Jan Winczorek

Wykorzystanie oprogramowania R i RQDA w jakościowo-ilościowej analizie treści orzeczeń Trybunału Konstytucyjnego

Przegląd Socjologii Jakościowej 10/2, 138-161

2014

Jan Winczorek
Uniwersytet Warszawski

Wykorzystanie oprogramowania R i RQDA w jakościowo-ilościowej analizie treści orzeczeń Trybunału Konstytucyjnego

Abstrakt Artykuł opisuje wykorzystanie pakietu statystycznego „R" z nakładką CAQDA „RQDA" w jakościowo-ilościowej triangulowanej analizie treści z kodowaniem swobodnym, analizą rzetelności kodowania i statystycznej istotności wyników, przeprowadzonej na próbie uzasadnień orzeczeń Trybunału Konstytucyjnego z lat 1986–2009. Przedstawia najważniejsze własności oprogramowania, założenia badania, jego metodologię i zastosowane procedury analizy danych, a także pytania epistemologiczne, które pojawiły się w efekcie jego realizacji.

Słowa kluczowe CAQDA, RQDA, R (oprogramowanie), analiza treści, rzetelność kodowania, Trybunał Konstytucyjny

Niniejszy artykuł przedstawia założenia metodologiczne, doświadczenia realizacyjne i niektóre wnioski badawcze powzięte w efekcie realizacji badania „Wzorce wykładni konstytucji w orzecznictwie Trybunału Konstytucyjnego w latach 1986–2009"[1]. Badanie to zostało zrealizowane w latach 2010–2012 przez sześcioosobowy zespół badawczy pracowników Wydziału Prawa i Administracji Uniwersytetu Warszawskiego[2] i miało charakter interdyscyplinarny. Wychodziło od ogólnych

Jan Winczorek, dr, adiunkt w Katedrze Socjologii Prawa Wydziału Prawa i Administracji Uniwersytetu Warszawskiego. Zajmuje się socjologią prawa, socjologią teoretyczną i teorią prawa. Opublikował między innymi książkę poświęconą prawu w teorii Luhmanna i kilkadziesiąt innych tekstów.

Adres kontaktowy:
Wydział Prawa i Administracji
Uniwersytet Warszawski
ul. Krakowskie Przedmieście 26/28
00-926 Warszawa
e-mail: janwin@janwin.info

[1] Realizacja badania opisywanego w artykule została dofinansowane z grantu NCN N N110 121237, którego realizacją kierował dr hab. Tomasz Stawecki, prof. UW. Przedstawienie wyników zawierać będzie praca zbiorowa pod redakcją prof. Tomasza Staweckiego imoją, która zostanie opublikowana w 2014 roku w wydawnictwie Wolters Kluwer. Niniejszy tekst koncentruje się wyłącznie na wątkach metodologicznych; ze względu na znaczną objętość nie jest możliwe przedstawienie w nim nawet cząstkowych wyników badania.

[2] W skład zespołu wchodzili: dr hab. Tomasz Stawecki, prof. UW, dr hab. Marcin Matczak, dr Wiesław Staśkiewicz, dr Krzysztof Kaleta, mgr Marcin Romanowicz, dr Jan Winczorek. Podział prac w ramach zespołu przedstawiał się następująco: J. Winczorek – założenia metodologiczne badania w głównej części, opracowanie narzędzi informatycznych, wykonanie prac badawczych na części materiału w objętości wynikającej z przyjętej metodologii badania, analiza statystyczna danych, raport z wyników; pozostali członkowie zespołu – współautorstwo (w mniejszej części) założeń metodologicznych, wykonanie prac badawczych na części materiału w objętości wynikającej z przyjętej metodologii badania.

obserwacji dokonywanych w teorii prawa i teorii polityki na temat pozycji sądownictwa konstytucyjnego w systemie prawnym i ustroju politycznym współczesnego społeczeństwa. Głównym obiektem zainteresowania były w nim metody dyskursywnego uzasadniania decyzji wykorzystywane przez polski Trybunał Konstytucyjny. Przedmiot badania stanowiły więc stwierdzenia wypowiadane w języku ściśle profesjonalnym, artykułowane za pomocą kategorii i schematów rozumowania słabo znanych poza światem prawniczym. Realizacji tego celu poznawczego służyła natomiast technika badawcza wykorzystywana częściej w naukach społecznych niż w naukach prawnych – jakościowo-ilościowa, wspomagana komputerowo analiza treści.

Artykuł jest podzielony na cztery części. Rozpoczynają go uwagi wstępne, w których przedstawione zostają założenia badania i najważniejsze inspiracje jego autorów. W części drugiej omawiane są główne cechy i funkcje oprogramowania (programu statystycznego wraz ze specjalizowaną nakładką *Computer-Assisted Qualitative Data Analysis*, dalej CAQDA) użytego do realizacji badania. W części trzeciej mowa jest o zastosowanych w badaniach procedurach analizy danych. W części czwartej przedstawiane są uwagi na temat granic produktywności narzędzi CAQDA w socjologicznych analizach orzecznictwa sądowego i ogólniejszych problemów epistemologicznych, jakie się z tym wiążą.

Uwagi wstępne

Zagadnienie metod wykładni prawa służących do uzasadniania decyzji sądowych stanowi, wraz z pokrewnymi kwestiami tak zwanej ideologii stosowania prawa oraz argumentacji i rozumowań prawniczych, jeden z centralnych tematów teorii prawa. W ramach tej dyscypliny wypracowano wiele normatywno-proskryptywnych doktryn wykładni uzasadniających rozmaite praktyki podejmowane w tym obszarze (przydatny przegląd stanowisk zawiera książka Feteris [1999]; w sprawie bardziej idiosynkratycznych koncepcji formułowanych w polskim kontekście zob. np. Wróblewski 1988; Zieliński 2008; Morawski 2010). Rzadziej spotyka się – choć i one są tworzone w naukach prawnych – deskryptywne teorie odnoszące się do rzeczywistego przebiegu procesów decyzyjnych, zakresu władzy sędziowskiej i granic jej dyskrecjonalności. Wykładnia i stosowanie prawa stanowi również obiekt zainteresowania nauk społecznych – na przykład w kontekście zjawiska „judycjalizacji polityki" dostrzeganego przez nauki polityczne (Stone Sweet 1999; Maravall 2003) czy złożoności czynników wpływających na ostateczną treść wyroku, obserwowanych przez socjologię czy antropologię prawa (za przykłady badań prowadzonych metodami jakościowymi mogą posłużyć prace Lautmanna [1972] czy Scheffera [2010]).

Sądy konstytucyjne, czy szerzej – instytucje orzekające o tak zwanej hierarchicznej zgodności aktów prawnych, są tu obiektem wzmożonego zainteresowania (szerokie omówienie literatury w kontekście polskim i zagranicznym można znaleźć w pracy pod red. Staweckiego i Winczorka [w druku]). Wynika ono, z jednej strony, ze szczególnego charakteru i szerokiego zakresu władzy, jaką dysponują te organizacje, oraz, z drugiej strony, z ich specyficznego instytucjonalnego usytuowania *vis à vis* innych organów władzy. Pierwsza z tych kwestii

wiąże się z – z natury rzeczy – szerokimi formalnoprawnymi uprawnieniami tych ciał i ich znaczną dyskrecjonalną władzą orzeczniczą, wynikającą z abstrakcyjnej i niekiedy niejasnej materii konstytucyjnej, na podstawie której są zmuszone wydawać swoje wyroki. Nie tylko więc sądy konstytucyjne są jedynymi, obok parlamentów, organami uprawnionymi do uchylania norm powszechnie obowiązującego prawa, ale także wykonują swoją władzę na podstawie bardzo nieostrych przesłanek. W związku z tym dysponują zdolnością określaną w nauce prawa mianem *kompetenz-kompetenz*, to znaczy faktyczną możliwością określania własnych uprawnień poprzez odpowiednią interpretację przepisów, na podstawie których orzekają. Dostrzegając ten stan rzeczy, wysuwa się często pod ich adresem apele o stosowanie się do ideału „powściągliwości sędziowskiej".

Gdy chodzi o drugie wspomniane zagadnienie, w obfitej literaturze przedmiotu wskazuje się na trudności, jakie teoria demokracji ma z uzasadnieniem istnienia trybunałów konstytucyjnych (Sadurski 2008: 21 i nast.). W tak zwanym skoncentrowanym modelu kontroli konstytucyjności, funkcjonującym dziś także w Polsce, trybunały konstytucyjne są ciałami quasi-sądowymi, niemieszczącymi się w klasycznym trójpodziale władzy. Orzekający w nich sędziowie pochodzą z nominacji politycznych, tak samo jak osoby wchodzące w skład władzy ustawodawczej i wykonawczej. Jednocześnie jednak zakres ich rozliczalności jest bardzo ograniczony, podobnie jak ma to miejsce w przypadku sędziów sądów powszechnych – które są zazwyczaj ciałami merytokratyczno-technokratycznymi. W związku z tym nie jest rzeczą jasną, na jakiej podstawie możliwe jest legitymizowanie decyzji podejmowanych przez trybunały.

W klasycznej dla Europy tak zwanej austriackiej doktrynie sądownictwa konstytucyjnego problem legitymizacji rozwiązuje się, podkreślając szczególność funkcji pełnionej przez trybunały konstytucyjne (zob. Wronkowska 2008; Kelsen 2009). Mają one być wyłącznie „negatywnym ustawodawcą", to znaczy ciałem uprawnionym jedynie do uchylania wadliwych (niekonstytucyjnych) norm prawa, ale nie tworzenia nowych. W rzeczywistości władza interpretowania prawa, z której korzystają trybunały, a także istotne problemy praktyczne (jak na przykład zagadnienie tzw. zaniechania ustawodawczego) sprawiają, że organy te nie uchylają po prostu norm prawnych, ale pełnią także rolę „pozytywnych ustawodawców", to znaczy wydając swoje orzeczenia, *de facto* tworzą nowe normy.

Zarówno jedna, jak i druga kwestia pozostaje w bliskich związkach z zagadnieniem sposobu uzasadniania decyzji sądowych podjętym w omawianych badaniach. Można mianowicie twierdzić, że problem legitymizacji daje się przynajmniej częściowo rozwiązać wtedy, gdy działalność trybunałów będzie miała charakter dyskursywny, to znaczy nie będzie opierać się tylko na czysto formalnych uprawnieniach do podejmowania określonych decyzji, ale będzie również zmierzała do przekonania ich audytorium o trafności wydawanych wyroków (jedną z bardziej znanych doktryn tego rodzaju przedstawił Alexy [1985; 1996]). Rozważając zagadnienie dyskrecjonalnej władzy trybunałów, zauważa się natomiast, że realizacja wymogu intersubiektywnej uzasadnialności wyroków w istotny sposób

ogranicza arbitralność podejmowanych decyzji, ponieważ nie każda decyzja może zostać przekonująco uzasadniona w danym kontekście kulturowym, aksjologicznym czy instytucjonalnym.

