

SYLWIA MĘCFAL 
Uniwersytet Łódzki

ANALIZA PRAKTYK ANONIMIZACJI DANYCH W POLSKICH JAKOŚCIOWYCH BADANIACH SOCJOLOGICZNYCH NA TLE LITERATURY PRZEDMIOTU ORAZ ZASAD ETYCZNO-PRAWNYCH

Streszczenie

Celem artykułu jest przedstawienie praktyk anonimizacji danych stosowanych w jakościowych badaniach społecznych polskich socjologów oraz stworzenie typologii stosowanych rozwiązań.

Aby sprawdzić, jak badacze prezentują sposoby anonimizacji danych, przeanalizowałam artykuły opublikowane w latach 2015–2019 w czterech ważnych polskich czasopismach socjologicznych.

Na podstawie analizy wyróżniłam różne sposoby anonimizacji danych stosowane przez autorów artykułów, które poddane dodatkowej analizie skupień utworzyły cztery strategie anonimizowania danych: anonimizacji domyślnej podstawowej, z wiodącą rolą pseudonimizacji, anonimizacji rozszerzonej o anonimizację miejsca oraz anonimizacji rozszerzonej o charakterystykę cech na poziomie zbiorowości, a nie jednostki. Powstałe typy anonimizacji stanowią odzwierciedlenie typów występujących w literaturze przedmiotu i stanowią potwierdzenie tego, że badacze w znacznej części stosują anonimizację jako procedurę naturalną i nieodłączną część badań (zgodnie z zapisami „Kodeksu etyki socjologa” PTS). Używanie prawdziwych tożsamości uczestników badania oraz autoryzacja stanowią rzadkość.

Wybierane strategie anonimizacji mają często postać prostej anonimizacji domyślnej, rzadziej stosowane są rozszerzone typy anonimizacji.

Słowa kluczowe: etyka badawcza, anonimizacja danych, poufność danych, autoryzacja danych, kodeksy etyczne

WPROWADZENIE

Zapewnienie poufności zbieranych danych oraz odpowiedniej ich anonimizacji stanowiło znaczącą trudność podczas prowadzonych przeze mnie badań terenowych na potrzeby rozprawy doktorskiej obronionej w 2017 roku. Zajął się złożonym problemem badawczym – relacjami pomiędzy kluczowymi aktorami społecznymi w małych społecznościach a dziennikarzami i właścicielami związanymi z głównymi tygodnikami lokalnymi w wybranych miejscowościach [Męcfal 2019]. Badania z wykorzystaniem metody studium przypadku prowadziłam w latach 2013–2014 w czterech niewielkich miejscowościach w województwie łódzkim. Początkowo planowałam anonimizację danych na poziomie indywidualnym – uczestnika badań. Bardzo szybko zrozumiałam, że anonimizacja powinna mieć jednak szerszy zakres, gdyż niektórzy uczestnicy badań pełnili takie funkcje formalne i nieformalne w małych społecznościach, które identyfikowałyby ich w momencie ujawnienia na przykład nazwy miejscowości. W związku z tym, aby chronić tożsamość badanych, zdecydowałam się także na anonimizację nazw miejscowości oraz na usuwanie z cytowanych fragmentów wywiadów charakterystycznych cech badanych osób, instytucji czy firm w danej miejscowości.

Takie postępowanie nie spotkało się ze zrozumieniem w niektórych środowiskach naukowych. Na jednej z konferencji z zakresu nauk o mediach (medioznawstwa) podniesiono kwestię braku wiarygodności zebranych przeze mnie danych i trudności odtworzenia kontekstu badania na podstawie gęstego opisu przypadku. Zdecydowanie bardziej moje działania akceptowane były w środowisku socjologicznym, przyjęto je wręcz jako naturalne.

Uczestnicy badań zostali poinformowani przeze mnie o zasadach poufności i anonimizacji ich wypowiedzi podczas pierwszego kontaktu z nimi i uzyskiwania ich zgody (ustnej) na udział w projekcie. Czasami, zanim zdążyłam przedstawić zasady anonimizacji, badani sami pytali o to, czy ich wypowiedzi będą publikowane z użyciem ich danych personalnych. Wówczas zapewniałam ich o poufności danych oraz informowałam o regułach anonimizacji. Nikt z badanych nie zgłosił zastrzeżeń do takiego postępowania, a wręcz zaobserwowałam, że było to działanie zgodne z oczekiwaniami badanych. Nieliczni badani nie zgodzili

się na nagrywanie (w sumie ze 137 wywiadów jedynie 10 rozmów nie zostało nagranych), jeden z badanych zgodził się na nagranie, ale w trakcie wywiadu dwukrotnie prosił mnie o wyłączenie dyktafonu, a w kilku przypadkach badani odpowiadali zdawkowo na zadawane przeze mnie pytania (nie łączę jednak tego braku współpracy z samym faktem nagrywania, a raczej ze świadomością badanych, że ich relacje z dziennikarzami lokalnymi oraz właścicielami tygodników miały charakter konfliktowy).

Na etapie analizy danych oraz tworzenia publikacji wielokrotnie zastanawiałam się, na ile decyzja o tak szerokiej anonimizacji była słuszna. Obawiałam się, że powstająca publikacja będzie niezrozumiała dla czytelnika ze względu na brak możliwości umieszczenia opisywanych sytuacji w kontekście. Jednak nie zdecydowałam się na powtórny kontakt z badanymi i zapytanie ich o możliwość udostępnienia nazwy miejscowości. Starłam się przytoczyć charakterystyki wybranych przypadków (miejscowości oraz tygodnika lokalnego), aby umieścić opisywane relacje w choć trochę szerszym kontekście.

Inspiracją do napisania artykułu dotyczącego praktyk polskich badaczy w zakresie anonimizacji było także moje uczestnictwo w kilku dyskusjach dotyczących etyki badań społecznych¹. Jedną z praktyk w badaniach społecznych, o której dyskutowano podczas tych spotkań, była autoryzacja wypowiedzi badanych, czasami połączona z anonimizacją danych [Surmiak 2018, 2019a], a czasami z opublikowaniem prawdziwych danych badanych [Kaźmierska 2018; Grzeszkiewicz-Radulska, Krzewińska 2015; Giermanowska, Mrozowicki, Róg-Ilnicka 2019]. Własne doświadczenia oraz obserwacja różnych praktyk badawczych skłoniły mnie do przeanalizowania artykułów naukowych z pięciu ostatnich lat (2015–2019) z czterech ważnych w środowisku polskich czasopism socjologicznych – „Studiów Socjologicznych”, „Polish Sociological Review”, „Przeglądu Socjologicznego” oraz „Przeglądu Socjologii Jakościowej” – pod kątem stosowanych przez autorów/autorki praktyk anonimizacyjnych. Wyniki tych analiz zostaną zaprezentowane w kontekście rozwiązań w zakresie anonimizacji opisywanych w literaturze przedmiotu, a także w kontekście reguł etyczno-prawnych obowiązujących w Polsce.

¹ Szczególnie uczestnictwo w dwóch panelach dyskusyjnych dotyczących etyki badań społecznych, organizowanych przez Oddział Łódzki Polskiego Towarzystwa Socjologicznego – „Etyka w badaniach społecznych – pomiędzy instytucjonalizacją a intuicją badacza” (13 maja 2019 roku) oraz „Etyka w jakościowych i ilościowych badaniach społecznych” (22 listopada 2019 roku), a także współorganizacja (z Adrianną Surmiak i Izabelą Ślęzak) i uczestnictwo w grupie tematycznej „Kwestie etyczne w badaniach społecznych” podczas XVII Ogólnopolskiego Zjazdu Socjologicznego we Wrocławiu (11–14 września 2019 roku).

POUFNOŚĆ I ANONIMIZACJA W LITERATURZE PRZEDMIOTU. ANONIMIZACJA DOMYŚLNA CZY PARTYCYPACYJNA?

W artykule będę odnosić się głównie do praktyk anonimizacji danych, które są częścią szerszego pojęcia – poufności danych. Anonimizacja jest jednym z głównych sposobów zapewnienia poufności danych, ale nie jedynym. Aby dotrzymać zobowiązania wobec uczestników badania o poufności przekazywanych badaczowi informacji, badacz powinien także zadbać o odpowiednie przechowywanie danych, ich archiwizację i udostępnianie czy ustalenie z badanym zasad wykorzystania danych w późniejszych publikacjach. Karen Kaiser [2009: 1635–1638] opisuje dwa podejścia do poufności – dominujące i alternatywne².

Podejście dominujące (zwane również w literaturze podejściem domyślnym [Saunders, Kitzinger, Kitzinger 2015; Nesper 2000] czy konwencjonalnym [Baez 2002]) polega na zastosowaniu reguły poufności w odniesieniu do każdego uczestnika badania, co przedstawiane jest przez badaczy jako ochrona badanych przed „krzywdą”, ochrona ich prywatności czy też wypełnianie standardów etycznych przez badacza. Tego rodzaju poufność planuje się już na etapie przygotowywania wniosku grantowego, a następnie na trzech etapach procesu badawczego: 1. zbierania danych (zapewnienie badanych, że ich dane będą zmienione, uzyskanie zgody na badanie, gdy badacze nie wiedzą jeszcze, jakie dane uzyskają; rzadko badacze wracają później do kwestii świadomej zgody na późniejszych etapach badania); 2. wstępnego opracowania danych/oczyszczania danych (np. podczas transkrypcji zastąpienie nazwisk pseudonimami, usunięcie innych danych identyfikujących); oraz 3. upowszechniania danych (np. rezygnacja z wykorzystania części danych w publikacji jako alternatywa dla anonimizacji). Jest to zatem pewna standardowa procedura, którą stosują badacze, a która może być traktowana jedynie jako formalność. W takim ujęciu nie jest ona przydatna do rozstrzygania konkretnych dylematów i problemów etycznych, które spotyka się w terenie [Guillemin, Gillam 2004].

