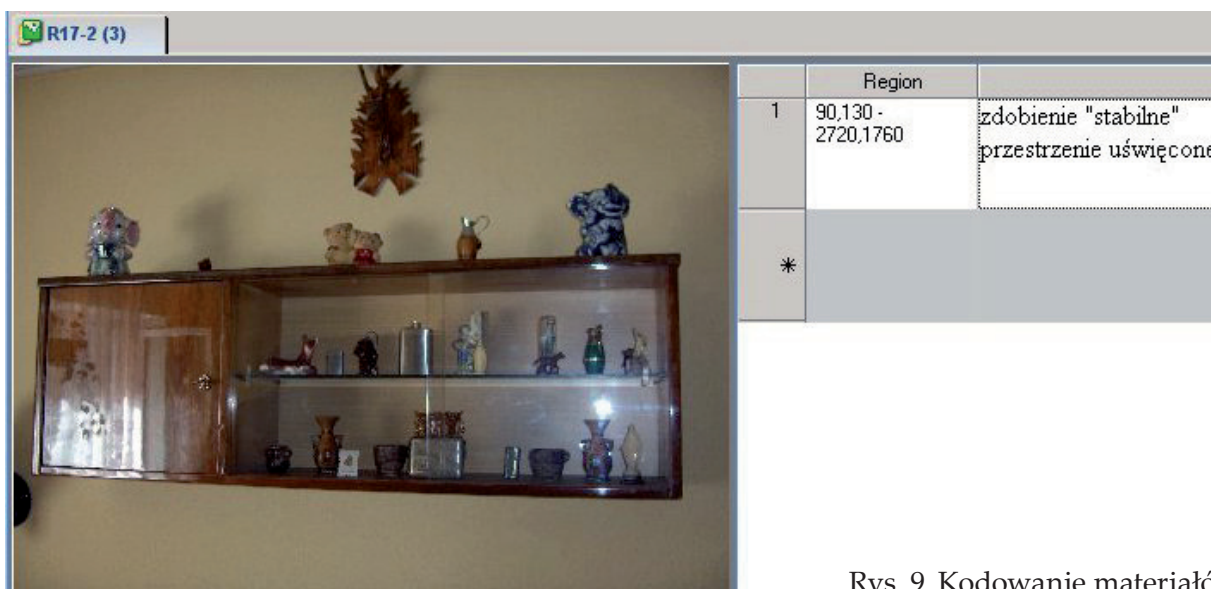
Rys. 8. Kodowanie wiersz po wierszu z wykorzystaniem przypisów

Funkcjonalność programu polega w tym miejscu, podobnie jak w przypadku materiałów tekstowych, na zaznaczeniu treściwego fragmentu, w tym przypadku określonego pola powierzchni zdjęcia i przypisaniu mu odpowiedniej treści kodowej. Została tu zastosowana zasada analogiczna do opisanej wyżej metody klasyfikowania tematycznie zróżnicowanych treści.

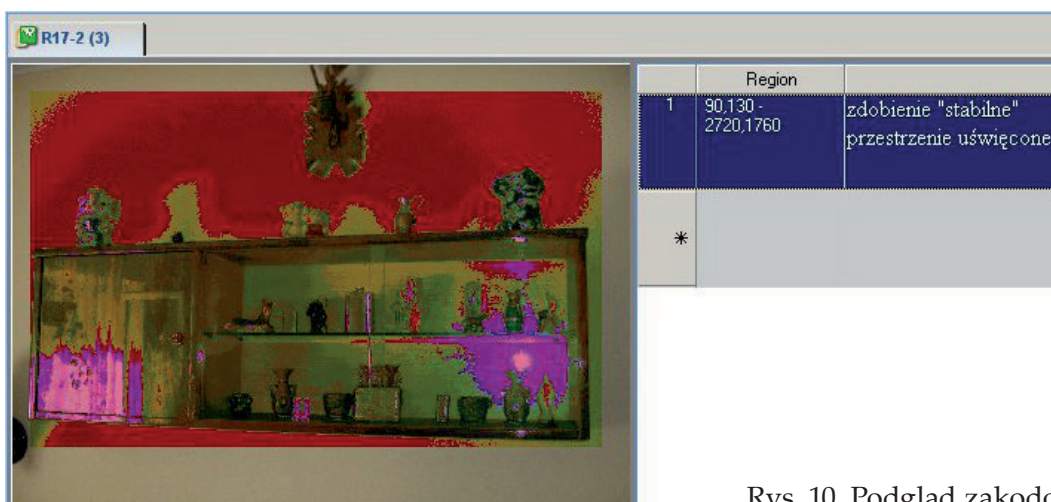
Przy pomocy myszki należy zaznaczyć kodowany fragment – tak jakby był to tekst, a następnie przy pomocy kombinacji klawiszy (Ctrl+Insert)

lub wskazując odpowiednią pozycję menu kontekstowego, wywołać możliwość wpisania treści kodu czy komentarza. Stworzona w ten sposób adnotacja zostanie trwale powiązana z zaznaczonym wcześniej fragmentem zdjęcia.

Przeglądając zgromadzone materiały wizualne, można dokonywać podglądu zakodowanych obszarów. Wskazanie przy pomocy myszki wybranej adnotacji będzie skutkowało zaznaczeniem odpowiadającego jej zakodowanego fragmentu. Przygotowane w ten sposób adnotacje mogą zo-



