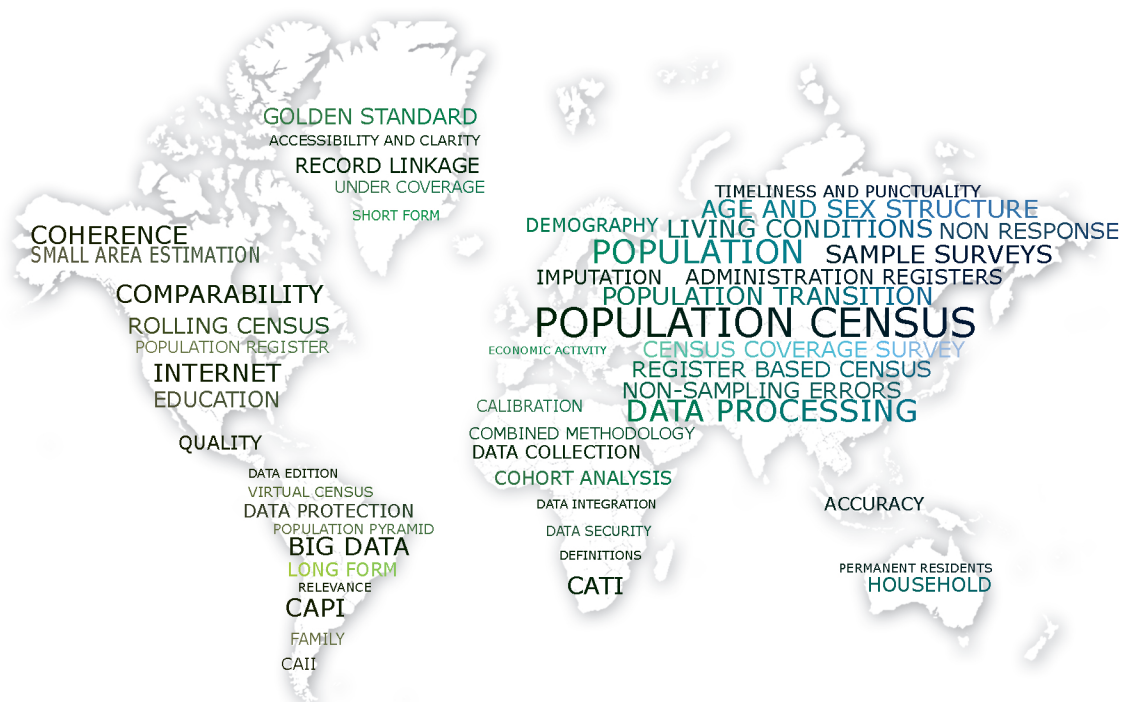


# ELŻBIETA GOŁATA



## KONIEC ERY TRADYCYJNYCH SPISÓW LUDNOŚCI

WYDAWNICTWO UEP



UNIWERSYTET  
EKONOMICZNY  
W POZNANIU

KONIEC  
ERY TRADYCYJNYCH  
SPISÓW LUDNOŚCI

ELŻBIETA GOŁATA

# KONIEC ERY TRADYCYJNYCH SPISÓW LUDNOŚCI

**WYDAWNICTWO UEP**



UNIWERSYTET  
EKONOMICZNY  
W POZNANIU

Poznań 2018

## KOMITET REDAKCYJNY

*Szymon Cyfert, Elżbieta Gołata (przewodnicząca), Jacek Lisowski, Ewa Małuszyńska,  
Jerzy Schroeder (sekretarz), Krzysztof Walczak, Ryszard Zieliński*

## RECENZENT

*Ireneusz Kuropka*

## PROJEKT OKŁADKI

*Boobry Group  
Marta Brzóstowicz*

## REDAKCJA

*Anna Grześ*

## KOREKTA

*Deal*

**Rozdziały 3 i 4 powstały w ramach projektu: Estymacja pośrednia w zakresie badania niepełnosprawności na podstawie NSP 2011**

**Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/11/B/HS4/01472**

© Copyright by Uniwersytet Ekonomiczny w Poznaniu  
Poznań 2018



Ta książka jest udostępniana na licencji Creative Commons – Uznanie autorstwa-Użycie niekomercyjne-Bez utworów zależnych 4.0 Międzynarodowe

**ISBN 978-83-7417-971-3; e-ISBN 978-83-66199-11-8**

**<https://doi.org/10.18559/978-83-66199-11-8>**

WYDAWNICTWO UNIwersYTETU EKONOMICZNEGO W POZNANIU

ul. Powstańców Wielkopolskich 16, 61-895 Poznań

tel. 61 854 31 54, 61 854 31 55

www.wydawnictwo.ue.poznan.pl, e-mail: wydawnictwo@ue.poznan.pl

adres do korespondencji: al. Niepodległości 10, 61-875 Poznań

Skład: Wydawnictwo eMPI<sup>2</sup>

*Reginaldo Cammarano*

Druk: UNI-DRUK Wydawnictwo i Drukarnia

ul. Przemysłowa 13, 62-030 Luboń

*Mojemu Mężowi*

---

## SPIS TREŚCI

---

<b>Wprowadzenie</b> .....	11
<b>1. Spis ludności jako źródło danych o ludności</b> .....	21
1.1. Spis ludności – historia, cele i zadania .....	21
1.2. Kontrowersje wobec podstawowych pojęć i własności .....	26
1.2.1. Definicja i podstawowe własności spisu .....	26
1.2.2. Kategorie ludności spisowej .....	30
1.2.3. Ludność rezydująca a faktyczna w NSP 2002 i NSP 2011 .....	35
1.2.4. Minimum spisowe .....	38
1.3. Zapotrzebowanie na dane spisu ludności .....	43
1.4. Kryteria oceny jakości badań statystycznych .....	49
<b>2. Transformacja metod przeprowadzania spisów</b> .....	53
2.1. Metody przeprowadzania spisu a źródła danych .....	53
2.1.1. Spis tradycyjny .....	61
2.1.2. Spis tradycyjny z coroczną aktualizacją danych .....	62
2.1.3. Spis kroczący .....	63
2.1.4. Spis oparty na rejestrach .....	64
2.1.5. „Spis wirtualny” .....	74
2.1.6. Metody mieszane .....	75
2.2. NSP 2011 – spis ludności metodą mieszaną .....	77
2.3. Alternatywne metody przeprowadzania spisów w rundzie 2010 .....	83
2.4. Spisy tradycyjne i nowej generacji w świetle międzynarodowych kryteriów jakości .....	90
2.4.1. Jakość spisów według kryteriów Eurostatu .....	90
2.4.2. Ekonomiczno-społeczne problemy jakości spisów .....	98
<b>3. Jakość spisów</b> .....	105
3.1. Metody oceny jakości spisów ludności .....	105
3.1.1. Klasyfikacje błędów i metod oceny jakości spisów .....	105
3.1.2. Spis kontrolny .....	111
3.2. Problemy jakości w spisach nowej generacji .....	116
3.3. Jakość spisów tradycyjnych w Polsce .....	118
3.4. Ocena jakości NSP 2011 .....	123

## 8 Spis treści

3.4.1. Ocena kompletności NSP 2011 z wykorzystaniem metod analizy demograficznej .....	123
3.4.2. Jakość szacunków dzieci i niemowląt – analiza kompletności urodzeń .....	129
3.4.3. Jakość szacunków – analiza kompletności na podstawie wskaźników „przetrwania” .....	138
3.4.4. Jakość szacunków osób starszych – analiza stanu i porządku wymiarowania stulatków .....	142
3.4.5. Ocena długiego formularza NSP 2011 .....	148
<b>4. Perspektywy spisów ludności .....</b>	<b>153</b>
4.1. Perspektywy spisu w świetle wytycznych ONZ i Eurostatu .....	153
4.2. Perspektywy spisów rundy 2020 w wybranych państwach .....	155
4.2.1. Stany Zjednoczone Ameryki .....	155
4.2.2. Wielka Brytania .....	158
4.3. NSP 2021 w perspektywie planów GUS .....	161
<b>Zakończenie .....</b>	<b>181</b>
<b>Aneks .....</b>	<b>187</b>
I. Informacje uzupełniające rozdział I: Spis ludności jako źródło danych o ludności .....	187
I.1. Definicje populacji według Rekomendacji ONZ .....	187
I.2. Rekomendacje dotyczące sposobu kwalifikacji ludności według miejsca zamieszkania w przypadku wątpliwości .....	189
I.3. Wyjaśnienia dotyczące szacunków dla populacji trudnych do spisania .....	190
I.4. Wyjaśnienie GUS dotyczące liczby mieszkańców Polski .....	192
I.5. Definicje ludności w NSP 2011 .....	194
I.6. Obszary tematyczne spisu, tematy i ocena stopnia zaspokojenia potrzeb, konsultacje społeczne przed spisem 2021 r. ....	195
I.7. Cechy zespołowe (1–3) oraz demograficzne (4–25) uwzględnione w poszczególnych spisach ludności w Polsce .....	196
I.8. Cechy społeczno-zawodowe uwzględniane w powszechnych spisach ludności w Polsce .....	197
I.9. Najczęściej zgłaszane dodatkowe potrzeby informacyjne i powiązania tematów, konsultacje społeczne przed spisem 2021 r. ....	199
I.10. Definicja jakości w statystyce publicznej .....	200
II. Informacje uzupełniające rozdział II: Transformacja metod przeprowadzania spisów .....	203
II.1. Rejestry administracyjne i systemy informacyjne wykorzystane w NSP 2011 .....	203
II.2. Zakres tematyczny formularza do badania reprezentacyjnego w NSP 2011 .....	205
II.3. Sposób wykorzystania centralnego rejestru ludności w statystyce ludności, wybrane państwa europejskie .....	208
III. Informacje uzupełniające rozdział III: Jakość spisów .....	209
III.1. Ocena kompletności NSP 2011: Rozbieżności szacunków bilansu ludności generacji 1916 i wcześniejszych na podstawie danych ewidencji ruchu naturalnego i stanów ludności według roczników urodzenia .....	209

III.2. Ocena kompletności NSP 2011: Liczba osób starszych według Bazy Danych Demografia GUS oraz szacunku bilansu na podstawie danych ewidencji zgonów według roku urodzenia w latach 2006–2016 .....	212
IV. Informacje uzupełniające rozdział IV: Perspektywy spisów ludności .....	215
IV.1. Wykaz pozastatystycznych systemów informacyjnych, rejestrów i baz danych stanowiących potencjalne źródło danych dla NSP 2021 .....	215
IV.2. Metryka zmiennej statystycznej według koncepcji spisu 2021 r. ....	216
IV.3. Planowana analiza jakości mikrodanych operacyjnych według koncepcji spisu 2021 r. ....	217
<b>Bibliografia</b> .....	219
<b>Spis tabel i rysunków</b> .....	235



---

## WPROWADZENIE

---

*The Official Statistics Olympic Challenge:  
Wider, Deeper, Quicker, Better, Cheaper<sup>1</sup>*

Sformułowane przez Holta (2007) motto tej książki jest wyjątkowo trafną syntezą politycznych i społecznych oczekiwań oraz presji wobec statystyki oficjalnej. Jest to pięć „olimpijskich” wyzwań stawianych statystykom przez współczesny świat, dotyczących tego, jak dostarczać informacji: **szerzej, głębiej, szybciej, lepiej i taniej**. Czy statystyka oficjalna, której zadaniem jest zapewnienie rzetelnego systemu informacji, jest w stanie sprostać tym oczekiwaniom? Zamiast odpowiedzi na to pytanie pojawiają się kolejne pytania, dotyczące przede wszystkim rozwoju metod umożliwiających realizację zadania i zapewnienia jakości otrzymanych w efekcie informacji. Właśnie dlatego, że nie ma prostej odpowiedzi, nie wolno zaprzestać jej poszukiwania, dążenia do poznania, konstruowania nowych rozwiązań w zmieniającym się otoczeniu i uwarunkowaniach. Idea ta przyświeca przedstawionej na łamach niniejszej książki dyskusji o możliwości realizacji owych wyzwań w odniesieniu do spisu ludności, jednego z największych i najważniejszego badania demograficznego i statystycznego, którego wyniki są podstawą systemu zarządzania demokratycznym państwem.

Nowoczesne technologie dynamizujące rozwój gospodarki oraz ich wszechstronna obecność w każdej sferze życia społecznego i kulturowego wywierają piętno również na rozwój statystyki. Dostarczają nowych, różnorodnych źródeł danych, diametralnie zmieniają metody ich pozyskiwania i gromadzenia, przetwarzania i analizowania, wręcz przeobrażając cały proces badania statystycznego. Komputeryzacja i cyfryzacja oraz wszechobecne technologie informacyjno-komunikacyjne (*Information and Communication Technologies, ICT*), przede wszystkim powszechny dostęp do internetu, zmieniają także sposób przeprowadzania spisu ludności.

---

<sup>1</sup> Tytuł wykładu prof. D. Tima Holta (2007) podczas seminarium z okazji setnego posiedzenia Committee on National Statistics of The National Academy of Sciences.

Zmiana w zakresie organizacji badań statystycznych ma charakter fundamentalny. Dotyczy modelowego ujęcia procesu badania statystycznego rozumianego w sposób kompleksowy, nie tylko jako gromadzenie danych (Wallgren i Wallgren, 2014). Przez wieki spis powszechny był podstawowym sposobem pozyskiwania informacji o całej populacji danego kraju. Rozwój metody reprezentacyjnej wprowadził możliwość ograniczenia badania jedynie do losowo wybranych jednostek z tej populacji. W drugiej połowie XX w. powszechne spisy ludności przeprowadzano zazwyczaj raz na 10 lat<sup>2</sup>, gdyż badanie to było zbyt kosztowne i czasochłonne. Obecnie jednak, poszukując informacji o populacji, zanim zaprojektowane zostanie badanie reprezentacyjne, w pierwszej kolejności dokonuje się kwerendy istniejących źródeł, przede wszystkim rejestrów administracyjnych, ale także innych, pozastatystycznych źródeł<sup>3</sup>. Wzrost wykorzystania rejestrów administracyjnych przez statystykę publiczną Wallgren i Wallgren (2014) określają mianem transformacji badań statystycznych w kierunku systemu opartego na rejestrach z systemu bazującego na gromadzeniu danych w wyniku indywidualnych wywiadów. Dane z tych źródeł są dostępne natychmiast, ale charakteryzuje je duża różnorodność, odmienna jakość i właściwości. Dlatego szczególnego znaczenia nabiera ich wnikliwa ewaluacja, edycja i integracja. Wykorzystanie rejestrów przez statystykę publiczną wymaga rozwiązania wielu problemów, m.in. integracji danych, estymacji pośredniej, imputacji, kalibracji, zastosowania technologii systemu informacji geograficznej (*Geographical Information System, GIS*), niestandardowych konstrukcji operatu losowania, badań internetowych, upowszechniania danych oraz ochrony ich poufności (Baffour, King, i Valente, 2013; Kena i in., 2015; Zhang, 2012). Oznacza to reorganizację statystyki publicznej, co może być postrzegane w kategoriach zmiany paradygmatu badań statystycznych (Zhang, 2015, 2013).

W ostatniej rundzie spisów około 2010 r. byliśmy świadkami historycznych przeobrażeń o niespotykanym zasięgu. Ogromne zmiany i modyfikacje dotyczyły każdego etapu spisu. Dążąc do redukcji kosztów, ale także poprawy jakości szacunków oraz terminowości i formy upowszechniania wyników, poszczególne państwa decydowały się na wprowadzanie nowoczesnych rozwiązań technologicznych (ICT) czy alternatywnych metod przeprowadzania spisu (UN, 2013b). Szeroko wykorzystywano technologie systemu informacji geograficznej, urządzenia mobilne, kwestionariusze internetowe, mapy cyfrowe, systemy nawigacji GPS (*Global*

<sup>2</sup> W niektórych państwach spisy sporządzano co 5 lat, tak było np. we Francji (por. rozdz.1.1) oraz w Kanadzie, gdzie ta praktyka nadal się utrzymuje. W Polsce w okresach międzyspisowych trzykrotnie przeprowadzono tzw. mikrospis. Były to spisy przeprowadzone metodą reprezentacyjną w latach: 1974, 1984 i 1995.

<sup>3</sup> Przykładem mogą być spisy abonentów telefonicznych, wykazy odbiorców energii elektrycznej, dane geodezyjne. Ponadto nowoczesne technologie informacyjno-komunikacyjne dostarczają niewiarygodnych ilości danych z dotychczas niedostępnych źródeł. 'Big data' są rzeczywistością, stanowiącą wyzwanie dla statystyki publicznej, przede wszystkim pod względem oceny reprezentatywności (Beręsewicz, 2017).

*Positioning System*, GPS), telefony komórkowe, umożliwiające monitorowanie prac w terenie. Coraz większe zainteresowanie skupiały alternatywne metody spisu wykorzystujące zasoby rejestrów administracyjnych. Metodologia spisu kroczącego została po raz pierwszy zastosowana we Francji. W Stanach Zjednoczonych po raz pierwszy badanie gospodarstw domowych *American Community Survey* (ACS) zastąpiło tradycyjne badanie reprezentacyjne gromadzące dane dla tzw. długiego formularza spisowego. Nie sposób wymienić wszystkich procedur spisowych i wskazać wszystkich poddanych modyfikacjom obszarów, jak również przewidzieć ich konsekwencji. Proces zmian nie jest zakończony i trwa nadal, dla ukazania skali wprowadzanych zmian warto podkreślić, że:

- tradycyjne spisywanie zastąpiono ekstrakcją danych z innych źródeł, głównie z rejestrów administracyjnych;
- zbieranie danych w terenie zastąpiono transferem za pośrednictwem internetu;
- kwestionariusz papierowy zamieniono na elektroniczny;
- wykorzystanie GIS umożliwiło gromadzenie i upowszechnianie informacji dla dowolnie zdefiniowanych jednostek przestrzennych (na podstawie współrzędnych geograficznych);
- wykorzystanie technik ICT usprawniło gromadzenie i upowszechnianie informacji;
- wykorzystanie informacji z wielu źródeł stymulowało rozwój i zastosowanie nowoczesnych metod:
  - imputacji,
  - statystycznej integracji danych,
  - kalibracji,
  - estymacji statystycznej wykorzystującej zmienne pomocnicze z dodatkowych źródeł;
- wykorzystanie informacji z wielu źródeł stymulowało poprawę jakości szacunków spisowych.

Niektóre państwa, np. Austria czy Niemcy, wprowadziły zmiany fundamentalne, kompleksowo obejmujące metodologię spisu od określenia źródeł danych, poprzez metody ich gromadzenia, po estymację i upowszechnienie wyników. Inne państwa zdecydowały się jedynie na pewne innowacje w zakresie techniki zbierania i upowszechniania danych (Wielka Brytania, Słowacja, Węgry). Niejednokrotnie towarzysząca tym modyfikacjom dyskusja przybierała bardzo gwałtowny charakter (Ditchburn, 2010; Grand, 2010; Slany, 2014)<sup>4</sup>, a formułowane wnioski sugerowały „koniec spisów ludności” (Baffour i in., 2013; Coleman, 2013; Poulain i Herm, 2013; Valente, 2010). Wobec powyższych konkluzji Kukutai, Thompson i McMillan (2015) w latach 1985–2014 przeprowadzili badanie praktyk

<sup>4</sup> Przykładem może tutaj być Kanada, gdzie decyzja rządu zmieniająca reprezentacyjny charakter badania gospodarstw domowych na dobrowolny spowodowała rezygnację dr. Munira Sheikha, prezesa Statistics Canada (Urzędu Statystycznego Kanady).

dotyczących spisów w 241 niezależnych krajach<sup>5</sup>, którego celem była próba odpowiedzi na pytanie, czy rzeczywiście jesteśmy świadkami „zmierzchu spisów”. We wnioskach autorzy stwierdzają, że liczba krajów przeprowadzających spis ludności w badanym okresie wzrosła, a zasadnicza transformacja metodologii dotyczy głównie państw europejskich. Jest ona rezultatem powszechnej dostępności danych, coraz większych możliwości organizacyjnych, w tym będących efektem innowacji technologicznych.

**Celem niniejszego opracowania jest przedstawienie obserwowanych w ostatnim czasie zmian metod przeprowadzania spisów ludności oraz ocena ich jakości. Ukazana zostanie różnorodność proponowanych rozwiązań, przy jednoczesnym wskazaniu ich możliwości i ograniczeń.**

Tak nakreślony cel ogólny opracowania składa się z następujących celów szczegółowych:

- dyskusji podstawowych pojęć i własności spisu ludności;
- wskazania konsekwencji przyjęcia określonej definicji ludności spisowej;
- klasyfikacji i prezentacji metod przeprowadzania spisów;
- oceny jakości spisów, ich możliwości i ograniczeń w świetle międzynarodowej definicji jakości badań statystycznych;
- przedstawienia rodzajów błędów i metod ewaluacji spisów ludności;
- oceny kompletności NSP 2011;
- próby odpowiedzi na pytanie o perspektywy spisu ludności w świetle dostępnych źródeł danych i rozwoju metodologii, z uwzględnieniem międzynarodowych standardów statystyki publicznej.

Głównym powodem zmian metodologii przeprowadzania spisu była chęć redukcji kosztów, ale wszystkie państwa zgodnie podkreślają także dążność do poprawy jakości wyników spisu. Ograniczenie kosztów spisu można osiągnąć w różny sposób, jednak niewątpliwie najważniejsze znaczenie w tej kwestii należy przypisać decyzji dotyczącej wykorzystywanych źródeł danych. Dlatego tak istotne jest rozstrzygnięcie dotyczące tradycyjnego spisowania całej populacji bądź ograniczenia się tylko do podstawowych pytań demograficznych (krótki formularz), podczas gdy odpowiedź na pytania dotyczące kwestii społecznych i ekonomicznych (długi formularz) jest opracowana na podstawie reprezentacyjnego badania gospodarstw domowych. Zasadnicza zmiana dotyczy w szczególności wykorzystania danych rejestrów administracyjnych i innych źródeł pozastatystycznych przy rezygnacji z tradycyjnego badania w terenie.

---

<sup>5</sup> Pojęcie „country” użyte w publikacji obejmuje suwerenne państwa, terytoria zależne i inne obszary ujęte w Kodeksie Handlu oraz Lokalizacji Transportowych Organizacji Narodów Zjednoczonych (Code for Trade and Transport Locations). Wydany przez Głównego Geodetę Kraju *Urządowy wykaz nazw państw i terytoriów niesamodzielnych* obejmuje 195 państw uznawanych przez Rzeczpospolitą Polską (tj. 193 państwa członkowskie Organizacji Narodów Zjednoczonych, Kosowo i Watykan) i 69 terytoriów niesamodzielnych (Zych i Czerny, 2015).