Martin Tolich [2004] argumentuje, że dzięki zastosowaniu takiej anonimizacji na ogół badacze uzyskują tylko „poufność zewnętrzną” (osoby spoza kręgu badanego faktycznie nie będą mogły rozpoznać tożsamości danej osoby), jed-

² Podejście alternatywne oparte jest na doświadczeniach z badań przeprowadzonych przez Autorkę wśród kobiet, które przeżyły raka piersi. W przypadku jednej badanej nie udało się badaczce skutecznie zastosować anonimizacji, gdyż jedna z informacji, którą otrzymała od uczestniczki badania (orientacja seksualna), okazała się bardzo identyfikująca w małej grupie badanych. Badaczka obawiała się, że złamie zasadę poufności wobec tej uczestniczki badania i nie użyła ostatecznie tego wywiadu w swoich publikacjach [Kaiser 2009].

nak często nie udaje się w ten sposób uzyskać „poufności wewnętrznej”, która chroniłaby uczestników badania przed rozpoznaniem w kręgu osób najbliższych.

Inną słabością podejścia dominującego jest brak partnerskiego traktowania badanych (to badacz decyduje, jakie dane zmienia i w jaki sposób) [Tolich 2004; Kaiser 2009; Wiles i in. 2012]. Często także zakłada się, że wszyscy badani chcą, aby ich dane zanonimizowano, a są badania pokazujące, że nie zawsze taka jest wola badanych [m.in. Corden, Sainsbury 2006; O'Reilly i in. 2012]. Ponadto przyjmuje się, że zapewnienie poufności sprawi, że badany bardziej zaufa badaczowi. Tymczasem, jeśli tak rzeczywiście się stanie i badacz otrzyma bardzo szczegółowe i gęste dane, może się okazać, że nie będzie mógł ich wykorzystać, gdyż trudno będzie je poddać anonimizacji (lub, co gorsza, dane zostaną wykorzystane ze złamaniem zasady poufności). Zakłada się też, że zanonimizowane i zmienione dane będą tak samo wartościowe jak te przed anonimizacją. Jednak często okazuje się, że bez pewnych danych biograficznych i kontekstowych trudno jest zrozumieć i zinterpretować narrację badanego (na te problemy w badaniach biograficznych zwraca uwagę Kaja Kaźmierska [2018]).

Podejście alternatywne [Kaiser 2009; Wiles i in. 2012] można inaczej określić jako podejście partycypacyjne czy podmiotowe. Jego celem jest zachowanie bogactwa danych, ale także uzgodnienie z badanymi, jak przekazane dane mogą być wykorzystane. Proponuje się w nim zatem, aby lepiej informować badanych o tym, jak użyte będą dane (kto będzie odbiorcą oraz w jaki sposób będą one publikowane i prezentowane) oraz aby podejmować próby szerszego dialogu z uczestnikami badania o tym, jak ich dane mogą być wykorzystane. Tu zwłaszcza podnosi się problem ponownego uzyskania świadomej zgody także po zebraniu materiału. Dobrym przykładem takiego postępowania w kontekście polskim jest badanie Doroty Jaworskiej i in. [2019] wśród uchodźców starających się o azyl w Polsce, gdzie przyjęto negocjacyjny model procesu badawczego (zaadaptowany od Mackenzie i in. [2007] oraz Darlinga [2014]), który „opiera się na założeniu, że badani mogą stać się relatywnie równoprawną stroną procesu, jak również mają możliwość jego kontrolowania. Przyjmuje się tutaj, że zgoda nie ma charakteru jednorazowego aktu, lecz osiągana jest przez ciągłe negocjacje” [Jaworska, Alieva, Boryczko 2019: 215]. Partnerskie i otwarte podejście do uczestników badania pozwoli badaczom uniknąć niezadowolonych badanych z tego, jak użyte zostały ich dane czy w jaki sposób zostały one zanonimizowane (wyraży niezadowolonych z zastosowania pseudonimizacji odnotowali w swoich badaniach Ann Corden i Roy Sainsbury [2005]). Warto jednak dodać, że zastosowanie podejścia „partycypacyjnego” jest możliwe jedynie wówczas, gdy badania prowadzone są w „pełnym świetle” [Wyka 1990, 1993]. Gdy jed-

nak badacz zajmuje się zjawiskami „trudnymi”, decyduje się na ukrycie swojej roli, nie ujawnia celu swoich badań, bądź ujawnia go częściowo, przyjmuje zupełnie inną koncepcję rzeczywistości badawczej – konfliktową [Douglas 1976; Męcfal 2012] – a co się z tym łączy, nie ma w zasadzie możliwości uzgadniania zasad poufności z uczestnikami badania.

ANONIMIZACJA CZY INTEGRALNOŚĆ DANYCH?

Benjamin Saunders i in. [2015] pokazują, że istnieje konflikt pomiędzy anonimizacją danych a ich integralnością [por. m.in. Wiles 2008; Kaiser 2009; Surmiak 2018]. Nie jest on możliwy do rozwiązania za pomocą uniwersalnych procedur anonimizacji czy też stosowania tzw. anonimizacji całościowej (*blanket anonymisation*), polegającej na anonimizowaniu wszystkich danych identyfikujących (imion badanych, nazw miejscowości, nazw instytucji, szczególnych sytuacji), powinien być rozstrzygany w odniesieniu do specyficznego kontekstu badania.

Kwestia ta jest szczególnie problematyczna w badaniach wizualnych [Jordan 2014]. Jak dalece badacz może ingerować w obraz, nie manipulując jednocześnie danymi? Sara Jordan [2014: 444] odróżnia „zarządzanie obrazem” od „manipulowania obrazem”. Zarządzanie obrazem obejmuje zmiany dotyczące rozmiaru, jasności, czytelności obrazu (kontrastu) w sposób, który nie zmienia celowo oryginalnego czy zamierzonego znaczenia obrazu. Te techniki nie powinny zmieniać sposobu przedstawiania jednostki na zdjęciu, nie powinny naruszać warunków zgody na badanie, jest zresztą mało prawdopodobne, że wpłyną na interpretację zdjęcia (wszelkie zmiany powinny być też przez badacza wyraźnie odnotowane). Manipulacja obrazem ma miejsce wtedy, gdy odbiorca/czytelnik oczekuje, że prezentowany obraz odnosi się do sytuacji rzeczywistej opisywanej przez badacza, tymczasem zmiany wprowadzone w obrazie powodują, że na przykład jakaś osoba jest na zdjęciu bardziej widoczna, na zdjęciu brakuje pewnych detali czy też dwa zdjęcia połączone są w jedno [Jordan 2014: 445].

Takie działanie badacza (zwłaszcza jeśli nie jest oznaczone) może prowadzić do „oszustwa” względem odbiorców, a także niezgody samych badanych na taką prezentację obrazu. Badaczka zaleca zatem dokumentowanie wszelkich zmian, konsultowanie ich ze specjalistami, a przede wszystkim informowanie uczestników badania o tym, jak będą prezentowane ich zdjęcia i uzgadnianie z nimi tego sposobu prezentacji. Konsultację z badanymi zaleca także Łukasz Rogowski [2019: 64], „tak aby to oni określili, czy faktycznie jest ona konieczna i czy nie zmienia znaczenia materiałów wizualnych”. Ważne jest też odpowiednie przygotowanie procesu anonimizacji takich materiałów. „Aby nie niszczyć

danych wizualnych poprzez anonimizację, trzeba o niej pomyśleć jeszcze przed zbieraniem danych” [Wiles i in. 2012: 47].

Badacze stosujący studium przypadku spotykają się także często z wyborem pomiędzy anonimizacją a integralnością danych. Anonimizacja jest prezentowana jako ostateczność, a nie opcja pierwszego wyboru. Robert Yin [2009: 181–182, 2015: 228–229] pokazuje, jak decyzja o anonimizacji może wpłynąć na raport ze studium przypadku – gdy pozbawia czytelników możliwości odniesienia go do kontekstu – i w jakich sytuacjach powinna ewentualnie być stosowana. Ujawnianie nazw przypadków daje odbiorcom możliwość przypomnienia sobie potencjalnych poprzednich informacji o danym przypadku, a zatem jeśli znamy nazwę przypadku (np. miejscowości czy instytucji), łatwiej jest nam sprawdzić informacje i ocenić przygotowane studium. Yin [2015] dopuszcza jednak stosowanie anonimizacji, kiedy przypadek jest kontrowersyjny (ochrona rzeczywistego przypadku oraz uczestników tego przypadku); anonimizacja jest też wskazana, jeśli mamy zasadne przypuszczenia, że nasz raport może wpłynąć na późniejszy stan/działania naszych badanych (taka argumentacja pozwoliła mi zdecydować o anonimizacji moich przypadków [Męćfal 2019]). Trzecią możliwością stosowania anonimizacji jest zamiar ukazania przypadku jako typu idealnego. Autor przedstawia też kilka rozwiązań, jego zdaniem kompromisowych: nieujawnianie tożsamości osób, ale ujawnianie nazwy przypadku; podanie tożsamości uczestników badania w wykazie źródeł, bez przyporządkowywania ich do konkretnej cytowanej wypowiedzi czy prezentowanych poglądów (z możliwością zmiany pewnych faktów w wypowiedziach); w wielokrotnych studiach przypadków rezygnacja z prezentowania pojedynczych przypadków, a opublikowanie jedynie raportu z analizy przekrojowej [Yin 2015: 229].