Rys. 9. Kodowanie materiałów wizualnych



Rys. 10. Podgląd zakodowanego fragmentu zdjęcia

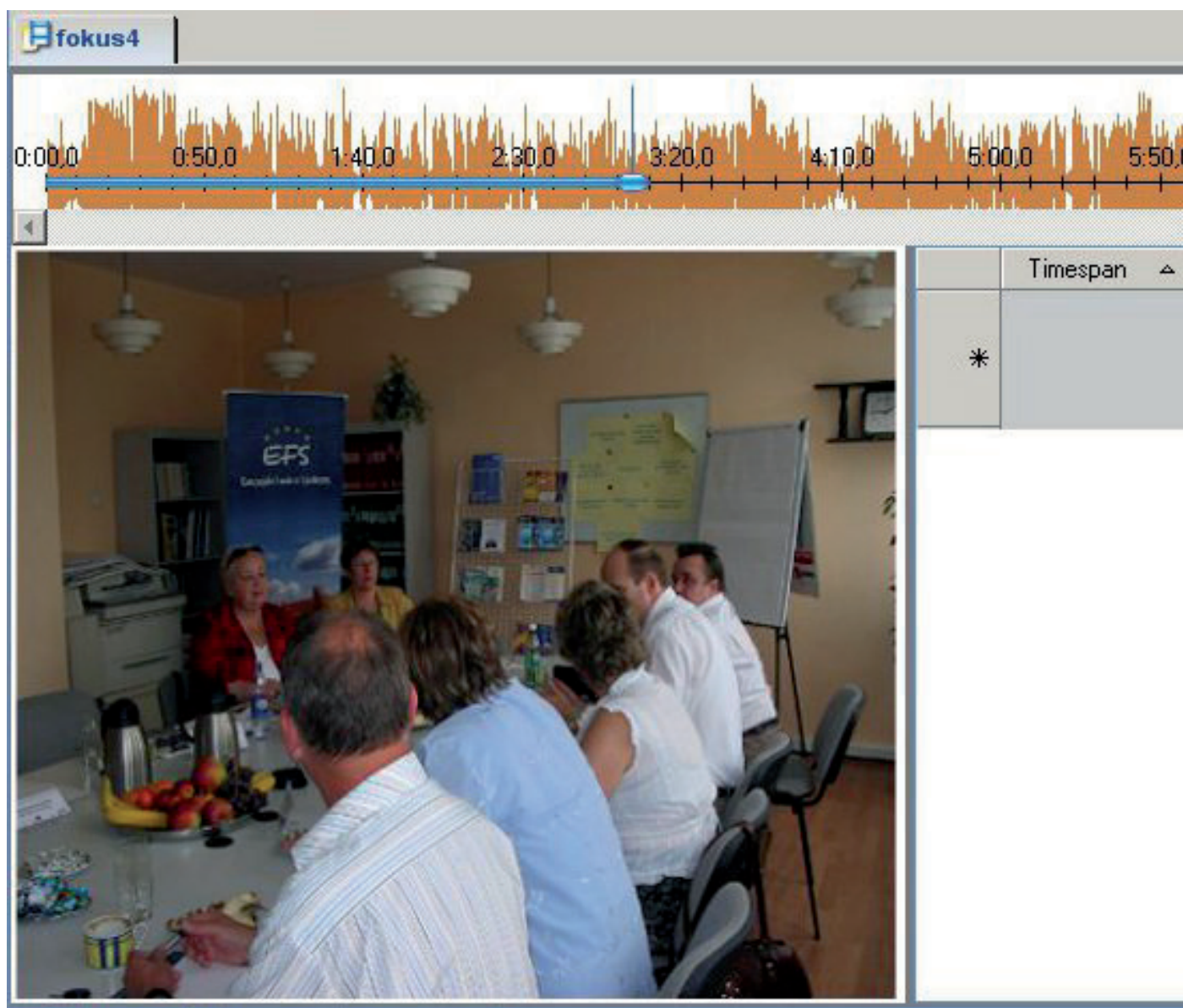
stać, podobnie jak porcje tekstu, przypisane do określonych kategorii kodowych i pełnić w ogólnym zbiorze danych równorzędną funkcję jak zakodowane fragmenty materiałów tekstowych.

Kodowanie materiałów audiowizualnych

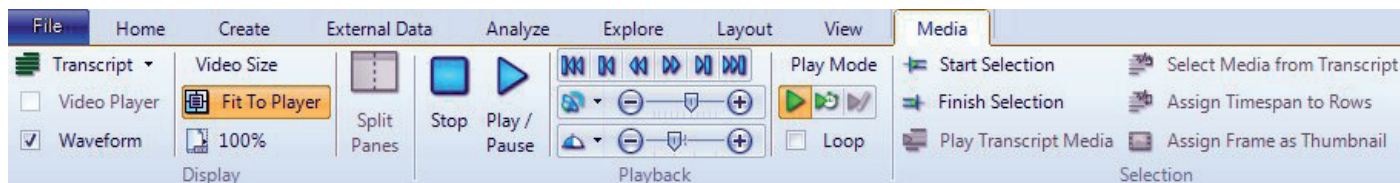
Możliwość kodowania materiałów nieedytowalnych została w NVivo (od wersji 8) rozszerzona na materiały multimedialne – nagrania audio i wideo. Sposób opracowywania tego typu źró-

deł zostanie zaprezentowany łącznie, bowiem różnice wyrażają się jedynie w nieco innym interfejsie dostępnym podczas pracy z dźwiękiem (widać jedynie falę ilustrującą widmo dźwięku), a innym podczas pracy z materiałem filmowym (oprócz widma dźwięku widać okienko z poruszającym się obrazem).

Podobnie jak w przypadku zdjęć w programie zastosowano narzędzia kodowania analogiczne do wykorzystywanych w pracy z tekstem. Meto-



Rys. 11. Materiały audiowizualne w NVivo



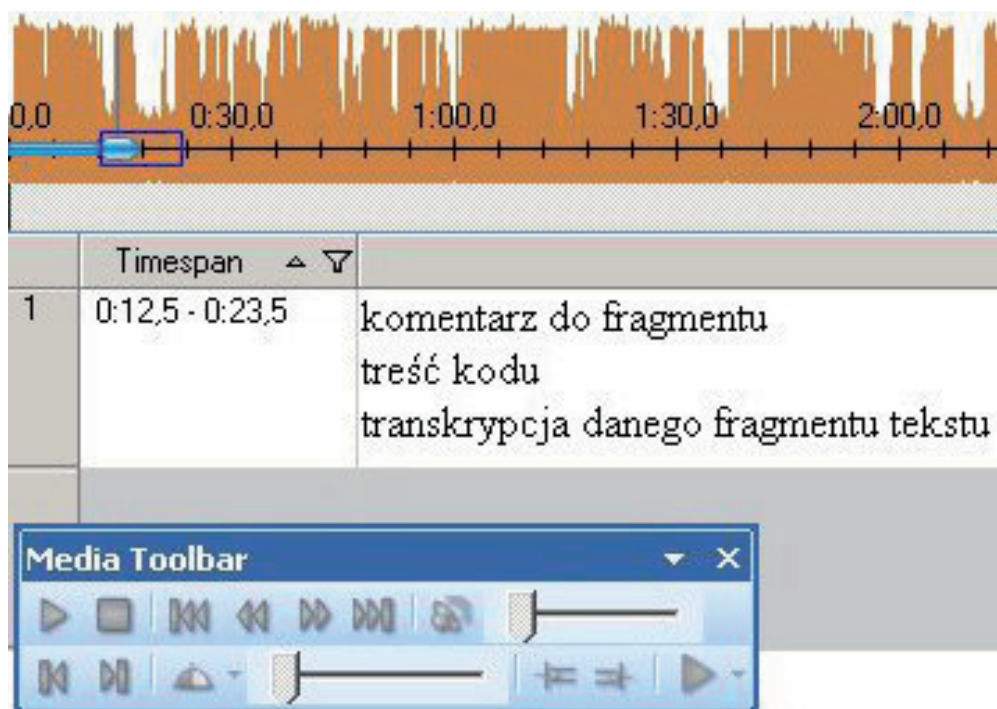
Rys. 12. Pasek narzędzi obróbki materiałów audiowizualnych

dy te nie mogą być w pełni wykorzystane w odniesieniu do materiałów multimedialnych – ze względu na ich dynamiczny charakter. Poddając obróbce materiały statyczne, takie jak teksty czy zdjęcia, można posługiwać się metodą zaznaczania i przypisywania wybranemu w ten sposób elementowi pożądanym treści. W przypadku materiałów dynamicznych zaznaczenie danego fragmentu nie może w prosty sposób polegać na wskazaniu pewnego obszaru. Na potrzeby obróbki źródeł audiowizualnych analityk dysponuje specjalnym paskiem narzędzi, który jest wyposażony w typowe przyciski pozwalające kontrolować proces odsłuchiwania, przewijania,

ustawiania prędkości odtwarzania oraz oznaczania wybranych fragmentów.