Wobec powyższego w pracy skoncentrowano się na omówieniu zmian wynikających z wykorzystania dla celów spisu ludności rejestrów administracyjnych oraz innych alternatywnych źródeł informacji. Konsekwencje takiego podejścia dyskutowano z perspektywy oceny jakości wyników spisu. Założono, że o jakości danych decyduje nie tylko metoda ich gromadzenia, lecz także, jeśli nie przede wszystkim, zgodność i poprawność pojęć i definicji, pozwalająca na realizację celu. Zwrócono uwagę, że procedura wykorzystania wielu źródeł danych wprowadza mechanizm badania zgodności, wzajemnej oceny i porównań, wymaga harmonizacji. Porównanie wyników spisu tradycyjnego z danymi takich rejestrów, jak PESEL, PIT czy ZUS<sup>6</sup>, ale także przykładowo z danymi gmin dotyczącymi zużycia energii elektrycznej, wody czy zbiórki śmieci, prawdopodobnie ukazałoby istotnie różne liczebności populacji. Dopiero wzajemna analiza wielu źródeł pozwoli wychwycić i wyjaśnić rozbieżności w różnych przekrojach (np. płeć, wiek, miejsce zamieszkania). Badanie, w którym przeprowadza się integrację danych z wielu źródeł, ma wbudowany mechanizm weryfikacji jakości. Sam ten fakt stanowi przewagę nad podejściem zakładającym bezgraniczne zaufanie tylko do jednego źródła.

Powyższy tok rozumowania pozwolił na sformułowanie tezy, że **spisy nowej generacji oparte na wielości źródeł są lepszej jakości aniżeli spis, który bazuje wyłącznie na jednym źródle danych. Lepsza jakość tych spisów wynika nie tylko z tego, że odpowiadają one zasadzie *eine Quelle, keine Quelle*<sup>7</sup>. Spisy nowej generacji to wykorzystanie informacji lepszych, tańszych i pewniejszych, pozyskanych metodą kija i marchewki (Paradysz, 2007b, s. 197) z użyciem najnowszych metod estymacji.**

Zgodnie z zasadą: „garbage in – garbage out” przyjęto, że o jakości wyników badań i analiz (w przypadku spisu – stanu ludności i struktur demograficznych) decyduje jakość danych wejściowych. Zasada ta obejmuje zarówno dane pochodzące z tradycyjnego spisowania, badań reprezentacyjnych, samospisu internetowego, jak i dane rejestrów administracyjnych. Ważne jest dążenie do poprawy jakości źródeł.

Jeśli chodzi o źródła statystyki publicznej, wiele uczyniono i nadal wiele się czyni w zakresie poprawy jakości wyników badań specjalnych. Dużo uwagi poświęca się projektowaniu badania, schematowi doboru próby, metodom estymacji oraz oceny błędów szacunku. Szczegółowe informacje w tym zakresie stają się obowiązkowym

---

<sup>6</sup> Powszechny Elektroniczny System Ewidencji Ludności (PESEL) to rejestr ludności utworzony w Polsce w 1968 r. (Okólski, 2004, s. 37) i prowadzony przez ministra właściwego do spraw informatyzacji (do końca 2015 r. przez ministra właściwego do spraw wewnętrznych); Personal Income Tax (PIT) to rejestr zeznań podatkowych od osób fizycznych prowadzony przez Ministerstwo Finansów; Centralny Rejestr Płatników Składek i Centralny Rejestr Ubezpieczonych prowadzi Zakład Ubezpieczeń Społecznych (ZUS);

<sup>7</sup> „... musimy się nauczyć pracować w warunkach wielu źródeł. W naukach społecznych i gospodarczych jedynie rzadko mamy do czynienia z jednym i niepodważalnym źródłem informacji. *Eine Quelle, keine Quelle*, jak mawiali nasi zachodni sąsiedzi” (Paradysz, 2007a). *Eine Quelle, keine Quelle* przetłumaczyć można następująco: *jedno źródło – żadne źródło*.

elementem każdej publikacji wyników badań prowadzonych przez GUS. Podkreślić należy także wysiłki podejmowane w celu niwelowania wpływu braku odpowiedzi i udziału w badaniu, w szczególności poprzez stosowanie odpowiednich metod imputacji i kalibracji. Niestety GUS nie publikował informacji w zakresie oceny jakości danych tradycyjnego spisu. Ewaluacje takie podejmowane były przede wszystkim przez demografów i statystyków. Jednak w powszechnej opinii wyniki spisu, jako badania pełnego, uznawane były za bezbłędne.

W przypadku rejestrów administracyjnych w zasadzie dopiero pozytywna ewaluacja upoważnia do wykorzystania zawartych w nich danych w spisie. Wymusza to niejako prace nad poprawą jakości rejestrów. Oczywiście prace takie mogą i z pewnością są prowadzone niezależnie, jednak dodatkowe wykorzystanie rejestrów jako źródła danych dla statystyki publicznej i wynikające z tego tytułu wymogi w zakresie wiarygodności nie tylko przyspieszą proces, ale uczynią go także bardziej wszechstronnym. Zakładając, że prace nad poprawą jakości rejestrów są prowadzone w formie międzyresortowych uzgodnień z uwzględnieniem różnych zastosowań danych, sformułowano hipotezę dodatkową, że **wykorzystanie zasobów rejestrów administracyjnych dla celów spisu ludności sprzyja poprawie ich jakości**. Poprawa jakości danych rejestrów pozytywnie oddziałuje także na jakość funkcjonowania służb administracji państwowej.

Podstawowym warunkiem rzetelnych analiz jest przede wszystkim właściwe określenie stanu i struktury ludności. Spis ludności jest tym badaniem demograficznym, którego głównym celem jest dostarczenie oszacowań struktur demograficznych. Z mocy prawa to dane spisu stanowią podstawę określania wysokości subwencji państwowych. Za publikowanymi przez GUS liczbami ludności stoją konkretne pieniądze, subwencje, dotacje czy dofinansowanie ze środków Unii Europejskiej, o które zabiegają różne instytucje, w tym organy administracji państwowej. Spisowy szacunek liczby ludności winien więc być rozpatrywany także jako problem administracyjny, samorządowy, instytucjonalny, finansowy<sup>8</sup>. Dane spisu ludności są też podstawą korekt liczby ludności w okresach międzyspisowych, oczekuje się więc, że będą one bardzo wiarygodnym źródłem danych. Nieprawidłowo oszacowane liczebność i struktura populacji są istotną przyczyną błędów prognoz i analiz demograficznych (Alho i Spencer, 2005; Bijak, Kicingier, i Kupiszewski, 2007; Keilman i Kučera, 1991). Jednak **konsekwencje złej jakości szacunków ludności dotyczą nie tylko badań demograficznych, ale wszystkich dziedzin życia, w tym gospodarczej, społecznej i politycznej, oraz wszystkich decyzji opartych na wynikach analiz, w których wykorzystuje**

---

<sup>8</sup> „W spisie 2011 r. szczególnie głośne były protesty burmistrza gminy Muszyna w Małopolsce, który twierdził, że w miasteczku mieszkają 4902 osoby, podczas gdy według szacunku GUS mieszkańców było 5115 osób. Różnica zaledwie 213 osób miała dla samorządu ogromne znaczenie – utraty prawie 1,5 mln zł subwencji oświatowej” (Rzeczpospolita 18.08.2014, <http://www.portalsamorzadowy.pl/lodzkie/edukacja/samorzady-chca-zmian-w-naliczaniu-subwencji-oswiatowej,62120.html>).

**się takie wskaźniki jak PKB *per capita*, dla których podstawą odniesienia jest liczba ludności.**

Tematyka spisu ludności i jego jakości należy do wyjątkowo obszernych. Jednak ze względu na ograniczone ramy opracowania oraz dążność do zachowania spójności wywodu niektóre z problemów zostaną zaledwie wzmiankowane i ujęte w sposób bardzo syntetyczny. Dotyczy to przede wszystkim charakterystyki przemian podstawowych procesów demograficznych. Autorka jest świadoma, że jednym z zasadniczych celów spisu jest przedstawienie jak najrzetelnieszego obrazu badanej populacji. Wnikliwy opis zmian demograficznych i społecznych jest przedmiotem innych opracowań. Ich znajomość jest oczywiście niezbędna do wiarygodnego, liczbowego opisu populacji, jej aktywności i warunków życia, jakim jest spis ludności. I w takim zakresie zostaną one uwzględnione. Wskazać tutaj można przede wszystkim problemy szacunku liczby ludności w spisie tradycyjnym, jak też wykorzystującym nowe źródła danych, które wynikają ze wzrostu mobilności społecznej, przeobrażeń dotyczących rodziny, związków nieformalnych, zamieszkiwania dzieci na przemian z jednym z rodziców, w rodzinie patchworkowej (Grotowska-Leder, 2014a; Grotowska-Leder i Rokicka, 2015; Kotowska, 2009b, 2010; Matysiak, 2014; Mynarska, Baranowska-Rataj, i Matysiak, 2014; Perelli-Harris i in., 2014; Slany, Ślusarczyk i Krzyżowski, 2014; Taranowicz i Grotowska, 2015). Wobec coraz bardziej zróżnicowanej populacji i intensywnych migracji powstaje pytanie, jak poprawnie oszacować liczbę ludności z uwzględnieniem posiadania przez jednych kilku miejsc zamieszkania oraz żadnego przez innych (osoby bezdomne). W jaki sposób dynamiczny rozwój miast, w tym suburbanizacja i przyciąganie mieszkańców przez ośrodki akademickie, gospodarcze i kulturowe, wpływają na poprawność szacunku stanu i struktury ludności (Gołata i Kuropka, 2016; Grotowska-Leder, 2002; Szymańska, 2013). Na ile tradycyjne ujęcie poziomu wykształcenia zgodnie z tzw. ramami kwalifikacyjnymi oddaje obraz społeczeństwa w rzeczywistości zdalnego czy ustawicznego kształcenia (Grotowska-Leder, 2014b). Jak wzrost natężenia migracji czy rozwój przedsiębiorczości, upowszechnianie innych form zatrudnienia, np. pracy zdalnej, wpływają na pomiar aktywności ekonomicznej ludności i rzetelną ocenę sytuacji na rynku pracy (Grotowska-Leder, 2014c; Kuropka i Pisz, 2008; Matysiak i Mynarska, 2013; Witkowski, 2002a, 2002b). W jaki sposób gwałtownie zmieniające się wykorzystanie nowoczesnych technologii komunikacyjnych łącznie z generowanymi przez nie informacjami może usprawnić proces gromadzenia danych, wzbogacić ich zakres, poprawić jakość. Jak zapobiec konsekwencjom nierównomiernego dostępu do ICT czy wręcz wykluczenia cyfrowego. Jak zapewnić bezpieczeństwo danych w warunkach eksplozji informatycznej (Beręsewicz i Szymkowiak, 2015; Cellary, 2015, 2016; Ruggles, 2014; Szreder, 2015; Wunsch i Gourbin, 2018; Yu i Chen, 2018). Jak przezwyciężyć problemy wynikające ze wzrostu odmów udziału w badaniach statystycznych, braku zaufania do instytucji administracji państwowej (Chipperfield, Brown, i Bell, 2017). Jaki powinien

być spis ludności, jak go przeprowadzić, by pomimo całej złożoności opisywanej populacji obraz, którego dostarczy, był jak najmniej zniekształcony?

Rozważania zawarte w pracy obejmują przede wszystkim spisy ludności przeprowadzone w okresie dwóch ostatnich dekad. Nie oznacza to jednak, że dyskutując cele, zadania i jakość spisów, pominięto czasy wcześniejsze, często wręcz odwołując się do historii. Opisując transformację metodologii spisów, przedstawiono przeobrażenia w wybranych państwach świata. Państwa te wybrano w sposób celowy, by jak najlepiej zaprezentować metody przeprowadzania spisów. Przykładowo w odniesieniu do spisu opartego na rejestrach odwołano się do państw o najdłuższych i najkrótszych doświadczeniach, tj. krajów skandynawskich i Austrii. Francja jest jedynym państwem, w którym przeprowadzany jest spis kroczący. Szczegółowa charakterystyka metody spisu mieszanego obejmuje natomiast spis ludności przeprowadzony w Polsce w 2011 r. (NSP 2011). Opisując metody przeprowadzania spisów, wykorzystano w szczególności informacje Dywizji Statystycznej Organizacji Narodów Zjednoczonych (ONZ), Europejskiej Komisji Gospodarczej ONZ i Eurostatu. Zastosowano dodatkowo podejście uwzględniające kwerendę stron internetowych narodowych instytutów statystycznych oraz przeprowadzono studia literaturowe, co umożliwiło prezentację przeobrażeń w wybranych państwach oraz zestawienia syntetyczne.

Dokonując oceny jakości spisu, ograniczono się tylko do Polski. Przeprowadzono badanie kompletności szacunków ludności w 2011 r. oraz dokonano porównania z oceną jakości spisu w 2002 r. (NSP 2002). Dla celów oceny kompletności spisu zastosowano różne metody analizy demograficznej, z wykorzystaniem jednego oraz kilku źródeł informacji. Badaniem kompletności objęto całą populację ze szczególnym uwzględnieniem niemowląt, dzieci oraz osób starszych. Spośród metod demograficznych zastosowano analizę „przetrwania między urodzeniem a spisem”, metodę „postarzania” zgodnie z prawdopodobieństwem przeżycia według odpowiednich dla danego wieku i roku kalendarzowego tablic trwania życia oraz analizę stanu i porządku wymierania stulatków. Oceniając jakość spisu 2011 r., wykorzystano dane poprzednich spisów, w szczególności NSP 2002, ewidencji ruchu naturalnego ludności w zakresie urodzeń i zgonów, bieżących bilansów ludności. W niezbędnym zakresie wzbogacono je danymi rejestrów administracyjnych oraz instytutów statystycznych innych państw, spisu w Anglii i Walii w 2011 r., co pozwoliło wprowadzić elementy tzw. statystyki lustrzanej.

Struktura opracowania została dostosowana do realizacji nakreślonego celu ogólnego oraz celów szczegółowych. Rozważania ujęto w czterech rozdziałach poświęconych odpowiednio (1) spisowi ludności jako podstawowemu badaniu demograficznemu, (2) transformacji metod przeprowadzania spisów ludności, (3) metodom ewaluacji spisów oraz (4) perspektywom spisów ludności.

Tematem pierwszego rozdziału jest spis ludności jako podstawowe badanie demograficzne. Rozdział rozpoczyna rozważania na temat celu przeprowadzania



spisu, zarówno w przeszłości, jak i w obecnych czasach, powszechnego dostępu do informacji. Następnie, z perspektywy definicji spisu, przedstawiono dyskusję podstawowych pojęć i własności. Zasadnicze znaczenie w tym zakresie przypisano definicji kategorii ludności spisowej oraz próbie identyfikacji spisywanej populacji. Omówiono obowiązujące w tym zakresie regulacje międzynarodowe oraz definicje ludności rezydującej, stale i faktycznie zamieszkałej. Określono różnice między poszczególnymi kategoriami ludności oraz konsekwencje wynikające z ich przyjęcia. W dalszej części rozdziału podjęto temat wykorzystania danych spisowych, wskazując na oczekiwania różnych instytucji gospodarczych i społecznych oraz administracji państwowej. Nie pominięto w tej kwestii także głosu środowiska naukowego, w tym demografów i statystyków. Rozważania kończy dyskusja kryteriów jakości stawianych badaniom statystycznym przez ich użytkowników.

Rozdział drugi poświęcono omówieniu transformacji metod przeprowadzania spisów ludności. W pierwszej kolejności przedstawiono klasyfikacje tych metod według rekomendacji ONZ. W tym zakresie zaproponowano nowe ujęcie porządkujące spisy w zależności od wykorzystanych źródeł oraz stosowanych metod gromadzenia danych. Jest to rozszerzenie ujęcia proponowanego przez ONZ poprzez uwzględnienie różnych pozastatystycznych źródeł danych, bez ograniczania się jedynie do rejestrów administracyjnych. Następnie scharakteryzowano poszczególne metody przeprowadzania spisów. W tym celu posłużono się rekomendacjami ONZ, ale zdecydowanie je rozszerzono, wykorzystując materiały źródłowe zamieszczone na stronach internetowych urzędów statystycznych państw stosujących poszczególne metody. Poczynając od tradycyjnego spisu, przedstawiono spis z coroczną aktualizacją danych, spis kroczący oraz oparty na rejestrach z uwzględnieniem spisu „wirtualnego”. Szczególną uwagę poświęcono metodom mieszanym, w tym metodzie przeprowadzenia spisu ludności w Polsce w 2011 r. Rozdział kończy omówienie spisów tradycyjnych i nowej generacji pod względem kryteriów jakości badań statystycznych. Uwzględniono tutaj definicję jakości przyjętą przez Eurostat i stosowaną w GUS, przedstawiając możliwości i ograniczenia poszczególnych metod w zakresie spełnienia międzynarodowych kryteriów.

Specyficzne metody ewaluacji spisów ludności jako szczególnego badania statystycznego i demograficznego są przedmiotem rozważań w rozdziale trzecim. Temat rozpoczynają rozważania dotyczące źródeł błędów w spisach ludności, sposobów ich identyfikacji oraz eliminacji. Wśród metod zalecanych w rekomendacjach międzynarodowych podstawowe znaczenie ma spis kontrolny, dlatego wnikliwie przedstawiono jego uwarunkowania i właściwości. Szczególną uwagę zwrócono na błędy pokrycia, które zilustrowano przede wszystkim przykładami ze spisów przeprowadzonych metodą tradycyjną. Odwołano się w tym zakresie do wyników badań polskich demografów (Jończy, 2010; Kordos, 2007, 2017; Paradysz, 2002a, 2002b, 2010; Śleszyński, 2004a, 2004b, 2005). Następnie podjęto próbę oceny kompletności spisu przeprowadzonego w 2011 r. metodą mieszaną.

Ostatni, czwarty rozdział pracy zawiera rozważania dotyczące przyszłości spisów ludności. Perspektywy tego badania starano się przedstawić w sposób wielowymiarowy, z uwzględnieniem źródeł danych, metod badawczych oraz rekomendacji międzynarodowych i planów Głównego Urzędu Statystycznego. Poczynając od wykorzystania źródeł administracyjnych w obecnie przeprowadzanych spisach, przedstawiono przykłady badań statystycznych wykorzystujących źródła pozastatystyczne, wraz z oceną proponowanych rozwiązań. Następnie omówiono wnioski wynikające z wypracowanych przez organizacje międzynarodowe wytycznych i standardów. Szczególną uwagę poświęcono podjętym przygotowaniom do spisu ludności w 2021 r. w Polsce, w tym wynikiem przeprowadzonych przez GUS konsultacji.

Oddając do druku tę książkę, pragnę wyrazić podziękowanie wszystkim Przyjaciołom za wsparcie, pomoc i mobilizację do podjęcia tej pracy i jej opublikowania. Jednocześnie z całą świadomością chcę podkreślić, że chociaż wiele osób przyczyniło się do powstania i poprawy tej książki, odpowiedzialność za wszelkie błędy i nieścisłości spoczywa na mnie. Nie zmienia to faktu, że jestem bardzo zobowiązana Profesorowi Ireneuszowi Kuropcie, Recenzentowi wydawniczemu, za wnikliwe uwagi, komentarze oraz sugestie zmian, które przyczyniły się do poprawy spójności i jakości opracowania. Szczególne podziękowanie kieruję pod adresem Profesora Jana Paradysza. Dziękuję za wielogodzinne, inspirujące dyskusje, podczas których pozwalałeś mi czerpać z zasobów Twojej wiedzy i wskazywałeś problemy czekające rozwiązania. Jesteś nie tylko ekspertem w zakresie prognoz demograficznych, ale masz też wyjątkowy dar przewidywania istotnych kierunków rozwoju. Pragnę wyrazić wdzięczność Profesor Irenie Elżbiecie Kotowskiej za poświęcony mi czas, za gorące rozmowy w letnie dni i jesienne wieczory, za wszelką inspirację i każde słowo rozsądku. Dziękuję Profesor Grażynie Dehnel i Doktorowi Tomaszowi Klimankowi, moim najbliższymi Współpracownikom, za to, że zawsze mogę na nich polegać. Lista osób, które przyczyniły się do powstania tej książki, czasem w sposób nieświadomy, jest długa. Dlatego dziękuję wszystkim Koleżankom i Kolegom z Katedry Statystyki, z Uniwersytetu Ekonomicznego w Poznaniu, z Urzędu Statystycznego w Poznaniu, z Głównego Urzędu Statystycznego, z Komitetu Nauk Demograficznych PAN oraz z Komitetu Statystyki i Ekonometrii PAN za to, że mogłam z nimi współpracować i dzięki kontaktom z nimi poznawać tajemnice demografii i statystyki. Jestem świadoma mojej ograniczonej wiedzy, jak i tego, że pomimo wszelkich starań w książce tej zaledwie poruszyłam niektóre z fundamentalnych problemów spisów ludności i transformacji, jakiej obecnie podlegają badania demograficzne statystyczne.

---

# 1. SPIS LUDNOŚCI JAKO ŹRÓDŁO DANYCH O LUDNOŚCI

---

## 1.1. Spis ludności – historia, cele i zadania

Słowo *census* wywodzi się od łacińskiego słowa *censere*, co oznacza szacować (Rao, 1994, s. 53). W starożytnym Rzymie *census* oznaczał spis obywateli i w tym znaczeniu słowo to nadal jest powszechnie używane, nie tylko w języku angielskim. W cesarstwie rzymskim spis był przeprowadzany raz na pięć lat, głównie w celach militarnych i podatkowych. Wyniki spisu wykorzystywano do oszacowania majątku ludności i zapisania jej do odpowiednich zgromadzeń. W przeciwieństwie do stosunkowo krótkiej historii statystyki jako dyscypliny naukowej historia spisów ludności liczy ponad 6 tys. lat i sięga starożytnych Chin, Japonii, Grecji, Persji, Egiptu czy cesarstwa rzymskiego. Pierwsze spisy przeprowadzono w Babilonii około 4 tys. lat p.n.e. jako badanie populacji w celu określenia zapotrzebowania na żywność. Z kolei Egipcjanie, około 2,5 tys. lat p.n.e., spisywali ludność w celu ustalenia zasobów siły roboczej, która mogła być wykorzystana do budowy piramid. Spisy ludności należą do największych i najważniejszych badań demograficznych i statystycznych<sup>9</sup>. Naturalne jest więc, że przez tysiąclecia historii zmieniały się metody przeprowadzania spisów oraz cel badania.

Nowożytne spisy mają historię dużo krótszą. Można je śledzić od połowy XVII w., ale jak zauważa Kędzelski (Kędzelski i Paradysz, 2006, s. 28 i nast.), ze względu na słabość administracji publicznej i niski poziom kultury statystycznej ludności, ich wyniki nie były wiarygodne. Jednak rozszerzył się zakres i przeznaczenie gromadzonych danych do „szerzej rozumianych potrzeb państwowych”. Pierwsze nowożytne spisy powszechne zaczęły się pojawiać w Europie na przełomie XVII i XVIII w., między innymi w Szwecji, Francji, Rosji i Prusach. Krywułt-Albańska (2012) podaje przykłady spisów ludności Ameryki na zlecenie Brytyjskiej Izby Handlowej oraz spisu ludności późniejszego Quebecu przeprowadzonego przez Francuzów w latach

---

<sup>9</sup> Związków demografii i statystyki jest bardzo wiele. Zwracają na nie uwagę między innymi Kurkiewicz (2010, s. 16) oraz Paradysz (Kędzelski i Paradysz 2006, s. 5). Wzajemnym relacjom między naukami społecznymi (w szczególności demografią) a teorią prawdopodobieństwa poświęcona jest także najnowsza książka Courgeau (Courgeau, 2012).