Badania opisywane w niniejszym artykule zostały pomyślane tak, aby udzielały odpowiedzi na pytanie, jak w rzeczywistości swoje decyzje uzasadniał polski Trybunał Konstytucyjny (dalej TK). Nie miały jednak ambicji konkluzywnego rozstrzygania problemów teorii prawa. Zmierzały natomiast do ustalenia rzeczywistych wzorców argumentowania w uzasadnieniach orzeczeń wydawanych przez TK[3] oraz zbadania ich zmienności w czasie i w zależności od dostępnych zmiennych wyjaśniających. Choć więc nie były wolne od założeń teoretycznych, to miały charakter zdecydowanie eksploracyjny, nie zaś eksplanacyjny i nie miały służyć do testowania hipotez.

Sprawia to, że zanim przejdzie się do przedstawienia bliższego opisu procedury badawczej i oprogramowania użytego do jej wdrożenia, trzeba poczynić dodatkowe uwagi wstępne. Jakkolwiek orzecznictwo sądowe stanowi stały obiekt badań nauk prawnych, to już jego badanie metodami nauk społecznych nie jest powszechne. Jest tak z przynajmniej dwóch powodów. Po pierwsze, tak zwane dogmatyczne nauki prawne (na przykład nauka prawa cywilnego czy prawa administracyjnego) zorientowane są przede wszystkim na analizę zawartości źródeł prawa w celu ustalenia jego właściwej treści w świetle zewnętrznie narzuconych założeń ontologicznych i aksjologicznych. Rekonstrukcje „linii orzeczniczych" albo odwołania do pojedynczych orzeczeń mają, w polskiej kulturze prawnej i w odniesieniu do polskich sądów, charakter zazwyczaj pomocniczy. Analiza orzecznictwa służy tu głównie porównaniu rozumowania stojącego za stosowaniem prawa z tym, za którym opowiada się badacz, albo ma wprost charakter krytyczny. Z tego powodu badania zmierzające do ustalenia charakterystyki wypowiedzi sądów w dłuższym przekroju czasowym przeprowadzane na systematycznych (zwłaszcza losowych) próbach i w sposób możliwy do uogólnienia na cały korpus orzeczeń danego typu, choć mogą być istotne dla teorii prawa, nie odpowiadają potrzebom poznawczym dogmatycznych nauk prawnych.

Po drugie, prowadzenie tego rodzaju badań jest trudniejsze niż analiza wielu materiałów innego rodzaju. Wypowiedzi sądów zmierzające do uzasadnienia decyzji wyrażonych w sentencjach wyroków cechuje wysoki poziom skomplikowania, niekiedy znaczna objętość[4] oraz pewna hermetyczność. Wynika to tyle ze stosowanej prawniczej terminologii i niekiedy niebagatelnej złożoności rozważanych zagadnień, co z usytuowania sądów w ustroju państwa i swoistej ekonomii orzekania, wpływającej na sposób przedstawiania swoich racji przez organy orzecznicze. Stwierdzenia wypowiadane w uzasadnieniach orzeczeń mają wprawdzie – z samej zasady – charakter argumentacyjny,

[3] Warto podkreślić, że czym innym jest sposób uzasadnienia wyroku sądowego przedstawiony w jego uzasadnieniu, a czym innym rzeczywisty przebieg procesów psychicznych i społecznych, które przyczyniły się do wydania wyroku w określonym kształcie. Omawiane badanie miało służyć, rzecz jasna, ustaleniu tej pierwszej okoliczności.

[4] W polskiej kulturze prawnej pisemne uzasadnienia orzeczeń bywają bardzo obszerne, a w przypadku Trybunału Konstytucyjnego liczą niekiedy nawet kilkadziesiąt stron.

bo zmierzają do uzasadnienia decyzji, którą podjął sąd, jednak repertuar wykorzystywanych w nich środków retorycznych jest ograniczony. Wpływają na to akceptowane *explicite* i *implicite* ideologie i metodologie orzekania oraz spodziewane reakcje stron postępowania i zapatrywania sądów wyższych instancji wyrażane podczas ewentualnej kontroli instancyjnej.

Analiza tego typu materiałów wymaga w związku z tym szczególnych kompetencji i odpowiednich narzędzi badawczych. Badacze podejmujący analizę orzecznictwa sądowego powinni w szczególności być w stanie w sposób kompetentny dekodować wypowiedzi sądów dokonywane w języku prawniczym, co wymaga znajomości doktryny prawa w konkretnym obszarze oraz powszechnie akceptowanych w danej kulturze prawnej metod argumentacji czy wnioskowań prawniczych. Jednocześnie, badacze powinni być w stanie czynić to w sposób refleksyjny, dostrzegając uwikłanie wypowiedzi sądów w społeczny kontekst. Ten drugi aspekt istotnie odróżnia socjologiczne badanie orzecznictwa od prawno-dogmatycznego namysłu nad jego treścią i powinien być uwzględniony przy konstruowaniu narzędzi badawczych. Jest to zarazem istotna bariera dla podejmowania tego typu badań przez prawników. Konieczność posiadania kompetencji prawniczych sprawia natomiast, że badania takie są trudne do realizacji przez socjologów i innych badaczy z nauk społecznych[5].

Można sądzić, że opisane względy sprawiają, że badania przedstawiane w niniejszym artykule mają charakter względnie nowatorski. W Polsce podobne przedsięwzięcia badawcze (w odróżnieniu od tak zwanych badań aktowych, skupiających się zazwyczaj na treściach decyzji sądowych, a nie ich uzasadnieniach) podejmowano tylko sporadycznie. Pod koniec poprzedniej dekady trzej autorzy wchodzący w skład zespołu realizującego omawiane tu badania przeprowadzili zbliżone studia nad orzecznictwem Trybunału Konstytucyjnego, Naczelnego Sądu Administracyjnego i Sądu Najwyższego (Winczorek, Stawecki, Staśkiewicz 2008). Inny członek zespołu badawczego prowadził badania orzecznictwa Naczelnego Sądu Administracyjnego (Galligan, Matczak 2005). Godna wspomnienia jest również praca nad pewnymi aspektami orzecznictwa Sądu Najwyższego (Wyrembak 2009).

Założenia metodologiczne badania zostały dostosowane do opisanych celów poznawczych. Badanie miało charakter jakościowy w tym względzie, że zakładało swobodne kodowanie tekstu uzasadnień orzeczeń Trybunału Konstytucyjnego zgodnie z jego interpretacją przez każdego z koderów[6]. Wynikał on również z przyjętej metody generowania kodów użytych w badaniu. Choć nie można uznać jej za ugruntowaną w ścisłym rozumieniu tego słowa, to wychodziła ona z założenia, że wytwarzanie kodów powinno mieć charakter możliwie indukcyjny, wynikać z kontaktu z badanymi

[5] Nie jest przypadkiem, że w badaniach sądów i orzecznictwa najczęściej podejmowanych w naukach społecznych poszukuje się mierzalnych wskaźników, za pomocą których można w sposób zewnętrzny opisywać działania sądów (np. częstości apelacji, liczby zdań odrębnych itp.). Bardzo rzadko, jeśli kiedykolwiek, wskaźniki takie dotyczą sposobu uzasadniania orzeczeń.

[6] Wszystkie osoby biorące udział w badaniu ukończyły studia prawnicze, a pięć z nich uzyskało także stopień doktora lub tytuł doktora habilitowanego nauk prawnych. Trzech uczestników badania wykonuje ponadto zawód prawniczy, a trzech ukończyło także studia w zakresie nauk społecznych.

treściami i stanowić przedmiot interakcji pomiędzy badaczami.

Ilościowy charakter badania przejawiał się w założeniu, że treści zaobserwowane w toku analizy powinny być mierzalne, a wynik uzyskany w badaniu próby orzeczeń powinien dać się uogólnić na całą ich populację w drodze wnioskowania statystycznego. Przyjęto także założenie (które odróżnia omawiane badania od podobnych badań realizowanych wcześniej w Polsce), że zbieranie danych należy przeprowadzić przy zachowaniu triangulacji personalnej, to znaczy w drodze równoległego kodowania tego samego materiału badawczego przez więcej niż jednego kodera. Ta ostatnia decyzja wynikała w szczególności ze wspomnianego faktu, że analizowana materia miała charakter szczególnie złożony, a także z wcześniejszych doświadczeń wspomnianych badaczy.

Oprogramowanie

Złożoność i obszerność materiału badawczego, wynikające z powyższych założeń, spowodowały konieczność zastosowania w badaniu wyspecjalizowanego oprogramowania wspomagającego analizę treści. Po rozpoznaniu dostępnych możliwości, zdecydowano się na użycie oprogramowania R w wersji 2.13 (R Core Team 2012) wraz z graficzną nakładką RQDA (nazwa stanowi akronim zwrotu „R-based Qualitative Data Analysis") w wersji 0.2.2 (Huang 2012). Nakładka ta pozwoliła na łatwe przeprowadzenie kodowania materiału badawczego za pomocą szerokiego zestawu kodów. System statystyczny R umożliwił natomiast dokonanie analizy statystycznej kodów wraz z analizą rzetelności kodowania.

Decyzja o takim doborze oprogramowania była podyktowana kilkoma okolicznościami. Po pierwsze, była to potrzeba dysponowania narzędziem pozwalającym zarówno na swobodne kodowanie złożonych treści za pomocą jakościowo określonych kodów, jak i na wykonanie złożonej analizy statystycznej, w tym także za pomocą procedur, które nie zostały zdefiniowane przez autorów oprogramowania. Po drugie, wynikała ona z faktu, że wymienione narzędzia są dostępne nieodpłatnie – autorzy udostępniają je na licencji GNU GPL – i że można z nich korzystać na kilku systemach operacyjnych. Po trzecie, nie bez znaczenia było to, że program RQDA ma względnie niskie „koszty wejścia". Ponieważ oferuje użytkownikowi tylko ograniczoną liczbę opcji, pozostawiając wszystkie funkcje analityczne w programie R, jest łatwy w obsłudze i pozwala na udział w badaniach – tylko po krótkim przeszkoleniu – także takim osobom, które nie miały wcześniej do czynienia z tego typu oprogramowaniem[7].

Program R stanowi jedno z najbardziej rozbudowanych i najbardziej uniwersalnych narzędzi analizy statystycznej dostępnych na komputery PC (jego uruchomienie jest możliwe między innymi na systemach Windows, Linux i MacOSX). Jego historia sięga początku lat dwutysięcznych (Ripley 2001), a program jest wciąż aktywnie rozwijany. Opiera się na zaawansowanym języku programowania, będącym implementacją języka S, używanego w oprogramowaniu komercyjnym[8], i zawierającym gotowe procedury pozwalające na prowadzenie bardzo

[7] Warto wspomnieć, że w przypadku nakładki RQDA, oprócz standardowej dokumentacji, nieodpłatnie dostępne są także instrukcje korzystania z programu w formie wideo.

[8] Pakiet S-PLUS firmy TIBCO. Obydwie implementacje języka S są ze sobą w dużym stopniu kompatybilne.

zaawansowanych analiz statystycznych we wszystkich dziedzinach wiedzy: zarówno w naukach przyrodniczych i technicznych, jak i społecznych czy lingwistyce[9]. Charakterystyczną cechą R – która dla niektórych użytkowników może być wadą – jest to, że w odróżnieniu od innych znanych pakietów statystycznych, takich jak SPSS czy Statistica, nie posiada on interfejsu graficznego. Praca z programem odbywa się natomiast interaktywnie, za pomocą komend wydawanych w wierszu poleceń (środowisku CLI), albo nieinteraktywnie, za pomocą wcześniej przygotowanych skryptów[10].