Helen Simons [2009: 103, 106–107] prezentuje anonimizację jako część standardowej procedury stosowanej w studiach przypadku. Zaleca badaczowi, w sytuacji, gdy niemożliwa jest anonimizacja (np. w przypadku osób publicznych czy charakterystycznych instytucji), uzyskanie zgody (*clearance*) na publikację wypowiedzi, które mogą zostać zidentyfikowane. Przyznaje, że sama na ogół stosuje anonimizację zarówno osób, jak i instytucji, w których prowadzi badania. Argumentuje jednak, że nie powinno to być robione w sposób odgórnie założony, a w sposób uzgodniony z uczestnikami badania. Popierana przez nią „perspektywa demokratyczna” większy nacisk kładzie na „procesualną” świadomą zgodę (a zatem powtarzalną na różnych etapach badania), na włączanie uczestników w proces badawczy (np. sprawdzanie interpretacji badacza), w tym negocjowanie z badanymi, jakie dane mogą zostać upublicznione i w jaki sposób.

Anonimizację jako obszar do negocjacji z uczestnikami badania postrzega także Robert E. Stake [1995]. W przykładowym studium przypadku prezentowanym w książce *The art of case study research* badacz stosuje pseudonimizację (zarówno osób, jak i przypadku jako całości). Autor opisuje także przypadek, kiedy brak zgody jednego z obserwowanych nauczycieli na opisanie jego zajęć bez anonimizacji spowodował, że cały przypadek musiał zostać zanonimizowany. Stake nie zdecydował się na usunięcie tego opisu z raportu [1995: 116], uznając, że jest on istotnym elementem przypadku (zajęcia tego nauczyciela miały nietypowy, aktywizujący charakter), a także zakładając, że ryzykowałby w ten sposób utratę części kontekstu prezentowanego przypadku.

TYPY ANONIMIZACJI I SZCZEGÓLNE PRZYPADKI JEJ BRAKU

W literaturze prezentowane są także różnorodne podejścia czy też typy anonimizacji. Adrianna Surmiak [2018: 10–11] wyróżnia podejście ochronne (*protective*) oraz zrównoważone (*balanced*). Wyniki jej badań prowadzonych z badaczami społecznymi i humanistycznymi (głównie socjologami i antropologami społecznymi, badającymi grupy podatne na zranienie) wskazują, że podejście ochronne, czyli stosowanie znacznie szerszego zakresu anonimizacji, było często pochodną anonimizowania „na wszelki wypadek”. Badacze wskazywali, że nie byli pewni, co może być uznane za dane „wrażliwe”, czy też obawiali się, że ich badani mogą nie zgodzić się z ich interpretacją czy określonym ich przedstawieniem w publikacji (np. woleli usunąć jakieś dane z tekstu, nawet jeśli oznaczałoby to utratę ich integralności). Podejście zrównoważone cechowało badaczy, którzy szukali złotego środka pomiędzy ochroną badanych a integralnością danych. Według nich anonimizacja ma sens jedynie wtedy, gdy istnieje realne ryzyko, że uczestnik badania może być zidentyfikowany. Zwracali też uwagę, że ryzyko to jest zmniejszone przez stosunkowo niewielki zasięg publikacji naukowych.

O dwóch orientacjach wobec anonimizacji wspominają Rose Wiles i in. [2008: 423]: pierwsza zakłada, że badacz jest odpowiedzialny za anonimizację, gdyż to on ma odpowiednią wiedzę, pozwalającą mu ocenić ryzyko dla badanych; druga wskazuje na odpowiedzialność badanych – to oni powinni ocenić, które dane przekazywane badaczowi powinny być wykorzystane i w jakim stopniu ich tożsamość powinna zostać zanonimizowana (czy też ujawniona) w relacji do tych danych. Zdaniem autorów coraz częściej można zauważyć, że niektórzy badani chcą, aby ich dane zostały upublicznione. W takiej sytuacji badacze pierwszej orientacji nadal publikują dane, anonimizując tożsamość badanych. Jest to główna zasada etyczna, którą stosują, wypływająca z poczucia odpowiedzialności [Wiles

i in. 2008: 425]. Natomiast badacze drugiej orientacji wolą współpracować ściśle z badanymi, aby ustalić zasady publikacji wyników. Dotyczy to na przykład uczestniczących badań w działaniu (*participatory action research*), w przypadku których badacze są zwolennikami wracania do badanych i pytania ich, czy byli zadowoleni z zastosowanych rozwiązań anonimizacyjnych. Wiles i in. [2008: 426] uważali, że jeśli badani chcą wystąpić pod własnym imieniem, to powinno się im na to pozwolić.

Saunders i in. [2015: 621–626] pokazali z kolei, jak w ich badaniach³ wyglądało poszukiwanie optymalnej (możliwie maksymalnej) anonimizacji przy jednoczesnym zachowaniu integralności danych. Zwrócili uwagę na sześć kluczowych obszarów, w zakresie których stosowali anonimizację i napotkali różnorodne trudności:

1. imiona i nazwiska uczestników badania (niektórzy badani chcieli, aby używać ich imion i imienia pacjenta, ale zdecydowano się na użycie pseudonimów – po negocjacjach z badanymi; użycie prawdziwych imion mogłoby powodować ujawnienie tożsamości innych osób; czasami stosowano zasadę „podwójnej zasłony dymnej” (*smoke screen strategy*), czyli do tego samego badanego stosowano dwa różne pseudonimy) [Saunders, Kitzinger, Kitzinger 2015: 621; Kaiser 2009];

2. miejsca (ze względu na niewielką liczbę szpitali specjalistycznych tego typu ujawniono tylko region, w którym znajdował się szpital, np. miasto w południowo-wschodniej Anglii, i nadano szpitalom kolejne numery);

3. religijne lub kulturowe pochodzenie (badacze zastępowali religię/pochodzenie kulturowe innym podobnym lub też stosowali terminy ogólne, np. termin „wyznanie” zamiast wskazywania konkretnej religii);

4. zawód (zastępowanie konkretnego zawodu terminami bardziej ogólnymi, np. „kierownik teatru” zastąpiony został przez sformułowanie „praca w obszarze sztuki”);

5. relacje rodzinne (na ogół nie były anonimizowane, z wyjątkiem relacji nieheteronormatywnej, zastępowano także takie terminy jak „ojczym”, „rodzice zastępczy”, „dziecko adopcyjne” bardziej ogólnymi – „ojciec”, „rodzice”, „dziecko”);

6. inne potencjalnie identyfikujące informacje (specyficzne dla tego badania – sposób, w jaki doszło do uszkodzenia mózgu pacjenta; badacze starali się używać generalnych stwierdzeń dotyczących sytuacji uszkodzenia mózgu, nie dodając szerokiego opisu tej sytuacji, czasem stosowana była „podwójna zasłona dymna”, np. zmiana płci pacjenta, inny pseudonim, zmiana relacji rodzinnej, gdy te szczegóły były istotne dla badania).

³ Przeprowadzali wywiady z rodzinami osób będących w stanie niewielkiej przytomności lub w stanie wegetatywnym.

BRAK ANONIMIZACJI?

Wiles i in. [2008] wspominają o coraz większej chęci badanych do upubliczniania swoich danych. Nie zaobserwowałam takich sugestii w badaniach prowadzonych przeze mnie – moi badani raczej oczekiwali anonimizacji⁴.

Niemniej o problemach z anonimizacją wspominają badacze biograficzni czy badacze historii mówionej, dla których szczegóły z życia „narratorów” są niezwykle istotne w interpretacji i często stanowią sedno ich badań. Pozostawienie tych detali może spowodować rozpoznanie tożsamości badanego, a ich usunięcie zubaża materiał [Moore 2012]. Badacze biograficzni starają się mimo to stosować anonimizację; identyfikujące szczegóły zostawiają dla siebie bądź odpowiednio je modyfikują [Świątek-Młynarska 2019]. Ponadto w rozważaniach badaczy biograficznych powraca inny ważny problem – kto jest właścicielem historii mówionej, którą podzielił się uczestnik badania. Badacz, anonimizując wywiad, odbiera badanemu autorstwo swojej historii [Kaźmierska 2018: 401]. Polskie Towarzystwo Historii Mówionej [2018] zaleca, aby uznawać autorstwo wszystkich osób biorących udział w tworzeniu danego materiału, przede wszystkim uczestnika badań, mimo stosowanych praktyk anonimizacyjnych. Jednak wówczas, kiedy uczestnicy badań są osobami publicznymi bądź są dobrze rozpoznawalni w określonym środowisku, będą bardzo łatwo identyfikowalni na podstawie szczegółów z ich biografii, ich funkcji, dzieł, które stworzyli i innych. W takiej sytuacji zdarza się, że badacze i badaczki decydują się na ujawnienie prawdziwych tożsamości badanych osób (za ich zgodą) [Giermanowska, Mrozowicki, Róg-Ilnicka 2019], a czasem także przekazują uczestnikom badania transkrypcje wywiadów do „autoryzacji” [Kaźmierska, Waniek, Zysiak 2016], a nawet włączają ich w proces interpretacji danych [Grzeszkiewicz-Radulska, Krzewińska 2015].

Anonimizacja danych jest zatem raczej powszechną praktyką, często traktowaną jako etyczna konieczność, jednak badacze stosunkowo rzadko stawiają sobie pytania o to, jakie mogą być skutki anonimizacji [por. m.in. Tolich 2004; Wiles i in. 2008; Kaiser 2009; Moore 2012; Kaźmierska 2014, 2018; Surmiak 2018, 2019a]. Niestosowanie anonimizacji, czy też stosowanie rozwiązań pośrednich, może zostać dodatkowo utrudnione przez wprowadzane ograniczenia instytucjonalne.

⁴ Można przypuszczać, że było to związane z kluczowymi rolami społecznymi, jakie pełnili w społecznościach lokalnych oraz z faktem, iż tematyka badania obejmowała zagadnienia drażliwe.