1665–66. Podstawowym zadaniem tych spisów było monitorowanie rozwoju kolonii zamorskich. Kukutai (Kukutai i in. 2015) podkreśla, że spis powszechny nigdy nie polegał jedynie na liczeniu populacji. W czasach kolonialnych stanowił mechanizm kontroli i był bardziej szczegółowy niż w metropolii, choć często o ograniczonym zasięgu terytorialnym.

Po wojnie o niepodległość w Stanach Zjednoczonych, na podstawie Konstytucji z 1787 r., spisy są przeprowadzane regularnie co 10 lat (od 1790 r.). Coraz powszechniejsze było organizowanie spisów od początku XIX w., we Francji od 1800 r.<sup>10</sup>, w Anglii od 1801 r. (w obu krajach powtarzane co 10 lat), a w Kanadzie od 1851 r. (z uwagi na potrzebę uzyskania wiedzy o liczbie ludności poszczególnych prowincji kraju, w celu przyznania im odpowiedniej do tej wielkości liczby miejsc w parlamencie). Szczególne znaczenie przypisuje się spisowi ludności przeprowadzonemu w 1846 r. w Belgii. Jest on powszechnie uznawany za pierwszy nowoczesny spis ludności, głównie dzięki wprowadzeniu zasady imiennego i jednodniowego gromadzenia informacji z określeniem momentu krytycznego. Jednocześnie uznano wyłącznie poznawczy charakter spisu, niezależny od jakichkolwiek bezpośrednich i osobistych konsekwencji. Zmiany te były niewątpliwą zasługą Adolfa Queteleta<sup>11</sup> (1796–1874) i stanowiły załączek podstawowych zasad przeprowadzania spisów ludności w latach 1853 i 1872 przez Międzynarodowy Kongres Statystyczny. Rolę pierwszego koordynatora spisów od chwili powstania w 1855 r. pełnił Międzynarodowy Instytut Statystyczny (Kędelski i Paradysz, 2006, s. 30).

Niestety, w tym samym czasie, kiedy zaczęto wdrażać praktyki charakterystyczne dla badań naukowych, spis ludności wpisywał się w dyskryminacyjne praktyki skierowane przeciw mniejszościom etnicznym i rasowym oraz ludności tubylczej (Kukutai i in., 2015; Snipp, 2003). Są to szczególnie ważne kwestie w społeczeństwach zróżnicowanych rasowo i etnicznie, gdzie abolicja nie rozwiązała problemów wolności i dyskryminacji rasowej. Przykładowo w Stanach Zjednoczonych dopiero ruch na rzecz praw obywatelskich wpłynął na poprawę kompletności spisu i zmniejszenie niedoszacowania mniejszości, co skutkowało wprowadzeniem zmian wzmacniających grupy etniczne, wzrostem tożsamości narodowej, dążnością do kultywowania tradycji i zwyczajów. Jak zauważa Snipp (2003), w latach 60. wprowadzenie ustawy o prawach wyborczych w USA<sup>12</sup> doprowadziło do uświadomienia sobie, że spis powszechny może być wykorzystany do uzyskania władzy

<sup>10</sup> W latach 1801–1936 spisy we Francji przeprowadzano co pięć lat, z trzema wyjątkami (1872 zamiast 1871 oraz 1916 i 1941, kiedy z powodu wojen ich nie było). Po II wojnie światowej spisy organizowano co sześć–dziewięć lat, na ogół jako otwarcie kadencji prezydenckiej, trwało to aż do 1999 r. (Biraben, 1963).

<sup>11</sup> Adolf Quetelet (1796–1874), belgijski astronom, matematyk, statystyk, ma ogromne zasługi dla zgromadzenia statystyków podczas Wystawy Światowej w 1853 r. w Londynie i powstania w 1885 r. Międzynarodowego Instytutu Statystycznego.

<sup>12</sup> Ustawy o prawie do głosowania *Voting Rights Act* oraz o prawach obywatelskich *Civil Rights Act* zostały w Stanach Zjednoczonych ustanowione w 1964 r.

politycznej. Znamiennym, choć wydającym się dalekim od kwestii dyskryminacji rasowej przykładem, może być kampania spisowa przeprowadzona w 2001 r. przez Polonię amerykańską pod hasłem: „Dajmy się policzyć, bo nie będą się z nami liczyć”<sup>13</sup>. Problem przyjęcia właściwej metodologii spisu dla rzetelnego i obiektywnego szacunku mniejszości narodowych jest nadal aktualny także w Polsce, choć dotyczy nieporównywalnej skali zjawiska (Gołata, 2015; Łodziński, 2012; Łodziński, Warmińska i Gudaszewski, 2015; Paradysz, 2010).

Pierwszy polski spis ludności odbył się w 1789 r. z inicjatywy Fryderyka Józefa Moszyńskiego na podstawie uchwalonej przez Sejm Czteroletni Konstytucji z 22 czerwca 1789 r. „Lustracya dymów i podanie ludności” uzasadniającej jego przeprowadzenie<sup>14</sup>. Celem spisu było określenie wielkości podatków na rzecz utrzymania stałej armii, co w konsekwencji oznaczało, że nie dotyczył on stanów uprzywilejowanych (nieopodatkowanych), takich jak duchowieństwo czy szlachta. Spis ten miał charakter sumaryczny, trwał dziewięć miesięcy i nie objął blisko połowy populacji (Berger, 2002; Cieślak, 1992, s. 77 i nast.; Kędelski i Paradysz, 2006, s. 29). Na ziemiach polskich spisy były przeprowadzane łącznie ze spisami w państwach zaborczych. Spisy przeprowadzane pod zaborem pruskim ujmowały szeroki zakres cech społeczno-demograficznych. Przeprowadzano tam również spisy zawodowe ludności, gromadząc informacje o zawodzie, wieku i stanie cywilnym. W Galicji od 1869 r. przeprowadzano regularne spisy imienne. W latach 1808, 1810, 1811 przeprowadzono spisy na terenie Księstwa Warszawskiego. Spis z 1810 r. był z założenia pierwszym spisem imiennym i powszechnym, ale choć trwał pięć miesięcy, nie objął 30% mieszkańców (Kędelski i Paradysz, 2006, s. 29). Do opracowania wyników tego spisu w 1810 r. powołano Biuro Statystyczne, pierwszą instytucję statystyczną na ziemiach polskich. W okresie międzywojennym przeprowadzono w Polsce dwa spisy ludności: 30.09.1921 i 9.12.1931. Bezpośrednio po zakończeniu II wojny światowej został przeprowadzony spis sumaryczny 14.02.1946, który dostarczył informacji o stanie i strukturze ludności według płci i wieku (kilka wybranych grup). Natomiast pełne, imienne spisy ludności przeprowadzono siedem razy: 3.12.1950, 6.12.1960, 8.12.1970, 7.12.1978, 7.12.1988, 20.05.2002 i 1.04.2011. Warto w tym miejscu podkreślić szczególną publikację opisującą spisy ludności Rzeczypospolitej Polskiej, przygotowaną przez Polskie Towarzystwo Demograficzne i Główny Urząd Statystyczny (Strzelecki i Toczyński, 2002).

<sup>13</sup> Przed spisem w 2001 r. Polonia przeprowadziła kampanię w celu uświadomienia ważności uczestniczenia w spisie, rozsyłając informacje następującej treści: „Zdobędziemy uznanie i siłę polityczną wówczas, gdy rząd federalny, stanowy i miejski przekona się o naszej liczebności. Wypełnij ankietę, podaj, że jesteś Polakiem(a) i wyślij ją”.

<sup>14</sup> Sejm Czteroletni uchwalił również obowiązek stałej rejestracji ruchu naturalnego ludności począwszy od 1 stycznia 1790 r. Zobowiązano władze kościelne do przesyłania do komisji skarbowych danych dotyczących rejestracji metryk ślubów, chrztów i pogrzebów oraz rejestracji wszystkich mieszkańców (por. <http://stat.gov.pl/porta1/edukacyjny/historia-statystyki/historia-spisow-powszechnych/pierwsze-spisy-ludnosci/#!/strona-4>).

Przez lata opracowano wiele metod przeprowadzania spisów, w różnych krajach mają one odmienną tradycję. Jednym z podstawowych celów spisu jest dostarczenie informacji o stanie i strukturze ludności według płci i wieku oraz jej rozmieszczenia terytorialnego na badanym obszarze (Kędelski i Paradysz, 2006, s. 28 i nast.). Określone przez spis stan i struktura ludności stają się nową podstawą<sup>15</sup> corocznych szacunków liczby ludności w kolejnych latach.

Roland Pressat (1966) w swoim klasycznym podręczniku ujmuje cele i zadania demografii, wskazując na pomiar stanu i struktury ludności oraz pomiar i analizę zdarzeń demograficznych. Jednak bez względu na ujęcie demografii jako statystyki stanu czy ruchu ludności, Pressat podkreśla konieczność badania owych zdarzeń „w związku ze strukturą ludności, w łonie której te wydarzenia nastąpiły” (Pressat, 1966, s. 17). Odnosząc powyższe zadania stawiane przed demografią do statystyki i spisu ludności jako narzędzi ich realizacji, warto podkreślić ich aktualność i potrzebę. Są one szczególnie ważne wobec istotnych zmian w reprodukcji ludności będących konsekwencją przejścia demograficznego<sup>16</sup> czy kolejnej fali gwałtownych przeobrażeń społecznych określanych mianem drugiego przejścia<sup>17</sup>. Zmiany te zostały zapoczątkowane w latach 60. ubiegłego wieku w krajach skandynawskich, następnie objęły Europę Zachodnią i Południową, a w latach 90. także Polskę i inne państwa Europy Środkowej i Wschodniej. Jesteśmy świadkami procesu opóźniania tworzenia rodzin i ich destabilizacji, wzrostu częstości rozwodów i separacji, upowszechniania związków nieformalnych, a w konsekwencji opóźnienia średniego wieku macierzyństwa i długotrwałego spadku płodności poniżej prostej zastępowalności pokoleń (Esping-Andersen i Billari, 2015; Kotowska, 2009b). W państwach, w których te zmiany rozpoczęły się później, ich przebieg był gwałtowniejszy (Kotowska, 2009a). Przemianom małżeńskości i płodności towarzyszy systematyczna i zauważalna poprawa oczekiwanej długości trwania życia przede wszystkim niemowląt, ale także osób dorosłych, głównie mężczyzn (Kuroпка, 2011; Wróblewska, 2006). Jednocześnie obserwowane starzenie od dołu i od góry piramidy powoduje znaczne przyspieszenie procesu i sprawia, że populacja Polski oraz innych państw Europy Środkowej dołącza do grupy starych demograficznie (Kurkiewicz, 2003, 2007, 2012; Pocięcha, 2003a, 2003b, 2007). Nie bez znaczenia jest fakt, że Polska jest krajem odpływu migracyjnego. Od początku okresu transformacji obserwujemy masową emigrację zarobkową, która po akcesji do Unii Europejskiej przyjęła

<sup>15</sup> Podstawa ta ulega zmianie po opublikowaniu wyników kolejnego spisu ludności.

<sup>16</sup> Przejściem demograficznym określa się proces zmian reprodukcji ludności na tle przeobrażeń społeczno-ekonomicznych opisany przez teorię przejścia demograficznego, za której twórców uznaje się Thompsona oraz Notesteina (Kędelski i Paradysz, 2006; Okólski, 1990; Okólski i Fihel, 2012).

<sup>17</sup> Intensywne zmiany w sferze małżeńskości i rozrodczości, prowadzące do niemal jednoczesnego w wielu krajach spadku płodności poniżej poziomu prostej zastępowalności pokoleń, są określane mianem koncepcji drugiego przejścia demograficznego (Lesthaeghe i van de Kaa, 1986; Okólski i Fihel, 2012; van de Kaa, 1987).

jeszcze znaczniejsze rozmiary (Kupiszewska i Kupiszewski, 2012; Kupiszewski, 2004, 2005; Kupiszewski i Kupiszewska, 2014). Prognozy ludności wskazują, że w połowie XXI w. populacje państw Europy Środkowej i Wschodniej będą w czołówce najstarszych, o najwyższym poziomie zaawansowania procesu starzenia (Bussolo, Koettl i Sinnott, 2015; EU, 2015; UN, 2015b; World Economic Forum, 2012). Proces starzenia się ludności Polski, ale także innych państw tego regionu, np. Ukrainy i Bułgarii, jest dodatkowo istotnie wzmocniony przez intensywną emigrację obejmującą najbardziej mobilne roczniki. Dynamika i złożoność ruchów migracyjnych sprawiają, że wiele zdarzeń związanych z migracją wymyka się ewidencji, która pomimo doskonalenia rejestrów, nie pokrywa wzmożonej mobilności ludności (Baffour i in., 2013; Bell i in., 2015; Kotowska, 2009a; Wiśniowski i in., 2013; Witkowski, 2002b). Sytuacja ta jest zatem ważnym argumentem za koniecznością korzystania z informacji spisowych<sup>18</sup>.

Doniosłość badań dotyczących ludności, szczególnie w dobie znaczących przemian procesów demograficznych, jest zrozumiała i powszechnie akceptowana. Podstawowym warunkiem rzetelnych analiz jest właściwe określenie stanu i struktury ludności. Dotyczy to wszelkich badań i analiz w skali makro i mikro. Jak podkreślają Kupiszewski i Bijak<sup>19</sup>, nieprawidłowe określenie stanu i struktury populacji wyjściowej należy do najistotniejszych przyczyn błędów i rozbieżności wszelkich szacunków (Kupiszewski i Bijak, 2012). Warto podkreślić, że konsekwencje tych rozbieżności dotyczą wszystkich dziedzin życia, w tym gospodarczej i społecznej, oraz wszystkich analiz, w których podstawą odniesienia jest ludność. Na rozbieżności takie wskazywali między innymi (Bijak i in., 2007; Jończy, 2010; Kędelski, 1990; Paradysz, 2002a, 2002b; Sakson, 2002; Śleszyński, 2004a, 2004b, 2005). Podobne wnioski formułują Keilman i Kučera (1991) czy Alho i Spencer (2005), którzy na podstawie przeprowadzonych badań stwierdzają, że niepoprawne szacunki liczby i struktury populacji są zasadniczym powodem błędów prognoz i analiz demograficznych. Należy podkreślić, że spisy ludności są podstawą szacunku wszystkich wskaźników makroekonomicznych, podstawą do tworzenia polityki społecznej, polityki mieszkaniowej, wyznaczania obwodów wyborczych, tworzenia operatów losowania do badań statystycznych, alokacji środków państwowych na systemy edukacji i opieki zdrowotnej itp.

---

<sup>18</sup> Opinię taką przedstawia m.in. Kotowska (2009a), pomimo podkreślenia zmiany perspektywy badań demograficznych: od struktur do procesów, w których wyjaśnieniu główny nacisk kładzie się na przyczyny zachowań indywidualnych.

<sup>19</sup> W opracowaniu dotyczącym szacowania „rzeczywistej” ludności Warszawy zespół kierowany przez Kupiszewskiego zwraca uwagę na znaczenie poprawności struktur zarówno dla analiz w ujęciu terytorialnym, jak i „problemowym” (np. w związku z systemem emerytalnym, rozwojem edukacji itp.). Autorzy opracowania od lat podkreślali konieczność uwzględniania kategorii ludności rezydującej, a za niedopuszczalne uważali konstruowanie prognoz demograficznych na podstawie ludności zameldowanej na pobyt stały (Bijak i in., 2007).

## 1.2. Kontrowersje wobec podstawowych pojęć i własności

### 1.2.1. Definicja i podstawowe własności spisu

Spis ludności jest najstarszym badaniem statystycznym o powszechnym zasięgu, przeprowadzanym w zasadzie we wszystkich państwach świata. Stąd jest zrozumiałe, że badanie to bywa różnie definiowane i w różny sposób przeprowadzane. Zarówno metody przeprowadzania spisów, jak i związane z nimi tradycje i zwyczaje były kształtowane przez wieki i ulegały wielu zmianom. Każdy podręcznik demografii podaje też nieco odmienną definicję badania, jakim jest spis ludności. Niemniej jednak autorzy wszystkich podręczników zwracają uwagę na wyczerpujący charakter spisu, co oznacza pełny charakter badania, tzn. że „obejmuje wszystkie osoby zamieszkałe na terytorium objętym spisem” (Kurkiewicz, 2010, s. 27). Celem spisu jest „ustalenie liczby, struktury i rozmieszczenia przestrzennego ludności w określonej jednostce terytorialnej i ściśle określonym momencie” (Kędełski i Paradysz, 2006).

Warto w tym miejscu podkreślić, że narodowe spisy powszechne są badaniami przeprowadzanymi w Polsce na podstawie ustaw sejmowych. Spis ludności stanowi podstawowy element każdego krajowego systemu statystycznego, a także systemu międzynarodowego. Zgodnie z zaleceniami ONZ to rządy poszczególnych państw ponoszą odpowiedzialność finansową i organizacyjną za dostarczenie oficjalnych danych przez organa statystyki publicznej (UN, 1958b). Tę odpowiedzialność rządu Kędełski określa mianem centralizacji (Kędełski i Paradysz, 2006). Rządy wyznaczają też odpowiednie organy odpowiedzialne za przeprowadzenie spisu. W przypadku ostatniego spisu w Polsce ustawa z 4 marca 2010 r. wskazała Prezesa Głównego Urzędu Statystycznego jako osobę odpowiedzialną za przeprowadzenie Narodowego spisu powszechnego ludności i mieszkań 2011 r.

Aby oddać właściwy charakter badania, zasadne wydaje się odwołanie do definicji i rekomendacji zawartych w publikacjach Biura Spisów Organizacji Narodów Zjednoczonych (ONZ), bowiem to właśnie ONZ od lat przygotowuje rekomendacje i zalecenia odnośnie do metod przeprowadzania spisów z uwzględnieniem regionalnej specyfiki i odrębności<sup>20</sup>. Według definicji ONZ (UN, 2015a, s. 4) spis ludności to kompleksowy proces planowania, gromadzenia, kompilacji, ewaluacji, upowszechniania oraz analizowania demograficznych, ekonomicznych i społecznych danych z uwzględnieniem najbardziej szczegółowych przekrojów terytorialnych w określonym czasie w odniesieniu do wszystkich osób w państwie. Definicja

<sup>20</sup> Pierwsze zestawienie zasad i rekomendacji dotyczących metodologii przeprowadzania spisów ludności opublikowano w 1958 r. Stanowiło ono odpowiedź na zgłoszoną przez Komisję Statystyczną Narodów Zjednoczonych potrzebę rozwijania międzynarodowych standardów i stanowiło „kamień węgielny” pierwszego Światowego Programu Spisów Powszechnych Ludności i Mieszkań. Wcześniejsze opracowanie (UN, 1949) zestawiało podstawowe pojęcia i miało być pomocne w przeprowadzeniu spisów 1950 r., jednak nie miało charakteru rekomendacji (por. rozdz. 2.1).



spisu wprowadza podstawowe cechy charakterystyczne dla tego badania (*essential features*), które są zawarte w Rekomendacjach (UN, 2015a). Są to mianowicie:

1. Indywidualne spisywanie (*individual enumeration*).
2. Powszechność w ramach wyodrębnionego terytorium (*universality within a defined territory*).
3. Jednoczesność (*simultaneity*).
4. Określona regularność (*defined periodicity*).
5. Możliwość przedstawienia danych dla małych jednostek terytorialnych (*capacity to produce small-area statistics*).

Warto wyjaśnić znaczenie podstawowych charakterystyk spisów, tak jak ich rozumienie wyjaśniono w ostatnich rekomendacjach ONZ. Nie tylko zmieniała się lista tych cech<sup>21</sup>, ale dodatkowo są one różnie interpretowane, szczególnie wobec istotnych zmian metodologicznych. Pierwsze wyraźne przedstawienie definicji spisu oraz sześciu jego zasadniczych cech zawarto w Rekomendacjach ONZ z 1958<sup>22</sup> (UN, 1958b), by w kolejnej edycji wyróżnić już tylko cztery cechy: (i) indywidualne spisywanie, (ii) powszechność, (iii) jednoczesność oraz (iv) określoną regularność. Te cztery cechy weszły do kanonu charakterystyk spisów ludności i w kolejnych dokumentach wyraźnie się do nich odwoływano. W związku ze zmianami metodologii spisów i wykorzystaniem źródeł administracyjnych pojawiło się wiele wątpliwości, czy nowe podejście spełnia wymogi określone przez podstawowe cechy spisów. W odniesieniu do Narodowego Spisu Powszechnego 2011 r. w Polsce (NSP 2011) formułowano zarzuty braku indywidualnego spisywania oraz powszechności spisu<sup>23</sup>. Wykładnia przedstawiona w ostatnich rekomendacjach wyraźnie te wątpliwości

<sup>21</sup> W niektórych opracowaniach można znaleźć też inne opisy, w których łącznie ujęto podstawowe cechy spisów z innymi ważnymi charakterystykami. Przykładowo: Kędelski (Kędelski i Paradysz, 2006), powołując się na wskazówki Biura Statystycznego ONZ, wymienia następujące cechy spisu: centralizacja, powszechność, imiennosc, jednoczesność, regularność i międzynarodowa porównywalność, statystyczne ujęcie wyników i zagwarantowanie tajemnicy statystycznej. Z kolei Holzer łączył zasadę imiennosci z wymogiem bezpośredniego pozyskiwania danych (Holzer, 2003). Natomiast Waszak (Waszak, 2002) jako piątą cechę – warunek spisu umożliwiający wskazanie „prawdziwego stanu ludności” – podaje „fachowość dokonania spisu”.

<sup>22</sup> W pierwszym dokumencie ONZ przedstawiającym zasady przeprowadzania spisów skoncentrowano się przede wszystkim na tematach spisowych i ich definicjach, kładąc nacisk na międzynarodową porównywalność oraz podkreślając potrzebę przeprowadzenia spisu we wszystkich państwach świata w 1950 r. (UN, 1949). Kolejny dokument wymienia już definicje i następujące podstawowe cechy spisów: odpowiedzialność finansową i organizacyjną rządów, precyzyjną definicję terytorium, powszechność, jednoczesność, indywidualne spisywanie, opracowanie i upowszechnianie wyników (UN, 1958b). Ten sam dokument zawierał także wskazówki w kontekście międzynarodowym dotyczące celu spisu, zakresu danych spisowych, regularności spisów, jednoczesnego ich przeprowadzania oraz międzynarodowej porównywalności.

<sup>23</sup> Zarzuty niespełnienia tych podstawowych wymogów przez Narodowy Spis Powszechny 2011 r. w Polsce pojawiły się zarówno w prasie, jak i w dyskusji naukowej, szczególnie w odniesieniu do takich drażliwych zagadnień, jak problemy etniczne i kwestie narodowościowe (Krywułt-Albańska, 2012; Krzyk, 2012; Łodziński, 2012).

wyjaśnia. Zwrócono jednak dodatkowo uwagę na inną, wydawałoby się oczywistą właściwość, jaką jest przedstawienie szacunków dla małych obszarów.