Istotną własnością R jest także modułowość, przejawiająca się w możliwości instalowania gotowych pakietów rozszerzających standardowy zestaw komend oraz łatwego tworzenia własnych zestawów poleceń. Pakiety rozszerzeń są przygotowywane zarówno przez węższą grupę deweloperów zajmujących się rozwijaniem programu R, jak i przez szerokie kręgi użytkowników, co pozwala na znalezienie gotowych pakietów dopasowanych do rozwiązywania nawet bardzo szczególnych problemów statystycznych. Obecnie (wrzesień 2013) w repozytorium pakietów CRAN, wykorzystywanym przez twórców programu R do dystrybucji oprogramowania, dostępnych jest 4861 takich pakietów. Na szczególną uwagę zasługują praktycznie nieograniczone możliwości wizualizacji danych i szerokie możliwości ich wymiany z innymi programami, oferowane przez oprogramowanie zgromadzone w repozytorium.

Opisywany w niniejszym artykule projekt badawczy wykorzystywał możliwości analizy oferowane przez R jedynie w niewielkim stopniu, ograniczając się do tworzenia tabel kontyngencji, analiz korelacyjnych, analizy logitowej i analiz rzetelności kodowania. Niemniej jednak elastyczność pakietu, przejawiająca się w możliwości łatwego definiowania nowych funkcji i dowolnego łączenia istniejących poleceń oraz w możliwości importowania i swobodnego przekształcania różnorodnych zestawów danych, stanowiła istotny czynnik umożliwiający realizację omawianego badania w formie wynikającej z przyjętych założeń.

Nakładka RQDA, która została wykorzystana w badaniu do bezpośredniego kodowania uzasadnień orzeczeń, stanowi autorski projekt chińskiego socjologa Ronggui Huanga. Do jej uruchomienia niezbędny jest funkcjonujący system R, zaś sama instalacja RQDA odbywa się w sposób zautomatyzowany poprzez wspomniane repozytorium CRAN (program sam pobiera niezbędne komponenty). Interesującą cechą omawianego oprogramowania jest także możliwość wykonania jego tak zwanej statycznej instalacji. Dzięki temu możliwe jest przygotowanie pakietu oprogramowania dla uczestników programu badawczego na nośnikach wymiennych (na przykład pendrive'ach) bez potrzeby jego instalowania na wielu komputerach, co ułatwia realizację badań i pozwala na rozpoczęcie pracy z programem

[9] Bardzo obszerna dokumentacja pakietu R jest dostępna na stronie internetowej projektu www.r-project.org (ostatni dostęp: 28 września 2013 r.).

[10] Dostępne jest jednak dodatkowe oprogramowanie ułatwiające korzystanie z R. Są to programy wspomagające tworzenie skryptów w języku R, działające w środowisku graficznym. Istnieje także możliwość instalacji oprogramowania R w środowisku WWW za pomocą odpowiednich nakładek na serwery internetowe, co pozwala na zdalne korzystanie z możliwości oferowanych przez ten pakiet. Wreszcie, istnieją również osobne programy statystyczne dysponujące interfejsem graficznym, porównywalnym z SPSS czy Statistica, i wykorzystujące pakiet R jako „silnik" analiz statystycznych.

także użytkownikom o mniejszych kompetencjach z zakresu obsługi komputera[11].

Większość danych wytworzonych w nakładce RQDA jest dostępna z poziomu programu R za pomocą specjalnego zestawu komend instalowanych wraz z tą nakładką. W założeniu twórcy, dane wytworzone w programie RQDA powinny podlegać analizie jakościowej lub prostej analizie ilościowej za pomocą przygotowanych przez niego gotowych narzędzi statystycznych, dostępnych z poziomu CLI programu R[12]. Ponieważ jednak projekty RQDA zapisywane są jako bazy danych w standardowym formacie programu SQLite, możliwe jest także ich łatwe przetwarzanie za pomocą typowego oprogramowania, w tym wyeksportowanie wytworzonych danych do formatów możliwych do odczytania przez inne programy statystyczne (np. formatu .csv). Możliwe jest również wczytanie danych z nakładki RQDA do systemu R z pominięciem standardowych procedur zaimplementowanych w samym RQDA.

Ta ostatnia możliwość była wykorzystywana w analizie danych wytworzonych w ramach omawianego projektu. Wynika to z faktu, że niezależnie od wielu możliwości analizy oferowanych przez oprogramowanie RQDA, nie jest ono bezpośrednio przystosowane do przeprowadzania badań wykorzystujących triangulację personalną i nie zawiera opcji równoległej pracy wielu koderów na tym samym materiale[13]. Z tego względu w badaniu wykorzystano pakiet dodatkowy „RSQLite" pozwalający na odczyt w programie R baz danych zapisanych w formacie SQLite. Za jego pomocą połączono zbiory danych wytworzone przez poszczególnych koderów, zachowując informację o autorstwie poszczególnych zapisów.

Oprogramowanie RQDA zostało zaprojektowane w taki sposób, aby pozwalało na kodowanie swobodne, to jest bez narzuconej z góry jednostki (czy segmentu) analizy. Kodowanie może się odbywać zarówno za pomocą kodów wytwarzanych *in vivo* przez samego kodera, jak i określonych wcześniej. Dzięki temu, a także za sprawą opcji grupowania kodów, tworzenia notatek i dzienników pracy możliwe jest prowadzenie za pomocą RQDA analiz o charakterze czysto jakościowym, w tym odwołujących się zwłaszcza do wytycznych metody ugruntowanej. Oprogramowanie RQDA zostało ponadto przystosowane do wytwarzania danych niezbędnych do analizy statystycznej typu QCA (*qualitative comparative analysis*, dalej QCA), to jest nieprobabilistycznej analizy danych ilościowych o niewielkiej liczbie obserwacji. Możliwe jest także opisywanie poszczególnych plików za pomocą dowolnej liczby

[11] Ta możliwość jest z zasady wykluczona w przypadku narzędzi badawczych udostępnianych na zasadach komercyjnych, w przypadku których liczba dokonywanych instalacji jest ściśle ograniczona do komputerów objętych licencją.

[12] Ich dokładny opis jest zawarty w dokumentacji programu, dostępnej na jego stronie internetowej http://rqda.r-forge.r-project.org/ (ostatni dostęp: 28 września 2013 r.).

[13] Istnieje natomiast opcja pracy na tym samym materiale bez stosowania triangulacji. W takim przypadku możliwe jest wykorzystanie zdalnej bazy danych MySQL umieszczonej na serwerze dostępnym z sieci Internet, z którą łączą się badacze korzystający z programu RQDA na swoich lokalnych komputerach. Takie rozwiązanie, w połączeniu ze wspominanymi wyżej nakładkami instalującymi R na serwerach WWW, może (po przezwyciężeniu dodatkowych trudności technicznych) pozwalać na automatyczne informowanie badaczy o wynikach badania w trakcie kodowania, co może przyczyniać się do wzrostu refleksyjności całego procesu. Warto zauważyć, że w takiej konfiguracji omawiane oprogramowanie nie narzucałoby praktycznie żadnych limitów w sprawie liczby koderów biorących udział w projekcie (ograniczeniem byłaby tu jedynie pojemność i szybkość działania bazy danych).

zmiennych numerycznych lub tekstowych, dających się następnie wykorzystać w klasycznej analizie ilościowej, na przykład jako zmienne niezależne.

Interfejs graficzny programu RQDA można opisać jako nieskomplikowany i przez to intuicyjny. Grupuje on opcje dostępne dla użytkownika w osiem kategorii: menu „Project", „Files" i „Settings" pozwalają na zarządzanie projektami i dodawanie do nich nowych plików do zakodowania. RQDA pozwala tu na importowanie jedynie plików tekstowych, a więc pliki innego rodzaju muszą zostać przekształcone na ten format przed rozpoczęciem kodowania[14]. Menu „Codes" pozwala na definiowanie kodów wykorzystywanych do kodowania zaimportowanych plików, wyświetlanie fragmentów tekstu zakodowanych za pomocą poszczególnych kodów oraz tworzenie notatek odnoszących się do konkretnych fragmentów kodowanego materiału. Samo kodowanie odbywa się poprzez zaznaczenie wybranego fragmentu tekstu i wybór odpowiedniego kodu z menu programu[15]. Menu „Code categories" zawiera opcje ułatwiające grupowanie kodów, pozwalając tym samym na budowanie teorii w ramach ugruntowanych strategii badawczych lub w celu uporządkowania zebranych danych. Ograniczeniem RQDA w tym zakresie jest niemożność tworzenia wielopoziomowych kategoryzacji kodów – program pozwala jedynie na grupowanie kodów za pomocą kategorii mieszczących się na tym samym poziomie klasyfikacji, a więc nie daje możliwości określenia ich hierarchii lub innych relacji. Menu „Cases" pozwala na pogrupowanie analizowanych plików w sposób dogodny dla użytkownika, na przykład połączenie ich w jedną jednostkę analizy na potrzeby badania QCA. Menu „Attributes" pozwala na zdefiniowanie zmiennych opisujących poszczególne pliki lub ich grupy (*cases*). Menu „File categories" pozwala na grupowanie plików w kategorie. Wreszcie, menu „Journals" umożliwia tworzenie notatek z prowadzonych prac z uwzględnieniem chronologii. Jak to zostało powiedziane, RQDA nie zawiera natomiast żadnych bardziej zaawansowanych opcji analizy danych dostępnych z poziomu interfejsu graficznego. Zgodnie z założeniem autora programu, ilościowa analiza danych ma być prowadzona z poziomu programu R. Podobnie jest z opcją wyszukiwania określonych fragmentów tekstu[16].

Najważniejszą cechą RQDA jako narzędzia analizy treści jest to, że pozwala ono na prowadzenie całkowicie swobodnego kodowania analizowanych tekstów. Badacz oznaczający jakiś fragment tekstu wybranym przez siebie kodem może uczynić to w odniesieniu do fragmentu tekstu o dowolnej długości, poczynając od 1 znaku. Możliwe jest przy tym jednoczesne (wielokrotne) zakodowanie tego samego fragmentu tekstu dowolną liczbą kodów. Zakodowane fragmenty mogą też pozostawać w dowolnych relacjach: pokrywania się, zawierania, krzyżowania lub rozłączności.

[14] Program obsługuje polskie znaki diakrytyczne w formacie utf8.

[15] Program dopuszcza przy tym edytowanie zawartości kodowanego pliku bez straty kodowania.

[16] Elastyczność oprogramowania R pozwala na jego wykorzystanie do zautomatyzowanego lub półautomatycznego kodowania analizowanych tekstów w oparciu o ich zadane cechy (na przykład przy wykorzystaniu wyszukiwań pełnotekstowych). W omawianym badaniu tej opcji jednak nie wykorzystywano; można sądzić, że wymagałoby to stworzenia własnej sekwencji komend w języku R. Autor oprogramowania RQDA przygotował również pakiet dodatkowy RQDAtm, służący do realizacji badań ilościowych na korpusie wypowiedzi. Rozszerza on możliwości R i RQDA o funkcje typowe dla badań lingwistycznych (analiza kolokacji, konkordancji itp.), jego ograniczeniem może być jednak niedostosowanie do badania tekstów w językach innych niż angielski.