ZASADY ETYCZNE I PRAWNE A ANONIMIZACJA

Zarówno polskie przepisy prawne, jak i zasady etyczne sformułowane dla socjologów przez Polskie Towarzystwo Socjologiczne kierują badaczy na ścieżkę stosowania anonimizacji. „Kodeks etyki socjologa” [2012]⁵ podejmuje kwestie anonimizacji, prywatności i poufności w punktach 19–23. Kodeks wyraźnie zaleca badaczom stosowanie anonimizacji i ochronę badanych i ich danych (mówi o „obowiązku poufności” [2012: pkt 20]). Badacz powinien, według zapisów, umieć przewidzieć ryzyko naruszenia zasad poufności i anonimowości i podjąć odpowiednie kroki, aby bezpiecznie przechowywać dane badawcze. Podawane są przykłady, jak należy to zrobić, ale są one nieliczne: „usuwanie identyfikatorów, użycie pseudonimów oraz inne środki techniczne zacierające związek między danymi a identyfikowalnymi na ich podstawie osobami” [2012: pkt 21].

Co ciekawe, punkt 22 kodeksu zawiera warunkowe przyzwolenie na odstępstwo od normy anonimizacyjnej. Na badaczu ciąży mniejszy obowiązek zapewniania poufności i ochrony prywatności, kiedy bada instytucje o charakterze publicznym (np. organy administracji czy urzędników); to odstępstwo nie obowiązuje, jeśli wcześniej zapewniono badanych o poufności danych. Niemniej i w tej sytuacji pojawia się wyjątek (choć sformułowany w terminach nieostrych): badacz może odstąpić od ochrony danych badanego, jeśli za takim działaniem „przemawia oczywisty i istotny interes publiczny” [2012: pkt 22].

Ostatni z punktów w sekcji „Anonimowość, prywatność i poufność” zwraca uwagę socjologów na badania w Internecie, które powinny być traktowane ze szczególną starannością, gdyż badacz może napotkać problemy z uzyskaniem świadomej zgody, może mieć trudności z oceną, jakie dane są publiczne, a jakie

⁵ Zdradzając kulisy powstawania kodeksu, można wspomnieć, że był on dyskutowany przez wiele lat i członkowie Polskiego Towarzystwa Socjologicznego nie mogli się zgodzić co do jego kształtu. Wynik głosowania nad obecnym kształtem kodeksu nie był jednomyślny, a były też głosy, aby nie poddawać go wcale pod głosowanie. Prawdopodobnie pokazuje to, że część społeczności jest raczej niechętna formalizacji zasad etycznych prowadzenia badań. Jak pisze Grażyna Woroniecka [2014: 249]: „Argument wysuwany najczęściej przeciwko tworzeniu kodeksu umocowany był w przeświadczeniu, że sam etos akademicki stanowi wystarczające zabezpieczenie przed naruszeniami istotnych norm etycznych, co sprawia, że ich kodyfikacja staje się zbędna. Dopiero burzliwie przebiegający proces usamodzielniania się komercyjnych firm badawczych w połączeniu z masowością kształcenia badaczy okazały się czynnikiem wymuszającym regulację o charakterze sformalizowanym”. Formalizacja taka została zapewne wymuszona także funkcjonowaniem w coraz bardziej zbiurokratyzowanym systemie instytucjonalnym, gdzie bardzo istotna stała się ewaluacja i kontrola działalności naukowej. Jednak uchwalenie kodeksu w 2012 roku stanowiło dla mnie przede wszystkim wyraz profesjonalizacji stowarzyszenia i jasny przekaz, że istnieją wyraźne granice w procesie badawczym, których żaden socjolog nie powinien przekraczać.

prywatne czy też może mieć kłopot z zapewnieniem bezpieczeństwa danych przekazywanych przez badanego [2012: pkt 23]. Poza sekcją o anonimowości w kodeksie znajduje się punkt 18, który dotyczy uczestników badań niejawnych i nakazuje badaczom w tego typu badaniach chronienie anonimowości uczestników. „Kodeks etyki socjologa” wzorowany był na kodeksach anglosaskich (zwłaszcza kodeksie Brytyjskiego Towarzystwa Socjologicznego – BSA – w jego poprzedniej wersji oraz Amerykańskiego Towarzystwa Socjologicznego – ASA), czym można tłumaczyć występowanie wielu podobnych elementów i wskazań etycznych, także tych dotyczących anonimizacji i poufności⁶. Warto jednak zauważyć, że w nowej wersji kodeksu BSA wspomina się o tym, że badanym powinno się wytłumaczyć, iż całkowita anonimizacja nie jest możliwa i badacze nie mogą składać nierealistycznych obietnic w zakresie poufności. Badacz ma obowiązek wyjaśnić badanym także, w jakim zakresie będą im udostępnione do wglądu i potencjalnych zmian transkrypcje wywiadów, notatki terenowe czy też pierwsze interpretacje badaczy [Kodeks etyczny BSA 2017: pkt. 26–28]. Żadna z tych uwag nie pojawia się w polskim kodeksie, który nie był aktualizowany od jego ogłoszenia w 2012 roku.

Ciekawe ujęcie poufności zawierają „Rekomendacje etyczne Polskiego Towarzystwa Historii Mówionej” z 2018 roku [m.in. pkt 5 i 6], które podkreślają możliwość anonimizacji (stosowania pseudonimów), a nie jej obowiązek. Dają też rozmówcy prawo do decydowania, jakie fragmenty rozmowy będą udostępniane, komu i w jakim zakresie. To, co powinno być chronione, to dane osobiste rozmówcy, które nie wynikają wprost z nagrania (np. adres, stan zdrowia, posiadany majątek, treść rozmów poza nagraniem).

O skutkach wprowadzania kodeksów etycznych w kontekście amerykańskim pisała Yvonna Lincoln [2009]. Jej zdaniem pociągnęło to za sobą zwiększenie uprawnień uniwersyteckich komisji etycznych i dało im narzędzie kontroli (nie mają one zatem służyć ochronie uczestników, a umożliwić wpływanie na autonomiczny proces badawczy). W kontekście polskim większe przełożenie na rzeczywistość badawczą mają zasady ustalane przez Narodowe Centrum Nauki (jednego z głównych grantodawców w Polsce) – najpierw ogłoszone w 2016 roku „Zalecenia Rady Narodowego Centrum Nauki dotyczące badań z udziałem ludzi” [2016: 1–2], które zobowiązały uczelnie do utworzenia komisji etycznych, opiniujących projekty w kilku typach badań z udziałem ludzi, a także zapowiedziane w kwietniu 2019 roku, i obecnie już zrealizowane, wprowadzenie do formularza

⁶ Niektóre aspekty kodeksów etycznych PTS, BSA, ASA i innych porównywała w swoim artykule Adrianna Surmiak [2019a].

wniosku o finansowanie projektu załącznika, „w którym wnioskodawca przedstawi skrócony plan zarządzania danymi badawczymi (*data management plan*), które powstaną w ramach realizacji projektu. Plan ten będzie uzupełniany na etapie składania raportu końcowego i oceniany po zakończeniu realizacji projektu” [Pismo dyrektora NCN w sprawie zarządzania danymi 2019⁷]. Te regulacje instytucjonalne mogą bezpośrednio wpływać na stosowane przez badaczy zasady anonimizacji i poufności danych (badacz będzie musiał na początku procesu badawczego, nie znając zebranego materiału, założyć określony typ postępowania z danymi), mogą wprowadzić formalizację uzyskiwania świadomej zgody (w tej chwili nie ma takiego obowiązku), a także skłaniać naukowców do szukania „sprawdzonych” rozwiązań i nieeksperymentowania w tym zakresie, gdyż może to się skończyć nieuzyskaniem środków grantowych.

Oprócz reguł etyczno-instytucjonalnych badania społeczne są prowadzone też w określonym kontekście prawnym, który może wpływać na postępowanie badaczy w zakresie poufności i anonimizacji danych. Ze względu na ograniczone miejsce nie będę go szczegółowo analizować. Wspomnę jedynie, że ma on dwa wymiary: 1. nakładający na badacza ochronę danych uczestników badania i mówiący o ewentualnych konsekwencjach, jakie może ponieść badacz, jeśli tego obowiązku nie spełni (regulują to: Ustawa o ochronie danych osobowych [2018], Konstytucja RP [1997: art. 47], Rozporządzenie Parlamentu Europejskiego i Rady Unii Europejskiej o ochronie danych [2016], Kodeks karny [1997: art. 266, par. 1]); 2. regulujący możliwość odstępstwa od ochrony danych lub nakazujący ujawnienie danych poufnych (np. jeśli badacz dowie się o przestępstwie, ma obowiązek jak każdy obywatel poinformować o nim odpowiednie służby) i konsekwencje, jakie mogą być zastosowane, jeśli badacz tego nie zrobi (regulują to Kodeks postępowania karnego [1997: art. 180, par. 1, art. 304, par. 1] oraz Kodeks karny [1997: art. 240]).

Zasady etyczne i prawne mogą zatem stawiać badacza w sytuacji wyboru pomiędzy ochroną badanego (głównie zasady o charakterze etycznym) a nakładanym obowiązkiem ujawnienia tożsamości badanego (zasady prawne). Badacze są podzieleni co do tego, które zasady mają pierwszeństwo. John Lowman i Ted Palys [2014] wyróżnili na podstawie swoich badań dwa podejścia: „najpierw prawo” (*law first*) oraz „najpierw etyka” (*ethics first*) [por. też Surmiak 2019a].

⁷ W piśmie czytamy, że: „Będzie to pierwszy etap wprowadzania przez Narodowe Centrum Nauki polityki otwartego dostępu do danych naukowych wytworzonych/wykorzystanych w ramach realizacji projektów finansowanych przez Centrum, a w niedalekiej perspektywie również do publikacji naukowych”.

Badacze wybierający pierwsze podejście uważają, że tak jak każdy obywatel mają obowiązek przestrzegania prawa, są zatem skłonni do złamania zasady poufności, jeśli zmuszą ich do tego reguły prawne. Zwolennicy drugiej opcji wybierają ochronę badanych, popierają niezależność badań akademickich i nie wypełniają prawnych zobowiązań, nawet jeśli miałyby się to wiązać z konsekwencjami dla nich samych. Nie są mi znane przypadki, gdy socjologowie w Polsce musieliby dokonywać takiego wyboru. Podobne rozważania nie pojawiły się też w analizowanych przeze mnie artykułach.