### 1. Indywidualne spisywanie

Określenie „spis powszechny” oznacza, że każda osoba i każda grupa mieszkań są oddzielnie zliczane i spisywane, zatem dla każdej jednostki jest tworzony odrębny rekord zawierający wszystkie jej charakterystyki. Tylko takie ujęcie umożliwia tworzenie zestawień (cross-klasyfikacji) pozwalających na badanie różnych relacji i zależności pomiędzy wyszczególnionymi w spisie charakterystykami. W określeniu „indywidualne spisywanie” zawiera się cecha bezpośredniości. Akcentowano ją w wielu podręcznikach (Holzer, 2003, s. 43)<sup>24</sup>. Pozyskanie danych z rejestru administracyjnego o każdej pojedynczej osobie spełnia wymóg bezpośredniości, gdyż wpis w rejestrze powstaje na podstawie dokumentów przez tę osobę złożonych, poświadczonych jej podpisem. Według Paradysza żaden z polskich spisów nie spełniał tego atrybutu w stopniu większym niż 40% (Paradysz, 2010). Pozyskanie informacji „z drugiej ręki” jest koniecznością w odniesieniu do dzieci i młodzieży poniżej określonego wieku, osób, z którymi nie można nawiązać kontaktu z powodu choroby i w innych przypadkach, kiedy są one dłużej nieobecne, a ich bliscy mogliby udzielić wiarygodnych informacji. Jednak każdorazowe uzyskanie informacji „z drugiej ręki” winno być odnotowane.

W Rekomendacjach (UN, 2015a) zawarto wyraźną informację, że wymóg indywidualnego spisywania można osiągnąć poprzez zbieranie danych w terenie, korzystając z danych zawartych w rejestrach administracyjnych lub kombinacji tych metod. Każda informacja pozyskana z dowolnego źródła (rejestr, badanie reprezentacyjne, samospis internetowy) w NSP 2011 była imienna oraz została połączona tworząc tzw. „złoty rekord”<sup>25</sup>.

### 2. Powszechność w ramach wyodrębnionego terytorium

Spis powszechny powinien obejmować precyzyjnie określone terytorium (na przykład całe państwo lub dobrze wyodrębnioną jego część). Powszechność oznacza spisanie każdej osoby obecnej i (lub) zamieszkującej w wyodrębnionym terytorium (w zależności od wymaganej definicji populacji). Spis ludności powinien także obejmować każdą grupę mieszkań, niezależnie od rodzaju.

W Rekomendacjach (UN, 2015a) podkreślono, że wymóg powszechności nie oznacza wykluczenia stosowania badań reprezentacyjnych w celu uzyskania

<sup>24</sup> Kiedyś mocno podkreślano bezpośredniość uzyskiwania wyników – prawdopodobnie w sensie rozróżniania spisu sumarycznego (np. 1946) od imiennego, ale z czasem zaczęto sobie zdawać sprawę z bezpośredniości w sensie samospisu konkretnej osoby.

<sup>25</sup> Złoty rekord w bazie danych spisowych to rekord łączący zmienne z różnych źródeł. Są to zarówno dane ze spisu pełnego (czyli pozyskiwane z rejestrów i systemów administracyjnych), internetu, jak też z badania reprezentacyjnego. Wszystkie zmienne są poddane szczegółowemu badaniu jakości uwzględniającemu synchronizację, harmonizację, w tym weryfikację pojęć, definicji, kategorii itp.

danych o określonych cechach. Oczywiście pod warunkiem, że badanie umożliwi uogólnienie wyników dla jednostek terytorialnych, dla których są one wymagane z uwzględnieniem odpowiedniego stopnia szczegółowości w tabelach krzyżowych. Ze względu na fakt, że NSP 2011 przeprowadzono w Polsce metodą mieszaną przy wykorzystaniu źródeł administracyjnych i badania reprezentacyjnego, pojawiły się zarzuty, że stracił on swoją podstawową cechę, jaką jest powszechność (Krywulł-Albańska, 2012; Krzyk, 2012; Łodziński, 2012). Nawiązując do etymologii słowa powszechny, „uniwersalny”, „cały”, „zupełny”, można je rozumieć jako objęcie badaniem (spisanie) całej populacji. Czy rzeczywiście spis przeprowadzony metodą mieszaną tego warunku nie spełnia? NSP 2011 był jak najbardziej powszechny, gdyż zgodnie z ustawą objął całą podlegającą badaniu ludność. Informacje dotyczące każdej osoby stale zamieszkałej (bez względu na fakt, czy w czasie spisu była obecna w kraju, czy nie), zostały pobrane z rejestrów administracyjnych.

### 3. Jednoczesność

Każda osoba i każda grupa mieszkań powinny być spisane w tym samym, ściśle określonym momencie, a zgromadzone dane powinny się odnosić do precyzyjnie zdefiniowanego okresu referencyjnego (tzw. moment krytyczny). Okres referencyjny nie musi być jednakowy dla wszystkich badanych cech. Dla większości danych będzie to moment spisu ludności, ale w niektórych przypadkach może to być okres poprzedzający spis powszechny.

Zasada jednoczesności wiąże się z tzw. momentem krytycznym spisu (stan i struktura ludności są ustalane na ściśle określony moment), który w NSP 2011 określono jako północ z 31 marca na 1 kwietnia 2011 r. Warto zwrócić uwagę, że nie tylko bezpośrednio gromadzone informacje dotyczyły tego momentu, ale wykonano także olbrzymią pracę nad synchronizacją stanu rejestrów administracyjnych z momentem krytycznym NSP 2011 (Bijak, 2008a, 2008b).

### 4. Określona regularność

Spisy powszechne powinny być realizowane regularnie, tak by mogły dostarczać porównywalnych danych w określonej sekwencji. Dane kolejnych spisów winny umożliwiać ocenę zmian demograficznych w przeszłości, dokładny opis aktualnej sytuacji oraz prognozę na przyszłość. ONZ zaleca przeprowadzenie spisu narodowego co najmniej raz na 10 lat. Niektóre państwa mogą uznać za konieczne przeprowadzanie spisów powszechnych częściej z powodu tempa poważnych zmian demograficznych i (lub) warunków mieszkaniowych.

Wyniki spisu ludności państwa mają większą wartość na poziomie krajowym, regionalnym i międzynarodowym, jeśli można je porównać z wynikami spisów powszechnych innych państw, które zostały przeprowadzone w przybliżeniu tym samym czasie. Dlatego też państwa powinny podjąć wszelkie starania, aby spis powszechny przeprowadzić w latach kończących się na „0” lub zbliżonych. Jest oczywiste, że względy prawne, administracyjne, finansowe i inne mogą uniemożliwić państwu dostosowanie się do standardowego wzorca międzyro-

dowego w zakresie czasie przeprowadzenia spisu. Ustalając datę spisu ludności, należy przyznać takim czynnikom krajowym większe znaczenie niż pożądanej równoczesności międzynarodowej.

NSP 2011 w Polsce został przeprowadzony zgodnie z Rekomendacjami (UNECE, 2006b) oraz w terminie wskazanym w rozporządzeniu Parlamentu Europejskiego i Rady nr 763/2008 dla krajów członkowskich Unii Europejskiej (Rozporządzenie, 2008).

### 5. Zdolność przedstawienia danych dla małych jednostek terytorialnych

Spis powszechny powinien dostarczać danych o liczbie i charakterystyce ludności i mieszkań dla najniższego odpowiedniego poziomu podziału geograficznego, zgodnego z warunkami danego państwa oraz dla małych grup ludności, przy jednoczesnym zachowaniu poufności informacji indywidualnych na temat każdej osoby. Dane NSP 2011 zostały oszacowane i opublikowane w przekrojach terytorialnych określonych w rekomendacjach międzynarodowych. Jednak nie oznacza to, że zostały spełnione wszystkie oczekiwania w zakresie danych spisowych, np. dotyczące niepełnosprawnych czy mniejszości etnicznych i narodowych. Procedury NSP 2011 ściśle przestrzegały zapewnienia poufności danych przy korzystaniu z rejestrów administracyjnych, w badaniu internetowym jak również podczas wywiadów bezpośrednich. Na podstawie operacyjnej bazy mikrodanych (OBM) w wyniku anonimizacji utworzona została analityczna baza mikrodanych (ABM) stanowiąca podstawę szacunków. GUS gwarantuje wykorzystanie informacji jedynie dla celów statystycznych.

Omawiając zalecenia i rekomendacje międzynarodowe dotyczące spisów ludności warto zaznaczyć, że mają one odmienny charakter dla różnych regionów świata w zależności od stopnia ich rozwoju. Nadrzędnym celem zaleceń formułowanych przez ONZ jest przygotowanie wskazówek metodologicznych, określenie zakresu tematycznego, zwiększenie porównywalności międzynarodowej. W każdym jednak przypadku podkreśla się decydującą rolę interesu krajowego. Poszczególne państwa decydują też o sposobie zbierania danych oraz zastosowanej metodologii. Jednak obok wyżej omówionych zaleceń wśród standardów międzynarodowych wymienia się takie kwestie, jak dbałość o potrzeby użytkowników, jakość danych, kompletność spisu, ochronę i zabezpieczenie danych, porównywalność międzynarodową, obciążenie respondentów, terminowe opracowanie wyników oraz implikacje finansowe i polityczne (Józwiak, 2002, s. 675).

#### 1.2.2. Kategorie ludności spisowej

Zgodnie z definicją, **spis ludności to okresowe badanie mające na celu uzyskanie pełnej informacji o stanie i strukturze ludności według podstawowych cech demograficznych, społecznych i ekonomicznych z uwzględnieniem najbardziej szczegółowych przekrojów terytorialnych** (por. rozdział 1.2.1). W tym miejscu warto raz jeszcze podkreślić, że przez pełną informację należy rozumieć dążność do

uwzględnienia wszystkich jednostek badanej populacji<sup>26</sup>. W ten sposób rozumiana i interpretowana jest także druga z podstawowych własności spisu określana jako „powszechność w ramach wyodrębnionego terytorium”. Zatem rozpoczynając dyskusję dotyczącą jakości szacunku liczby ludności, należy przede wszystkim zdefiniować przedmiot badania oraz rozpoznać dostępność i kompletność informacji statystycznej.

W Rekomendacjach ONZ (UN, 2015a; UNECE, 2015) znajdujemy zalecenie, że najbardziej pożądane są informacje o liczbie i rozkładzie ludności rezydującej (*usually resident population*), gdyż miejsce zamieszkania jest bardzo ważnym wskazaniem miejsca, w którym ludność oczekuje świadczenia (dostępności) różnego rodzaju usług. A zatem jest to jedna z najważniejszych informacji spisowych, niezbędna w planowaniu i polityce społecznej i gospodarczej. W Rekomendacjach zamieszczono definicję ludności rezydującej oraz tzw. ludności faktycznej – obecnej (*present population*) łącznie z wieloma wyjaśnieniami zasad postępowania w sytuacjach szczególnych<sup>27</sup>. Obok definicji ludności w Rekomendacjach wskazano także na konieczność zapewnienia zgodności z międzynarodowymi standardami.

Kwestie definicji ludności na potrzeby spisów rundy 2010 w państwach Unii Europejskiej zostały określone rozporządzeniem Parlamentu Europejskiego i Rady Nr 763/2008 z dnia 9 lipca 2008 r. (Rozporządzenie, 2008). Z kolei podstawę prawną przeprowadzenia spisu w Polsce stanowiła Ustawa z dnia 4 marca 2010 r. o narodowym spisie powszechnym ludności i mieszkań w 2011 r. (Ustawa, 2010) wraz z aktami wykonawczymi. Dyskutując regulacje w zakresie statystyki ludności dla potrzeb statystyki krajów Unii Europejskiej w rozporządzeniu Parlamentu Europejskiego i Rady (UE) nr 1260/2013 ustanowiono wspólne ramy prawne opracowywania, tworzenia i rozpowszechniania statystyk europejskich dotyczących ludności i zdarzeń ruchu naturalnego (UE, 2013). Wprowadzona w artykule 2 tego rozporządzenia definicja ludności jest następująca (UE, 2013) (Dziennik Urzędowy Unii Europejskiej z 10.12.2013: L 330/41):

Art. 2 pkt c) „ludność zamieszkująca” oznacza wszystkie osoby mające miejsce zamieszkania w państwie członkowskim w czasie odniesienia;

Art. 2 pkt d) „miejsce zamieszkania” oznacza miejsce, w którym osoba zazwyczaj spędza czas przeznaczony na odpoczynek, niezależnie od czasowych nieobecności związanych z wypoczynkiem, urlopem, odwiedzinami u przyjaciół i krewnych, interesami, leczeniem medycznym lub pielgrzymkami religijnymi. Jedynie następujące osoby uważa się za osoby mające miejsce zamieszkania w danym obszarze geograficznym:

<sup>26</sup> Pojęcie badania pełnego jest zazwyczaj stosowane dla określenia kompletności populacji statystycznej, ale może być także stosowane w celu podkreślenia kompleksowego ujęcia zakresu merytorycznego.

<sup>27</sup> Również w przypadku definicji ludności rezydującej może się pojawić wiele sytuacji niejednoznacznych. Z uwagi na znaczenie, jakie w tym opracowaniu przypisuje się precyzyjnemu definiowaniu poszczególnych kategorii ludności, w aneksie I.1, I.2 i I.3 zamieszczono oryginalne wyjaśnienia sytuacji dyskusyjnych zawarte w Rekomendacjach (UN, 2015a).

- (i) osoby, które mieszkały w swoim miejscu zamieszkania nieprzerwanie przez okres co najmniej 12 miesięcy przed czasem odniesienia; lub
- (ii) osoby, które przybyły do swojego miejsca zamieszkania w ciągu 12 miesięcy przed czasem odniesienia z zamiarem przebywania tam przez co najmniej rok<sup>28</sup>.

Szczegółowe, jednolite warunki wykonania rozporządzenia 1260/2013 ustanowione zostały Rozporządzeniem Wykonawczym Komisji (UE) NR 205/2014 z dnia 4 marca 2014 r. Polska statystyka publiczna dostosowała się do powyższych regulacji unijnych i w wyjaśnieniu opublikowanym 28 października 2014 r.<sup>29</sup> GUS poinformował, że wypełniając zobowiązania wynikające z art. 4 rozporządzenia 1260/2013 przekazał Radzie UE informację o liczbie ludności uwzględniającą kryterium 12 miesięcy nieobecności/przebywania. Szacowana liczba ludności Polski według stanu na dzień 1.01.2014 r. wyniosła 38 018 tys. i została przytoczona przez Eurostat pod adresem: [http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo\\_urespop&lang=en](http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_urespop&lang=en).

Definiując populację dla potrzeb NSP 2011, określono, że będzie ona obejmować osoby stale zamieszkałe na obszarze Polski, bez względu na fakt, czy przebywały w kraju w czasie spisu, czy też były za granicą, oraz osoby przebywające czasowo (GUS, 2011b, 2012a, 2012b, 2013b)<sup>30</sup>. W publikacjach spisowych GUS przedstawia przede wszystkim informacje dotyczące tradycyjnie stosowanej w polskich spisach kategorii ludności faktycznej, wprowadzając jedynie ogólne dane o ludności rezydującej (GUS, 2012b)<sup>31</sup>.

Kategoria ludności rezydującej została w polskiej statystyce publicznej uwzględniona już podczas spisu ludności w 2002 r. i obejmuje:

- stałych mieszkańców, z wyjątkiem osób przebywających poza miejscem zamieszkania przez okres co najmniej 12 miesięcy bez względu na ich miejsce przebywania w kraju czy za granicą;
- osoby przebywające czasowo przez okres co najmniej 12 miesięcy, przybyłe z innego miejsca w kraju lub z zagranicy (GUS, 2012a, s. 32).

<sup>28</sup> „Jeżeli ustalenie okoliczności opisanych w pkt (i) lub (ii) nie jest możliwe, wówczas za „miejsce zamieszkania” może być uznane miejsce formalnego zamieszkania lub miejsce zameldowania, z wyjątkiem zastosowania do celów art. 4.”(UE, 2013)

<sup>29</sup> Tekst oświadczenia był dostępny na stronie GUS w dniu 6 listopada 2014, <http://stat.gov.pl/index.php>. Z uwagi na różne aktualizacje strony w chwili obecnej autorka nie potrafi podać linku do oświadczenia, ale jego treść została zamieszczona w aneksie I.4.

<sup>30</sup> Dokładne definicje ludności zgodnie z publikacjami GUS przedstawiono w aneksie I.5.

<sup>31</sup> Należy wyjaśnić, że od 2014 r. GUS regularnie wywiązuje się z Rozporządzenia 1260/2013 (UE), publikując corocznie informację o ludności rezydującej. Dodatkowo zostały uzupełnione szacunki od 2009 r., które są dostępne na stronie: <http://stat.gov.pl/obszary-tematyczne/ludnosc/ludnosc/rezydenci-ludnosc-rezydujaca,19,1.html>.

Natomiast kategoria ludności faktycznie zamieszkałej uwzględnia:

- stałych mieszkańców (osoby zameldowane na pobyt stały w gminie lub deklarujące w niej stały pobyt), z wyjątkiem osób przebywających poza miejscem zamieszkania przez okres powyżej 3 miesięcy w kraju oraz wszystkie osoby przebywające za granicą (bez względu na okres ich nieobecności);
- osoby przebywające czasowo przez okres powyżej 3 miesięcy, przybyłe z innego miejsca w kraju (nie są uwzględniani przebywający czasowo w Polsce cudzoziemcy) (GUS, 2012a, s. 32).

Analizując definicje ludności przyjęte w NSP 2011, wyraźnie można zauważyć różnicę między ludnością rezydującą i faktyczną, polegającą na ujęciu osób przebywających za granicą w kategorii ludności stałej. W przypadku ludności faktycznie zamieszkałej długość czasu pobytu za granicą nie ma znaczenia. Natomiast w ujęciu ludności rezydującej pojawia się kryterium co najmniej 12 miesięcy<sup>32</sup>.

Oceniając szacunki ludności w spisach powszechnych Paradyś (2007a) wyróżnia dodatkowo kategorię ludności zameldowanej na pobyt stały i ludności faktycznie zamieszkałej według deklaracji złożonych przez respondentów rachmistrzowi spisowemu. Według Paradyśa (2007a) można wyróżnić cztery rodzaje ludności: (i) zameldowaną na pobyt stały, (ii) faktycznie zamieszkałą<sup>33</sup>, (iii) „deklaratywną”<sup>34</sup> ludność stałą” oraz (iv) „deklaratywną ludność faktycznie zamieszkałą”. Definicja ludności faktycznej na potrzeby NSP 2011 odpowiada więc czwartej kategorii wyodrębnionej przez Paradyśa. Zauważa on, że pomimo zmian systemowych określanych mianem transformacji konsekwencje podziału ludności na zameldowaną na stałe oraz na pobyt czasowy powyżej dwóch miesięcy<sup>35</sup> nadal mają swój wymiar statystyczny i nie straciły wiele na znaczeniu<sup>36</sup>.

Informacjom uzyskiwanym za pośrednictwem rachmistrza spisowego w tradycyjnej metodzie przeprowadzania spisów można przypisać szereg właściwości, których konsekwencje dla określenia stanu ludności i jej struktury są niedoceniane i trudne do oszacowania. Przede wszystkim większość uzyskanych informacji to deklaracje osób spisywanych, które nie podlegają żadnej weryfikacji. W większości przypadków informacje dotyczące wszystkich mieszkańców są pozyskiwane od jednej bądź kilku osób obecnych w mieszkaniu, a nie bezpośrednio od każdego

---

<sup>32</sup> Problem ujęcia tej kwestii będzie jeszcze przedmiotem rozważań w dalszej części niniejszego opracowania.

<sup>33</sup> Jest to suma zameldowanych na pobyt stały oraz salda zameldowań i wymeldowań na pobyt czasowy ponad trzy miesiące.

<sup>34</sup> Paradyś wprowadza pojęcie deklaratywnej ludności faktycznej bądź stałej dla podkreślenia, że szacunek tej wielkości jest oparty na deklaracji spisywanych osób bez weryfikacji stanu prawnego.

<sup>35</sup> Definicja ludności zameldowanej na pobyt czasowy została zmieniona. Do 2005 r. włącznie uwzględniano pobyt czasowy trwający ponad dwa miesiące, obecnie są to trzy miesiące i dłużej.

<sup>36</sup> Zdaniem Paradyśa (Paradyś, 2007a), różnice między liczbą zameldowanych na pobyt stały i czasowy zmniejszyły się znacznie, od kiedy zaczęły znikać hotele robotnicze, ale im niższy jest poziom terytorialnego podziału kraju, tym ich znaczenie jest większe.

z respondentów. W efekcie pojawia się tzw. obciążenie „drugiej ręki”, którego konsekwencje mogą być bardzo znaczne<sup>37</sup>, ale niestety są ignorowane<sup>38</sup>. Informacje są uzyskiwane nie tylko z „drugiej ręki”, ale mają charakter deklaracyjny. Zarówno o osobach obecnych, jak i nieobecnych pozyskujemy informacje według deklaracji osoby udzielającej odpowiedzi. Nie jest znany ani prawny, ani faktyczny stan cywilny badanych osób, gdyż nikt nie sprawdza aktu małżeństwa/rozvodu. Podobnie informacje o poziomie wykształcenia czy statusie na rynku pracy pozyskane przez rachmistrza spisowego mają wyłącznie deklaracyjny charakter.

Z kolei informacje dotyczące kwestii drażliwych mogą być gromadzone jedynie za zgodą zainteresowanych i według ich deklaracji, przy czym odpowiedź nie jest obowiązkowa<sup>39</sup>. Pozyskiwanie takich danych za pośrednictwem osób trzecich narusza prawo wolności sumienia i wyznania oraz nieujawniania informacji. Przykładowo, informacje zebrane o osobach niepełnosprawnych mogą być uzyskane wyłącznie bezpośrednio od zainteresowanych i bez weryfikacji orzeczeń o niepełnosprawności. Podobnie pytania dotyczące kwestii etnicznych, wyznania, narodowości czy języka ojczystego są sprawą osobistą każdego respondenta i mogą być pozyskane jedynie w sposób bezpośredni za zgodą i według indywidualnych deklaracji.

Problem informacji z tzw. „drugiej ręki” jest jednym z zasadniczych, które wymagają należytej uwagi i właściwych rozstrzygnięć na etapie planowania spisu. Gromadzenie danych z „drugiej ręki” jest powszechnie praktykowane i akceptowane. Kuriozalny wydaje się fakt, że nawet w środowisku naukowym nie budzi zastrzeżeń fakt zbierania danych na tematy drażliwe od osób trzecich<sup>40</sup>, co świadczy o niezrozumieniu istoty badanych zagadnień. Na marginesie powyższych rozważań warto

<sup>37</sup> Szerzej na ten temat w rozdz. 3.3.