Wewnętrzny format zapisu danych wykorzystywany w programie RQDA zapewnia użytkownikowi prowadzącemu analizę dostęp do wszystkich danych wytworzonych podczas kodowania. W szczególności z poziomu programu R (lub innego programu statystycznego) możliwe jest odczytanie następujących informacji umieszczonych w bazie danych: kod użyty do oznaczenia danego fragmentu tekstu, plik, w którym znajduje się zakodowany fragment, początek zakodowanego fragmentu jako odległość w znakach od początku kodowanego tekstu, koniec zakodowanego fragmentu (analogicznie), zakodowany tekst, wartości zmiennych (atrybutów) ustalonych dla poszczególnych plików. Pozwala to na zachowanie pełnej wiedzy o zakodowanych fragmentach tekstu także podczas analizy ilościowej i również po wyeksportowaniu danych stworzonych w RQDA do innego niż R programu statystycznego.

Metodologia badania orzecznictwa Trybunału Konstytucyjnego

Opisane cechy oprogramowania R i RQDA pozwoliły na przeprowadzenie badań uzasadnień orzeczeń Trybunału Konstytucyjnego w sposób zgodny z potrzebami poznawczymi, posiadanymi zasobami i założeniami wyjściowymi. Własności narzędzia informatycznego nie nałożyły zwłaszcza większych ograniczeń na konceptualizację badań. Zastosowana procedura badawcza była kilkuetapowa. W pierwszej kolejności przeprowadzono pilotaż polegający na swobodnym kodowaniu – przez wszystkich sześciu badaczy biorących udział w projekcie – pięciu dobranych celowo orzeczeń Trybunału Konstytucyjnego. Na tym etapie badania każdy z koderów pracował na identycznym materiale, przeprowadzając jego niezależne kodowanie za pomocą samodzielnie wytwarzanych kodów *in vivo* – w trakcie lektury kodowanych orzeczeń. Jedynym ustalonym *explicite* ograniczeniem dla inwencji badaczy było tu uzgodnione założenie, że celem badania jest ustalenie wzorców argumentacji wykorzystywanych przez Trybunał.

W ten sposób wytworzono 312 kodów o przecinających się zakresach i różnym stopniu wykorzystania przez różnych koderów i w różnych orzeczeniach. Następnie przeprowadzono kilka sesji uzgodnieniowych zmierzających do ustalenia listy kodów wspólnej dla zespołu badawczego oraz do ujednolicenia znaczeń przypisywanych poszczególnym kodom przez poszczególnych badaczy.

W efekcie tych działań wytworzono listę 50 kodów, z których większość opatrzono notatkami opisującymi ich uzgodnioną treść. Kody ustalone na tym etapie były dwojakiego rodzaju. Kody pierwszego poziomu (w liczbie 37) opisywały argumenty wysuwane przez Trybunał Konstytucyjny w uzasadnieniach orzeczeń. Kody drugiego poziomu (13) służyły do tego, by scharakteryzować rodzaj tekstu prawnego, do którego odnosiły się kody pierwszego poziomu. Rozróżnienie to miało służyć przede wszystkim ustaleniu, czy i w jakim stopniu wzorce argumentacji stosowane w odniesieniu do różnych aktów prawnych różnią się od siebie[17]. Listę kodów przedstawia Aneks.

[17] W konsekwencji każdy fragment uzasadnienia orzeczenia TK, odnoszący się do jakiegoś aktu prawnego, był kodowany co najmniej dwukrotnie – przynajmniej raz kodem pierwszego poziomu i raz kodem drugiego poziomu.

Jego lektura pozwoli łatwo zauważyć, że niezależnie od zaakceptowanej przez wszystkich badaczy indukcyjnej metody wytwarzania kodów – zgodnie z którą mogły one być określone w sposób dowolny, ale przede wszystkim na podstawie lektury uzasadnień orzeczeń – dominują wśród nich kategorie pojęciowe dobrze znane teorii prawa i wykorzystywane w dyskursie prawniczym. Ich status metodologiczny jest też zróżnicowany – są to zarówno kody odnoszące się do tradycyjnych, znanych od setek lat rozumowań czy topik prawniczych, jak i do koncepcji wypracowanych w ramach konkretnych doktryn wykładni. Sprawiło to, że typologia sposobów argumentowania, wyrażona za pomocą kodów, nie była skonstruowana według jednolitego schematu i w konsekwencji nie była logicznie zupełna. Niektóre jej elementy nie były też rozłączne, a granice ich zakresów były niekiedy nieostre.

Na drugim etapie badania zrealizowano główną część analizy treści. W jej trakcie – za pomocą ustalonych kodów – przeprowadzono badanie 150 uzasadnień orzeczeń TK z lat 1986–2009[18]. Orzeczenia zostały dobrane w taki sposób, aby były reprezentatywne w dwóch wymiarach. Po pierwsze, chodziło o dwa istotne okresy w historii działalności TK w Polsce, to znaczy czas przed 1997 rokiem (kiedy obowiązywały poprzednie przepisy konstytucyjne) i po 1997 roku (kiedy w życie weszła nowa konstytucja). Po drugie, założono reprezentatywność ze względu na rodzaj orzeczenia określony jego sygnaturą[19]. Tak dobrane orzeczenia poddano zautomatyzowanej obróbce za pomocą prostego narzędzia informatycznego, usuwając z nich wszystkie elementy poza uzasadnieniem stanowiska TK[20]. Łączna objętość tekstów poddanych badaniu (po obróbce) wynosiła 3 604 822 znaków (średnio 24 032 znaki na uzasadnienie), co stanowi 7,2% całego korpusu orzeczeń Trybunału po obróbce i 4,2% przed jej dokonaniem.

Na tym etapie badania kody nie podlegały już dalszym modyfikacjom. Badacze nie dodawali również nowych, a jedynie kodowali uzasadnienia orzeczeń w sposób swobodny, bez zdefiniowanej jednostki analizy i z wykorzystaniem możliwości wielokrotnego kodowania tego samego fragmentu tekstu różnymi kodami. Ta ostatnia możliwość wynikała ze wspomnianego faktu, że nie wszystkie kody były

[18] Według bazy danych „Orzecznictwo Trybunału Konstytucyjnego 1986–2009" wydanej przez Biuro Trybunału Konstytucyjnego z datą 3 maja 2010 r., w badanym okresie TK wydał łącznie 5521 orzeczeń, z czego 1908 rozstrzygały problem prawny co do meritum (pozostałe orzeczenia nosiły sygnatury rozpoczynające się od „T", a więc dotyczyły tzw. wstępnej kontroli dopuszczalności orzekania). Za operat stanowiący podstawę doboru próby przyjęto orzeczenia należące do tej pierwszej kategorii. Dobór orzeczeń do próby w ramach ustalonych kategorii miał charakter losowy.

[19] Orzeczenia TK są oznaczane następującymi sygnaturami: K – orzeczenie odnoszące się do niezgodności z konstytucją przepisów ustaw, Sk – orzeczenie wydane w efekcie skargi konstytucyjnej, P – orzeczenie wydane w efekcie pytania prawnego sądu, U – orzeczenie odnoszące się do niezgodności z przepisami rangi ustawowej lub konstytucją aktów podustawowych (rozporządzeń), S – wyroki sygnalizujące potrzebę nowelizacji prawa, W – wyroki wypowiadające powszechnie wiążącą interpretację prawa (do 1997 roku).

[20] Orzeczenia TK kończące sprawę co do meritum składają się najczęściej z czterech części. Część pierwsza to sentencja wyroku, w której Trybunał orzeka o uchyleniu lub utrzymaniu w mocy konkretnych przepisów prawa. W tej części zawarta jest również informacja o składzie sędziowskim, który wydał wyrok, stronach postępowania itd. Kolejne trzy części składają się na uzasadnienie wyroku. W części pierwszej uzasadnienia Trybunał przedstawia pisemne stanowisko strony, która przedstawiła wniosek, pytanie prawne lub skargę konstytucyjną oraz stanowiska innych podmiotów. W części drugiej zawarte jest streszczenie stanowisk podmiotów biorących udział w postępowaniu, przedstawione na rozprawie. Dopiero trzecia część uzasadnienia zawiera autorskie wypowiedzi Trybunału. Przedmiotem badania była tylko owa trzecia część, co wynika z faktu, że pozostałe elementy uzasadnienia stanowią zazwyczaj mechaniczne powtórzenie treści zawartych w pismach przedstawianych przez strony postępowania. Gdy dalej w niniejszym tekście jest mowa o „uzasadnieniu orzeczenia", chodzi właśnie o tę część.

w swoich zakresach całkowicie rozłączne i same podlegały interpretacji, z faktu, że stosowano kody znajdujące się na dwóch poziomach, jak i faktu, że wypowiedzi TK były złożone i niekiedy nie dawały się opisywać za pomocą tylko jednego kodu.

Zastosowano ponadto triangulację personalną. Polegała ona na podzieleniu 150 uzasadnień orzeczeń pomiędzy członków zespołu badawczego w taki sposób, aby każde orzeczenie zostało zakodowane przez dwóch badaczy i jednocześnie, by poszczególne pary badaczy powtarzały się z równą częstością. W efekcie, każdy z koderów analizował 50 orzeczeń, z czego każde 10 było analizowane również przez innego z pozostałych pięciu koderów. W ten sposób uniknięto tworzenia „stałych par" koderów, co mogłoby się przyczynić do zmniejszenia intersubiektywnej sprawdzalności wyniku.

W efekcie przeprowadzonych działań badawczych przypisano kody do 15 080 fragmentów uzasadnień orzeczeń (tj. do około 50 fragmentów na każde uzasadnienie na jednego kodera). Objętość zakodowanych fragmentów tekstu wynosiła 14 896 778 znaków, co oznacza, że wybrane do badania uzasadnienia orzeczeń zostały pokryte kodami przez każdego kodera średnio około dwukrotnie (tj. każdy znak w każdym uzasadnieniu był przypisany do co najmniej dwóch fragmentów tekstu określonych przez kodera i oznaczonych różnymi kodami). Pozwala to sądzić, że przeprowadzona analiza miała charakter kompletny, to znaczy objęła całość materiałów przeznaczonych do badania, ale jednocześnie, że możliwość wielokrotnego kodowania tych samych fragmentów tekstu różnymi kodami przez tego samego kodera nie była często wykorzystywana.

Na tym etapie badania do zbioru danych dołączono też informacje o poszczególnych orzeczeniach, zakodowane jako atrybuty poszczególnych plików. Były to następujące zmienne: sygnatura orzeczenia, rok jego wydania, rodzaj rozstrzygnięcia Trybunału (zgodność, niezgodność, brak niezgodności poszczególnych przepisów z konstytucją lub innymi aktami albo umorzenie postępowania)[21], liczba zdań odrębnych, wielkość składu sędziowskiego. Dodatkowo dokonano syntetycznej oceny zgodności rozstrzygnięcia Trybunału z przypuszczalnymi celami przedłożonego mu wniosku, kodując zgodność lub niezgodność orzeczenia z żądaniem wnioskodawcy za pomocą zmiennej zerojedynkowej[22]. Podczas analizy informacje te posłużyły jako zmienne wyjaśniające zaobserwowaną zmienność wzorców argumentacji TK.