METODA BADAWCZA I ZASADY ANALIZY

W celu uzyskania odpowiedzi na pytanie, jakie strategie anonimizacji stosują polscy przedstawiciele nauk społecznych (głównie socjologowie) publikujący w polskich czasopismach socjologicznych, przeanalizowałam artykuły naukowe z czterech uznanych czasopism: „Studiów Socjologicznych”, „Polish Sociological Review”, „Przeglądu Socjologii Jakościowej” oraz „Przeglądu Socjologicznego”⁸, które ukazały się w latach 2015–2019 (oraz pierwszego numeru „Przeglądu Socjologii Jakościowej” z 2020 roku, gdyż został już opublikowany w trakcie przeprowadzania analiz). Pominęte zostały artykuły autorów zagranicznych, recenzje książek czy eseje recenzyjne, które nie stanowiły materiału, który mógłby przynieść informacje pozwalające na udzielenie odpowiedzi na główne pytanie problemowe. Numery z lat wybranych do analizy zawierają zarówno artykuły powstałe po reformie szkolnictwa wyższego z 2012 roku, jak i te, które zostały napisane już w zmieniających się warunkach instytucjonalno-biurokratycznych pod wpływem ustawy Prawo o szkolnictwie wyższym i nauce z lipca 2018 roku (tzw. Ustawa 2.0). Obie zmiany prawne wprowadziły do polskiej nauki nastawienie na mierzalny efekt publikacji, większą kontrolę i biurokratyzację. Mogły także wpłynąć na wybory etyczne w badaniach i ich większą zależność od rozwiązań instytucjonalnych niż od indywidualnej decyzji badacza [Surmiak 2019b].

We wszystkich wybranych czasopismach naukowych opublikowano w tym czasie 626 artykułów, z czego 175 artykułów (28%) zawierających opis badań jakościowych⁹ lub też odnoszących się do etycznych kwestii badań jakościowych

⁸ W analizie pominęłam „Qualitative Sociology Review”, gdyż czasopismo to zawierało w ciągu ostatnich pięciu lat w większości artykuły autorów zagranicznych.

⁹ Stosowane przez autorów metody badawcze to: metoda biograficzna (w tym wywiady narracyjne) – 15%, wywiady pogłębione (IDI), wywiady częściowo ustrukturyzowane – 39%, techniki obserwacji – 2%, analiza danych zastanych – 5%, metody mieszane – 25% (np. wywiady

zostało włączonych do analizy¹⁰. Podjęłam decyzję o niewłączeniu do analizy artykułów dotyczących badań ilościowych, gdyż tylko dwa artykuły z tej puli odnosiły się do zastosowanych zasad anonimizacji, pozostałe zaś nie wyjaśniały tego w żaden sposób, a jedynie prezentowały zagregowane wyniki. Pomięłam także artykuły o charakterze rozważań teoretycznych.

W badaniach zastosowałam jakościową analizę zawartości (z elementami ilościowymi), za jednostkę analizy uznając pojedynczy artykuł naukowy (tzw. dzieło, wypowiedź). Artykuły poddane zostały najpierw wstępnej analizie – pierwsze czytanie – po której zdecydowałam o wyborze artykułów dotyczących badań jakościowych. Drugie czytanie wybranych artykułów miało na celu zakodowanie ich do interesujących mnie kategorii analitycznych: obecność anonimizacji, jaki rodzaj anonimizacji jest stosowany (opisowo), czy uzgadniano z badanymi sposób anonimizacji (czy stosowano anonimizację domyślną, czy alternatywną/partycypacyjną – kategoria dodana na późniejszym etapie analiz), obecność jakiegokolwiek formy autoryzacji, jaka to forma (opisowo), czy obecna jest szersza refleksja etyczna, jakie tematy są dyskutowane (krótki opis), czy badano drażliwe kwestie, jakie to były kwestie (opisowo), stosowane metody badawcze, rok publikacji, czasopismo, autorzy artykułu (kodowanie miało charakter klasyfikacyjno-rejestracyjny – szersze wypisy z artykułów). Na następnym etapie wypisy z artykułów w zakresie typów anonimizacji zostały sklasyfikowane tak, iż powstało osiem nowych kategorii analitycznych odpowiadających stosowanym przez autorów sposobom anonimizacji: pseudonimizacja, anonimizacja za pomocą nadania materiałom ogólnych symboli, stosowanie anonimizacji, ale podawanie cech badanych na poziomie jednostkowym, anonimizacja polegająca na podaniu cech badanych jedynie na poziomie zbiorowości, anonimizacja rozszerzona – zmienianie szczegółów w wypowiedziach badanych lub ich usuwanie, anonimizacja zdjęć/filmów, anonimizacja nazw miejsc (miejscowości lub instytucji), anonimizacja z wykorzystaniem jedynie opisu danych bez podawania cytatów z wywiadów. Wyniki tych analiz zaprezentowane są w kolejnej części artykułu.

łączone z obserwacją), zogniskowane wywiady grupowe (FGI) – 3%, badania etnograficzne/studium przypadku/partycypacyjne – 10%, inne – 1%.

¹⁰ W „Polish Sociological Review” – 30 artykułów (na 153), w „Studiach Socjologicznych” – 47 artykułów (na 177), w „Przeglądzie Socjologii Jakościowej” – 76 artykułów (na 161) oraz w „Przeglądzie Socjologicznym” – 22 artykuły (na 135).

TYPY STRATEGII ANONIMIZACYJNYCH – WYNIKI ANALIZ

94% przeanalizowanych artykułów wykorzystywało przynajmniej jeden ze sposobów anonimizowania danych, w 3% tekstów można zaobserwować częściową anonimizację (np. niezanonimizowane są wypowiedzi osób publicznych, podane są na tyle szczegółowe cechy badanych, że trudno mówić o anonimizacji), a jedynie 3% (5 artykułów) nie zostało zanonimizowanych. Wśród tych pięciu artykułów można wskazać dwa przypadki autoryzowania wywiadów i uzyskania zgody na publikację prawdziwego nazwiska uczestników badania; jeden artykuł, dotyczący analizy blogów, gdzie autorki posługują się jednocześnie imionami (i nazwiskami, jeśli są znane) oraz pseudonimami blogerów, których używają oni sami na blogach i podają prawdziwe linki do blogów; jeden artykuł będący analizą metodologiczną i etyczną wywiadu biograficznego; wreszcie ostatni, w którym autorka korzysta z historycznych dokumentów osobistych pochodzących z różnych archiwów; jeśli znany jest autor takiego dokumentu, autorka podaje imię i nazwisko autora, nie anonimizuje tych materiałów.

Widać zatem, że anonimizacja stosowana jest powszechnie (zgodnie z tendencją, którą można odnaleźć w literaturze przedmiotu) i tylko w nielicznych przypadkach badacze wybrali inny sposób prezentacji danych. Jedynie 7% tekstów zawiera informację o tym, że anonimizacja była uzgadniana z uczestnikami badania, pozostałe (93%) albo nie zawierają takiej informacji, albo autorzy przedstawiają badanym zasady anonimizacji jako reguły, według których będą analizowane ich wypowiedzi, a badani jedynie wyrażają na nie zgodę. Odnosząc te wyniki do prezentowanych wcześniej typów anonimizacji, można wywnioskować, że tylko 7% autorów stosowało rozwiązania charakterystyczne dla „anonimizacji alternatywnej/partycypacyjnej”, a pozostali badacze postępowali raczej zgodnie z zasadami „anonimizacji domyślnej” (albo nie informowali w artykule o podjęciu dialogu z badanymi). Nieco częściej (w 27% artykułów) pojawiają się elementy refleksji etycznej czy opisy dylematów etycznych, np. refleksja etyczna i metodologiczna związana z badanym problemem, deklaracje badaczy, że uczestnicy badań zostali zapewnieni o poufności zbieranych danych, rzadziej występują szczegółowe informacje dotyczące uzyskiwania zgody na badania (badacze przekazują badanym informacje, że nie muszą czuć się zobowiązani do udzielenia zgody czy też że mogą wycofać się z badania); w niektórych artykułach wspomina się o zasadzie niekrzywdzenia badanych, informuje się o ograniczeniu opisu respondenta w celu zachowania jego anonimowości, pojawiają się refleksje etyczne związane z prowadzeniem badań w „trudnym” terenie badawczym, a w jednym z tekstów badacze prezentują krok po kroku działania, jakie podjęli,

aby proces badawczy był jak najbardziej etyczny, a badani mieli zapewnioną jak największą podmiotowość (Dorota Jaworska i współautorzy [2019] proponują kompletny program etyczny badań prowadzonych z grupą podatną na zranienie, jaką są uchodźcy, w tym także dzieci).

Jak wyglądają konkretne strategie anonimizacyjne stosowane przez autorów analizowanych artykułów? (tabela 1). W 140 artykułach sposób anonimizacji jest opisany wprost lub można go dość łatwo odtworzyć. Jednak rzadko pojawia się wyjaśnienie badaczy, dlaczego zastosowany został określony sposób anonimizacji. Najczęściej stosowana (75% przypadków) jest anonimizacja poprzez zastępowanie imienia i nazwiska badanego (względnie nazwy instytucji) ogólnym symbolem, np. numerem wywiadu czy inicjałami badanego. 65% artykułów zawiera, mimo stosowanej anonimizacji, opis istotnych dla badania cech każdego z badanych (na poziomie jednostkowym często do opisu cytatów służą dodatkowe symbole). 29% tekstów charakteryzuje badanych tylko jako grupę. 1/5 artykułów zamiast klasycznej anonimizacji stosuje pseudonimizację – imiona i nazwiska badanych zastępowane są fałszywymi imionami bądź pseudonimami. 14% autorów dodatkowo anonimizuje także nazwy miejscowości. Anonimizacja rozszerzona, polegająca na zmianie szczegółów w cytowanych fragmentach wywiadów czy też nieumieszczeniu identyfikujących danych w publikacjach, stanowi jedynie 6% badanych przypadków. W 4% tekstów nie są zamieszczane w ogóle cytaty z zebranych materiałów, a wyniki prezentuje się w postaci opisu¹¹. W czterech artykułach, wykorzystujących metody wizualne, autorzy stosowali dodatkowo anonimizację zdjęć lub filmów.