<sup>38</sup> Autorce nie jest znana żadna próba szacunku tych błędów, jeśli nie liczyć ustalania różnic między źródłami (np. porównanie liczby niemowląt z liczbą dzieci urodzonych w ostatnim roku). Brakuje jakiegokolwiek rozeznania wielkości odsetka prawidłowo raz spisanych osób, szacunku wielokrotnie spisanych oraz liczby osób (rodzin bądź gospodarstw domowych) w ogóle niespisanych. Fakt niezrozumienia znaczenia problemu oraz jego ignorowania, także w środowisku naukowym, został dobitnie ujawniony podczas dyskusji na konferencji naukowej: „Depopulacja a nowy paradygmat rozwoju miast i regionów”, zorganizowanej przez Komitety naukowe i problemowe Polskiej Akademii Nauk: Badań nad Migracjami, Nauk Demograficznych, Nauk Geograficznych, Nauk o Pracy i Polityce Społecznej, Prognoz Polska 2000+, Przestrzennego Zagospodarowania Kraju, Wydziały: Ekonomiczno-Socjologiczny, Nauk Geograficznych i Zarządzania Uniwersytetu Łódzkiego i Komisję Geografii Osadnictwa i Ludności Polskiego Towarzystwa Geograficznego, Łódź-Spała, 20–21–22 kwietnia 2017 r.

<sup>39</sup> Zgodnie z zapisami Ustawy z dnia 4 marca 2010 r. o narodowym spisie powszechnym ludności i mieszkań w 2011 r. udzielanie odpowiedzi na pytania dotyczące niepełnosprawności odbywało się na zasadzie dobrowolności. Decyzję taką podjęto ze względu na specyfikę i wrażliwość tematu. Wzbudziła ona wiele kontrowersji w środowisku naukowym z powodu konsekwencji odmowy udziału w badaniu (Slany, 2014). Na pytania odpowiadały osoby dorosłe, które wyraziły na to zgodę, natomiast na temat dzieci informacji udzielali rodzice lub opiekunowie.

<sup>40</sup> W dyskusji podczas konferencji „Depopulacja a nowy paradygmat rozwoju miast i regionów” Łódź-Spała 20–22 kwietnia 2017 r. padło m.in. wymowne stwierdzenie: „przecież ojciec wie, jakiego wyznania, czy jakiej narodowości jest córka”.



podkreślić, że znaczenie zasady bezpośredniego pozyskiwania danych było podejmowane w polskiej literaturze. Problem konsekwencji informacji z „drugiej ręki” jest od wielu lat przedmiotem badań Paradysza, który wielokrotnie opisywał wynikające z tego powodu błędy spisów (Paradysz, 1992, 2002a, 2002b, 2010). Sytuację, w której jedna osoba w gospodarstwie domowym udziela informacji w imieniu wszystkich pozostałych, określał on efektem „respondenta grupowego”. Holzer łączył zasadę imienności spisów (indywidualnego spisywania) z zasadą bezpośredniego pozyskiwania danych (Holzer, 2003). Podkreślał on znaczenie maksymalnego zapewnienia możliwości zebrania danych bezpośrednio od osób podlegających badaniu. Holzer dopuszczał jednak w wyjątkowych sytuacjach możliwość zebrania danych od najbliższego otoczenia osób nieobecnych w momencie spisywania. Przedstawione powyżej rozważania sugerują, że w rzeczywistości szacunki ludności przedstawione jako wyniki spisów mogą znacznie odbiegać od założonych, zgodnie z przyjętymi kryteriami, wielkości ludności faktycznej i rezydującej.

### 1.2.3. Ludność rezydująca a faktyczna w NSP 2002 i NSP 2011

Pomijając trudne do oceny konsekwencje niezachowania zasady bezpośredniości odpowiedzi, poniżej zostanie przedstawiona relacja między szacunkami ludności rezydującej oraz faktycznej na podstawie danych dwóch ostatnich spisów ludności w Polsce (NSP 2002 i 2011). Jednak ze względu na istotne konsekwencje zastosowanej metody szacunku w dalszych rozważaniach będą stosowane liczne odwołania<sup>41</sup>. Przede wszystkim warto podkreślić, że w publikacjach NSP 2002 została zamieszczona jedna tablica zawierająca informacje nt. szacunków ludności rezydującej (GUS, 2003, tab. 45). Natomiast w publikacjach NSP 2011 dane dotyczące ludności rezydującej opublikowano w wynikach wstępnych (GUS, 2012b)<sup>42</sup>. Dalsze porównania są możliwe dzięki wypełnieniu przez GUS zobowiązania wynikającego z art. 4 Rozporządzenia 1260/2013<sup>43</sup>. Celem tych porównań jest unaocznienie rozbieżności między definicją stosowaną w dotychczasowych spisach oraz zalecaną w rekomendacjach międzynarodowych. Ludność rezydująca, czyli zazwyczaj zamieszkała przez okres co najmniej 12 miesięcy<sup>44</sup> (*usually resident*), jest powszechnie uznawana za bardziej wiarygodną, dającą szacunek możliwie najbliższy rzeczywistości.

<sup>41</sup> Relacja między ludnością rezydującą i faktyczną była przedmiotem dyskusji we wcześniejszych opracowaniach autorki (Gołata, 2012, 2013a, 2016a, 2016b).

<sup>42</sup> O takiej kwalifikacji świadczą informacje zamieszczone na stronie internetowej Urzędu (<http://stat.gov.pl/spisy-powszechne/nsp-2011/nsp-2011-wyniki-wstepne/>).

<sup>43</sup> Odpowiednie dane są dostępne na stronach internetowych Eurostatu: <https://ec.europa.eu/CensusHub2/query.do?step=selectHyperCube&qhc=false> oraz GUSu: <http://stat.gov.pl/obszary-tematyczne/ludnosc/ludnosc/rezydenci-ludnosc-rezydujaca,19,1.html>.

<sup>44</sup> Osobę zmieniającą zwyczajowe miejsca zamieszkania na okres co najmniej roku statystyki ONZ ujmują jako migranta (EU, 1998). Definicja ta może być stosowana w analizie migracji zewnętrznych i wewnętrznych.

Tabela 1.1. Ludność faktyczna i rezydująca w Narodowych Spisach Powszechnych 2002 i 2011

Kategoria ludności	Ogółem	Mężczyźni	Kobiety
Ludność faktyczna w tys. (20.5.2002)	38 230	18 516	19 714
Ludność faktyczna w tys. (31.3.2011)	38 501	18 432	20 067
Ludność faktyczna, różnica w latach (2011–2002)	271	–84	353
Ludność faktyczna, dynamika zmian (2002 = 100)	100,7	99,5	101,8
Ludność rezydująca w tys. (20.5.2002)	37 620,1	18 237,5	19 382,6
Ludność rezydująca w tys. (31.3.2011)	37 244	17 830	19 413
Ludność rezydująca, różnica w latach (2011–2002)	–376,1	–407,5	30,4
Ludność rezydująca, dynamika zmian (2002 = 100)	99,0	97,8	100,2
NSP 2002, różnica absolutna: ludność rezydująca a faktyczna	–609,9	–278,5	–331,4
NSP 2002, ludność rezydująca jako procent ludności faktycznej	98,4	98,5	98,3
NSP 2011, różnica absolutna: ludność rezydująca a faktyczna	–1257	–602	–654
NSP 2011, ludność rezydująca jako procent ludności faktycznej	96,7	96,7	96,7

Uwaga: Dla około 1,6 tys. osób nie ustalono informacji o płci, stąd rozbieżności między danymi ogółem i według płci.

Źródło: (Gołata, 2012) na podstawie: (GUS, 2003, tab. 45; GUS, 2012b, s. 7, 38).

Zastosowanie definicji ludności rezydującej w czasie spisu powszechnego w 2002 r. umożliwiło oszacowanie liczby ludności Polski jako 37 620 085 osób, tj. o 610 tysięcy niższej (1,6%) od populacji faktycznie zamieszkałej (tab. 1.1). Wstępne wyniki spisu ludności z 2011 r. wskazują, że liczba rezydentów to 37 244 tys. osób. Oznacza to, że różnica między ludnością faktyczną a rezydującą uległa podwojeniu, wynosząc w 2011 r. 1 mln 257 tys. osób<sup>45</sup>, co stanowi 3,3% ludności faktycznej. W przeciwieństwie do opartej na ludności faktycznej informacji o wzroście liczby ludności Polski w latach 2002–2011 o 271 tys., podanej w publikacji (GUS, 2012a, 2012b), należy zauważyć, że liczba ludności rezydującej zmniejszyła się o ponad 376 tys. Inaczej rozbieżności te kształtują się według płci, gdzie w okresie między spisami zanotowano zmniejszenie liczby mężczyzn według kryterium faktycznego o 84 tys., ale według kryterium ludności rezydującej spadek ten wyniósł ponad 407 tys. Natomiast wzrost liczby kobiet wyniósł odpowiednio: 353 tys. bądź „zaledwie” 30 tys.

Przedstawiając relację między ludnością rezydującą a faktyczną, należy unikać formułowania wniosków na podstawie wstępnych wyników spisów 2002 r. i 2011 r. Stąd jeszcze krótkie odwołanie do informacji o liczbie rezydentów opublikowanej w połowie 2014 r. na stronie Eurostatu. Porównanie to jest tym bardziej istotne, że długo wyczekiwane szacunki ludności rezydującej okazały się zupełnie odmienne od oczekiwanych (Kupiszewski i Bijak, 2012). Zdziwienie wzbudził przede wszystkim o ponad 63% niższy ostateczny szacunek ludności rezydującej w porównaniu

<sup>45</sup> Pierwsze publikacje wyników spisu z 2011 r. w grudniu 2011 r. podawały różnicę między ludnością rezydującą a faktyczną jako 1 mln 140 tys., natomiast publikacja z 23 marca 2012 r. podaje liczbę 1 mln 257 tys. Jest to prawdopodobnie związane z różnicą między szacunkami ludności faktycznej: 38 325 tys. w grudniu 2011 r. oraz 38 501 tys. w marcu 2012 r.

z szacunkiem wstępnym (tab. 1.2). Oczywiście rozbieżności między szacunkami wstępnymi a ostatecznymi są zrozumiałe, ale tak znaczne wymagają przynajmniej próby wyjaśnienia. Nie może bowiem pozostać bez komentarza wynosząca blisko 1,5 mln osób różnica w szacunku ludności w wieku produkcyjnym w porównaniu do niespełna 400 tys. Przede wszystkim rozbieżności te nie pozostają bez konsekwencji dla gospodarki, a w ujęciu regionalnym mają jeszcze większe znaczenie<sup>46</sup> (Gołata, 2012, 2013a, Jończy, 2010, 2017; Józefowski i Rynarzewska-Pietrzak, 2010).

Tabela 1.2. Ludność faktyczna i rezydująca w Narodowym Spisie Powszechnym 2011

Kategoria ludności/ płeć-wiek	Ogółem	Przedprodukcyjny	Produkcyjny			Poprodukcyjny
			razem	mobilny	niemobilny	
Ludność faktyczna NSP 2011						
Polska	38 511 824	7 202 273	24 797 408	15 418 861	9 378 547	6 512 143
Mężczyźni	18 643 870	3 691 430	12 973 926	7 821 177	5 152 749	1 978 514
Kobiety	19 867 954	3 510 843	11 823 482	7 597 684	4 225 798	4 533 629
Rezydenci marzec 2012 w tys.						
Polska	37 244	7 202	23 352	14 308	9 044	6 686
Mężczyźni	17 830	3 691	12 120	7 196	4 924	2 018
Kobiety	19 413	3 511	11 232	7 112	4 120	4 669
Rezydenci lipiec 2014						
Polska	38 044 565	7 143 974	24 401 728	15 102 110	9 299 618	6 498 862
Mężczyźni	18 420 389	3 661 812	12 783 311	7 673 032	5 110 279	1 975 265
Kobiety	19 624 176	3 482 162	11 618 417	7 429 078	4 189 339	4 523 597
Różnica ludność faktyczna – rezydująca (marzec 2012) w tys.						
Polska	1 267 824	273	1 445 408	1 110 861	334 547	-173 857
Mężczyźni	813 870	430	853 926	625 177	228 749	-39 486
Kobiety	454 954	-157	591 482	485 684	105 798	-135 371
Różnica ludność faktyczna – rezydująca (lipiec 2014)						
Polska	467 259	58 299	395 680	316 751	78 929	13 281
Mężczyźni	223 481	29 618	190 615	148 145	42 470	3 249
Kobiety	243 778	28 681	205 065	168 606	36 459	10 032

Źródło: Gołata (2012, 2016a) na podstawie: GUS (2012b, s. 7, 38; 2013b) oraz <https://ec.europa.eu/CensusHub2/query.do?step=selectHyperCube&qhc=false>.

Różnice w stanie ludności faktycznej i rezydującej GUS (2012a) tłumaczy faktem, że wśród Polaków przebywających za granicą są osoby mieszkające poza krajem przez wiele lat, które utraciły już status polskich rezydentów. Podejście uwzględniające w definicji ludności faktycznej osoby przebywające za granicą (bez względu na okres ich nieobecności) oraz nieuwzględniające cudzoziemców powoduje istotne

<sup>46</sup> W *Raporcie* opublikowanym 26 lipca 2012 r. nie zamieszczono żadnych informacji o ludności rezydującej, a zatem nic nie wiadomo o możliwych korektach uwzględniających rezultaty spisu kontrolnego (GUS, 2012a, s. 32).

znieszczenie stanu i struktury ludności, a w konsekwencji współczynników demograficznych oraz wszystkich analiz. Konsekwencje szacunków liczby ludności według kryterium faktycznego, podczas gdy rejestracja statystyczna całego ruchu naturalnego ludności (urodzeń, zgonów, małżeństw, rozwodów i migracji) jest prowadzona według kryterium zameldowania na pobyt stały, są przedmiotem badań Paradysza (2002a, 2002b, 2007a, 2010).

Upowszechnienie szacunków ludności rezydującej jest istotne ze względu na potrzebę uporządkowania sposobu obliczania współczynników demograficznych. Jednym z podstawowych warunków poprawnego szacunku współczynników jest odpowiedniość kategorii licznika i mianownika. Sytuacja ujednoczenia wielkości licznika i mianownika współczynników demograficznych wystąpiła w polskiej praktyce statystycznej w połowie lat 80. (Kędelski i Paradysz, 1984). Polegała ona na odniesieniu liczby urodzeń według stałego miejsca zamieszkania matki do liczby kobiet również według stałego, a nie jak wcześniej faktycznego miejsca zamieszkania. Obecnie współczynniki demograficzne uwzględniają zarówno zdarzenia, jak i stan ludności, której analizowane zdarzenia dotyczą, według stałego miejsca zamieszkania<sup>47</sup>. Oznacza to uwzględnianie w mianowniku również wszystkich tych osób, które od lat przebywają za granicą, ale nie wymeldowały się ze swojego „stałego” miejsca zamieszkania. Jak wynika z danych NSP 2011, osób takich może być nawet ponad 1 milion 250 tys. Na niespójność metodologiczną systemu rejestracji zdarzeń demograficznych według stałego miejsca zamieszkania oraz odnoszenia go do ludności stałej, której korekta jest dokonywana na podstawie ludności faktycznej, zwracali uwagę również Bijak, Kicinger i Kupiszewski (2007, s. 13–14). Rozbieżności szacunku współczynników demograficznych jako konsekwencji ujęcia ludności faktycznej i rezydującej były też przedmiotem opracowania Gołaty (2013a).

Publikując wyniki NSP 2011, GUS poinformował, że zgodnie z rozporządzeniem Parlamentu Europejskiego i Rady Nr 763/2008 z dnia 9 lipca 2008 r. (Rozporządzenie, 2008) to ludność rezydująca jest obowiązującą kategorią w porównaniach międzynarodowych i począwszy od 2011 r. corocznie – w najbliższych kilku latach – będą upowszechniane dwa stany ludności: rezydującej i faktycznej (GUS, 2012b, s. 26). Rzeczywiście, GUS wywiązał się z tego zobowiązania. Jednak dalekie od realizacji są oczekiwania dotyczące uporządkowania sposobu obliczania współczynników demograficznych, zlikwidowania wskazanej wyżej niespójności metodologicznej oraz innych konsekwencji.

#### 1.2.4. Minimum spisowe

W XXI w. zapotrzebowanie na informacje jest większe niż kiedykolwiek w przeszłości. Ale też żyjemy w wieku informacji, może raczej w dobie danych, megadanych czy big data. Stąd wybór tematów spisowych jest zadaniem trudnym i wymaga przemyślanego,

<sup>47</sup> Według definicji podanej dla ludności według kryterium faktycznego – por. rozdział 1.2.3.

racjonalnego podejścia. Ważnym czynnikiem determinującym zakres tematyczny spisu jest porównywalność danych dla różnych państw. Zalecenia ONZ wskazują tematy podstawowe i uzupełniające, oddzielnie dla spisu ludności i dla spisu mieszkań. Ponadto zawierają szereg szczegółowych wyjaśnień dotyczących pojęć i definicji umożliwiających harmonizację i zapewnienie zgodności szacunków i klasyfikacji w ujęciu międzynarodowym oraz sprostania wymogom standaryzacji danych w różnych przekrojach terytorialnych. Pozyskanie danych według uzgodnionych list umożliwia monitorowanie rozwoju społeczno-ekonomicznego populacji poszczególnych państw oraz realizację strategicznych celów organizacji międzynarodowych.

W zaleceniach ONZ jest publikowana lista tematów podstawowych (*core topics*) oraz uzupełniających (*additional topics*). Lista ta jest aktualizowana przed każdą rundą spisów. Obejmuje ona cechy geograficzne, demograficzne, ekonomiczne, dotyczące wykształcenia oraz gospodarstwa domowego i rodziny. Listę tematów określonych dla spisów rundy 2020 przedstawiono w tabeli 1.3 (UN, 2015a).

Tabela 1.3. Tematy podstawowe i dodatkowe według rekomendacji ONZ dla spisów ludności rundy 2020

Lp.	Tematy spisu	Rodzaj cech
	A. Charakterystyka geograficzna oraz migracji wewnętrznych <i>Geographic and internal migration characteristics</i>	
1	Miejsce stałego zamieszkania / <i>Place of usual residence</i> (par. 4.52–4.57)	◆
2	Miejsce przebywania w momencie krytycznym spisu / <i>Place where present at time of census</i> (par. 4.58–4.63)	◆
3	Miejsce urodzenia / <i>Place of birth</i> (par. 4.4–4.71)	◆
4	Czas zamieszkiwania / <i>Duration of residence</i> (par. 4.72–4.74)	◆
5	Poprzednie miejsce zamieszkania / <i>Place of previous residence</i> (par. 4.75–4.76)	◆
6	Miejsce zamieszkania w ściśle określonym momencie w przeszłości / <i>Place of previous residence at a specified date in the past</i> (par. 4.77–4.81)	◆
7	Ludność ogółem / <i>Total population</i> (par. 4.82–4.88)	◉
8	Miejscowość / <i>Locality</i> (par. 4.89–4.91)	◉
9	Miasto i wieś / <i>Urban and rural</i> (par. 4.92–4.100)	◉
	B. Charakterystyka migracji międzynarodowych <i>International migration characteristics</i>	
1	Miejsce urodzenia / <i>Country of birth</i> (par. 4.105–4.109)	◆
2	Obywatelstwo / <i>Country of citizenship</i> (par. 4.110–4.115)	◆
3	Nabyte obywatelstwo / <i>Acquisition of citizenship</i> (par. 4.116)	■
4	Rok lub okres przybycia / <i>Year or period of arrival</i> (par. 4.117–4.120)	◆
	C. Charakterystyka rodziny i gospodarstwa domowego <i>Household and family characteristics</i>	
1	Stosunek do głowy gospodarstwa domowego / <i>Relationship to the reference person of household</i> (par. 4.129–4.139)	◆
2	Skład rodziny I gospodarstwa domowego / <i>Household and family composition</i> (par. 4.140–4.147)	◇
3	Status rodziny I gospodarstwa domowego / <i>Household and family status</i> (par. 4.148)	■

Lp.	Tematy spisu	Rodzaj cech
	D. Charakterystyka demograficzna i społeczna <i>Demographic and social characteristics</i>	
1	Płeć / <i>Sex</i> (par. 4.150)	◆
2	Wiek / <i>Age</i> (par. 4.151–4.162)	◆
3	Stan cywilny / <i>Marital status</i> (par. 4.163–4.171)	◆
4	Charakterystyki etniczno-kulturowe / <i>Ethnocultural characteristics</i> (par. 4.172–4.173)	■
5	Wyznanie / <i>Religion</i> (par. 4.174–4.178)	■
6	Język / <i>Language</i> (par. 4.179–4.182)	■
7	Etniczność / <i>Ethnicity</i> (par. 4.183–4.187)	■
8	Autochtoni / <i>Indigenous peoples</i> (par. 4.188–4.192)	■
9	Niepełnosprawność / <i>Disability status</i> (par. 4.193–4.213)	◆
	E. Charakterystyka płodności i umieralności <i>Fertility and mortality</i>	
1	Dzieci żywo urodzone / <i>Children ever born alive</i> (par. 4.228–4.233)	◆
2	Dzieci żyjące / <i>Children living</i> (par. 4.234–4.236) ◆	◆
3	Data urodzenia ostatniego żywego dziecka / <i>Date of birth of last child born alive</i> (par. 4.237–4.240)	◆
4	Urodzenia w ostatnich 12 miesiącach / <i>Births in the past 12 months</i> (par. 4.241–4.243)	Δ
5	Zgony dzieci urodzonych w ostatnich 12 miesiącach / <i>Deaths among children born in the past 12 months</i> (par. 4.244–4.246)	Δ
6	Wiek, czas trwania pierwszego małżeństwa / <i>Age, date or duration of first marriage</i> (par. 4.247–4.248)	■
7	Wiek matki w momencie urodzenia pierwszego żywego dziecka / <i>Age of mother at birth of (date or time when) first child born alive</i> (par. 4.249)	■
8	Zgony w gospodarstwie domowym w ostatnich 12 miesiącach / <i>Household deaths in the past 12 months</i> (par. 4.250–4.254)	◆
9	Osierocenie przez matkę lub ojca / <i>Maternal or paternal orphanhood</i> (par. 4.255–4.257)	■
	F. Charakterystyki poziomu wykształcenia <i>Educational characteristics</i>	
1	Umiejętność czytania i pisania / <i>Literacy</i> (par. 4.258–4.264)	◆
2	Uczęszczanie do szkoły / <i>School attendance</i> (par. 4.265–4.271)	◆
3	Osiągnięte wykształcenie / <i>Educational attainment</i> (par. 4.272–4.280)	◆
4	Dziedzina wykształcenia, posiadane kwalifikacje / <i>Field of education and training, and educational qualifications</i> (par. 4.281–4.288)	■
	G. Charakterystyki ekonomiczne <i>Economic characteristics</i>	
1	Status na rynku pracy / <i>Labour force status</i> (par. 4.307–4.338)	◆
2	Charakter zatrudnienia / <i>Status in employment</i> (par. 4.339–4.351)	◆
3	Zawód / <i>Occupation</i> (par. 4.352–4.355)	◆
4	Przemysł / <i>Industry</i> (pars. 4.356–4.359)	◆
5	Miejsce pracy / <i>Place of work</i> (par. 4.360–4.365)	■
6	Sektor zatrudnienia / <i>Institutional sector of employment</i> (par. 4.366–4.368)	■
7	Czas pracy / <i>Working time</i> (par. 4.369–4.375)	■
8	Udział w produkcji własnej / <i>Participation in own-use production of goods</i> (par.4.376–4.381)	◆
9	Dochód / <i>Income</i> (paras. 4.382–4.386)	■

cd. tabeli 1.3

Lp.	Tematy spisu	Rodzaj cech
	H. Charakterystyki związane z rolnictwem <i>Agriculture</i>	
1	Działalność rolna na rachunek własny / <i>Own-account agriculture production</i> (par. 4.389–4.392)	■
2	Charakterystyka wszystkich prac rolnych w ostatnim roku / <i>Characteristics of all agricultural jobs during the last year</i> (par. 4.393–4.396)	■

Legenda:

◆ Temat podstawowy – informacje bezpośrednio gromadzone

◉ Temat podstawowy – informacje wynikowe

■ Temat dodatkowy

△ Temat dodatkowy – informacje wynikowe pozyskane z tematu podstawowego

Uwaga: Tłumaczenie własne autorki; w zestawieniu nie uwzględniono tematów określonych dla spisu mieszkań.