Treści wygenerowane w efekcie opisanej procedury nie były poddawane dalszej obróbce czy modyfikacjom, a ich analiza – stanowiąca trzecią część procedury badawczej – miała charakter ściśle ilościowy. W szczególności, na podstawie tak wygenerowanych kodów, nie budowano teorii w rozumieniu metodologii ugruntowanej. Nie próbowano też uzgadniać sposobu kodowania wybranych fragmentów pomiędzy badaczami. Przeciwko takim działaniom przemawiała znaczna objętość kodowanego materiału, rezygnacja z kodowania *in vivo* podczas zasadniczego etapu analizy, a także doświadczenia związane z sesjami uzgodnieniowymi

[21] Ponieważ w każdym orzeczeniu możliwe jest wydanie dowolnej kombinacji decyzji w odniesieniu do poszczególnych przepisów stanowiących przedmiot orzekania, w badaniu kodowano tę kwestię jako liczbę rozstrzygnięć poszczególnego rodzaju w danym orzeczeniu.

[22] Tę część prac badawczych przeprowadził student WPiA UW, pan Jędrzej Maśnicki.

na etapie pilotażowym, które dowiodły znacznej pracochłonności takich działań w sześcioosobowej grupie badaczy.

Dla celów analizy statystycznej wykorzystano dwie syntetyczne miary: względną objętość kodu w całym korpusie uzasadnień orzeczeń oraz jego częstość. Ta pierwsza zmienna została obliczona jako stosunek sumy objętości wszystkich fragmentów zakodowanych danym kodem (w znakach) do sumy objętości wszystkich fragmentów zakodowanych wszystkimi kodami danego poziomu. Wartości tej drugiej ustalono jako stosunek liczby użyć danego kodu we wszystkich orzeczeniach do sumy użyć wszystkich kodów we wszystkich orzeczeniach. Tak zdefiniowane miary posłużyły do stworzenia tabeli kontyngencji poszczególnych kodów i analiz korelacyjnych (za pomocą miary właściwej dla skal nominalnych, tj. V Cramera) oraz (po przeprowadzeniu stosownego przekształcenia kodów ze skali nominalnej na zerojedynkową) także do przeprowadzenia analizy logitowej. Dla ustalenia istotności wyników zawartych w tabelach kontyngencji użyto testu Chi^2 [23]. Jego zastosowanie dowiodło, że wszystkie ustalenia powzięte w toku analizy były istotne statystycznie, zazwyczaj na poziomach istotności $p<0,01$. Przeprowadzono także analizę rozkładu kodów (tj. zróżnicowania korpusu ze względu na kody w poszczególnych orzeczeniach) oraz współwystępowania kodów pierwszego poziomu.

Warto zauważyć, że pewną komplikację w analizie danych i prezentacji wyników stanowi konieczność uwzględnienia w nich triangulacji personalnej. Ponieważ każde z analizowanych uzasadnień orzeczeń było kodowane niezależnie przez dwóch koderów, niecelowe było podawanie wyników badania w miarach bezwzględnych (tj. liczby wystąpień poszczególnych kodów w korpusie). Takie działanie znacznie zawyżyłoby rzeczywiste występowanie poszczególnych kodów w korpusie, prowadząc do błędnych wniosków.

Rozwiązaniem tego problemu może być podawanie wspomnianych miar względnych. Takie działanie może być jednak obciążone błędem wynikającym z różnic w sposobie kodowania tego samego materiału przez różnych koderów, zwłaszcza związanych z różną szczegółowością czy „gęstością" kodowania. Może się mianowicie zdarzyć, że jeden z pary koderów ma tendencję do wyodrębniania dłuższych segmentów tekstu, zaś drugi do dzielenia go na krótsze cząstki. W takiej sytuacji (nawet gdy w parze koderów zostanie zachowana całkowita zgodność w kwestii kodu użytego do oznaczenia danego fragmentu) syntetyczna miara częstości kodów będzie w większym stopniu odzwierciedlać sposób kodowania drugiego z koderów, zaś miara długości – pierwszego. Z tego względu w badaniu wykorzystano proste wagi analityczne, zmniejszające wpływ stylu kodowania poszczególnych koderów na wynik. Zostały one ustalone dla każdego analizowanego uzasadnienia, osobno dla miar względnej objętości i częstości. Wagi te obliczono jako różnicę między liczbą 1 a ilorazem całkowitej objętości (liczebności) fragmentów tekstu zakodowanych w danym orzeczeniu przez danego kodera oraz sumą objętości (liczebności) fragmentów tekstu zakodowanych w danym orzeczeniu przez obydwu koderów, którzy je

[23] Analizę istotności przeprowadzono dla bezwzględnych wartości obydwu wspomnianych miar.

kodowali. Porównanie wyników uzyskanych za pomocą danych ważonych i nieważonych nie prowadziło jednak do zmiany zasadniczych wniosków płynących z badania. Można to w jakimś stopniu wiązać z zastosowaną techniką triangulacji, przewidującą brak stałych par koderów.

Zapewne najbardziej interesujące zagadnienie analityczno-metodologiczne, które pojawiło się podczas analizy danych, wiąże się z kwestią rzetelności kodowania (*intercoder reliability*). Zastosowanie triangulacji personalnej w naturalny sposób prowadzi mianowicie do potrzeby oceny zgodności pomiędzy sposobem kodowania tych samych fragmentów tekstu przez różnych koderów (Neuendorf 2002: 141 i nast.; Krippendorff 2004a: 211 i nast.). W systemie R obliczenie statystycznych miar rzetelności nie nastręcza większego problemu, ponieważ dostępne pakiety dodatkowe dostarczają gotowych procedur pozwalających wyznaczyć rozmaite współczynniki, takie jak Kappa Cohena, współczynnik ICC czy Alfa Krippendorfa (w badaniu użyto pakietu dodatkowego „irr"). W badaniu zdecydowano się wykorzystać, jako posiadającą najbardziej korzystne własności matematyczne, Alfę Krippendorfa (Krippendorff 2004b).

Wszystkie jednak miary tego rodzaju, znane autorowi analizy – i wszystkie zaimplementowane w środowisku R – zakładają porównanie kodów zastosowanych do określonej z góry jednostki analizy, to znaczy opierają się na porównaniu kodów użytych przez różnych koderów do oznaczenia tych samych fragmentów tekstu. Jak to zostało powiedziane, w przypadku omawianych badań taka jednostka analizy nie istniała. Każdy z pary koderów kodujących to samo orzeczenie mógł zastosować każdy z 50 kodów na oznaczenie fragmentów tekstu pozostających w dowolnych relacjach, co uniemożliwiało bezpośredni pomiar rzetelności.

Literatura odnosząca się do pomiaru rzetelności kodowania (na przykład Carey, Morgan, Oxtoby 1996; Lombard, Snyder-Duch, Campanella Bracken 2002; Krippendorff 2004a: 211 i nast.; 2004b; 2008; Carey i in. 2004 i tam cytowane prace) nie zawiera wielu propozycji rozwiązania tego problemu. Jedną z możliwości stanowi oddzielenie segmentacji tekstu od przypisywania kodów do segmentów (Campbell i in. 2013), to znaczy zastosowanie procedury, w której jeden z badaczy przeprowadza pełne, swobodne kodowanie, a pozostali jedynie kodują wskazane przez niego segmenty, ale nie biorą udziału w ich definiowaniu. Wadą tego rozwiązania jest to, że nie pozwala ono na dokonanie triangulacji samego schematu segmentacji, a więc to, iż w tym względzie w istocie nie różni się od sytuacji, gdy jednostka analizy jest z góry ustalona.

Inna autorka (Kurasaki 2000) proponuje stosowanie w takiej sytuacji losowego próbkowania porównywanych fragmentów tekstu i stosowania miar rzetelności tylko do nich. Rozwiązanie to, choć jest niewątpliwie eleganckie i pozwala na ominięcie problemu segmentacji kodowanego tekstu, jest także obarczone pewnymi wadami. Powoduje zwłaszcza, że zagadnienie rzetelności kodowania nabywa charakteru podwójnie probabilistycznego[24].

[24] Już w sytuacji kompletnego pomiaru rzetelności kodowania w badaniach na materiale stanowiącym próbę pojawia się problem wnioskowania na temat (hipotetycznej) rzetelności kodowania w całej populacji. Zastosowanie próbkowania będzie, rzecz jasna, obniżać ufność pomiaru rzetelności.

Prowadzi to do dalszych komplikacji, takich jak na przykład wpływ wielkości próbkowanego fragmentu i liczby próbek na istotność oszacowania rzetelności. To z kolei wpływa na pewność wyników całego badania, które same mają charakter probabilistyczny (można twierdzić, że za miarę ufności badania należałoby przyjąć iloczyn poziomu ufności wyniku i poziomu ufności rzetelności kodowania)[25].

Nie chcąc rozstrzygać takich problemów – jako wtórnych wobec głównego tematu badań – zastosowano rozwiązanie polegające na wtórnej (sztucznej) segmentacji już zakodowanego tekstu. Ponieważ rozwiązanie takie nie jest często spotykane, przeprowadzono swojego rodzaju eksperyment metodologiczny, polegający na dokonaniu pomiaru rzetelności kodowania przy wykorzystaniu kilku różnych schematów segmentacji tekstu i zestawieniu ze sobą uzyskanych wyników.

W pierwszym wariancie tej analizy za segment tekstu poddawany porównaniu w parze koderów, którzy go kodowali, uznano całe uzasadnienie orzeczenia. Działanie to opierało się na przeświadczeniu, że można opisać dane uzasadnienie za pomocą wartości objętości względnej i częstości każdego z kodów użytych w tym orzeczeniu, traktując owe wartości jako „zmierzoną" przez *i*-tego kodera „intensywność" występowania tego kodu w danym uzasadnieniu[26]. Dokonując takich obliczeń dla każdego kodera, kodu i uzasadnienia, uzyskano tabelę zawierającą 15 000 *data points* (50 orzeczeń * 50 kodów * 6 koderów). Następnie porównano „intensywność" występowania poszczególnych kodów w tych samych orzeczeniach kodowanych przez poszczególne pary koderów, za stan idealnej zgodności kodowania uznając taką sytuację, w której „intensywność" występowania danego kodu w danym orzeczeniu byłaby taka sama według *i*-tego i *j*-tego kodera kodującego to orzeczenie[27]. Jako miarę rzetelności wykorzystano Alfę Krippendorfa w wariancie przewidzianym dla skali ilorazowej.