W czasie odsłuchiwania czy przeglądania materiału należy wskazać początek zaznaczenia (*Start Selection* lub klawisz F11) i koniec (*Finish Selection* lub klawisz F12). Wyodrębnionemu w ten sposób fragmentowi – podobnie jak we wskazanym wyżej przykładzie ze zdjęciem – można przypisać odpowiednie treści, wykorzystując polecenie *Insert Row* (Ctrl+Insert).

Opatrzony adnotacją fragment może zostać z powodzeniem zakodowany obok analogicznych porcji tekstu czy fragmentów zdjęć. Mechanizm komentowania fragmentów nagrań może zostać



Rys. 13. Przypisywanie fragmentom nagrań treści kodo-
wych

wykorzystany do transkrypcji materiału. Program został wyposażony w moduł wspomagający przepisywanie, do pewnego stopnia automatyzujący proces przełączania się pomiędzy funkcjami odsłuchiwania i wpisywania tekstu.

Przygotowywanie not

Pisanie not teoretycznych i metodologicznych (*memos*) towarzyszy analitykowi niemal od początku procesu badawczego. W programie wyznaczono odpowiednią przestrzeń, w której przechowywane są opracowywane noty. Program NVivo jest wyposażony w edytor tekstu oferujący podstawową funkcjonalność i wystarczający na potrzeby artykułowania własnych pomysłów i spostrzeżeń. Opracowane noty mogą być eksportowane do standardowych plików tekstowych (.rtf, .doc, .docx).

Kluczowym udogodnieniem jest możliwość podpinania, linkowania do wybranych fragmentów noty cytatów ze zgromadzonego materiału badawczego. Oprócz typowych odniesień

do ilustracji tekstowych badacz ma możliwość podłączania do swoich not fragmentów multimedialnych. W starszych wersjach programu (od 2 do 8) tworzenie odnośników pomiędzy notami a cytataми było czynnością niezwykle niewygodną i czasochłonną, wymagającą kilkakrotnego otwierania dokumentów oraz poszukiwania wybranego fragmentu. Zadanie ulegało dodatkowej komplikacji, kiedy cytata miał pochodzić nie tyle z konkretnego źródła – na przykład transkrypcji – ale z *listingu* kategorii kodowej, zawierającej informacje na dany temat z licznych źródeł. Najnowsza wersja programu (9) zawiera rozwiązanie tego problemu⁷. W programie należy otworzyć w osobnych zakładkach (dzieje się to automatycznie) dokument lub kategorię, z którego mają pochodzić podłączane cytaty oraz notę, którą chcemy wzbogacić ilustracją (rys. 14).

⁷ Na przestrzeni kilku lat korzystania z programu autor opracowania wielokrotnie zwracał się do twórców programu z problemem podłączania treści pozostających w obszarze kategorii do fragmentów noty, proponując możliwe rozwiązanie. Zostało ono wprowadzone w bardzo wygodnej formie w dziewiątej, najnowszej wersji programu.

The screenshot shows the NVivo interface. At the top is a table titled "Memos" with columns: Name, Nodes, References, Created On, and Created By. The table lists several memos, with "idea projektu" selected. Below the table, the content of the selected memo is displayed, including a date and time stamp "2010-11-08 23:22 (R14)" and a highlighted text fragment: "Treści materiałów źródłowych mogą być swobodnie 'podpinane' jako cytaty lub ilustracje (See Also Links) pod określone fragmenty not."

Name	Nodes	References	Created On	Created By
biografia mieszkania	0	0	2010-02-26 09:30	MB
dziennik metodologiczny	0	0	2006-04-01 22:23	MB
gotowość do zmian	0	0	2010-02-25 11:49	MB
idea projektu	0	0	2010-02-24 23:28	MB
ideal - przedmiot praktyczny, oszczędn	0	0	2010-02-27 00:28	MB
konteksty teoretyczne	0	0	2007-04-18 13:04	MB

2010-11-08 23:22 (R14)

Treści materiałów źródłowych mogą być swobodnie "podpinane" jako cytaty lub ilustracje (See Also Links) pod określone fragmenty not.

Rys. 14. Podłączanie cytatu pod fragment noty

W celu podpięcia cytatu pod treść noty należy zaznaczyć i skopiować wybrany cytat w dokumencie źródłowym (lub zestawieniu zakodowanych fragmentów w ramach danej kategorii), następnie w treści noty należy dokładnie wskazać – zaznaczając myszką – miejsce, do którego ma zostać podpięty. W powstałe zaznaczenie należy kliknąć prawym przyciskiem myszy i wybrać polecenie *Paste As See Also Link*. W tekście noty pojawi się podświetlenie w kolorze różowym wskazujące, iż w danym miejscu zostały ulokowane odniesienia do innych dokumentów. Powiązania tego typu można tworzyć między wszystkimi dokumentami, danymi i notami, które funkcjonują w obrębie projektu NVivo. Podczas eksportu noty do postaci pliku tekstowego program wyświetla monit z zapytaniem, jakiego rodzaju odnośniki mają zostać wyeksportowane razem z tekstem noty.

Wskazana wyżej procedura opracowywania not w powiązaniu z fragmentami ilustracyjnymi znajduje zastosowanie zarówno w przypadku projektów badawczych opartych o zasady metodologii teorii ugruntowanej, jak i innych, zakładających odmienne strategie wykorzystywania zgromadzonego materiału.