Źródło: (UN, 2015a, s. 188–189).

Każda z zalecanych do pozyskania w spisie cech jest szczegółowo omówiona w Rekomendacjach (UN, 2015a), a podane w tabeli numery paragrafów ułatwiają znalezienie odpowiednich wyjaśnień.

Obok spełnienia zobowiązań międzynarodowych w kontekście harmonizacji i porównywalności wybór tematów i konkretnych charakterystyk wymaga wcześniejszego rozpoznania oczekiwań i potrzeb użytkowników. Uwzględniając złożoność i odpowiedzialność za decyzje dotyczące zakresu informacji gromadzonych w trakcie spisów, rekomendacje ONZ są poprzedzone dyskusją czynników, które te decyzje kształtują. Są to: priorytet krajowy, porównywalność międzynarodowa, przydatność, określenie zasobów oraz źródeł alternatywnych. Respektując istotność tych czynników i przywiązując zasadniczą wagę do potrzeb zgłaszanych przez różne gremia, GUS przeprowadził konsultacje społeczne<sup>48</sup>. Ich celem było zgromadzenie opinii na temat potrzeb i oczekiwań względem danych spisu powszechnego 2021 r. Konsultacje zostały przeprowadzone w formie interaktywnej ankiety dostępnej na stronach GUS w dniach 17 listopada 2016 r. – 21 stycznia 2017 r. Ankieta została również bezpośrednio skierowana do kluczowych odbiorców, zwłaszcza administracji publicznej, ośrodków naukowych, instytucji wykorzystujących dane spisowe do realizacji swoich zadań statutowych. Wyniki konsultacji zostały opracowane i opublikowane na stronie internetowej GUS w raporcie ogólnym oraz w 11 raportach tematycznych (GUS, 2017b, 2017c).

Przeprowadzając konsultacje społeczne przed spisem 2021 r. w odniesieniu do zakresu tematycznego, założono, że w ramach ankiety uczestnicy określą swoje potrzeby przez odniesienie się do zakresu tematycznego spisu zaproponowanego przez statystykę publiczną, z uwzględnieniem jednostek podziału terytorialnego kraju

<sup>48</sup> Konsultacje takie w niektórych krajach, jak np. w Kanadzie, Stanach Zjednoczonych, Nowej Zelandii czy Wielkiej Brytanii, są stałym elementem strategii realizacji masowych badań.

oraz podstawowych zestawień danych (GUS, 2017b). Określając zakres tematyczny spisu, GUS odwołał się potrzeb krajowych wynikających z aktów prawnych oraz strategii i programów rządowych, a także potrzeb międzynarodowych wynikających z rozporządzenia WE 763/2008 (UE, 2008) w sprawie spisów i innych aktów wykonawczych określających zakres informacji przekazywanych do Eurostatu. Wyodrębniono w ten sposób 11 następujących obszarów tematycznych (tab. 1.4)

Tabela 1.4. Obszary tematyczne wyróżnione w konsultacjach społecznych przed NSP 2021

Numer	Obszar tematyczny
1	Stan i charakterystyka demograficzna ludności
2	Ludność w obiektach zbiorowego zakwaterowania oraz osoby bezdomne
3	Edukacja
4	Aktywność ekonomiczna osób
5	Dojazdy do pracy
6	Niepełnosprawność
7	Kraj urodzenia i obywatelstwo
8	Migracje wewnętrzne i zagraniczne
9	Narodowość, język i wyznanie
10	Gospodarstwa domowe i rodziny
11	Stan i charakterystyka zasobów mieszkaniowych (mieszkania i budynki)

Źródło: (GUS, 2017b).

Wyodrębnione obszary różnią się od listy tematów w zaleceniach ONZ. Z jednej strony uwzględniają ludność w gospodarstwach zbiorowego zamieszkania i bezdomność, niepełnosprawność czy dojazdy do pracy jako odrębne grupy tematyczne. Z drugiej strony nie uwzględniono np. diety kobiet. Należy dodać, że uwzględniona jako jedenasta grupa w zestawieniu GUS charakterystyka zasobów mieszkaniowych jest opisana w rekomendacjach ONZ w odrębnym rozdziale. Jest zrozumiałe, że proponowany przez GUS zakres tematyczny spisu 2021 r. nie jest identyczny z zawartym w rekomendacjach ONZ. Dla lepszego zobrazowania zakresu informacyjnego poszczególnych grup tematycznych GUS opisał je, podając dodatkowo zestaw kluczowych zmiennych. Respondenci wskazywali, czy tematy uwzględnione w danym obszarze zaspokajają ich potrzeby oraz określali zakres informacji, który powinien znaleźć się w danym obszarze. W ten sposób określone zostały wskaźniki zaspokojenia potrzeb informacyjnych (aneks I.6, por. też tab. 1.5). W większości tematów odsetek respondentów deklarujących spełnienie oczekiwań informacyjnych wynosił od blisko 80% do wyraźnie powyżej 90%. Ponadto zostały zebrane informacje na temat oczekiwań respondentów w zakresie niezaspokojonych potrzeb. Wśród zgłoszonych dodatkowych tematów (uzupełniających zmiennych) są m.in. związki nieformalne i diety w zakresie charakterystyki demograficznej ludności, informacje na temat kształcenia ustawicznego w odniesieniu do



edukacji, aktywność ekonomiczna cudzoziemców, środki lokomocji i utrudnienia związane z infrastrukturą w ramach dojazdów do pracy czy jakość życia osób niepełnosprawnych.

### 1.3. Zapotrzebowanie na dane spisu ludności

Powodów ważności danych spisowych można wskazywać bardzo wiele. Baffour, King i Valente (2013) czy Kotowska (2009a), podkreślając znaczenie spisu oraz jego jakości, wskazują następujące przyczyny: zaufanie do danych spisowych i ich wykorzystanie do planowania gromadzenia danych w przyszłości. Jakość danych spisowych jest postrzegana jako kwestia zasadnicza, gdyż są one traktowane jako referencyjne dla całego systemu statystyki państwowej. Dane spisów są podstawą dla opracowywania prognoz ludności, służą one też do konstrukcji operatów losowania w badaniach reprezentacyjnych. Wyniki spisów są traktowane jako wzorzec – standard odniesienia w badaniach porównawczych oceniających jakość innych źródeł danych (benchmark). Spis jest też ważnym źródłem danych dla małych populacji, nie tylko w znaczeniu przestrzeni, ale także innych domen. Spis ludności stanowi zasadniczą część narodowego i międzynarodowego systemu statystyki oficjalnej, gdyż dostarcza szacunków liczby ludności. Informacje te są wykorzystywane przez różne instytucje i organizacje krajowe i międzynarodowe w rachunkach narodowych, badaniach porównawczych czy w celu konstrukcji prognoz ludności. Próbując ocenić, w jaki sposób tematy spisowe były uwzględniane w dotychczasowych spisach w Polsce, można się posłużyć przedstawionymi w aneksie I.7 oraz I.8 zestawieniami zaczerpniętymi z pracy Krywułt-Albańskiej (2012), w których autorka skorzystała z opracowań Holzera (2003). W tabelach porównano dane gromadzone w polskich spisach ludności od 1921 r. Należy przede wszystkim zauważyć wyraźne i systematyczne poszerzanie zakresu zagadnień uwzględnianych w spisach. Zmiany rzeczywistości gospodarczej oraz sytuacji społecznej powodowały eliminowanie jednych tematów i dodawanie innych.

Zapotrzebowanie na dane spisowe zostało poddane weryfikacji w ramach konsultacji społecznych przed spisem 2021 r. (GUS, 2017b). W przeprowadzonym badaniu GUS otrzymał ponad 5200 ważnych ankiet, przy czym blisko 45% to odpowiedzi z różnych instytucji, a ponad 55% od indywidualnych użytkowników. Wyniki konsultacji wskazują na znaczne wykorzystanie danych spisowych. Potrzeby respondentów względem danych spisowych wynikały głównie z realizowanych przez nich zadań, takich jak kreowanie na szczeblu lokalnym polityki demograficznej i społecznej, tworzenie i ewaluacja dokumentów strategicznych, funkcjonowanie i monitorowanie różnych systemów usług (np. zdrowotnych, komunikacyjnych), planowanie budownictwa mieszkaniowego. Wskazywane przez respondentów cele wykorzystania danych spisowych zostały ujęte w osiem kategorii (tab. 1.5).

Tabela 1.5. Cel wykorzystania danych spisu ludności deklarowany przez organizacje i osoby indywidualne według obszarów tematycznych, konsultacje społeczne przed spisem 2021 r.

Cel wykorzystania danych	Odsetek wskazań celu w obszarze tematycznym										
	Stan i charakterystyka demograficzna ludności	Ludność w obiektach zbiorowego zakwaterowania oraz osoby bezdomne	Edukacja	Aktywność ekonomiczna	Dojazdy do pracy	Niepełnosprawność	Kraj urodzenia i obywatelstwo	Migracje wewnętrzne i zagraniczne	Narodowość, język i wyznanie	Gospodarstwa domowe i rodziny	Stan i charakterystyka zasobów mieszkaniowych
Planowanie, budowa i monitorowanie strategii, opracowywanie polityk	54	58	54	60	34	55	30	50	40	52	18
Podejmowanie decyzji w obszarze działalności jednostki	53	58	46	54	29	55	46	50	42	52	18
Zaspokojenie potrzeb informacyjnych	52	55	50	55	33	47	53	54	58	51	17
Korzystanie z funduszy europejskich	46	39	47	49	22	49	28	31	25	37	12
Spełnienie wymogów wynikających z przepisów prawnych	44	50	43	39	19	48	47	41	38	42	15
Poszerzenie wiedzy i edukacja	41	41	50	55	26	36	38	42	51	39	12
Działalność naukowo-badawcza i rozwojowa	24	26	30	44	18	21	24	27	32	24	7
Inne	2,4		1,9	2,2	1,2		1,9	1,5	1,8	2	0,5

Źródło: (GUS, 2017c).

Częstość wskazań wykorzystania danych według celów była zróżnicowana w poszczególnych obszarach tematycznych (GUS, 2017c). Najczęściej wskazywanymi przez ponad 50% respondentów celami były: planowanie, budowa i monitorowanie strategii, opracowywanie polityki, podejmowanie decyzji w obszarze działalności jednostki i zaspokojenie potrzeb informacyjnych. Szczegółowe opisy potrzeb informacyjnych na określonym poziomie terytorialnym, w tym niższym niż gmina, zostały zaprezentowane w raportach tematycznych. W większości obszarów ponad połowa respondentów deklarowała korzystanie z danych spisu w kilku różnych celach. W raporcie z konsultacji znajdujemy też informację, że zaproponowane w poszczególnych obszarach zmienne w relatywnie wysokim stopniu spełniają oczekiwania respondentów – od 82,1% w obszarze Dojazdy do pracy do 94,1% w obszarze Gospodarstwa domowe i rodziny (GUS, 2017b, s. 16–17) (aneks I.9).

Warto podkreślić, że zainteresowanie respondentów dotyczyło w znacznej mierze dostępności danych na najniższym poziomie podziału administracyjnego kraju, tj. dla gmin. Odpowiedzi takiej udzieliło blisko 55% respondentów we wszystkich obszarach tematycznych (GUS, 2017b, s. 20, 31–32). Wiąże się to z pewnością z licznym udziałem w konsultacjach jednostek instytucjonalnych szczebla gminnego – 2500, powiatowego – 380 i wojewódzkiego – 16. W konsultacjach zaproponowano także możliwość wskazania jeszcze bardziej szczegółowego poziomu dostępności danych. Korzystając z nich, respondenci zgłaszali potrzebę pozyskania danych na niższych poziomach agregacji, np. dla miejscowości, miast, wsi, dzielnic, osiedli, sołectw, rejonów statystycznych czy obwodów spisowych. Nieliczne zgłoszenia dotyczyły potrzeb danych według siatki gridowej, współrzędnych geograficznych, ulic, obrębów geodezyjnych, obwodów szkolnych, a także diecezji i parafii. Sytuacje takie dotyczyły osób prowadzących prace badawczo-naukowe uwzględniające kryteria agregacji przestrzennej.

W konsultacjach respondenci mieli okazję wyrazić opinię dotyczącą możliwości wykorzystania alternatywnych źródeł informacji (GUS, 2017b, s. 20). Wskazywano tutaj przede wszystkim na systemy i rejestry administracyjne, bazy danych pozostające w gestii instytucji oraz organizacji międzynarodowych. Jednak blisko 80% użytkowników stwierdziło, że nie widzi możliwości zastąpienia spisu ludności i mieszkań danymi z innych źródeł. Wyjątkiem są dane z obszaru Stan i charakterystyka demograficzna ludności, dla którego ponad 80% respondentów wskazało taką możliwość.

Wśród respondentów biorących udział w konsultacjach byli także pracownicy nauki, demografowie i statystycy wykorzystujący dane spisu przede wszystkim w badaniach naukowych. W swoich analizach ważność danych spisu ujmują oni także z nieco innej perspektywy (Kotowska, 2009a; Paradysz, 2010). Paradysz (2010) podkreśla potrójne znaczenie spisów ludności: a) jako źródła struktur demograficznych, b) jako źródła informacji nieobjętych systemem rejestracji bieżącej, c) jako źródła informacji w statystyce regionalnej. Paradysz zwraca uwagę, że oprócz struktury ludności według wieku w analizie demograficznej niezbędna jest znajomość struktury ludności według innych niż wiek argumentów czasowych oraz struktur jednostek bardziej złożonych niż osoby. Przykładowo autor wymienia trwanie małżeństwa i odstępny między urodzeniami. Zauważa, że „zmiana jednostki statystycznej z prostej (osoba) na bardziej złożone (rodziny, małżeństwa, związki konsensualne bądź gospodarstwa domowe) bez wątpienia ubogaca analizę rozwoju ludności, ale także powoduje wzrost wymagań odnośnie do struktur demograficznych oraz rodzaju źródeł statystycznych. Istniejące dotychczas rejestry administracyjne jeszcze się w ogóle do tego celu nie nadają. W takim przypadku spis lub mikrospis jest niezastąpiony” (Paradysz, 2010, s. 50). Podobną opinię przedstawia także Kotowska (2009a), wskazując na unikalne cechy informacji pochodzących ze spisów powszechnych, które powodują, że nie można ich zastąpić danymi z innych źródeł. Z zasad organizacji spisu, ich zakresu tematycznego, stosowanych międzynarodowych definicji

i klasyfikacji wynika porównywalność danych spisowych. Autorka zauważa, że niepodważalnym walorem danych spisowych jest także to, że stanowią one unikalne połączenie skali makro (dane dotyczące całej populacji) i mikro (o poszczególnych osobach) oraz umożliwiają jednoczesne i łączne przeprowadzanie spisu ludności i mieszkań, co stwarza możliwość pełnej diagnozy sytuacji mieszkaniowej w Polsce.

W Polsce, tak w organach ustawodawczych, w administracji państwowej, w środowisku naukowym, wśród demografów i statystyków, nie wspominając o przedstawicielach innych dyscyplin naukowych, jak i w społeczeństwie, dane spisu ludności powszechnie przyjmowano i nadal się przyjmuje jako bezdyskusyjne i prawdziwe (Paradysz, 2010). Zapewne właśnie z tego względu spisowy szacunek liczby ludności jest punktem odniesienia w różnych regulacjach ustawowych i rządowych. Jako przykład<sup>49</sup> można tutaj podać związek między wynikami NSP i realizacją ustawowych uprawnień mniejszości narodowych<sup>50</sup>. Otrzymanie statusu mniejszości, jak również posługiwanie się językiem regionalnym, wymaga spełnienia szeregu warunków, ale gwarantuje też szereg praw (Ustawa, 2005, art. 2 ust. 3, art. 19 ust. 2). Obejmują one między innymi swobodne posługiwanie się językiem, jego naukę bądź naukę w języku mniejszości oraz różnorodną działalność mającą na celu ochronę, zachowanie i rozwój tożsamości kulturowej mniejszości, z czym się wiąże określone dotacje z budżetu gminy i państwa. W szczególności uprawnienia te dotyczą gmin, w których mniejszość narodowa czy posługująca się językiem regionalnym stanowi co najmniej 20% mieszkańców, a zgodnie z art. 14 ustawy liczba ta winna być urzędowo ustalona, jako wynik ostatniego spisu powszechnego.

W ten sposób statystyka publiczna została zobowiązana do wskazania gmin posiadających określone uprawnienia. Podejście uzależniające formę realizacji ustawowych uprawnień określonych grup ludności od wyników spisu jednoznacznie wskazuje na znaczenie tego badania. Trzeba jednak zauważyć, że podejście to stawia statystykę publiczną wobec nacisków zainteresowanych grup, wystawiając na próbę obiektywności i sprostania zasadom Europejskiego Kodeksu Praktyk Statystycznych<sup>51</sup> obowiązującego w ramach Europejskiego Systemu Statystycznego<sup>52</sup>.

<sup>49</sup> Przykłady można mnożyć, wskazując m.in. na subwencje oświatowe (por. przypis 5).

<sup>50</sup> Ustawa o mniejszościach narodowych i etnicznych oraz o języku regionalnym z dnia 6 stycznia 2005 r. (Dz. U., Nr 17 poz. 141, 2005).

<sup>51</sup> Europejski Kodeks Praktyk Statystycznych został przyjęty przez Komitet ds. Programu Statystycznego dnia 24 lutego 2005 r. i określa 15 zasad wraz z odpowiednimi wskaźnikami w zakresie niezależności, wiarygodności i odpowiedzialności krajowych i wspólnotowych organów statystycznych. Zasada pierwsza to zasada niezależności zawodowej. Niezależność zawodowa władz statystycznych od innych departamentów lub organów politycznych, regulacyjnych lub administracyjnych, a także od podmiotów sektora prywatnego, zapewnia wiarygodność statystyk europejskich. Dostęp na stronie internetowej w dniu 23 października 2017 r. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52005DC0217:PL:NOT>.

<sup>52</sup> Europejski System Statystyczny jest tworzony przez – Eurostat oraz krajowe urzędy statystyczne i inne organy odpowiedzialne za opracowywanie, tworzenie i rozpowszechnianie statystyk europejskich.

Oceniając znaczenie spisu ludności w kontekście międzynarodowych badań statystycznych, podobnie jak przy omawianiu własności badania, słuszne się wydaje odwołanie do ustaleń międzynarodowych (UN, 2015a). Po pierwsze, uwagę zwraca kolejność, w jakiej autorzy publikacji przedstawili rekomendacje i zalecenia dotyczące zakresu merytorycznego oraz metodologii spisu. Otóż bezpośrednio po uwagach wprowadzających, jeszcze przed definicją i właściwościami badania, zamieszczono rozdział omawiający znaczenie spisu – *Essential roles of the census*. Rozwiązanie to jest jak najbardziej słuszne, gdyż dopiero dobrze uzasadnione badanie może stanowić podstawę wydatkowania środków publicznych. Warto w tym miejscu zauważyć, że urząd statystyczny Stanów Zjednoczonych, będący główną agencją statystyczną w USA, nosi nazwę Biura Spisów (The United States Census Bureau). Fakt ten podkreśla nadrzędną rolę spisu ludności wśród badań statystycznych<sup>53</sup>.

Spisowy szacunek liczby ludności winien więc być rozpatrywany wielowymiarowo, nie tylko w kategoriach badań naukowych w zakresie statystyki, analizy demograficznej i społecznej, ale również jako problem administracyjny, samorządowy, instytucjonalny, finansowy. Reasumując rozważania dotyczące znaczenia spisów ludności i zasadności ich przeprowadzania, można wskazać na istotną rolę spisu, która wynika z następujących powodów (Kotowska, 2009a; Paradysz, 2010; UN, 2015a):

**1. Powszechny spis ludności i mieszkań jest relatywnie dobrze udokumentowanym źródłem informacji, których wiarygodność warunkuje podejmowanie racjonalnych decyzji będących podstawą efektywnego zarządzania gospodarką i społeczeństwem.**

Zadaniem spisu ludności i mieszkań jest gromadzenie, przetwarzanie i rozpowszechnianie dokładnych, aktualnych i szczegółowych danych statystycznych o populacji osób, rodzin i gospodarstw domowych, jej strukturze według cech demograficznych i rozkładzie przestrzennym. Pod koniec XIX w. Międzynarodowy Kongres Statystyczny zalecił okresowe przeprowadzanie spisów we wszystkich państwach na całym świecie. Od 1958 r. ONZ opracowuje zasady i rekomendacje oraz wprowadza ogólnoswiatowe programy dotyczących spisów powszechnych ludności i mieszkań.

**2. Spis ludności i mieszkań jest badaniem organizowanym według zasad będących wynikiem międzynarodowych uzgodnień, badaniem o ustalonym zakresie tematycznym, uzgodnionych definicjach i klasyfikacjach, co stanowi o porównywalności jego wyników.** Dodatkowo powtarzalność spisów w regularnych odstępach sprawia, że dane spisów od lat stanowią unikatową bazę informacji o rozwoju populacji prawie całego świata<sup>54</sup>.

<sup>53</sup> Biuro Spisów USA wchodzi w skład Ministerstwa Handlu (US Department of Commerce), a jego dyrektor jest mianowany przez prezydenta USA. Od 30 czerwca 2017 r. funkcję tę sprawuje dr Ron S. Jarmin.

<sup>54</sup> Warto dodać, że informacje takie są gromadzone w ramach projektu IPUMS-International w Minnesota Population Center. IMUS dostarcza danych spisowych z całego świata zintegrowanych

3. **Spis ludności i mieszkań jest podstawowym źródłem danych o stanie i strukturze ludności w ujęciu przestrzennym według podstawowych cech demograficznych i społecznych.**
4. **Spis ludności i mieszkań jest podstawowym źródłem danych o stanie i strukturze rodzin, gospodarstw domowych i gospodarstw zbiorowych w ujęciu przestrzennym.**
5. **Spis ludności i mieszkań ma zasadnicze znaczenie w systemie administracji publicznej.**

Wyniki spisu powszechnego są wykorzystywane jako referencyjne w celu zapewnienia sprawiedliwej dystrybucji i przydzielania funduszy rządowych regionom i innym jednostkom administracyjnym w zakresie edukacji, ochrony zdrowia. Dane spisu służą jako podstawa wyznaczenia okręgów wyborczych. Spisowy szacunek liczby ludności jest podstawą odniesienia w różnych regulacjach ustawowych, w szczególności dotyczących subwencji państwowych w zakresie polityki społecznej przejawiającej się w różnych formach działalności administracji samorządowej (np. subwencja oświatowa czy uprawnienia mniejszości narodowych). Ustanowienie publicznego konsensusu w sprawie priorytetów polityki społecznej państwa byłoby prawie niemożliwe do osiągnięcia, gdyby odpowiednie regulacje nie zostały oparte na liczbie ludności określonej przez spisy powszechne. Również szerokie grono innych użytkowników, w tym sektora przedsiębiorstw, środowisk akademickich, społeczeństwa obywatelskiego i osób fizycznych, korzysta z wyników spisu powszechnego.