TABELA 1. Stosowane sposoby anonimizacji i częstość ich występowania

Sposoby anonimizacji	Częstość występowania w % (N=175)
Anonimizacja za pomocą kodu/symbolu	75%
Cechy badanych na poziomie jednostkowym	65%
Cechy badanych na poziomie zbiorowym	29%
Pseudonimizacja	20%
Anonimizacja nazw miejscowości/instytucji	14%
Usuwanie/zmianianie identyfikujących fragmentów	6%
Opis wyników bez cytatów	4%
Anonimizacja materiałów wizualnych	3%

Źródło: analiza własna. Wyniki nie sumują się do 100%, gdyż niektórzy autorzy stosowali więcej niż jeden sposób anonimizacji.

¹¹ Nie wszyscy autorzy tego typu artykułów deklarują wprost, iż jest to ich sposób anonimizowania danych. Niestosowanie cytatów i prezentowanie wyników w sposób opisowy traktuje się jako rodzaj anonimizacji danych w literaturze przedmiotu, stąd też takie zaklasyfikowanie tych artykułów przeze mnie.

Aby sprawdzić, w jaki sposób stosowane formy anonimizacji łączą się ze sobą w typy anonimizacji, przeprowadziłam dodatkowo analizę skupień (metodą dwustopniowego grupowania, odpowiednią do stosowania dla zmiennych nominalnych – jakościowych), w której uwzględniłam opisane powyżej sposoby anonimizacji oraz fakt uzgadniania z badanymi kwestii anonimizowania danych. W wyniku tego działania powstały cztery skupienia – typy anonimizacji, które po prześledzeniu profili skupień nazwałam odpowiednio (tabela 2): 1. typ: anonimizacja domyślna podstawowa (59 przypadków); 2. typ: wiodąca pseudonimizacja (40 przypadków); 3. typ: anonimizacja rozszerzona o anonimizację miejsca (27 przypadków)¹²; 4. typ: anonimizacja rozszerzona – charakterystyka na poziomie zbiorowości (33 przypadki). Każdy z tych typów, z wyłączeniem pierwszego, zawierał w sobie kilka dodatkowych sposobów anonimizacji, które jednak nie wpłynęły znacząco na kształt skupienia ze względu na niewielką liczebność.

TABELA 2. Typy anonimizacji stosowane w artykułach

Cecha	Typ (skupienie 1) anonimizacja domyślna podstawowa	Typ (skupienie 2) wiodąca pseudonimizacja	Typ (skupienie 3) anonimizacja rozszerzona o anonimizację miejsca	Typ (skupienie 4) anonimizacja rozszerzona, charakterystyka na poziomie zbiorowości
Pseudonimizacja	0	31	1	0
Anonimizacja za pomocą kodu/symbolu	59	0	27	33
Cechy badanych na poziomie jednostkowym	59	20	19	6
Cechy badanych na poziomie grupowym	0	13	0	33
Usuwanie/zmianie identyfikujących fragmentów	0	6	4	0
Anonimizacja zdjęć, filmów	0	1	3	0
Anonimizacja miejsca (miejsowości, instytucji)	0	6	13	3
Opis wyników bez cytatów	0	4	0	2

¹² Do typu trzeciego zostały zaklasyfikowane moje badania, o których wspominałam we wprowadzeniu.

Cecha	Typ (skupienie 1) anonimizacja domyślna podstawowa	Typ (skupienie 2) wiodąca pseudonimizacja	Typ (skupienie 3) anonimizacja rozszerzona o anonimizację miejsca	Typ (skupienie 4) anonimizacja rozszerzona, charakterystyka na poziomie zbiorowości
Uzgadnianie z badanym sposobu anonimizacji	0	2	4	1
Razem	59	40	27	33

Źródło: analiza własna. Liczba przypadków nie sumuje się do 175, gdyż 16 przypadków nie zostało uwzględnionych w modelu z powodu braków danych (tj. braku anonimizacji, wystąpienia częściowej anonimizacji, bądź artykuł nie miał charakteru empirycznego, a był refleksją metodologiczno-etyczną). Każda ze zmiennych w modelu miała dwie wartości (1 – występuje, 2 – nie występuje).

Pierwszy, domyślny typ anonimizacji, odpowiada według mnie strategiom anonimizacji opisywanym w literaturze pod taką samą nazwą. W tej grupie znajdują się artykuły, w których nazwisko badanego zostało zastąpione symbolem bądź kodem, a dodatkowo autor scharakteryzował każdego z badanych według istotnych cech. Obejmuje on największą liczbę artykułów, co potwierdza także fakt dominacji tego typu anonimizacji. Można przypuszczać, że stosowanie anonimizacji domyślnej, traktowanej jako nieodłączna część procedury badawczej, wpłynęło na fakt, iż badacze raczej nie uzasadniają, dlaczego stosują ten typ anonimizacji.

„Wiodąca pseudonimizacja” łączy się albo z podawaniem cech badanych na poziomie jednostkowym, albo na poziomie grupy, a dodatkowo w niektórych przypadkach występują także inne sposoby anonimizacji, jak: anonimizacja nazwy miejscowości, usuwanie/zmienianie fragmentów cytatów, opis wyników bez przytaczania cytatów, a także uzgadnianie z badanym sposobu anonimizacji. Pseudonimy stosowane są głównie w artykułach zawierających opis badań biograficznych, gdy liczba badanych nie jest duża, a życiorysy badanych są analizowane w sposób pogłębiony (czasami dodatkowo celowo zmieniane są szczegóły w biografjach badanych). Spotyka się je także w badaniach grup podatnych na zranienie (np. dzieci, uchodźców, osób z chorobami psychicznymi) oraz w badaniach grup/institucji „trudnych” (np. grupy religijne, kościoły). Warto zauważyć, iż w większości tego typu strategii anonimizacyjnych to badacze nadają pseudonimy badanym, a rzadziej badani decydują, pod jakim pseudonimem wystąpią (taką swobodę otrzymali badani między innymi od Marty Olcoń-Kubickiej [2016], choć okazało się to później kłopotliwe przy prezentacji wyników, gdyż badani wybierali często te same pseudonimy, tj. najbardziej popularne imiona [por. Mizielińska i in. 2018]).

Dla trzeciego typu charakterystyczna jest także anonimizacja za pomocą kodu oraz poprzez podawanie cech jednostkowych badanych; od typu pierwszego odróżnia ją fakt stosowania anonimizacji miejsca. Zwykle anonimizacja miejscowości dotyczy badań prowadzonych w społecznościach lokalnych (wówczas nazwa miejsca najczęściej zapożyczana jest od nazwy regionu). Czasem dodatkowo występuje drażliwa tematyka badania (np. relacje polsko-żydowskie, relacje polityczne). Anonimizowane są także nazwy instytucji, np. szkół, uczelni czy też agencji towarzyskich. Kilka artykułów zaklasyfikowanych do tego typu cechowało także uzgadnianie przez autora sposobu anonimizacji z badanymi, zmienianie identyfikujących szczegółów czy anonimizowanie materiałów wizualnych.

W ostatnim typie autorzy także stosowali anonimizację za pomocą symbolu/kodu, jednak podawali cechy badanych głównie na poziomie zbiorowości, a nie jednostki. Jedynie w kilku tekstach zastosowane były również inne sposoby anonimizacji (anonimizacja miejsca oraz opis wyników bez cytatów). W niektórych artykułach zaliczonych do tego typu występują metody mieszane (np. ilościowe czy analiza danych zastanych), co mogło wpłynąć na brak szczegółowej charakterystyki materiału jakościowego. Dodatkowo warto zauważyć, że autorzy tych artykułów w większości nie wyrazili wprost, że charakterystyka grupy, a nie jednostki, stosowana jest przez nich świadomie i celowo jako strategia anonimizacyjna.

Można zatem zaryzykować stwierdzenie, że w odróżnieniu od typu pierwszego pozostałe trzy typy anonimizacji są odmianami anonimizacji alternatywnej, stosowanymi raczej ostrożnie i oszczędnie. Mało znaczący aspekt tych strategii stanowi choćby uzgadnianie sposobu anonimizacji z uczestnikami badania (m.in. w badaniach Surmiak [2018], badaniach par intymnych opisywanych przez Miździelińską i in. [2018], badaniach uchodźców [Jaworska, Alieva, Boryczko 2019] czy badaniach kompetencji wizualnych Polaków [Rogowski 2019]).