Elementy automatyzacji procesu porządkowania danych

Pakiet NVivo został wyposażony w narzędzia wspomagające (ale nie zastępujące!) proces analizy danych. Narzędzia te dotyczą procesu porządkowania i sortowania danych (autokodowanie) oraz tworzenia kategorii na podstawie wyszukiwania słów kluczowych. Kodowanie danych nie jest procesem jednorodnym. Propozycja autorów programu, aby czynności automa-

tycznie wykonywane przez program kojarzyć z kodowaniem może budzić uzasadnione wątpliwości. Kodowanie będące elementem procedur metodologii teorii ugruntowanej stanowi bowiem zaprzeczenie automatyzmu. Automataczne rozwiązania zaimplementowane w NVivo mogą być traktowane raczej jako narzędzia porządkowania lub sortowania danych. Nie należy ich stawiać na równi z identyfikowaniem znaczeń. W tym miejscu ujawnia się ograniczenie oprogramowania, które jednocześnie wskazuje na właściwą rolę analityka danych (Lofland, Snow, Anderson, Lofland 2009).

Autokodowanie

Materiał gromadzony w czasie badań socjologicznych często może zostać uporządkowany w formie komputerowych plików tekstowych z ustaloną strukturą formalną. Przykładem mogą być tu zapisy wywiadów jakościowych opartych o jednolity scenariusz. Aby przeprowadzić procedurę autokodowania, zapisy takich wywiadów powinny zawierać jednakowo sformułowane, uporządkowane pytania lub tytuły odpowiednich fragmentów oraz treści wypowiedzi znajdujące się pomiędzy owymi pytaniami lub tytułami. Wykorzystując style formatowania akapitów dostępne w popularnych edytorach tekstu (nagłówek, tekst podstawowy itp.), należy uporządkować dane pochodzące z wywiadów. Elementem wspólnym wszystkim zapisom (treść pytań, tytuły wątków) należy nadać status nagłówka (lub innego stylu nadrzędnego względem tekstu podstawowego). Tak przygotowane materiały poddane procedurze autokodowania zostaną posortowane do osobnych kategorii. Nazwy kategorii będą pochodziły od treści, jakie zawierały wspólne elementy zapisów. W pro-

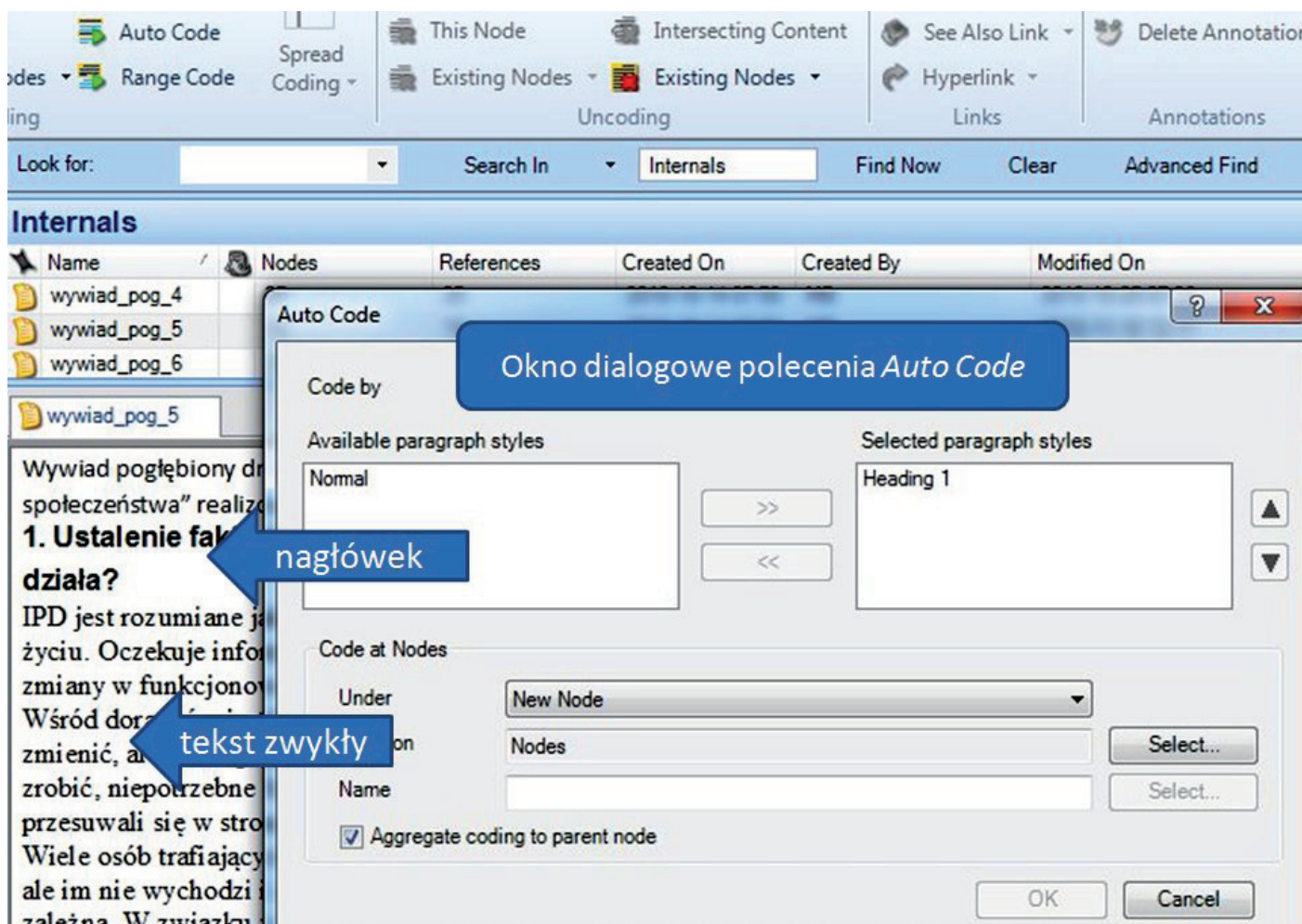
gramie należy wybrać polecenie *Auto Code*, a następnie wskazać, które dokumenty mają zostać posortowane oraz jaki styl ma zostać potraktowany jako wyróżnik tematycznie odmiennych treści. Uzyskany w ten sposób zbiór danych będzie użyteczny na potrzeby jakościowej analizy poszczególnych zagadnień poruszanych w ramach wywiadów.

Powyższa ilustracja przedstawia formalne zróżnicowanie zawartości zapisu wywiadu na większe nagłówkowe (np. pytania lub poruszone zagadnienia) oraz odpowiedzi (tekst zwykły, normalny). Wybierając opcję autokodowania, zadaniem użytkownika programu jest wskazanie

w oknie dialogowym, który ze stylów akapitów ma zostać potraktowany jako wyróżnik poszczególnych wątków do zakodowania.