6. **Spis powszechny odgrywa zasadniczą rolę w odniesieniu do wielu analiz demograficznych, społecznych i gospodarczych (benchmarking).**
7. **Spis ludności i mieszkań jest jednym z podstawowych elementów krajowego systemu statystycznego, m.in. :**

- **standardem odniesienia w planowaniu i organizacji innych badań,**
- **podstawą organizacji badań reprezentacyjnych,**
- **podstawą konstrukcji operatu losowania,**
- **źródłem danych o zmiennych pomocniczych w statystyce małych obszarów.**

Statystyki spisowe są wykorzystywane w konstrukcji operatu losowania dla badań reprezentacyjnych. Bez operatów losowania i charakterystyk populacji pochodzących z spisu ludności krajowe systemy statystyczne napotykały trudności z zapewnieniem rzetelnych oficjalnych statystyk dla rządów i ogółu społeczeństwa. Obecnie krajowe systemy statystyczne niemal wszystkich państw są

---

w czasie i przestrzeni, które umożliwiają prowadzenie badań porównawczych populacji osób, rodzin i społeczności. Dane i usługi są dostępne bezpłatnie za pośrednictwem strony internetowej <https://www.ipums.org/index.shtml>. Podobną rolę pełni Baza Danych Eurostatu ze spisu 2011 r.: CensusHub2, której informacje są dostępne na stronie internetowej <https://ec.europa.eu/CensusHub2/query.do?step=selectHyperCube&qhc=false>. Baza obejmuje dane ze spisów przeprowadzonych przez kraje Unii Europejskiej i EFTA według ujednoliconej metodologii. Są w niej zawarte porównywalne informacje dotyczące ludności, warunków życia i mieszkań.

oparte na badaniach reprezentacyjnych, niezależnie od pojawienia się nowych, pozastatystycznych źródeł danych, takich jak big data.

**8. Spis ludności i mieszkań jest źródłem danych demograficznych i społecznych nieobjętych systemem rejestracji bieżącej.**

Sprawozdawczość statystyczna w zakresie ewidencji ruchu naturalnego i wędrownego ludności dostarcza szeregu informacji o zaistnieniu zdarzeń, takich jak urodzenia, zawarcie związku małżeńskiego, zgony czy rozwody. Jednak ewidencja nie rejestruje takich zdarzeń, jak powstanie związku nieformalnego, dojazd do pracy czy niepełnosprawność biologiczna. Informacji takich nie zawierają także rejestry administracyjne. Ze względu na charakter badania ewidencja zdarzeń nie dostarcza także informacji o strukturach, np. rodzin, gospodarstw domowych, osób bezdomnych, grup narodowościowych i etnicznych. W takich sytuacjach spis jest jedynym źródłem danych.

**9. Spis powszechny jest zasadniczym źródłem danych w statystyce regionalnej.**

Podstawową cechą spisu powszechnego jest generowanie wiarygodnych, nieobciążonych błędem losowania statystyk dla małych obszarów i mało licznych populacji/subpopulacji. Użyteczność statystyk dotyczących małych obszarów jest z natury rzeczy oczywista, wynika także z przydatności w szacunkach statystyk dotyczących dowolnie zdefiniowanej jednostki przestrzennej. Na przykład w planowaniu lokalizacji szkoły konieczne jest posiadanie danych dotyczących rozkładu dzieci w wieku szkolnym według rejonów szkolnych, które niekoniecznie muszą odpowiadać jednostkom administracyjnym. Podobnie dane spisowe dla małych obszarów mogą być wykorzystane w aproksymacji statystyk dla regionów naturalnych/geograficznych (na przykład zlewisk wodnych lub stref wegetacyjnych). Ponieważ dane spisu ludności mogą być zestawione w tabelach dla dowolnej jednostki geograficznej, możliwe jest podanie wymaganych statystyk w niezwykle elastyczny sposób. Ta wszechstronna cecha spisu ludności jest również nieoceniona dla sektora prywatnego, np. w planowaniu biznesowym i analizach rynkowych.

**10. Spis ludności stanowi podstawę konstrukcji prognoz demograficznych.**

Projekcje populacji są jednym z najważniejszych wyników analitycznych opartych na danych spisu ludności. Prognozy przedstawiające szacunek liczby ludności w przyszłości mają kluczowe znaczenie dla wszystkich segmentów sektora publicznego i prywatnego.

## 1.4. Kryteria oceny jakości badań statystycznych

Spełnienie diskutowanych w poprzednim podrozdziale oczekiwań użytkowników danych spisu ludności, w tym organów administracji rządowej i samorządowej, instytucji społecznych, ośrodków naukowych i indywidualnych odbiorców, wymaga,

by dane te były rzetelne i spełniały określone przez użytkowników kryteria jakości. Badanie jakości należy do trudniejszych zagadnień. Jakość w sensie uniwersalnym jest czasami przyjmowana jako pojęcie pierwotne, niedefiniowalne (Borys, 2012). Borys podkreśla znaczenie intuicyjnych, a także filozoficzno-logicznych podstaw odróżnienia tego, co ilościowe i jakościowe w kategorii jakości. Zwraca uwagę, że definicje jakości można przydzielić do jednej z dwóch interpretacji: niewartościującej (opisowej) oraz wartościującej (preferencyjnej). Interpretacje te wskazują dwa nurty badań jakościowych o odmiennych celach: rozróżniania jakości i oceny stanów. Pierwszy dotyczy jakości w sensie uniwersalnym, która jest wszędzie – świat to jedna wielka różnorodność. Drugi nurt uniwersalnej wartościującej definicji jakości Borys (2012) przedstawia, przytaczając klasyczne określenie Platona, że jakość to stopień osiągniętej przez obiekt doskonałości. Świadomość ograniczeń wynikających z różnego pojmowania jakości, określenia jej definicji, kryteriów oceny, interdyscyplinarności kwalitologii i wreszcie sama ewaluacja nakazują ostrożność wnioskowania i powściągliwość ocen.

W dyskusji dotyczącej jakości spisów ludności Baffour, King i Valente (2013) w pierwszej kolejności zwracają uwagę na ich ważność, znaczenie i rolę, jaką dane spisu odgrywają w funkcjonowaniu państwa w sensie organizacji wewnętrznej i w kontekście międzynarodowym. Autorzy cytowanego opracowania podkreślają, że nadrzędnym celem spisu jest pozyskiwanie i dostarczanie danych, a co za tym idzie, generowanie informacji wykraczających poza to, co jest już dostępne. Dlatego ocena jakości spisu przekłada się na ocenę jakości wytwarzanych informacji. Przyjmując, że jakość spisu wynika z potencjału informacyjnego dostarczanych danych, wskazują oni na znaczenie danych dla użytkowników oraz to, jaką nowe dane mają użyteczność ponad to, co oferują istniejące źródła (Baffour i in., 2013, s. 409). Również Kenett i Shmueli (2016) dyskutują koncepcję jakości informacji (*Information Quality, InfoQ*), którą definiują jako potencjał danych do osiągnięcia konkretnego celu naukowego bądź praktycznego poprzez zastosowanie empirycznej analizy danych.

Podobną opinię prezentuje Lynn (2001, 2003), który główny nacisk kładzie na narodową użyteczność i międzynarodową porównywalność. Lynn podkreśla trudność określenia standardów jakości w badaniach międzynarodowych. Zauważa, że uwarunkowania określone w skali kraju, i w tym sensie nie wpływające na jakość badania, w odniesieniu do badań międzynarodowych mają duże znaczenie. Wśród podstawowych czynników o takim znaczeniu wskazuje na regulacje prawne oraz wynikające z nich ograniczenia dotyczące badań statystycznych, uwarunkowania historyczne i tradycję w zakresie organizacji badań, dostępność (w tym aktualność i kompletność) operatów losowania, różnice kulturowe, wynikające z przyjętych norm społecznych, zróżnicowania pod względem struktury narodowej i etnicznej, języków mniejszości. Lynn dyskutuje pięć różnych strategii osiągania jakości



w kontekście międzynarodowym, przede wszystkim w odniesieniu do badań reprezentacyjnych<sup>55</sup>. Przedstawione kwestie zachowują aktualność także w odniesieniu do spisów ludności zarówno ze względu na transformację metodologii spisów, jak i potrzebę porównań międzynarodowych.

Przyjmując logikę wspinania się od danych do informacji, a następnie do wiedzy, nad którą jest mądrość (Ostasiewicz, 2012), oraz uznając, że dane same w sobie nie mają jakości, lecz jedynie potencjał, który ujawnia się w związku z ich użyciem w konkretnym celu (Baffour i in., 2013; Lynn, 2003), można zauważyć potrzebę wielowymiarowego podejścia do oceny jakości. Brackstone (1999, 2003) podkreśla, że jakość nie jest łatwo definiowaną koncepcją i stała się terminem nadużywanym w ostatnich latach. Tym bardziej dla urzędu statystycznego kwestia zaufania do publikowanych przez tę instytucję informacji jest kwestią przetrwania. Jeśli informacje publikowane przez instytucje statystyki publicznej stają się podejrzane, podważana jest wiarygodność urzędu i jego reputacja jako niezależnego, obiektywnego źródła wiarygodnych informacji. Dlatego dbałość o jakość powinna być główną troską krajowych urzędów statystycznych. Brackstone (1999, 2003) definiuje więc jakość w taki sposób, by uwzględnić te wszystkie aspekty wyników badań statystycznych, które odzwierciedlają ich przydatność dla klientów. Określił on sześć następujących wymiarów jakości: (i) trafność, (ii) dokładność, (iii) terminowość, (iv) dostępność, (v) interpretowalność i (vi) spójność.

Pojęcie jakości w sensie użyteczności rozwinęło się statystyce oficjalnej i organizacjach międzynarodowych, takich jak ONZ, Międzynarodowy Fundusz Walutowy czy Organizacja Współpracy Gospodarczej i Rozwoju (OECD). Choć wyróżniane komponenty jakości i ich klasyfikacje różnią się<sup>56</sup>, jakość jest definiowana bardzo podobnie w statystyce publicznej w wielu państwach, instytucjach i urzędach międzynarodowych. W Dywizji Statystycznej ONZ podejmowano też próbę harmonizacji w tym zakresie (Vries, 2002).

Eurostat przyjął definicję jakości zgodną z normą ISO 8402-1986 jako ogół cech i właściwości produktu lub usługi, który decyduje o zdolności zaspokojenia potrzeb zadeklarowanych lub domyślnych. W definicji wyróżniono sześć następujących komponentów (Eurostat, 2000, 2003):

- przydatność – stopień spełnienia wymagań i potrzeb użytkowników (zakres i klasyfikacje);
- dokładność – bliskość do wartości rzeczywistych;

<sup>55</sup> W tym miejscu pominięto uwagi dotyczące organizacji badań reprezentacyjnych, wpływu schematu losowania próby (*deff*), techniki gromadzenia danych, odmów, sposobu rekrutacji ankierów i wiele innych kwestii.

<sup>56</sup> Statistics Canada jako jeden z pierwszych urzędów wprowadził system zapewniania jakości w 1997 r. (Statistics Canada, 2017), Statistics Netherlands wyróżnia pięć komponentów (Statistics Netherlands, 2014), a Statistics Sweden – cztery (Biemer, Trewin, Bergdahl i Japac, 2014).

- terminowość i punktualność – czas między wydarzeniem a publikacją informacji; rozbieżność między planowaną a rzeczywistą datą publikacji;
- dostępność i przejrzystość – kanały dystrybucji, metadane, formy publikacji;
- porównywalność – czasowa, geograficzna i porównywalność dziedzin;
- spójność – zgodność informacji z różnych źródeł w różnych klasyfikacjach i ujęciach.

Także GUS prowadzi prace nad podniesieniem jakości statystyki dla lepszego zaspokojenia potrzeb użytkowników, zmniejszenia obciążeń respondentów oraz obniżenia kosztów tworzenia statystyk. Definicja jakości przyjęta w polskiej statystyce publicznej jest oparta na definicji jakości Europejskiego Systemu Statystycznego i określona na podstawie sześciu komponentów wskazanych w definicji Eurostatu (aneks I.10).

W dokumentach Eurostatu znajdujemy też oczywiste stwierdzenie, że istnieje wyraźny związek między jakością statystyki a dostępnymi zasobami, wraz z wyraźną sugestią, by podczas oceny jakości uwzględniać także efektywność kosztów (Eurostat, 2003). Zatem można postawić postulat wypracowania kompromisu między poszczególnymi komponentami jakości, m.in.: terminowością i dokładnością, dokładnością i porównywalnością w przestrzeni<sup>57</sup>, przydatnością i porównywalnością w czasie, przydatnością i dokładnością. Również Baffour i in. (2013), odnosząc się do oceny jakości spisu, wskazują, że działania mające na celu poprawę jednego wymiaru jakości mogą skutkować negatywnym oddziaływaniem na inny. Znamiennym przykładem może być program przygotowań do spisu 2020 r. w USA wyrażony w książce *Envisioning the 2020 Census*, której jeden z zasadniczych rozdziałów zatytułowano: *Planning the 2020 Census: Cost and Quality* (Brown, Cohen, Cork, Citro i Lawrence, 2010).

Jak podkreślają autorzy, w konsekwencji żadna ocena nie musi obejmować wszystkich komponentów, by realnie ocenić jakość (Baffour i in., 2013). Jednak od razu zastrzegają, że nie ma optymalnego ani uniwersalnego rozwiązania, nawet w sytuacji podobnego budżetu i oczekiwań dotyczących jakości. Zauważają, że sytuacja, w której państwa wdrażają bardzo wiele odmiennych technik przeprowadzania spisu, wymaga starannej oceny jakości, identyfikacji i zrozumienia mocnych i słabych stron spisu. Co więcej, różne państwa i różni odbiorcy danych są mniej lub bardziej skłonni do kompromisu w zakresie jednych komponentów jakości kosztem innych. Jako jeden z ważniejszych efektów autorzy wskazują terminowość, którą wielu przedsiębiorców przedkłada nad idealną dokładność. Zatem nie tylko efekty w zakresie poszczególnych komponentów jakości, ale także indywidualna ocena potrzeb i preferencji są czynnikami stanowiącymi o jakości spisu.

---

<sup>57</sup> Porównywalność w przestrzeni, szczególnie w odniesieniu do małych obszarów, wymaga zastosowania bardziej zaawansowanych metod estymacji, które mimo to nie zawsze zapewniają dokładność wymaganą na poziomie krajowym.

---

## 2. TRANSFORMACJA METOD PRZEPROWADZANIA SPISÓW

---

### 2.1. Metody przeprowadzania spisu a źródła danych

W dbałości o jakość i porównywalność danych demograficznych, w szczególności dotyczących stanu i struktury ludności, Organizacja Narodów Zjednoczonych publikuje opracowania zawierające wytyczne i rekomendacje dotyczące spisów powszechnych. Pierwszym tego typu opracowaniem był raport przygotowany jako wdrożenie uchwały przyjętej przez Radę Gospodarczą i Społeczną ONZ podczas czwartej sesji (rezolucja 41 z 29 marca 1947 r.). Uchwała ta zobowiązywała Sekretarza Generalnego ONZ do przedstawienia państwom członkowskim, przygotowującym spis ludności w 1950 r., propozycji dotyczących uzyskania porównywalnych danych (UN, 1949). W 1953 r. został zainaugurowany przez ONZ Światowy Program Spisów Ludności (World Population Census Programme) i od tego czasu regularnie opracowywane są rekomendacje i wytyczne, także z uwzględnieniem specyfiki regionalnej (UN, 1949, 1958a, 1958b, 1969, 1980, 1990, 1997, 2008, 2015a; UNECE, 2015, 2006b). Zatem w zasadzie od początku istnienia ONZ Dywizja Statystyczna odgrywa kluczową rolę w koordynacji prac metodologicznych, opracowywaniu zaleceń, udzielaniu pomocy technicznej w operacjach spisowych oraz w organizacji konferencji i seminariów, a także w upowszechnianiu wyników spisów w skali międzynarodowej.

Klasyfikację metod przeprowadzania spisów zawarto po raz pierwszy w Rekomendacjach Konferencji Statystyków Europejskich z 2006 r. (UNECE, 2006b). Przed poprzednią rundą spisów 2010 r.<sup>58</sup> wyróżniono cztery podejścia:

- tradycyjne, bazujące na operacjach w terenie z wykorzystaniem długiego i krótkiego kwestionariusza z ewentualnym uzupełnieniem cech na podstawie badania reprezentacyjnego;
- tradycyjne, z coroczną aktualizacją na podstawie badań reprezentacyjnych;

---

<sup>58</sup> Runda spisów 2010 r. (inaugurowana przez Radę Gospodarczą i Społeczną ONZ) obejmowała spisy przeprowadzone w latach 2005–2014. Kolejna runda spisów obejmuje dekadę 2015–2024.

- wykorzystujące rejestry i inne źródła administracyjne;
- złożone, łączące źródła administracyjne z badaniami próbkowymi (UNECE, 2006b).

W dokumentach Organizacji Narodów Zjednoczonych dla wszystkich państw świata obok wyżej wskazanych kategorii wyróżnia się dodatkowo podejście „kroczące”, oparte na próbie rotacyjnej, oraz wskazuje na możliwe inne, alternatywne metody przeprowadzania spisu (UN, 2008).

Doświadczenie poprzedniej rundy spisów ludności i mieszkań wskazało, że wiele państw zdecydowało się na stosowanie alternatywnych metod. Biuro Spisów ONZ przewiduje, że w kolejnej rundzie spisów 2020 większość państw nadal zastosuje podejście tradycyjne. Jednocześnie jednak wskazuje się na wzrastającą liczbę państw, które decydują się na zastosowanie alternatywnych podejść. Przede wszystkim podkreśla się korzystanie z wielu różnych źródeł oraz stosowanie nowoczesnych technologii w celu gromadzenia danych. W tej sytuacji ONZ podjęło wyzwanie dotyczące uporządkowania i klasyfikacji stosowanych metod oraz przedstawienia ich dokładniejszego opisu. Dlatego w aktualnych rekomendacjach do najbliższego spisu przedstawiono podwójną klasyfikację, według sposobu gromadzenia danych oraz wykorzystania zasobów rejestrów administracyjnych (tab. 2.1)<sup>59</sup>. W zestawieniu ujęto tylko te opcje, które były już wykorzystywane w dotychczasowych spisach lub, jak zaznaczono, ich zastosowanie jest bardzo prawdopodobne. Autorzy podkreślają także, że tabela nie zawiera wszystkich możliwych kombinacji, w tym rozwiązań teoretycznych, które dotychczas nie zostały wdrożone przez żadne państwo. Ponadto w Rekomendacjach zamieszczono podstawową charakterystykę wyszczególnionych podejść metodologicznych (tab. 2.3) (UN, 2015a).

Odwołując się do klasyfikacji przedstawionej w rekomendacjach ONZ (tab. 2.1), uwzględniono pięć sposobów gromadzenia danych oraz trzy kategorie wykorzystania rejestrów administracyjnych w spisie. W zakresie metod pozyskiwania danych uwzględniono: (i) badanie pełne: spisywanie w terenie, (ii) spis kroczący, (iii) doraźne badania reprezentacyjne, (iv) regularne badania reprezentacyjne oraz (v) brak spisywania w terenie, spis w pełni oparty na rejestrach. Natomiast jeśli chodzi o źródła danych, wskazano w zasadzie jedynie na wykorzystanie rejestrów administracyjnych. Uwzględniono tutaj trzy możliwe sytuacje: (i) niekorzystanie z rejestrów, tzn. gromadzenie danych w terenie w badaniu pełnym lub reprezentacyjnym, (ii) wykorzystanie podstawowych rejestrów ludności i mieszkań (oraz utworzonego w wyniku ich integracji rejestru gospodarstw domowych), (iii) wykorzystanie zintegrowanych źródeł administracyjnych.

---

<sup>59</sup> Informacje porządkujące metodologię przeprowadzania spisu są także publikowane na stronie internetowej ONZ <https://unstats.un.org/unsd/demographic/sources/census/alternativeCensusDesigns.htm> (22.07.2017 r.)

Tabela 2.1. Metody przeprowadzania spisu ludności<sup>a</sup>

Sposób gromadzenia danych <i>type of data collection</i>	Rejestry administracyjne jako źródło informacji w spisie ludności <i>Use of registers as census data sources</i>		
	rejestry nie są wykorzystywane – gromadzenie danych w terenie – badanie pełne lub reprezentacyjne <i>no registers (fully field enumeration based or sample surveys)</i>	podstawowe rejestry (ludności, gospodarstw domowych, mieszkań) <i>base registers (individuals, households, dwellings)</i>	zintegrowane źródła administracyjne <i>integrated administrative sources</i>
Badanie pełne – spisywanie w terenie <i>full field enumeration</i>	badanie pełne spisywanie w terenie – podejście tradycyjne <i>full field enumeration only (traditional census)</i>	podstawowe rejestry i spisywanie w terenie – badanie pełne <i>base registers and full field enumeration</i>	zintegrowane źródła administracyjne i spisywanie w terenie – badanie pełne <i>integrated administrative sources and full field enumeration</i>
Spis kroczący (badanie ciągłe) <i>rolling surveys (continuous surveys)</i>	badanie pełne spisywanie w terenie i spis kroczący <i>full field enumeration and rolling surveys</i>	nie występuje <i>not available</i>	nie występuje <i>not available</i>
	spis kroczący <i>rolling census</i>		
Doraźne badania reprezentacyjne <i>ad hoc sample surveys</i>	nie występuje <i>not available</i>	podstawowe rejestry i doraźne badania reprezentacyjne <i>base registers and ad hoc sample surveys</i>	zintegrowane źródła administracyjne i doraźne badania reprezentacyjne <i>integrated administrative sources and ad hoc sample surveys</i>
Regularne badania reprezentacyjne <i>existing sample surveys</i>	nie występuje <i>not available</i>	podstawowe rejestry i regularne badania reprezentacyjne <i>base registers and existing sample surveys</i>	zintegrowane źródła administracyjne i regularne badania reprezentacyjne <i>integrated administrative sources and existing sample surveys</i>
Brak spisywania w terenie, spis w pełni oparty na rejestrach <i>no field enumeration (fully register based)</i>	nie występuje <i>not available</i>	nie występuje <i>not available</i>	spis w pełni oparty na rejestrach <i>fully register based</i>

<sup>a</sup> Ponieważ w tabeli przedstawiono poszczególne kategorie w tłumaczeniu autorki, zdecydowano także pozostawić oryginalną wersję.