Drugi wariant segmentacji polegał na tym, że analizę przeprowadzono na arbitralnie wydzielonych 1000-znakowych fragmentach uzasadnień orzeczeń, którym przypisano występujące w nich kody. W przypadku, gdy w danym fragmencie występował więcej niż jeden kod, traktowano je jak dwa lub więcej osobne fragmenty (nieliczne fragmenty niezakodowane żadnym kodem wyłączano z analizy). Następnie dokonano porównania wszystkich odpowiadających sobie par fragmentów uzasadnień orzeczeń, zakodowanych przez różnych koderów, stwierdzając, czy występuje pomiędzy nimi zgodność zastosowanego kodu i czy w ogóle obydwaj koderzy zakodowali ten fragment jakimkolwiek kodem. Na tej podstawie obliczono Alfę Krippendorfa dla każdego kodu i każdego kodera (porów-

[25] Teoretycznym rozwiązaniem może być tu zastosowanie „pełnego próbkowania", to znaczy porównywanie kodów zastosowanych przez różnych koderów w tym samym tekście dla każdego znaku. Rozwiązanie takie będzie – gdy chodzi o stosowany algorytm – zbliżone do drugiej z metod segmentacji, opisanych poniżej. Jego wadą jest jednak to, że prowadzi do wygenerowania dużej ilości danych (trójwymiarowej tabeli o rozmiarach x = długość porównywanego tekstu, y = liczba kodów, z = liczba koderów).

[26] Z zastrzeżeniem, że suma wartości zarówno objętości względnej, jak i częstości dla wszystkich kodów mogła dla każdego orzeczenia wynosić więcej niż 1 ze względu na wielokrotne kodowanie.

[27] Trzeba podkreślić, że tego rodzaju procedura opiera się na porównaniu zgodności wartości intensywności kodów dla danego orzeczenia, nie zaś współwystępowania samych kodów.

nując sposób kodowania *i*-tego kodera i wszystkich pozostałych).

W trzecim wariancie pomiaru rzetelności przyjęto, że sposób kodowania porównuje się pomiędzy wszystkimi fragmentami tekstu pochodzącymi z tego samego uzasadnienia i zakodowanymi przez różnych koderów, jeśli fragmenty te pozostają w jednej z następujących relacji: pokrywanie się (fragmenty zakodowane przez różnych koderów rozpoczynają się w tym samym miejscu orzeczenia i kończą się w tym samym miejscu), zawieranie się (jeden z fragmentów w całości obejmuje drugi), przecinanie się (jeden z fragmentów rozpoczyna się w miejscu, które zawiera tekst drugiego) lub bliskość (tekst obydwu fragmentów jest rozłączny, ale początek jednego z nich znajduje się w odległości nie mniejszej niż 50 znaków od końca drugiego)[28]. W tym celu wykorzystano Alfę Krippendorfa w wariancie przewidzianym dla skali nominalnej.

Konsekwencją zastosowania dwóch pierwszych z wymienionych schematów segmentacji tekstu i metod obliczania rzetelności było to, że porównanie rzetelności mogło odnosić się jedynie do poszczególnych kodów, nie zaś do wszystkich zakodowanych fragmentów. Tylko zastosowanie trzeciego schematu segmentacji pozwoliło na obliczenie całościowej rzetelności kodowania dla fragmentów tekstu zakodowanych przez *i*-tego kodera i wszystkich pozostałych koderów, i wszystkich kodów.

Uzyskane w efekcie każdej z procedur wartości Alfy Krippendorfa odbiegały od poziomów uznawanych za satysfakcjonujące – 0,8 i wystarczające – 0,67 (por. Krippendorff 2004a: 241; 2008). W przypadku pierwszych dwóch mechanizmów segmentacji tekstu, uzyskane miary były też bardzo zróżnicowane i wyraźnie zależne od częstości użycia i objętości poszczególnych kodów. W przypadku niektórych kodów i niektórych koderów występowały ujemne wartości Alfy Krippendorfa. W przypadku innych kodów uzyskano wartości przekraczające 0,67. Zaobserwowano również takie sytuacje, w których niektórzy z koderów uzyskiwali niskie lub bardzo niskie wartości Alfy, zaś inni, w odniesieniu do tych samych kodów – wysokie lub bardzo wysokie. Gdy wreszcie chodzi o wartości uzyskane w efekcie trzeciego z opisanych algorytmów, to pozostawały one w przedziale 0,29–0,41.

Wyniki te można interpretować na kilka sposobów. Po pierwsze, mogą one świadczyć o rzeczywistych różnicach w interpretacji tych samych fragmentów tekstu przez różnych koderów. Wydaje się to prawdopodobne zwłaszcza w świetle znacznego poziomu skomplikowania kodowanego tekstu. Po drugie, można sądzić, że miary te stanowią w pewnej mierze artefakt wynikający ze względnie niskiej częstości niektórych kodów (ich wąskiego zakresu) lub sposobu przeprowadzenia obliczeń rzetelności. Po trzecie, można uważać, że mogą być one związane z nieostrością samych kodów lub różnicami w ich interpretacji.

Dla poprawy rzetelności wyników badania i choć częściowego wyjaśnienia tych wątpliwości, w toku

[28] Warto zauważyć, że w języku R implementacja takiego sposobu segmentacji może okazać się bardzo zasobochłonna i w efekcie powolna. Najprostszy (z punktu widzenia standardowej implementacji informatycznej) algorytm segmentacji, który został zastosowany w omawianych badaniach, opierający się na zagnieżdżonych pętlach *for*, okazał się bardzo powolny (jego wykonanie na danych z badania trwało kilka godzin na komputerze z czterordzeniowym procesorem) i powinien zostać przebudowany w taki sposób, aby wykorzystywał działania na wektorach. Ma to związek z ogólniejszą architekturą systemu R, która zakłada właśnie taki mechanizm operowania na danych.

dalszej analizy zdecydowano się na przeprowadzenie agregacji kodów (to jest połączenie ze sobą wybranych kodów w kody o szerszym zakresie i w konsekwencji objętości zakodowanego tekstu i częstości wystąpień równej sumie objętości i wystąpień kodów „składowych"). Ponieważ wynik tego działania zależy od przyjętego schematu agregacji, który z kolei może być zależny od indywidualnych preferencji osoby definiującej ten schemat, agregację zrealizowano zgodnie ze specjalnym algorytmem.

Przedstawiał się on następująco. W pierwszej kolejności, każdy z sześciu koderów biorących udział w badaniu przedstawił własny schemat agregacji kodów pierwszego poziomu. Następnie, dla każdego zaproponowanego schematu agregacji przeprowadzono analizę rzetelności kodowania za pomocą trzeciego z wymienionych wyżej algorytmów, modyfikując przy tym założenie dotyczące warunku bliskości kodów, tak aby uwzględniał trzy warianty. Odległość pomiędzy porównywanymi fragmentami wynosiła w nich odpowiednio 5, 50 i 1000 znaków[29]. Wreszcie, jako kryterium wyboru docelowego schematu agregacji przyjęto wartości Alfy we wszystkich trzech wariantach, po uśrednieniu wyników uzyskanych dla wszystkich sześciu koderów objętych badaniem.

Zgodnie z oczekiwaniami, agregacja przyniosła istotną poprawę wskaźników rzetelności kodowania. W najlepszym z badanych schematów agregacji Alfa Krippendorfa wynosiła od 0,5 do 0,7, w zależności od kodera. Ostatecznie zredukowano liczbę kodów pierwszego poziomu z 37 do 12, a kodów drugiego poziomu – z 13 do 3. Można sądzić, że taka procedura agregacji stanowi w istocie swoisty odpowiednik sesji uzgodnieniowych i, w pewnym zakresie, funkcjonalny ekwiwalent budowania teorii ugruntowanej w drodze tworzenia grup kodów[30].

CAQDA a socjologiczna analiza orzecznictwa sądowego

Niezależnie od zasadniczej potrzeby uzyskania wysokich miar rzetelności, warto jednak zauważyć, że wykorzystanie triangulacji personalnej w badaniach opierających się na opisywanym schemacie prowadzi do interesującego dylematu. Nie jest mianowicie zupełnie jasne, czy sytuacja występowania całkowitej czy bardzo wysokiej zgodności pomiędzy koderami jest stanem zawsze pożądanym. Odpowiedź na to pytanie wydaje się oczywista w przypadku pomiaru rzetelności kodowania prostych tekstów (na przykład krótkich odpowiedzi na pytania otwarte w badaniach ankietowych). W takim przypadku interpretacja wypowiedzi badanych przez poszczególnych kode-

[29] Odległość między fragmentami może mieć znaczenie dla wartości Alfy Krippendorfa, bo im większa jest akceptowana odległość, tym więcej fragmentów tekstu jest porównywanych pod względem sposobu zakodowania. W konsekwencji, z jednej strony, wraz ze wzrostem akceptowanej odległości wzrasta liczba zaobserwowanych niezgodności pomiędzy parami kodów zastosowanych przez różnych koderów, ale też zwiększa się prawdopodobieństwo wystąpienia par zgodnych. Opisany dobór odległości stanowi swojego rodzaju eksperyment metodologiczny, związany z nieostrością wypowiedzi podlegających kodowaniu i nieostrością samych zastosowanych kategorii. W szczególności przyjęcie odległości 1000-znakowej może wydawać się przesadne, jednak opierało się na przeświadczeniu, że dostrzeżone przez jednego z koderów aspekty wypowiedzi TK mogą przez drugiego z nich zostać zaobserwowane w innych miejscach tekstu, w tym dość odległych.

[30] Jakkolwiek w omawianych badaniach dalsza analiza danych miała charakter ilościowy, otwiera to interesujące możliwości wypracowania komputerowo wspomaganej metodologii jakościowego budowania teorii na podstawie materiału zgromadzonego przez kilkuosobowy zespół badawczy, bez potrzeby pracochłonnego „interakcyjnego" uzgadniania wykorzystanych kodów. Możliwe jest również zastosowanie takiej procedury agregacji kodów, która maksymalizowałaby rzetelność kodowania w sposób czysto matematyczny, bez potrzeby formułowania konkurencyjnych propozycji agregacji przez samych badaczy. Osobną kwestią stanowiłaby tu, oczywiście, łatwość czy nawet możliwość interpretacji tak wygenerowanych danych zagregowanych.

rów powinna być tak zbieżna, jak to tylko możliwe. W przypadku bardziej złożonych tekstów, twierdzenie to można jednak kwestionować, twierdząc, że triangulacja powinna polegać właśnie na uwzględnieniu w badaniu *różnych* perspektyw na ten sam tekst.

Przypadek kodowania wypowiedzi Trybunału Konstytucyjnego jest tego dobrą ilustracją. Jak to zostało powiedziane, uzasadnienia orzeczeń TK charakteryzują się bardzo dużą złożonością i mogą podlegać różnorodnym interpretacjom, z których trudno jest wybrać interpretację ostateczną czy „prawidłową". Istnieje ku temu kilka powodów. W tym samym fragmencie swojego uzasadnienia TK może wykorzystywać kilka typów argumentów czy metod wykładni prawa jednocześnie i może to robić z różną intensywnością. Niekiedy akcenty położone na poszczególne kwestie mogą być bardzo subtelne. Dlatego różni badacze, którzy mogą przeprowadzać kodowanie na różnym poziomie szczegółowości, mogą dostrzegać wszystkie lub tylko niektóre aspekty tego samego fragmentu tekstu, co będzie prowadzić do rozbieżności mierzalnych za pomocą miar rzetelności. Nawet w sytuacji, w której kodowanie tego samego fragmentu nie jest wielokrotne (tj. na jego opisanie użyto tylko jednego kodu, a nie kilku), te same wypowiedzi mogą też nie być konkluzywne i w konsekwencji mogą być kodowane za pomocą różnych kodów. Zatem w przypadku zrealizowania badania pozbawionego triangulacji, uzyskane dane będą silnie zniekształcone pod wpływem indywidualnej percepcji badacza. W konsekwencji, w procedurze badawczej wykorzystującej triangulację, wyższy poziom niezgodności pomiędzy badaczami może być przejawem nie tyle rzeczywistych odmienności pomiędzy nimi, co nieostrości wypowiedzi poddanych badaniu.