Tworzenie kategorii na podstawie wyszukiwania słów kluczowych

Nie wszystkie analizy jakościowe wymagają gruntownego przeszukiwania materiału. Zdarzają się sytuacje, w których obróbka danych może sprowadzać się do poszukiwania określonych słów wraz z kontekstem ich występowania. Procedura taka może stanowić analizę samą w sobie lub też sekwencję bardziej złożonego przedsięwzięcia. Zaproponowany w następnych

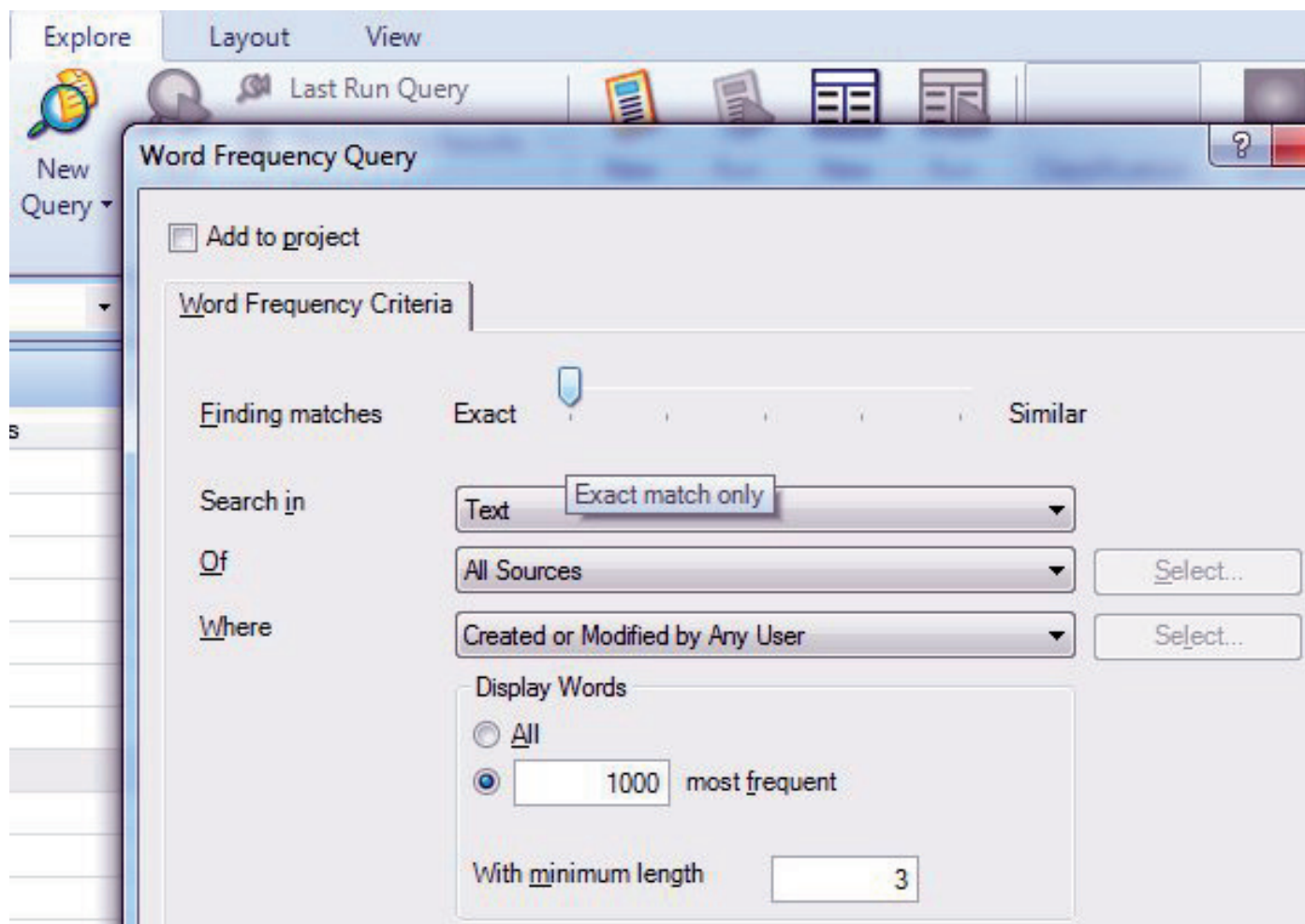


Rys. 15. Autokodowanie

akapitach model wykorzystania NVivo będzie obejmował następujące elementy: (1) utworzenie kwerendy ukierunkowanej na zliczenie słów występujących w dokumentach źródłowych, (2) wybranie słowa lub słów kluczowych, (3) przeprowadzenie wyszukiwania słów kluczowych wraz z kontekstem ich występowania, (4) automatyczne zakwalifikowanie rezultatów wyszukiwania jako kategorii.

Pierwszym etapem jest przegląd słów ze wszystkich źródeł lub z wybranej na podstawie założonych przez badacza kryteriów wyszukiwania grupy dokumentów (np. na podstawie kategorii, atrybutów lub dowolnej kombinacji informacji,

w które analityk wyposażył swoje dane). Zakładając, że dane nie zostały w żaden sposób uporządkowane, można przeprowadzić kwerendę wszystkich zgromadzonych źródeł. W odpowiednim oknie dialogowym, wywołanym przy pomocy polecenia *New Query*, a następnie *Word Frequency Query*, należy wskazać, które źródła mają zostać poddane analizie, jeśli w procesie gromadzenia danych (np. przeprowadzanie wywiadów) brała udział więcej niż jedna osoba, można wskazać wszystkich lub konkretnego badacza, którego opracowanie zostanie poddane zliczeniu słów oraz ograniczyć minimalną liczbę liter, jaką mają mieć zliczane wyrazy.



Rys. 16. Kwerenda ukierunkowana na zliczenie słów

Rezultatem tej operacji jest tabela przedstawiająca częstości występowania poszczególnych słów. Obok tej podstawowej opcji program został wyposażony w inne sposoby prezentacji wyników. Formy takie, jak chmura słów kluczowych (*Tag Clouds*) czy mapa pojęciowa lub dendrogram (*Cluster Analysis*) stanowią graficzne odzwierciedlenie częstotliwości występowania poszczególnych wyrażen i mogą stać się źródłem interesujących wniosków. Zamieszczone niżej ilustracje 17 i 18 przedstawiają wyniki zliczenia słów w różnych formach.