W zaprezentowanych typach anonimizacji nie znalazły się artykuły, w których autorzy decydowali się na ujawnienie tożsamości badanych (zostały wyłączone podczas analizy jako braki danych, gdyż same zmienne dotyczyły sposobu anonimizowania). Jednak w całym materiale jedynie w dwóch artykułach autorzy i autorki zdecydowali się na ujawnienie tożsamości badanych za ich zgodą¹³ [Grzeszkiewicz-Radulska, Krzewińska 2015 oraz Giermanowska, Mrozowicki, Róg-Ilńska 2019]. W jednym z artykułów autorka stosowała autoryzację łącznie

¹³ Taka decyzja autorek i autorów tych dwóch artykułów była połączona z refleksją dotyczącą pozorności anonimizowania uczestników badania w tych konkretnych przypadkach, gdyż w środowisku naukowym stosunkowo łatwo można ich (i ich dorobek) zidentyfikować.

z technikami anonimizacji, czego zresztą oczekiwali uczestnicy badania i chętnie z tego korzystali [Surmiak 2019b]. Jako pewne przejawy autoryzacji danych potraktowałam próby powrotu badaczy do badanych w celu przeprowadzenia powtórnego wywiadu, uzupełnienia danych czy dopytania badanych o zgodę na archiwizację danych, jednak w sumie takie praktyki zaobserwowałam tylko w kolejnych siedmiu artykułach. Analiza pokazuje zatem, że strategie anonimizacyjne cechuje dość duży stopień uniwersalizacji (można przypuszczać, że dotyczy to nie tylko tego aspektu procesu badawczego). Jednym z prawdopodobnych wyjaśnień może być fakt podążania za rozwiązaniami stosowanymi w krajach (głównie) anglosaskich, skoro także „Kodeks etyki socjologa” opiera się na tych rozwiązaniach. Innym możliwym wytłumaczeniem może być wymaganie od badaczy dostosowania się do instytucjonalnych „reguł gry”. Oba przypuszczenia wymagają oczywiście weryfikacji w dalszych badaniach.

PODSUMOWANIE

Biorąc pod uwagę wyniki przedstawionych analiz, można zastanowić się, w jakim celu i dla kogo badacze stosują techniki anonimizacji. Czy tylko po to, aby chronić badanych przed przewidywaną krzywdą? Czy także w celu ochrony interesów instytucjonalnych i swoich własnych? Ten pierwszy cel anonimizacji jest deklarowany stosunkowo często. Tego drugiego celu nie sygnalizuje żaden z autorów analizowanych artykułów. Niekiedy trudno w ogóle w artykułach odszukać cel stosowanej anonimizacji, gdyż, jak można przypuszczać, autorzy traktują tę procedurę jako oczywisty element ich warsztatu metodologicznego i etycznego, zatem rzadko kiedy poddają to postępowanie refleksji [por. m.in. Nespór 2002; Wiles i in. 2008; Moore 2012; Wiles i in. 2012; Taylor 2015, Surmiak 2018]. Traktowana jest ona wręcz jako pewien mechanizm, który w zasadzie działa sam i jest oddzielony od działań i sprawczości badacza [Mondada 2014].

W niektórych czasopismach zagranicznych stosuje się praktykę dołączania krótkiej informacji etycznej (*ethical statement*) na końcu artykułu, w której badacz wyraźnie określa, jakie procedury etyczne stosował i dlaczego. Według mnie jest to dobra praktyka, która mogłaby skłonić badaczy do większej refleksyjności w zakresie etyki badawczej, co z kolei mogłoby prowadzić do większej różnorodności w tym obszarze i w efekcie do zmiany podejścia także do anonimizacji (zarówno wśród badaczy, jak i badanych). Być może wówczas, zgodnie z rekomendacją Jana Nespóra [2000], automatyczna i domyślna anonimizacja zostałaby odrzucona, a badacze w sposób bardziej przejrzysty przedstawialiby, co kierowało nimi, kiedy decydowali się na ujawnianie bądź ukrywanie tożsamości

badanych. Większa refleksyjność i różnorodność podejść stoją jednak w pewnej sprzeczności z dążeniem do wprowadzania nowych regulacji w obszarze badań społecznych (np. komisje etyczne, wprowadzanie zasad tzw. otwartego dostępu), które są podporządkowane głównie perspektywie legalistyczno-pozytywistycznej [van den Scott 2020] i nie uwzględniają specyfiki badań społecznych.

Można też zastanowić się nad konsekwencjami stosowania wyróżnionych strategii anonimizacji przez autorów artykułów. Stosowanie tak szerokiej anonimizacji może utrudniać powrót do badanych i poddanie ich ocenie wniosków z badania, gdyż taka procedura może doprowadzić do złamania, zagwarantowanej wcześniej, zasady poufności. Decyzja o szerokiej anonimizacji to także pewna przeszkoda w późniejszej archiwizacji danych. W przypadku moich badań nie zapytałam badanych o zgodę na archiwizację (wówczas taka praktyka nie była powszechna), a zatem aby zdeponować niezanonimizowane transkrypcje wywiadów w wybranym archiwum, musiałabym wrócić do badanych i poprosić ich o taką zgodę, a także zastrzec, iż możliwe do udostępnienia innym badaczom są tylko dane zanonimizowane. Nawet jeśli badacz pomyśli o uzyskaniu zgody na archiwizację zanonimizowanych danych, to i tak mogą okazać się one w niewielkim stopniu przydatne dla innego badacza, który nie ma szerokiej wiedzy o kontekście badania, a nie będzie w stanie uzyskać jej z danych [por. Kaźmierska 2014]. Paul Quinn [2017: 5–8] używa w tym kontekście pojęcia „marnowania danych” (*waste data*). Anonimizując, powodujemy, że dane na przykład w obszarze zdrowia są niemożliwe do dalszego wykorzystania. Tymczasem dyrektywa unijna dopuszcza przetwarzanie danych, gdy badacze posiadają zgodę jednostki bądź gdy badania prowadzone są w interesie publicznym. O tym, jak ważna jest wymiana tego typu danych, mogliśmy się przekonać podczas pandemii koronawirusa, gdy naukowcy z różnych zespołów na świecie udostępniali dane i wyniki badań, dzięki czemu mieli dostęp do znacznie większej liczby przypadków.

Oprócz problematyczności ponownego wykorzystania danych poddanych szerokiej anonimizacji Niamh Moore [2012] podkreśla również kwestię podważania anonimizacji w niektórych podejściach badawczych, np. w badaniach wizualnych czy badaniach małych społeczności (zaliczyć tutaj można zapewne także badania realizowane z wykorzystaniem studium przypadku). Zatem tam, gdzie anonimizacja byłaby najbardziej przydatna – na poziomie lokalnym – może stwarzać najwięcej problemów i trudno uznać ją za skuteczną strategię ochrony tożsamości badanych, zwłaszcza w sytuacji, gdy inne osoby ze społeczności są współpracownikami czy bliskimi danej osoby (tu wracamy do problemu poufności wewnętrznej [Tolich 2004]).

Przeprowadzona analiza pokazuje jednak, że refleksje nad tymi zagadnieniami są niemal nieobecne w publikacjach w polskich czasopismach socjologicznych, co może w pewnym stopniu potwierdzać obawy van den Scott [2020] o kształtowaniu się generacji badaczy, która, funkcjonując w określonym otoczeniu instytucjonalnym, nie będzie w stanie przeciwstawić się kolejnym regulacjom, przyjmując je jako nieodłączną część procesu badawczego.

BIBLIOGRAFIA

- Baez Benjamin.** 2002. "Confidentiality in qualitative research: Reflections on secrets, power and agency". *Qualitative Research* 2(1): 35–58. DOI: 10.1177/1468794102002001638.
- Corden Anne, Roy Sainsbury.** 2005. Research participants' views on use of verbatim quotations. Social Policy Research Unit, University of York. <http://www.york.ac.uk/inst/spru/pubs/pdf/verb.pdf> [access: 12.07.2020].
- Darling Jonathan.** 2014. "Emotions, encounters and expectations: The uncertain ethics of 'the field'". *Journal of Human Rights Practice* 6(2): 201–212. DOI: 10.1093/jhuman/huu011.
- Douglas Jack.** 1976. *Investigative social research: Individual and team field research*. Sage Library of Social Research. Beverly Hills, California: Sage Publications.
- Giermanowska Ewa, Adam Mrozowicki, Joanna Róg-Ilnicka.** 2019. „Socjologia pracy w Polsce: Społeczno-gospodarcze i polityczne problemy instytucjonalizacji subdyscypliny”. *Przegląd Socjologiczny* 68(3): 11–41.
- Grzeszkiewicz-Radulska Katarzyna, Aneta Krzewińska.** 2015. „O tym, jak napisać biografię naukową (łódzka szkoła metodologiczna)”. *Przegląd Socjologiczny* 64(4): 27–49.
- Guillemin Marilyns, Lynn Gillam.** 2004. "Ethics, reflexivity, and 'ethically important moments' in research". *Qualitative Inquiry* 10(2): 261–280. DOI: 10.1177/1077800403262360.
- Jaworska Dorota, Khedi Alieva, Marcin Boryczko.** 2019. „Badania wśród uchodźców i przymusowych migrantów – rozważania metodologiczne i etyczne”. *Przegląd Socjologii Jakościowej* 15(2): 202–227. DOI: 10.18778/1733-8069.15.2.11.
- Jordan Sara R.** 2014. "Research integrity, image manipulation, and anonymizing photographs in visual social science research". *International Journal of Social Research Methodology* 17(4): 441–454. DOI: 10.1080/13645579.2012.759333.
- Karen Kaiser.** 2009. "Protecting respondent confidentiality in qualitative research". *Qualitative Health Research* 19(11): 1632–1641. DOI: 10.1177/1049732309350879.
- Każmierska Kaja.** 2014. „Autobiograficzny wywiad narracyjny – kwestie etyczne i metodologiczne w kontekście archiwizacji narracji”. *Studia Socjologiczne* 3(214): 221–238.
- Każmierska Kaja, Katarzyna Waniek, Agata Zysiak.** 2016. *Opowiedzieć Uniwersytet. Łódź akademicka w biografjach wpisanych w losy Uniwersytetu Łódzkiego*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Każmierska Kaja.** 2018. „Doing biographical research – ethical concerns in changing social contexts”. *Polish Sociological Review* 3(203): 393–412.
- Kodeks etyczny Amerykańskiego Stowarzyszenia Socjologicznego (ASA).** 2018. American Sociological Association. https://www.asanet.org/sites/default/files/asa_code_of_ethics-june2018a.pdf [access: 12.05.2019].