Warto również zauważyć, że uzasadnienia orzeczeń są wytwarzane we względnie złożonym procesie społecznym, co może negatywnie odbijać się na ich komunikatywności i ścisłości. Skład sędziowski rozstrzygający daną sprawę może, w przypadku TK, liczyć nawet kilkunastu orzeczników. Wprawdzie za przygotowanie pisemnego uzasadnienia wyroku jest formalnie odpowiedzialny jeden sędzia (w szczególnie złożonych sprawach może być ich więcej), ale w rzeczywistości w procesie tym udział bierze więcej osób. Wstępne wersje dokumentu są przygotowywane przez asystentów, a ostateczna treść uzasadnienia jest przedmiotem negocjacji (jeśli nie przetargu) pomiędzy poszczególnymi sędziami wchodzącymi w skład panelu orzekającego. Z tego względu ostateczny dokument może zawierać niespójności wewnętrzne, luki i różnego rodzaju niejasności.

Co być może jest jeszcze bardziej istotne, kody wykorzystane w opisywanych badaniach nie są zupełnie ostre. Jest to efekt wspomnianej wielości podejść teoretycznych, które zostały, *nolens volens*, zaimportowane do badania w efekcie zastosowania opisanej wyżej procedury wytwarzania kodów przez badaczy będących z wykształcenia prawnikami i specjalizujących się w teorii prawa. Posługiwali się oni, w naturalny dla nich sposób, kategoriami znanymi tej dziedzinie wiedzy, a nie pojęciami całkowicie sztucznymi czy wytworzonymi w ramach teorii socjologicznej[31]. W konsekwencji, kody te są

[31] Jest rzeczą wątpliwą, czy tego rodzaju kategorie teoretyczne, na obecnym etapie rozwoju socjologii prawa i socjologii teoretycznej, w ogóle istnieją. Warto tu odwołać się do obserwacji Bruno Latoura, który kwestionując możliwość ich istnienia, w następujący sposób scharakteryzował swoją pracę nad działalnością francuskiej Conseil d'État: „[w]ydaje się, że w prawie nie ma żadnego stopniowania: albo się w nim jest bez reszty, albo wcale i mówi się o innych rzeczach. Jego warunki fortunności (*conditions of felicity*)

niekiedy obciążone istotnym ładunkiem znaczeniowym.

Wreszcie, wiele z zastosowanych kategorii ma charakter raczej profesjonalnej wiedzy habitualnej, skodyfikowanej i zinstytucjonalizowanej w ramach profesji prawniczych, czy wręcz ideologicznej w sensie Mannheima, niż wiedzy ściśle teoretycznej, to znaczy dającej się sprowadzić do w pełni eksplikowalnych założeń epistemologicznych, ontologicznych czy deontologicznych. W efekcie niektóre przypadki niektórych typów wykładni mogą dać się zakwalifikować również jako przypadki innych typów wykładni. Przykład tego stanu rzeczy mogą stanowić niektóre typy tak zwanej wykładni językowej.

Istotę zagadnienia stanowi tu fakt, że w badaniach uzasadnień orzeczeń sądowych uniknięcie stosowania tego typu kategorii może nie być możliwe, a jeśli się to stanie, może prowadzić do negatywnych skutków poznawczych. Możliwe jest, rzecz jasna, zaprojektowanie takich narzędzi badawczych, które będą ścisłe w tym sensie, że ich interpretacja *in abstracto* nie będzie rodziła sporów pomiędzy koderami. Można to uczynić choćby porzucając metodę kodowania *in vivo* i definiując kody w sposób arbitralny i dedukcyjny lub wykorzystując do ich wytwarzania osoby niemające wykształcenia prawniczego i niedysponujące odpowiednim kapitałem symbolicznym. Rodzi to jednak zagrożenie, że analiza orzecznictwa sądowego z wykorzystaniem takich kategorii mogłaby nie być w ogóle możliwa lub nie doprowadziłaby do zrozumiałych wyników. W każdym razie, z definicji, opierałaby się na zerwaniu ze zbiorem kategorii znanych dyskursowi, do którego się odnosi, a byłaby dokonywana w pojęciach trudnych do translacji. Z tego względu jej wyniki mogłyby być niekomunikowalne w tym dyskursie.

Dlatego można byłoby sądzić, że uzyskanie wysokich miar rzetelności w badaniu na podobnym materiale zrealizowanym zgodnie z opisywaną procedurą badawczą wskazywałyby na zbliżone profesjonalne habitusy uczestników badania. W konsekwencji, osoba pochodząca spoza środowiska badaczy, kierując się inną ideologią czy zapatrywaniami teoretycznymi, mogłaby zinterpretować te same fragmenty orzeczeń zupełnie inaczej. Istnienie rozbieżności pomiędzy koderami może natomiast wskazywać na to, że wynik badania jest ostatecznie bardziej przekonujący, ponieważ koderzy odnosili się do szerszego zbioru przekonań i możliwości interpretacyjnych niż tylko jeden, partykularny punkt widzenia. Ten wniosek byłby zapewne tym bardziej przekonujący, im większa i bardziej zróżnicowana byłaby grupa badaczy[32].

Najbardziej może interesujący wniosek, jaki płynie z tych obserwacji, dotyczy jednak ideologicznego charakteru kategorii pojęciowych wykorzystanych w omawianym badaniu. Zaobserwowane w badaniu istotne rozbieżności w kodowaniu tych samych fragmentów tekstu przez różnych badaczy ujawniają mianowicie to, że pozornie jednoznaczne kategorie stosowane

i niefortunności są wyjątkowo ostre. Etnograf zauważył to wielokrotnie, uświadamiając (…) sobie skalę własnej niemożności, tego by nawet po latach studiów, stopniowo zbliżyć się do wypowiedzi prawniczych. Mówienie językiem prawa było dla niego nieosiągalne nie tylko z braku słów i pojęć, ale z braku wszystkiego, absolutnie wszystkiego. Aby stwierdzić coś po prawniczemu, musiałby zostać radcą stanu. (…) Aby opisać Prawo w sposób przekonujący, trzeba do niego już wcześniej wskoczyć" (2010: 255 [tłum. własne]).

[32] W tym kontekście, jako swoisty postulat metodologiczny, można byłoby sugerować swoistą kalibrację wykorzystywanych narzędzi poprzez syntetyczny pomiar różnic i podobieństw między habitusami badaczy w zakresie kodowanego materiału. W przypadku opisywanego tu schematu badawczego, wymagałoby to przeprowadzenia dodatkowej fazy badania: po ustaleniu listy kodów, należałoby przeprowadzić drugą fazę pilotażu, polegającą na kodowaniu dobranych celowo, tych samych orzeczeń TK.

przez prawników do opisywania podejmowanych w tym środowisku działań umysłowych w rzeczywistości są dość rozmyte i w znacznym zakresie podlegają interpretacji. Fakt ten pozostaje niezauważalny dopóty, dopóki nie przeprowadzi się systematycznego zestawienia użyć takich kategorii, dokonanych przez kompetentnych przedstawicieli tego zawodu.

Ta z kolei obserwacja daje się uogólnić poza problematykę socjologicznych badań nad orzecznictwem sądowym. Może mianowicie prowadzić do wątpliwości dotyczących wykorzystania wiedzy *insiderskiej* – czy szerzej, *insiderskiego* habitusu – w badaniach wykorzystujących metodologię postępowania badawczego, obcą temu habitusowi. Zmuszając koderów-*insiderów* do rezygnacji ze zwykłego sposobu użycia własnych kategorii i zastępując go systematyczną metodologią badań społecznych, niejako wyostrza się i unaocznia różnice pomiędzy rozumieniem tych samych pojęć przez różnych przedstawicieli tego samego środowiska. Innymi słowy, metodologia badań, z jej systematycznymi procedurami analizy treści, zwłaszcza wymuszonymi przez narzędzia informatyczne w CAQDA, występuje w roli analogicznej do Garfinkelowskiego wandalizmu interakcyjnego (Garfinkel 2007: 58 i nast.).

Taka alienacja badaczy od własnych, środowiskowych pojęć może prowadzić do ujawnienia ideologicznej natury tych kategorii, a efekt ten daje się mierzyć za pomocą współczynników rzetelności kodowania. Jest to rzecz jasna problem, który wymaga rozważenia za każdym razem, kiedy próbuje się zastosować *insiderskie* kategorie wypracowane w ramach grup zawodowych do prowadzenia *outsiderskich* badań praktyk dyskursywnych prowadzonych w takich grupach. Obserwację tę można ująć w następujący trylemat. Po pierwsze, można sądzić, że zjawisko alienacji będzie tym bardziej widoczne, im bardziej rygorystyczne będą procedury analizowania i pomiaru rzetelności kodowania. Zastosowanie oprogramowania CAQDA niewątpliwie może się do tego przyczynić, unaoczniając różnice znaczeniowe zapoznawane w toku zwykłego dyskursu. Z drugiej strony, wypracowanie metodologii badania, która nie będzie prowadzić do alienacji, może negatywnie ciążyć na wiarygodności uzyskanych wyników. Wreszcie, po trzecie, zastosowanie kategorii zupełnie obcych badanemu dyskursowi może nie przynieść spodziewanych efektów poznawczych.

Wnioski

Zaprezentowane obserwacje i doświadczenia powinny przekonywać, że realizacja badań o opisanym, szerokim zakresie i przedstawionych założeniach poznawczych nie byłaby możliwa bez wykorzystania oprogramowania CAQDA. Narzędzia wykorzystane w badaniu pozwoliły zwłaszcza na jego przeprowadzenie w formule mieszanej, jakościowo-ilościowej analizy treści z kodowaniem swobodnym i triangulacją. Jakkolwiek możliwość kwantyfikacji wyników nie musi stanowić sama w sobie zalety, to realizacja badań w skali przedstawionej w artykule, i przy ograniczonych zasobach, jakie były do dyspozycji ich autorów, jest trudna do wykonania tylko w formie jakościowej. Na uwagę zasługuje tu zwłaszcza możliwość obliczania rzetelności kodowania przy zachowaniu swobodnego charakteru kodowania oraz duża elastyczność zastosowanego oprogramowania, gdy chodzi o sposób prezentacji wyników w formie syntetycznych wskaźników.

Nie ulega też wątpliwości, że opisywane badania nie wyczerpały całego zakresu form badania i możliwości analizy oferowanych przez zastosowane w nich oprogramowanie. Za opcje godne bliższego rozpoznania należy zwłaszcza uznać możliwość automatyzacji niektórych działań badawczych, prowadzenia badań w większych i rozproszonych zespołach, a także wykorzystania środowiska WWW do prezentacji wyników *in vivo*.