Oczywistym jest w tym miejscu, iż ilościowe zestawienie słów oparte o związki korelacyjne występujące między nimi ma ograniczoną wartość dla analizy jakościowej, jeśli jednak będzie źródłem interesujących spostrzeżeń, należy

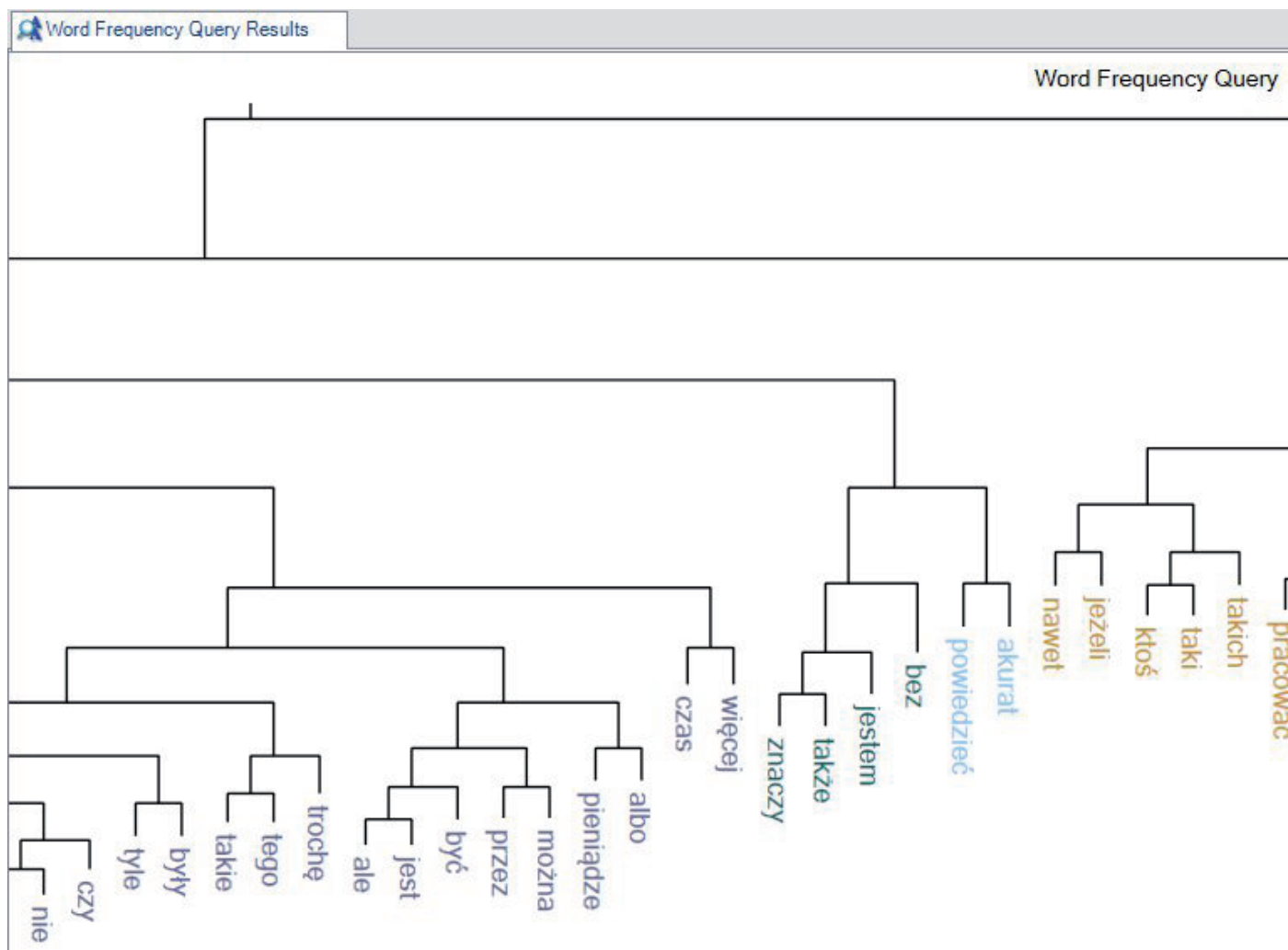
uwzględnić je jako element procedury poznawania danych.

Kolejnym etapem eksploracji danych będzie przeprowadzenie wyszukiwania fragmentów danych, które zawierają określoną sekwencję słów kluczowych. Słowa te oraz związki między nimi są dla badacza czytelne dzięki uprzedniej analizie graficznych prezentacji. Jakie możliwości wyszukiwania ma do dyspozycji analityk? Zakładając, że w badaniu praktyk unikania podejmowania pracy przez osoby bezrobotne⁸ pojawiają się kluczowe słowa „nie”, „chęć”, „pracy” i „pracować”, można przeprowadzić analizę po-

⁸ Dane pochodzą z badania przeprowadzonego przez Pracownię Realizacji Badań Socjologicznych Uniwersytetu Gdańskiego w 2007 roku, dotyczącego niepodejmowania pracy przez osoby bezrobotne w powiecie malborskim województwa pomorskiego.



Rys. 17. Przykład prezentacji wyników kwerendy z zastosowaniem wariantu *Tag Clouds*



Rys. 18. Przykład prezentacji wyników kwerendy z zastosowaniem wariantu *Cluster Analysis*

równawczą wypowiedzi zawierających stwierdzenia „chcę pracować” oraz „nie chcę pracować”. Porównanie wybranych zbiorów wypowiedzi pozwoli na uchwycenie wartościowych informacji. Celem artykułu jest prezentacja procedury, dlatego też nawiązanie do przedsięwzięcia badawczego pełni funkcje ilustracyjne.