- Kodeks etyczny Brytyjskiego Stowarzyszenia Socjologicznego (BSA).** 2017. British Sociological Association. https://www.britisoc.co.uk/media/24310/bsa_statement_of_ethical_practice.pdf [access: 12.05.2019].
- Kodeks etyki socjologa.** 2012. Polskie Towarzystwo Socjologiczne. <https://pts.org.pl/wp-content/uploads/2016/04/kodeks.pdf> [dostęp: 12.05.2019].
- Kodeks karny.** 1997. Dziennik Ustaw 88(553). <http://prawo.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU19970880553> [dostęp: 16.07.2020].
- Kodeks postępowania karnego.** 1997. Dziennik Ustaw 89(555). <http://prawo.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU19970890555> [dostęp: 16.07.2020].
- Lincoln Yvonna S.** 2009. Komisje etyczne i konserwatyzm metodologiczny. W: *Metody badań jakościowych* (tom 1), N.K. Denzin, Y.S. Lincoln (red.), 245–268. Warszawa: Wydawnictwo Naukowe PWN.
- Lowman John, Ted Palys.** 2014. “The betrayal of research confidentiality in British sociology”. *Research Ethics* 10(2): 97–118.
- Mackenzie Catriona, Christopher McDowell, Eileen Pittaway.** 2007. “Beyond ‘do no harm’: The challenge of constructing ethical relationships in refugee research”. *Journal of Refugee Studies* 20(2): 299–319. <https://doi.org/10.1093/jrs/fem008>.
- Męćfal Sylwia.** 2012. „Problemy badań terenowych – wybrane kwestie metodologiczne, praktyczne oraz etyczne w badaniu zjawisk «trudnych»”. *Przegląd Socjologiczny* 61(1): 156–178.
- Męćfal Sylwia.** 2019. *Prasa lokalna w relacjach z kluczowymi aktorami społecznymi. Studia przypadków*. Kraków/Łódź: Universitas/Wydawnictwo Uniwersytetu Łódzkiego.
- Mizielińska Joanna, Agata Stasińska, Magdalena Żadkowska, Mateusz Halawa.** 2018. „Dylematy etyczne w badaniu pary intymnej. Doświadczenia z pracy badawczej”. *Studia Socjologiczne* 3(230): 71–100. DOI: 10.24425/122473.
- Mondada Lorenza.** 2014. “Ethics in action: Anonymization as a participant’s concern and a participant’s practice”. *Human Studies* 37(2): 179–209.
- Moore Niamh.** 2012. “The politics and ethics of naming: Questioning anonymisation in (archival) research”. *International Journal of Social Research Methodology* 15(4): 331–340. DOI: 10.1080/13645579.2012.688330.
- Nespor Jan.** 2000. “Anonymity and place in qualitative inquiry”. *Qualitative Inquiry* 2000(6): 546–569. <http://qix.sagepub.com/cgi/content/abstract/6/4/546>.
- Olcoń-Kubicka Marta.** 2016. “Financial arrangement as a reflection of household order”. *Polish Sociological Review* 4: 477–494.
- O’Reilly Michelle, Khalid Karim, Helen Taylor, Nisha Dogra.** 2012. “Parent and child views on anonymity: ‘I’ve got nothing to hide’”. *International Journal of Social Research Methodology* 15(3): 211–223.
- Pismo dyrektora NCN w sprawie zarządzania danymi naukowymi.** 2019. Narodowe Centrum Nauki. https://www.ncn.gov.pl/sites/default/files/pliki/2019_04_03_pismo_dyrektora_NCN_zarzadzanie_danymi_naukowymi.pdf [dostęp: 12.07.2020].
- Quinn Paul.** 2017. “The anonymisation of research data – a pyrrhic victory for privacy that should not be pushed too hard by the EU Data Protection Framework?”. *European Journal of Health Law* 24(4): 1–21.
- Rekomendacje etyczne Polskiego Towarzystwa Historii Mówionej.** 2018. Polskie Towarzystwo Historii Mówionej. <http://pthm.pl/etyka/> [dostęp: 12.07.2020].

- Rogowski Łukasz.** 2019. „Etyczne dylematy badań biograficznych z wykorzystaniem fotografii: perspektywa badacza i badanych”. *Przegląd Socjologii Jakościowej* 15(2): 58–74. DOI: 10.18778/1733-8069.15.2.04.
- Rozporządzenie Parlamentu Europejskiego i Rady UE o ochronie danych.** 2016. Rozporządzenie Parlamentu Europejskiego i Rady UE 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych). <https://uodo.gov.pl/pl/404/539> [dostęp: 16.07.2020].
- Saunders Benjamin, Jenny Kitzinger, Celia Kitzinger.** 2015. “Anonymising interview data: Challenges and compromise in practice”. *Qualitative Research* 15(5): 616–632. DOI: 10.1177/1468794114550439.
- Simons Helen.** 2009. *Case study research in practice*. Thousand Oaks, CA: Sage Publications.
- Stake Robert E.** 1995. *The art of case study research*. Thousand Oaks, London, New Delhi: Sage Publications.
- Surmiak Adrianna.** 2018. “Confidentiality in qualitative research involving vulnerable participants: Researchers’ perspectives”. *Forum Qualitative Sozialforschung/Forum Qualitative Social Research* 19(3): 55 paragraphs, art. 12. DOI: 10.17169/fqs-19.3.3099.
- Surmiak Adrianna.** 2019a. “Should we maintain or break confidentiality? The choices made by social researchers in the context of law violation and harm”. *Journal of Academic Ethics*. DOI: 10.1007/s10805-019-09336-2.
- Surmiak Adrianna.** 2019b. „Komisje etyczne dla badań społecznych w Polsce. Perspektywa socjologów i antropologów społeczno-kulturowych”. *Studia Socjologiczne* 4 (235): 157–182. <https://doi.org/10.24425/sts.2019.126163>.
- Świątek-Młynarska Paulina.** 2019. „Problemy etyczne w relacji badacza z osobami badanymi (na przykładzie badań biograficznych z osobami starszymi)”. *Przegląd Socjologii Jakościowej* 15(4): 204–223. DOI: 10.18778/1733-8069.15.4.10.
- Taylor Rebecca.** 2015. “Beyond anonymity: Temporality and the production of knowledge in a qualitative longitudinal study”. *International Journal of Social Research Methodology* 18(3): 281–292. DOI: 10.1080/13645579.2015.1017901.
- Tolich Martin.** 2004. “Internal confidentiality: When confidentiality assurances fail relational informants”. *Qualitative Sociology*. 27(1): 101–106. DOI: 10.1023/B:QUAS.0000015546.20441.4a.
- Ustawa o ochronie danych osobowych.** 1997 (2018). Dziennik Ustaw 133 (883). <http://prawo.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=wdu19971330883> (Dziennik Ustaw 1000. <http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WDU20180001000/O/D20181000.pdf>. [dostęp: 16.07.2020].
- van den Scott Lisa-Jo K.** 2020. Sociology and ethics. Doing the right thing. In: *Handbook of research ethics and scientific integrity*, R. Iphofen (ed.), 769–782 (chapter 44). Cham, Switzerland: Springer.
- Wiles Rose, Graham Crow, Sue Heath, Vikki Charles.** 2008. “The management of confidentiality and anonymity in social research”. *International Journal of Social Research Methodology* 11(5): 417–428. DOI: 10.1080/13645570701622231.
- Wiles Rose, Amanda Coffey, Judy Robinson, Sue Heath.** 2012. “Anonymisation and visual images: Issues of respect, ‘voice’ and protection”. *International Journal of Social Research Methodology* 15(1): 41–53. DOI: 10.1080/13645579.2011.564423.

- Woroniecka Grażyna.** 2014. Kodeks etyczny w socjologii – wybrane problemy. W: *Idee – Historia – Społeczeństwo*, M. Fuszara, W. Pawlik (red.), 240–254. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Yin Robert K.** 2009. *Case study research. Design and methods* (Fourth Edition). Thousand Oaks, CA: Sage Publications.
- Yin Robert K.** 2015. *Studium przypadku w badaniach naukowych. Projektowanie i metody*, tłum. J. Gilewicz Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Wyka Anna.** 1990. „Ku nowym wzorom badań społecznych w Polsce. Cechy badań jakościowych w ostatnich latach”. *Kultura i Społeczeństwo* 34(1): 161–173.
- Wyka Anna.** 1993. *Badacz wobec doświadczenia*. Warszawa: Wyd. IFiS PAN.
- Zalecenia Rady Narodowego Centrum Nauki dotyczące badań z udziałem ludzi.** 2016. Narodowe Centrum Nauki. https://www.ncn.gov.pl/sites/default/files/pliki/2016_zalecenia_Rady_NCN_dot_etyki_badan.pdf [dostęp: 12.07.2020].

Sylwia Męcfal

THE ANALYSIS OF ANONYMISATION PRACTICES IN POLISH QUALITATIVE SOCIOLOGICAL RESEARCH IN THE LIGHT OF THE LITERATURE OF THE SUBJECT AND ETHICAL AND LEGAL REGULATIONS

Abstract

The aim of the article is to present the data anonymisation practices used in qualitative research by Polish sociologists and to combine those practices into types of data anonymisation. In order to realise this goal, I analysed the articles published in four major Polish sociological journals between 2015 and 2019.

The analysis led to distinguishing various ways of data anonymisation which then were additionally combined into types of anonymisation by cluster analysis. Four types were identified: basic default anonymisation, leading pseudonymisation, anonymisation extended by location anonymisation, and anonymisation extended by providing the features of participants at a group level (not at an individual level). These types of anonymisation reflect the types of anonymisation that are described in the literature and confirm that researchers treat anonymisation as a natural and inherent part of social research (as stated in the “Sociologist’s code of ethics” by the Polish Sociological Association). Researchers rarely use the real identities of research participants and the procedure of authorisation (clearance). The anonymisation strategies in the analysed articles are often the default ones and are rather simplified; extended anonymisation is used less often.

Keywords: research ethics, anonymisation, confidentiality, authorisation of data, codes of ethics