Negatywną stroną zestawu oprogramowania zastosowanego w badaniach jest konieczność posiadania umiejętności obsługi pakietu R, co wymaga znajomości podstawowych technik programistycznych. Wykorzystanie wszystkich możliwości oferowanych przez ten system, a zwłaszcza tworzenie procedur obliczeniowych, które byłyby efektywne pod względem wykorzystania zasobów komputerów, na jakich są uruchamiane (co może być istotne na przykład w przypadku dużych badań realizowanych przez rozproszony zespół badawczy albo środowiska WWW), wymaga natomiast pogłębionej wiedzy takiego rodzaju.

Stosunkowo duże wymagania pakietu R nie powinny jednak przysłonić faktu, że opisywany projekt dowiódł względnej łatwości korzystania z graficznej nakładki RQDA także przez użytkowników, którzy nie mają większego doświadczenia w pracy z oprogramowaniem tego rodzaju. Wreszcie, godną podkreślenia zaletą opisywanego oprogramowania jest jego dostępność na wolnej licencji i związana z tym nieodpłatność.

Prowadzone badania, oprócz otrzymania względnie nowatorskich wyników i uzyskania praktycznej wiedzy o stosunkowo mało znanym narzędziu badawczym, doprowadziły także do zadania interesujących pytań metodologicznych. Dotyczą one tak strony czysto technicznej badań – obliczania rzetelności kodowania w swobodnej analizie treści, jak i zagadnień epistemologicznych, związanych z możliwościami badania wyspecjalizowanego dyskursu profesjonalnego w kategoriach znanych temu dyskursowi i mu obcych. Obydwie te kwestie trudno uznać za ostatecznie rozwiązane w toku opisywanych badań, co dowodzi, że powinny być one przedmiotem dalszego namysłu.

Bibliografia

Alexy Robert (1985) *Theorie der Grundrechte*. Baden-Baden: Nomos.

------ (1996) *Theorie der juristischen Argumentation: Die Theorie des rationalen Diskurses der juristischen Begriffs*. Frankfurt am Main: Suhrkamp.

Campbell John L. i in. (2013) *Coding In-Depth Semistructured Interviews: Problems of Unitization and Intercoder Reliability and Agreement*. „Sociological Methods and Research", vol. 42(3), s. 294–320.

Carey James W., Morgan Mark, Oxtoby Margaret J. (1996) *Intercoder Agreement in Analysis of Responses to Open-Ended Interview Questions: Examples from Tuberculosis Research*. „Cultural Anthropology Methods", vol. 8(3), s. 1–5.

Carey James W. i in. (2004) *Reliability in Coding Open-Ended Data: Lessons Learned from HIV Behavioral Research*. „Field Methods", vol. 16(3), s. 307–331.

Feteris Eveline T. (1999) *Fundamentals of Legal Argumentation: a Survey of Theories on the Justification of Judicial Decisions*. Hague: Kluwer.

Galligan Denis J., Matczak Marcin (2005) *Strategie orzekania sądowego o wykonywaniu władzy dyskrecjonalnej przez sędziów sądów administracyjnych w sprawach gospodarczych i podatkowych*. Warszawa: Ernst & Young.

Garfinkel Harold (2007) *Studia z etnometodologii.* Przełożyła Alina Szulżycka. Warszawa: Wydawnictwo Naukowe PWN.

Huang Ronggui (2012) *RQDA: R-Based Qualitative Data Analysis* [dostęp 28 września 2013 r.]. Dostępny w Internecie ‹http://rqda-.r-forge.r-project.org›.

Kelsen Hans (2009) *Istota i rozwój sądownictwa konstytucyjnego.* Przełożył Bolesław Banaszkiewicz. Warszawa: Trybunał Konstytucyjny.

Krippendorff Klaus (2004a) *Content Analysis An Introduction to Its Methodology*. Thousand Oaks: Sage.

------ (2004b) *Reliability in Content Analysis: Some Common Misconceptions and Recommendations*. „Human Communication Research", vol. 30(3), s. 411–433.

------ (2008) *Testing the Reliability of Content Analysis Data* [w:] Krippendorff Klaus, Bock Mary Angela, eds., *The Content Analysis Reader*. Thousand Oaks: Sage, s. 350–357.

Kurasaki Karen S. (2000) *Intercoder Reliability for Validating Conclusions Drawn from Open-Ended Interview Data.* „Field Methods", vol. 12, s. 179–194.

Latour Bruno (2010) *The Making of Law: An Ethnography of the Conseil d'Etat*. Cambridge: Polity Press.

Lautmann Rüdiger (1972) *Justiz – Die Stille Gewalt.* Frankfurt am Main: Suhrkamp.

Lombard Matthew, Snyder-Duch Jennifer, Campanella Bracken Cheryl (2002) *Content Analysis in Mass Communication Assessment and Reporting of Intercoder Reliability.* „Human Communication Research", vol. 28(4), s. 587–604.

Maravall Jose Maria (2003) *The Rule of Law as a Political Weapon* [w:] Przeworski Adam, Maravall José María, eds., *Democracy and Rule of Law*. Oxford: Oxford University Press, s. 261–301.

Morawski Lech (2010) *Zasady wykładni prawa*. Toruń: TNOiK „Dom Organizatora".

Neuendorf Kimberly A. (2002) *The Content Analysis Guidebook*. Thousand Oaks: Sage.

R Core Team (2012) *R: A Language and Environment for Statistical Computing* [dostęp 28 września 2013 r.]. Dostępny w Internecie ‹http://www.R-project.org›.

Ripley Brian D. (2001) *The R Project in Statistical Computing*, „MSOR Connections. The Newsletter of the LTSN Maths, Stats & OR Network", vol. 1(1), s. 23–25.

Sadurski Wojciech (2008) *Rights before Courts: A Study of Constitutional Courts in Postcommunist States of Central and Eastern Europe*. Dordrecht: Springer.

Scheffer Thomas (2010) *Adversarial Case Making. An Ethnography of English Crown Court Procedure*. Leiden: Brill.

Stawecki Tomasz, Winczorek Jan, red., (w druku) *Wzorce wykładni konstytucji w Polsce i państwach Europy Środkowej – doktryna i praktyka*. Warszawa: Wolters Kluwer.

Stone Sweet Alec (1999) *Judicialization and the Construction of Governance*. „Comparative Political Studies", vol. 32(2), s. 147–184.

Winczorek Jan, Stawecki Tomasz, Staśkiewicz Wiesław (2008) *Between Polycentrism and Fragmentation: The Impact of Constitutional Tribunal Rulings on the Polish Legal Order*. Warszawa: Ernst & Young.

Wronkowska Sławomira (2008) *Kilka uwag o „prawodawcy negatywnym"*. „Państwo i Prawo", t. 64(10), s. 5–21.

Wróblewski Jerzy (1988) *Sądowe stosowanie prawa*. Warszawa: PWN.

Wyrembak Jarosław (2009) *Zasadnicza wykładnia znamion przestępstw: pozycja metody językowej oraz rezultatów jej użycia*. Warszawa: Wolters Kluwer.

Zieliński Maciej (2008) *Wykładnia prawa: zasady, reguły, wskazówki*. Warszawa: LexisNexis.

Aneks

Tabela 1. Lista kodów wraz ze schematem ich agregacji. Kursywą zaznaczono kody drugiego poziomu.

KOD PO AGREGACJI	KODY PRZED AGREGACJĄ
argument językowy	argumenty semantyczne – argument semantyczny inny argumenty semantyczne – argument z autonomicznego znaczenia pojęć argumenty semantyczne – argument z języka naturalnego argumenty semantyczne – argument z języka prawnego argumenty semantyczne – argument z niejasności przepisu przywołanie tekstu normatywnego
argument systemowy	argumenty systemowe – argument z metod regulacji i gałęzi prawa argumenty systemowe – argument z typu przepisu argumenty systemowe – argument z założeń idealizacyjnych systemu prawa argumenty systemowe – argument systemowy inny argumenty systemowe – argumenty ze struktury systemu prawa argument z faktu trwałości lub zmiany prawa argument z formalnego obowiązywania prawa klasyczne argumenta prawnicze
argument systematyczny	argumenty systemowe – argumenty z budowy aktu prawnego
argument funkcjonalny	argument z faktów powszechnie znanych argument z istoty instytucji prawnej argument ze skutków społecznych argument ze zmiany społecznej
argument celowościowy i intencjonalny	argument z ratio legis odwołanie do prac legislacyjnych
argument komparatystyczny	argumenty komparatystyczne – argument komparatystyczny wewnętrzny argumenty komparatystyczne – argument komparatystyczny zewnętrzny
argument z własnego autorytetu	argumenty z orzecznictwa – argument z własnego orzecznictwa TK – bez uzasadnienia argumenty z orzecznictwa – argument z własnego orzecznictwa TK z uzasadnieniem *krytyka lub odrzucenie argumentu*
argument z autorytetu innego sądu	argumenty z orzecznictwa – argument z orzecznictwa sądów polskich argumenty z orzecznictwa – argument z orzecznictwa ETPCz argumenty z orzecznictwa – argument z orzecznictwa ETS argumenty z orzecznictwa – argument z orzecznictwa innego sądu zagranicznego
argument z doktryny	argument z doktryny

argument z dystrybucji władzy	argument z dyskrecjonalności organów władzy argument z zakresu władztwa TK
argument z interesu (publicznego, grupowego, indywidualnego)	odwołanie do interesów grupowych lub grup społecznych
argument z zasad prawa lub wartości prawnych i pozaprawnych	argument z zasady interpretowanej argument z zasady rozstrzygającej interpretację argument z proporcjonalności i ważenia (zasad, wartości) odwołanie do wartości i innych norm pozaprawnych
nie-argumenty (fragmenty uzasadnień wyłączone z analizy)	*argumentum ad rem* *elementy orzeczenia – konkluzja rozumowania* *elementy orzeczenia – powtórzenie* *elementy orzeczenia – zdanie odrębne* *narzędzia retoryczne*
konstytucja	*typy aktów prawnych – konstytucja*
krajowe akty normatywne inne niż konstytucja	*typy aktów prawnych – non-state law* *typy aktów prawnych – prawo europejskie pierwotne* *typy aktów prawnych – prawo europejskie wtórne* *typy aktów prawnych – prawo międzynarodowe publiczne* *typy aktów prawnych – ustawa* *typy aktów prawnych – akty nieustawowe*

Źródło: opracowanie własne.

Cytowanie

Winczorek Jan (2014) *Wykorzystanie oprogramowania R i RQDA w jakościowo-ilościowej analizie treści orzeczeń Trybunału Konstytucyjnego.* „Przegląd Socjologii Jakościowej", t. 10, nr 2, s. 138–161 [dostęp dzień, miesiąc, rok]. Dostępny w Internecie: ‹www.przegladsocjologiijakosciowej.org›.

Usage of Software Packages R and RQDA in a Qualitative-Quantitative Content Analysis of Verdicts of Polish Constitutional Tribunal

Abstract: The paper describes utilization of R statistical software with RQDA CAQDA package in a mixed qualitative-quantitative triangulated content analysis with free coding, intercoder reliability analysis, and statistical significance analysis, performed on a sample of justifications of verdicts of Polish Constitutional Tribunal, issued between 1986 and 2009. It presents the most important features of software used, assumptions of the study, its methodology and procedures of data analyses, as well as epistemic questions which arose in the course of its execution.

Keywords: CAQDA, RQDA, R (software), content analysis, intercoder reliability, Polish Constitutional Tribunal