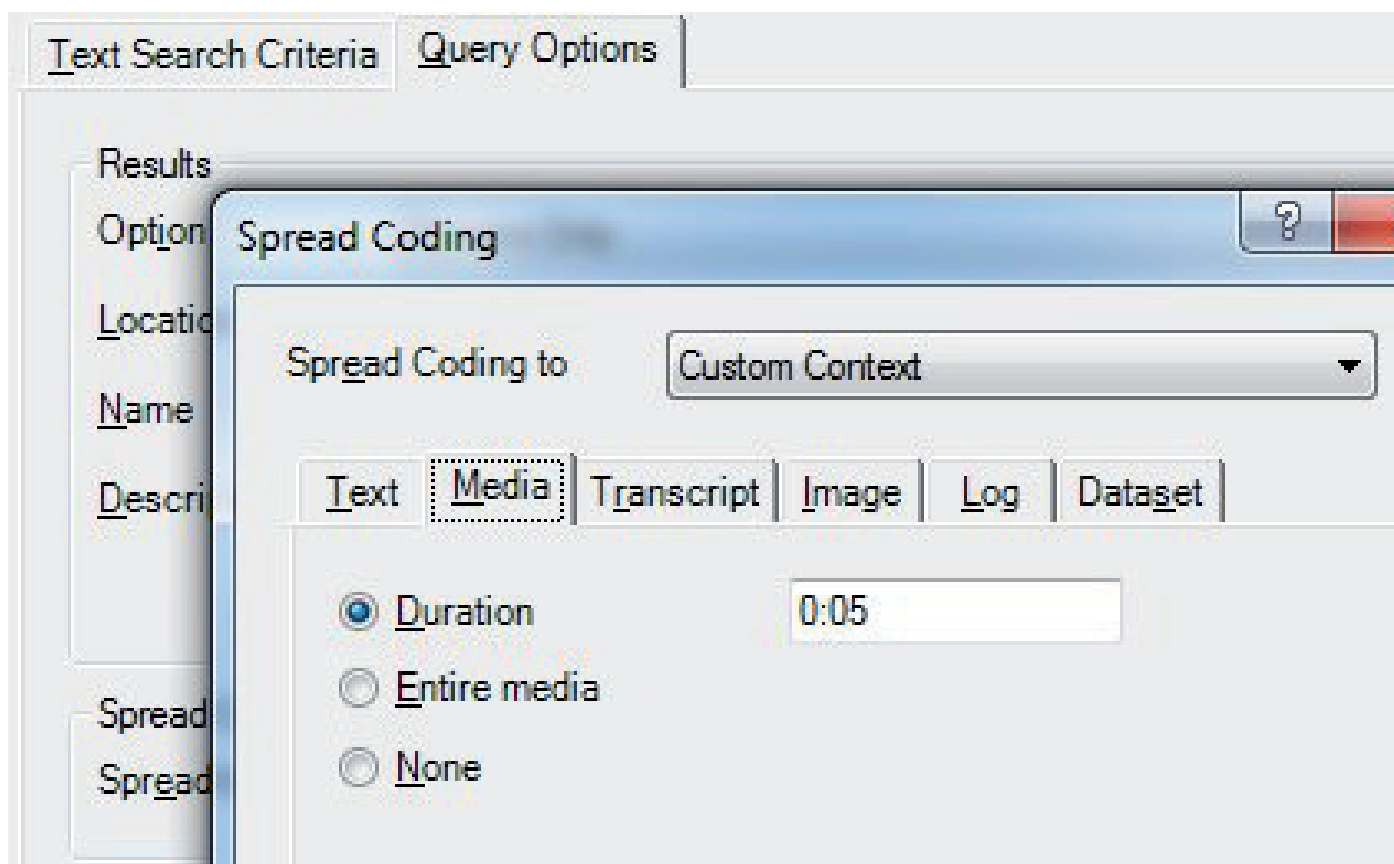
Wywołując okno dialogowe polecenia *New Query*, a następnie *Text Search*, analityk musi być przygotowany do wskazania istotnych opcji wyszukiwania. Przede wszystkim należy w odpowiednim miejscu wpisać poszukiwane słowa lub wyrażenie oraz zaznaczyć, jak program ma

je traktować (szukać dokładnie takiego wyrażenia, pomijać wskazane słowa, szukać łącznie, w przybliżeniu itp.). Następnie należy wskazać zakres materiału, który ma być objęty przeszukiwaniem (określona grupa badanych, wyróżnienie materiałów na podstawie etapu badań czy określone cechy przypisane badanym jako atrybuty – płeć, wiek, wykształcenie). Przeszukiwanie może być też ukierunkowane na analizę zawartości poszczególnych kategorii, które badacz wyodrębnił podczas kodowania. Kolejne decyzje dotyczyć będą specyfiki kontekstu występowania poszukiwanego wyrażenia, który program powinien zaprezentować łącznie z po-

szukiwanym tekstem. Do podstawowych opcji należą: kontekst bezpośredni, rozległy lub też całe źródło, w którym występuje dane stwierdzenie. Jednakże najciekawsze możliwości oferuje wariant poszukiwań odnoszący się do kontekstu sprecyzowanego wedle upodobań i pomysłu analityka (*Custom Context*).

W zależności od rodzaju danych, jakimi dysponuje badacz (teksty, zdjęcia, materiały audio-wizualne, arkusze danych) oraz sposobu ich opracowania istnieje możliwość precyzyjnego wskazania, jaki zakres danych tworzących tło występowania poszukiwanego elementu ma zostać dołączony do rezultatów. Warto rozpatrzyć kilka przypadków. W odniesieniu do materiałów tekstowych będzie to konkretna liczba słów towarzyszących, otaczający wyrażenie akapit

lub styl akapitu oraz cały dokument. W odniesieniu do materiałów audiowizualnych, przy założeniu, że została dokonana choćby jego częściowa transkrypcja, a poszczególne fragmenty wypowiedzi są podlinkowane do nagrania, badacz może wskazać długość nagrania, jaka ma zostać dołączona do rezultatów wyszukiwania. Program dokona przeszukania transkrypcji oraz przedstawi odnośniki do odpowiednio przygotowanego fragmentu nagrania. W przypadku zdjęć wykorzystywanych w badaniach miarą służącą określeniu, jaką porcję obrazu program ma dołączyć do wyników, będzie wskazanie procentowe – jaki fragment (procent powierzchni) zdjęcia, oprócz poszukiwanego fragmentu (opisanego czy zakodowanego), ma zostać dołączony do wyniku wyszukiwania.



Rys. 19. Precyzowanie sposobu wyszukiwania słów kluczowych

Zakładając, że w toku badania badacz zgromadził ogromne zasoby materiałów wizualnych oraz że zostały one rzetelnie zakodowane i opisane, istnieje możliwość ich precyzyjnego przeszukiwania pod określonym kątem.

Zaznaczając w oknie dialogowym opcję *Add to Project*, rezultat wyszukiwania zostanie dołączony do zgromadzonego materiału badawczego i będzie mógł być traktowany jako jedno ze źródeł dalszych poszukiwań lub wręcz jako samodzielna kategoria. Sposób wykorzystania zależy oczywiście od pomysłowości badacza oraz nośności wypracowanego materiału.

Podsumowanie

Zaproponowany wyżej model konstruowania i rekonstruowania danych w toku procedur analitycznych nie wyczerpuje możliwości, jakie oferuje najnowsze oprogramowanie typu CA-QDAS. Szczegółowe funkcje oraz ukryte możliwości stają się dostępne podczas prowadzonych analiz. Inną okazją do poznawania nowych wymiarów funkcjonalności jest wprowadzanie do analizowanego zbioru danych nowych, formalnie zróżnicowanych materiałów. Powstający w ten sposób jakościowy kolaż wymagał będzie od analityka nie lada pomysłowości. Prezentując własne doświadczenia terenowe, autor ma świadomość, iż opisał jedynie ich część. Fakt ten wynika przede wszystkim z formalnych ograniczeń, jakie narzuca forma artykułu, co stanowi swoistą egzemplifikację problemu stratnej kompresji danych jakościowych, prezentowanego w niniejszym opracowaniu. Zadaniem różnorodności form i modeli procedur kodowania jest minimalizowanie tego zjawiska.

W artykule nie omówiono mechanizmów tworzenia modeli i diagramów oraz baz danych, czytelnych dla pakietów wykorzystywanych w analizach ilościowych (np. SPSS, Statistica), na podstawie zestawień atrybutów dokumentów źródłowych (np. informacje o respondentach czy przypadkach) oraz częstości występowania poszczególnych kategorii kodowych.

Procedury te nie wykraczają poza standardową obsługę typowych biurowych programów użytkowych. Posługiwanie się modulem *Models* pozwala na wizualizowanie sieci powiązań pomiędzy kodami, dokumentami i notami. Budowanie graficznych reprezentacji formułowanych w toku badań twierdzeń jest w programie NVivo rozwiązane w sposób niezwykle przejrzysty i pozwalający na jego intuicyjne opanowanie. W tym przypadku intuicyjność w połączeniu z pomysłowością może dawać niezwykle rezultaty. Eksportowanie danych do pakietów statystycznych powinno pozostawać w ścisłym związku z planowanymi obliczeniami. Zdecydowana większość informacji dotyczących zgromadzonych przypadków jest w badaniach jakościowych rejestrowana na nominalnym lub porządkowym poziomie pomiaru. Co za tym idzie wykorzystywane parametry statystyczne powinny być stosownie dobrane. Procedura eksportu zestawień atrybutów poszczególnych dokumentów czy kategorii nie wykracza poza wybór opcji *Export data*, a następnie zapisanie ich w postaci pliku czytelnego dla arkuszy statystycznych (np. plik o rozszerzeniu .dat).

Dokonując wyboru narzędzia dostosowanego do własnych potrzeb oraz przede wszystkim wygodnego (pracy z materiałem należy poświęcić wiele godzin, a często nawet dni i tygodni),

warto przetestować kilka programów. Jedni badacze będą się czuć znakomicie, wykorzystując NVivo, inni będą preferować narzędzie Atlas.ti, dla innych potencjał oferowany przez standardowe edytory tekstu w połączeniu z arkuszami kalkulacyjnymi oraz tradycyjnymi metodami opracowywania materiału będzie wystarczający. Doświadczenie autora podpowiada, iż wybór dobrego dla siebie programu powinien opierać się na własnych doświadczeniach (przetestowanie nawet kilkunastu programów i wykorzy-

stanie kilku z nich w niewielkim projekcie badawczym – chodzi o zastosowanie różnych narzędzi do obróbki tego samego materiału) oraz na doświadczeniach osób już korzystających z tego typu narzędzi. Użytkownicy poszczególnych programów dzielą się swoimi doświadczeniami na forach internetowych poświęconych CAQDAS, tam też można zdobyć podstawowe informacje na temat możliwości narzędzi, czyli tego, co program potrafi, a czego na pewno nie.

Bibliografia

Babbie Earl (2003) *Badania społeczne w praktyce*. Przełożyli Witold Betkiewicz i in. Warszawa: Wydawnictwo Naukowe PWN.

Bieliński Jacek, Iwińska Katarzyna, Rosińska-Kordasiewicz Anna (2007) *Analiza danych jakościowych przy użyciu programów komputerowych*. „ASK”, nr 16, s. 89–114.

Charmaz Kathy (2009) *Teoria ugruntowana. Praktyczny przewodnik po analizie jakościowej*. Przełożyła Barbara Komorowska. Warszawa: Wydawnictwo Naukowe PWN.

Coffey Amanda, Holbrook Beverley, Atkinson Paul (1999) *Qualitative Data Analysis: Technologies and Representations* [w:] Alan Bryman, Robert G. Burgess, eds., *Qualitative Research*, vol. III. London, Thousand Oaks, New Delhi: Sage Publications, s. 165–182.

Gibbs Graham (2011) *Analizowanie danych jakościowych*. Przełożyła Maja Brzozowska-Brywczyńska. Warszawa: Wydawnictwo Naukowe PWN.

Glaser Barney B., Holton Judith (2004) *Przemodelowanie teorii ugruntowanej*. „Przegląd Socjologii Jakościowej”, t. 6, nr 2 [dostęp dzień października 2010]. Dostępny w Internecie: http://www.qualitativesociologyreview.org/PL/Volume13/PSJ_6_2_Glaser_Holton.pdf.

Kaufmann Jean-Claude (2010) *Wywiad rozumiejący*. Przełożyła Alina Kapciak. Warszawa: Oficyna Naukowa.

Lofland John, Snow David A., Anderson Leon, Lofland, Lyn H. (2009) *Analiza układów społecznych. Przewodnik metodologiczny po badaniach jakościowych*. Przełożyły Anna Kordasiewicz, Sylwia Urbańska, Monika Żychlińska. Warszawa: Wydawnictwo Naukowe Scholar.

Lonkila Markku (1995) *Grounded theory as an emerging paradigm for computer-assisted qualitative data analysis* [w:] Udo Kelle, ed., *Computer-Aided Qualitative Data Analysis*. London: Sage, s. 41–51.

Miles Matthew B., Huberman Michael (2000) *Analiza danych jakościowych*. Przełożył Stanisław Zabielski. Białystok: Transhumana.

Riessman Catherine K. (1993) *Narrative Analysis*. Newbury Park: Sage Publications.

Seale Clive (2008) *Wykorzystanie komputera w analizie danych jakościowych* [w:] David Silverman, red., *Prowadzenie badań jakościowych*. Przełożyła Joanna

Ostrowska. Warszawa: Wydawnictwo Naukowe PWN, s. 233–256.

Trutkowski Cezary (1999) *Analiza treści wspomagana komputerowo. Badanie społecznych reprezentacji polityki*. „ASK”, nr 8, s. 113–133.

Cytowanie

Brosz Maciej (2012) *Komputerowe wspomaganie badań jakościowych. Zastosowanie pakietu NVivo w analizie materiałów nieustrukturyzowanych*. „Przegląd Socjologii Jakościowej”, t. 8, nr 1, s. 98–125 [dostęp dzień, miesiąc, rok]. Dostępny w Internecie: <<http://www.przegladsocjologiijakosciowej.org>>.

Computer Aided Qualitative Research Using NVivo in unstructured data analysis

This paper concerns using NVivo software in qualitative data analysis. Main subject refers to the data reduction accompanying the process of qualitative data analysis. Using software does not necessarily cause the uncontrolled modifications of data, thereby, the loss of relevant aspects of collected data. The latest version of CAQDAS (i.e., NVivo 8, 9) enables the possibility of coding on barely altered sources. The paper presents examples of coding procedures on texts, pictures, audio-visual recordings. Additionally, the paper includes description of some techniques aiding the coding process.

Keywords: CAQDAS, loss of information, data reduction, coding procedures, NVivo, grounded theory