Źródło: (UN, 2015a, s. 18).

Pierwsza grupa odzwierciedla sytuację klasyczną, w której jedynym źródłem danych dla spisu są informacje zgromadzone podczas przeprowadzonych w terenie wywiadów. Druga grupa dotyczy wzbogacenia źródeł o dane z dwóch podstawowych rejestrów. W ostatniej, trzeciej kategorii ujęto zintegrowane źródła

administracyjne. W Rekomendacjach uwzględnia się połączenie źródeł administracyjnych z danymi dotyczącymi biznesu, podatków, wykształcenia, zatrudnienia i innych odpowiednich systemów. Zatem przedstawiona w Rekomendacjach klasyfikacja źródeł danych wykorzystywanych w spisach ludności w niejawnym sposób przedstawia takie, które już w rundzie 2010 r. były wykorzystywane i których wykorzystanie jest planowane w najbliższej przyszłości (GUS, 2017d).

Z tego powodu przedstawione w niniejszym opracowaniu zestawienie (tab. 2.2) zawiera modyfikacje uwzględniające szerszy zakres źródeł informacji. Dotyczy ono

Tabela 2.2. Metody<sup>a</sup> przeprowadzania spisu ludności a źródła danych i sposób ich gromadzenia

Źródła danych i sposób ich gromadzenia	Źródła danych i sposób ich gromadzenia					
	badanie pełne, spisywanie w terenie <sup>b</sup>	spis kroczący, badanie ciągłe	doraźne badania reprezentacyjne	regularne badania reprezentacyjne	podstawowe rejestry ludności i mieszkań	zintegrowane źródła administracyjne i pozaadministracyjne
Badanie pełne, „spisywanie w terenie” <sup>b</sup>	ST	SK	ST	ST	SM	SM
Spis kroczący (badanie ciągłe)	SK	SK	SK	SK	SM	SM
Doraźne badania reprezentacyjne	ST	SK	Doraźne badania reprezentacyjne nie jest spisem	Kombinacja taka nie jest spisem	SM	SM
Regularne badania reprezentacyjne	ST	SK	Kombinacja taka nie jest spisem	Regularne badania reprezentacyjne nie jest spisem	SM	SM
Podstawowe rejestry ludności i mieszkań	SM	SM	SM	SM	Podstawowy rejestr nie jest spisem <sup>c</sup>	SR
Zintegrowane źródła administracyjne i pozaadministracyjne	SM	SM	SM	SM	SR	SR

<sup>a</sup> Metoda przeprowadzania spisu: spis tradycyjny [ST], spis kroczący [SK], spis metodą mieszaną [SM], spis oparty na rejestrach administracyjnych [SR].

<sup>b</sup> Jako „spisywanie w terenie” przyjmuje się także zastosowanie CAPI, CATI, CAII.

<sup>c</sup> Określenie, czy podstawowy rejestr ludności spełnia wymogi spisu, jest uwarunkowane przede wszystkim tym, czy obejmuje minimum spisowe, dostarcza danych porównywalnych w skali międzynarodowej oraz czy możliwa jest aktualizacja.

przede wszystkim rozszerzenia ostatniej, trzeciej kategorii o źródła pozaadministracyjne, takie jak: systemy informacyjne dotyczące odbiorców energii elektrycznej, wody, systemy informacyjne dostawców usług telekomunikacyjnych zawierające wykaz odbiorców oraz inne, szczególne zasoby danych tworzone w ramach działalności przedsiębiorców.

Mając na względzie stwierdzenie autorów Rekomendacji (UN, 2015a), że w klasyfikacji metod przeprowadzania spisów (tab. 2.1) nie ujęto wszystkich możliwych kombinacji, zaproponowano nieco inne ujęcie (tab. 2.2), bardziej przejrzyste i pozwalające na eliminację pewnych nieścisłości zawartych w Rekomendacjach. Otóż przede wszystkim należy odpowiedzieć na pytanie, która z możliwych kombinacji źródeł i metod gromadzenia danych nie spełnia warunków spisu ludności. Z pewnością dotyczy to badań reprezentacyjnych, tak doraźnych, jak i regularnych, czy też ich kombinacji. Nie spełniają one warunków powszechności i możliwości przedstawienia danych dla małych jednostek terytorialnych. Należałoby też wyjaśnić, czy podstawowy rejestr ludności i mieszkań spełnia warunki spisu. Decyzja zdaje się zależeć od tego, czy rejestr obejmuje minimum spisowe i czy uzyskane dane odpowiadają standardom międzynarodowym, a także jakie są możliwości aktualizacji na moment krytyczny spisu. Obecnie podstawowe rejestry ludności i mieszkań, w państwach stosujących podejście oparte na rejestrach, pełniły rolę rejestru głównego. Jednak wymogi zakresu informacyjnego spisu były spełnione dopiero w połączeniu z innymi źródłami.

Wszystkie pozostałe kombinacje źródeł informacji i metod gromadzenia danych zdają się spełniać podstawowe warunki formułowane wobec spisów ludności (por. rozdz. 1.2.1). Liczba kombinacji ukazuje ogromną różnorodność możliwych podejść. Przy całej otwartości podejścia zaproponowanego w Rekomendacjach trzy kwestie wymagają podkreślenia. Pierwsza dotyczy wielu wariantów spisu tradycyjnego w kombinacjach z krótkim i długim formularzem oraz wykorzystaniem danych badań reprezentacyjnych. Zostało to uwidocznione w tabeli 2.2 oraz opisane w tabeli 2.3. Podobnie wiele jest możliwych kombinacji w przypadku spisu kroczącego. Druga kwestia godna podkreślenia, aczkolwiek w tej dyskusji może się wydawać oczywista, to zastąpienie charakterystycznego dla spisu tradycyjnego bezpośredniego spisywania w terenie rejestrem ludności. Często za „spisywanie w terenie” przyjmuje się także zastosowanie wspomaganych komputerowo technik gromadzenia danych, jak np. CAPI, CATI, CAII<sup>60</sup>. W tabeli 2.2 wyraźnie uwidacznia

---

<sup>60</sup> Są to wspomagane komputerowo techniki gromadzenia danych: CAPI – wywiad bezpośredni przeprowadzony z respondentem przy użyciu kwestionariusza w formie elektronicznej zamiast papierowej i gromadzeniu danych na urządzeniach mobilnych (CAPI – Computer Assisted Personal Interview); CATI – wspomagany komputerowo wywiad prowadzony za pośrednictwem telefonu (CATI – Computer Assisted Telephone Interview); CAII – metoda zbierania informacji od respondentów poprzez samodzielne wypełnienie formularzy elektronicznych za pośrednictwem internetu (CAII – Computer Assisted Internet Interview).

Tabela 2.3. Opis metod przeprowadzania spisów ludności

Podejście metodologiczne <i>Approach</i>	Opis podejścia <i>Description</i>
<p>Badanie pełne spisywanie w terenie – podejście tradycyjne <i>Full field enumeration (traditional census)</i></p> <p>Badanie pełne spisywanie w terenie i badanie kroczące <i>Full field enumeration and rolling surveys</i></p>	<p>W spisie tradycyjnym informacje dotyczące osób fizycznych i gospodarstw domowych są zbierane przy użyciu kwestionariusza spisowego bezpośrednio w terenie (możliwe jest zastosowanie takich metod, jak wywiad telefoniczny, pocztowy, poprzez internet lub kombinacja różnych form). Niektóre informacje mogą być gromadzone poprzez doraźne (lub ciągłe – spis kroczący) badanie reprezentacyjne przy użyciu długiego formularza, podczas gdy tzw. krótki formularz jest stosowany w badaniu pełnym. Zwyczajowo długi i krótki kwestionariusz wykorzystuje się razem w ramach jednego spisu w tym samym czasie. Wyniki badania reprezentacyjnego (długi formularz) są traktowane jako wyniki spisu.</p> <p>W wyniku modelowania rezultaty badania pełnego oraz kumulacja wyników badania ciągłego (spis kroczący) są wykorzystywane do oszacowania rocznych (lub w innym przedziale czasu) szczegółowych danych dla różnych poziomów agregacji terytorialnej. W tym podejściu wyniki badań częściowych mogą być kumulowane w czasie w celu częstszych szacunków dla najmniejszych jednostek podziału terytorialnego.</p> <p><i>Information on census topics concerning individuals and households is collected by census questionnaire directly from respondents using enumerators or other modes of data collection (e.g. telephone interview, mail out/mail back, Internet), or by applying a combination of different modes of data collection. In a traditional census, full field enumeration may include an in-built ad hoc survey by use of the long form or can be combined with a rolling survey. It is common that a short form is used together with a long form. Both forms are utilized during the same time frame of the census. While the long form estimates are not based on full coverage, they are regarded as census output. Through modelling, the full field enumeration with the continuous cumulative survey is used to generate yearly (or other interval) estimates of detailed characteristics for different geographic levels. With this approach, the sample can be cumulated over time to produce statistics at the lowest levels of geographic detail to provide more frequent and relevant data</i></p>
<p>Spis kroczący <i>Rolling census</i></p>	<p>Informacje o osobach i gospodarstwach domowych są gromadzone w wyniku badania ciągłego oraz kumulowane, obejmują obszar całego państwa w określonym czasie (zazwyczaj kilku lat, w przeciwieństwie do jednego dnia tzw. momentu krytycznego lub krótkiego okresu). Wyróżnia się dwa podstawowe parametry spisu kroczącego: długość okresu gromadzenia informacji (która jest powiązana z wymaganą częstością aktualizacji). Drugi parametr to wielkość próby, która jest powiązana z wielkością najmniejszych jednostek podziału terytorialnego, dla których są wymagane szacunki spisowe.</p> <p><i>Information on individuals and households is collected through a continuous cumulative survey covering the whole country over a period of time (generally years), rather than a particular day or short period of enumeration. The two main parameters of a rolling census are the length of the period of enumeration (which is linked to the frequency of updates required) and the sampling rate (which depends on the geographic levels required for dissemination purposes)</i></p>



Podejście metodologiczne <i>Approach</i>	Opis podejścia <i>Description</i>
Metodologia złożona <i>Combined methodology</i>	<p>Informacje o osobach i gospodarstwach domowych są gromadzone poprzez łączenie danych z jednego lub kilku badań reprezentacyjnych i badania pełnego (spisywanie tradycyjne) z danymi rejestrów administracyjnych lub statystycznych. Dane z rejestrów są wykorzystywane nie tylko jako operat losowania czy dla wsparcia działań operacyjnych w terenie, ale również bezpośrednio jako źródło danych dla niektórych informacji spisowych. W niektórych przypadkach dane rejestrowe są wykorzystywane do wstępnego wypełniania kwestionariuszy, które są weryfikowane lub poprawiane podczas gromadzenia danych. Badania doraźne są stosowane w celu dostarczenia informacji dla tych zmiennych (tematów) spisowych, które nie są dostępne na podstawie źródeł administracyjnych lub do dostosowywania danych z rejestrów o niewystarczającej jakości.</p> <p><i>Information on individuals and households is collected by combining data collected from one or more surveys or full field enumeration with administrative or statistical registers. Data from registers are employed not only as a frame or to support field operations, but directly as a data source for some census information. In some cases, register data are used to prefill the questionnaires to be verified or corrected during data collection. Ad hoc sample surveys are used to provide information on census topics not available from administrative sources or to adjust data that are of poor quality in registers</i></p>
Podstawowe rejestry i regularne badania reprezentacyjne <i>Base register and existing sample surveys</i>	<p>Informacje o osobach i gospodarstwach domowych są zbierane z istniejących źródeł administracyjnych, z których najważniejsze znaczenie mają rejestry: ludności, gospodarstw domowych i mieszkań. Są one powiązane na poziomie jednostkowym z informacjami pochodzącymi z istniejących badań reprezentacyjnych. Dane nie są gromadzone bezpośrednio w terenie. Istniejące badania częściowe obejmują badania prowadzone w okresie międzyspisowym dotyczące różnych tematów, takich jak badanie siły roboczej i badanie warunków i jakości życia.</p> <p><i>Information on individuals and households is collected from existing administrative sources, namely different kinds of registers, of which the following are of primary importance: individuals, households and dwellings. These are linked at the individual level with information from existing sample surveys. No field data collection will take place. Existing sample surveys include intercensal sample surveys on different topics, such as the labour force survey and the living standards survey</i></p>
Spis w pełni oparty na rejestrach <i>Fully register based</i>	<p>Informacje o osobach i gospodarstwach domowych są gromadzone z istniejących źródeł administracyjnych, przede wszystkim rejestrów, z których najważniejsze znaczenie mają rejestry: ludności, gospodarstw domowych i mieszkań. Są one łączone na poziomie jednostkowym z informacjami pochodzącymi z innych źródeł administracyjnych lub statystycznych, takich jak rejestry przedsiębiorstw, podatków, oświaty, zatrudnienia i inne odpowiednie rejestry.</p> <p><i>Information on individuals and households is collected from existing administrative sources, namely different types of registers, of which the following are of primary importance: individuals, households and dwellings. These are linked at the individual level with information taken from other administrative or statistical sources, such as business, tax, education, employment and other relevant registers</i></p>

się sytuacja analogicznych kombinacji spisu tradycyjnego (krótki formularz) oraz rejestrów administracyjnych (podstawowy rejestr ludności) z badaniami reprezentacyjnymi. Trzecią kwestią jest połączenie badania pełnego z danymi rejestrów administracyjnych, co w tabeli 2.2. ujęto jako jeden z wariantów metody mieszanej. Wymagająca podkreślenia wydaje się możliwość łącznego wykorzystania wszystkich rodzajów źródeł, co uwzględni zarówno ujęcie „spisu wirtualnego” (nieuwzględnione *explicite* w tab. 2.1), a także planowana metoda przeprowadzenia spisu w Polsce w 2021 r. (por. rozdz. 4.3).

Kompleksową ewolucję metod przeprowadzania spisu, począwszy od ujęcia tradycyjnego, z uwzględnieniem corocznych aktualizacji, spisu krocącego, opartego na rejestrach po spis „wirtualny”<sup>61</sup>, przedstawiono także w pracy (Baffour i in., 2013). Autorzy odwołują się do przykładów wybranych państw, podkreślając w każdym przypadku istotne cechy charakterystyczne. I tak spis tradycyjny przedstawiają na podstawie Wielkiej Brytanii i Malezji. Modyfikację tego ujęcia, polegającą na ciągłej aktualizacji, omawiają na przykładzie Stanów Zjednoczonych. Francja jest najczęściej wskazywanym państwem stosującym w praktyce spis kroczący. Ze względu na wieloletnie doświadczenie zrozumiałe jest przedstawienie spisów opartych na rejestrach w państwach skandynawskich, w tym przypadku Szwecji, a w odniesieniu do spisu „wirtualnego” – w Holandii.

Tabela 2.4. Liczba państw według metody przeprowadzania spisu ludności w latach 1985–2014

Metoda przeprowadzenia spisu	Runda 1990 (1985–1994)	Runda 2000 (1995–2004)	Runda 2010 (2005–2014)
Liczba państw, które nie przeprowadziły spisu <sup>a</sup>	31 (13%)	34 (14%)	12 (5%)
Liczba państw, które przeprowadziły spis	208 (86%)	207 (86%)	227 (94%)
Metoda tradycyjna	200 (83%)	194 (81%)	188 (78%)
Metoda alternatywna	8 (3%)	13 (5%)	39 (16%)
Inna metoda	2 (1%)	0 (-)	2 (1%)
Razem	241	241	241

<sup>a</sup> Pozycja ta uwzględnia państwa, które zgromadziły dane z rejestrów, ale formalnie nie wygenerowały spisu ludności.

Źródło: (Kukutai i in., 2015, s. 12).

Metody przeprowadzania spisów ludności wykorzystane w ostatniej rundzie 2010 są także przedmiotem badania przeprowadzonego przez Kukutai i in. (2015). Autorzy przedstawiają historię spisów z perspektywy celu gromadzenia danych oraz stosowanej metody, obecnie i w przeszłości. Wyniki badania obejmującego 241 krajów całego świata wskazują, że spis ludności, jako badanie statystyczne,

<sup>61</sup> W pracy zdecydowano, podobnie jak w opracowaniu (Baffour i in., 2013), wyróżnić określenie spis „wirtualny”. Autorka uważa je za bardzo niezręczne. Określenie to sugeruje, że jest to spis nierzeczywisty, tymczasem jego wyniki są jak najbardziej realne.

jest przeprowadzany w coraz większej liczbie państw, a jego alternatywne metody zyskują popularność przede wszystkim w Europie (tab. 2.4).

Statystyki ONZ<sup>62</sup> podają, że w poprzedniej rundzie spisów powszechnych, do 1.12.2014 r., spisanych zostało 6 391 500 187 osób, co zgodnie z szacunkami ONZ stanowiło 93% ludności całego świata zamieszkującej w 214 państwach.

### 2.1.1. Spis tradycyjny

Metoda tradycyjna (klasyczna) jest nadal najczęściej stosowaną metodą przeprowadzania spisów. Jest to złożony system gromadzenia, przetwarzania, ewaluacji, analizy i upowszechniania informacji demograficznych, ekonomicznych i społecznych od poszczególnych osób i gospodarstw domowych dotyczących ściśle określonego czasu i terytorium. Sposób gromadzenia danych uwzględnia zarówno wypełnienie kwestionariusza spisowego przez respondentów, jak i wywiad prowadzony przez ankieterów. Ankieterzy mają wyznaczony obszar, w którym zbierają informacje od wszystkich mieszkańców w zazwyczaj krótkim czasie, tak by spełniony był warunek powszechności i jednoczesności spisu. Mogą być stosowane zarówno tzw. krótkie, jak i długie formularze spisowe. Krótki formularz obejmuje zwykle całą populację i zawiera pytania dotyczące cech podstawowych. Natomiast długi formularz jest stosowany w badaniu reprezentacyjnym dla tzw. dużej próby i zawiera szereg szczegółowych pytań o różnym zakresie tematycznym (np. badanie dzietności przeprowadzane łącznie ze spisem ludności w 1970, 1988 i 2002 r. czy badanie migracji długookresowych w NSP 2002). Badania wykorzystujące obydwa formularze: krótki i długi, są przeprowadzane jednocześnie. Chociaż długi formularz wykorzystywany jest w badaniu reprezentacyjnym, jego wyniki są określane mianem spisowych.

Spisy przeprowadzane metodą tradycyjną są bardzo zróżnicowane, m.in. pod względem formy gromadzenia danych. Przykładowo można wymienić stosowany w Polsce wywiad przeprowadzany przez rachmistrza spisowego, formularze przesyłane pocztą, pozostawiane w administracji budynku, wywiad telefoniczny czy internetowy.

Spis klasyczny niewątpliwie dostarcza informacji o całej populacji w określonym czasie i umożliwia ich prezentację dla szczegółowych jednostek terytorialnych. W spisach tradycyjnych w Polsce publikowane były dane dla gmin. Tradycyjny spis jest szczególnie pożądanym w państwach o strukturze federalnej, w których jednym z wymogów dla planowania i alokacji funduszy jest jednoczesna informacja dla małych jednostek administracyjnych. Wymóg ten jest także ważny dla potrzeb określenia granic obwodów wyborczych. Jednocześnie tradycyjne spisy powszechne są postrzegane jako najbardziej skomplikowana, złożona i kosztowna forma

---

<sup>62</sup> Informacje dostępne na stronie internetowej 5 sierpnia 2017 r. <https://unstats.un.org/unsd/demographic/sources/census/wphc/default.htm>.

zbierania danych. Dodatkowo, podobnie jak w przypadku innych badań, udział w spisie, choć obowiązkowy, wymaga pełnej świadomości i zgody społeczeństwa<sup>63</sup>.

Ze względu na złożoność i koszt spisy tradycyjne są zazwyczaj przeprowadzane raz na dziesięć lat (czasami co pięć lat), a więc dane spisowe po kilku latach tracą na aktualności. Spis ludności metodą tradycyjną jest przeprowadzany nadal w większości państw świata, np. w Kanadzie, Wielkiej Brytanii, Rosji czy Kolumbii. W Kanadzie spisy są przeprowadzane raz na pięć lat. Moment krytyczny ostatniego spisu wyznaczony był na 10 maja 2016 r. W spisie obowiązywał zarówno krótki, jak i długi formularz. Krótki formularz (10 pytań) był powszechny, a długi (36 stron) obejmował co czwarte gospodarstwo domowe i był określony jako Narodowe Badanie Gospodarstw Domowych (*National Household Survey*<sup>64</sup>). Obydwa formularze można było wypełnić w internecie. Drogą mailową zostały rozesłane 16-znakowe kody umożliwiające wypełnienie długich formularzy online. Statistics Canada (Kanadyjski Urząd Statystyczny) zachęcał obywateli do jak najszybszego wypełnienia formularzy online<sup>65</sup>.

### 2.1.2. Spis tradycyjny z coroczną aktualizacją danych

Spis tradycyjny z coroczną aktualizacją danych, jak sugeruje nazwa, prezentuje podejście klasyczne, którego modyfikacja polega na koncentracji na szacunku liczby ludności i podstawowych cech demograficznych w roku spisowym. Natomiast duże badania gospodarstw domowych w wymiarze społeczno-ekonomicznym są przeprowadzane każdego roku przez całą dekadę, zastępując tzw. długi formularz spisowy powtarzaniem każdego roku badaniem reprezentacyjnym.

W spisie z coroczną aktualizacją przyjmuje się, że skumulowana roczna wielkość próby odpowiada tej, jaką przyjęto by w badaniu spisowym dla długiego formularza, przeprowadzanym raz na 5 bądź 10 lat. Dla zapewnienia odpowiedniej jakości danych dla małych obszarów losowana jest większa próba, której wyniki są odpowiednio ważone zgodnie ze schematem losowania oraz kalibrowane w celu redukcji błędów wynikających z odmowy udziału w badaniu i korekty wynikającej z oceny kompletności. Ostatecznie stosowany jest system ważenia, który zapewnia szacunki odpowiadające standardom spisu tradycyjnego przeprowadzanego jako badanie okresowe.

Powody organizacji spisu klasycznego z coroczną aktualizacją można określić dwojako. Przede wszystkim jest to dostarczenie danych spisowych z większą częstotliwością niż raz na 10 lat<sup>66</sup>. Z drugiej strony chodzi o zmniejszenie ryzyka

<sup>63</sup> Pomimo ustawowo określonych kar za odmowę udziału w spisie autorka nie zna przypadków jej egzekwowania w Polsce.

<sup>64</sup> Szczegółowe informacje dostępne są na stronie internetowej: <https://unstats.un.org/unsd/demographic/sources/census/wphc/Canada/CAN-2016-05-10.pdf#page=2&zoom=auto,-31,449>

<sup>65</sup> Już pierwszego wieczoru odnotowano internetowe wypełnienie około 700 tys. kwestionariuszy, a drugiego 1,5 mln. Tego obciążenia nie wytrzymał system komputerowy, który zawiesił się na 45 minut.

<sup>66</sup> W przeszłości spis ludności lub mikrospis dostarczał tzw. próby „matki” do losowania prób „córek” dla badań specjalnych w ciągu najbliższych kilku lat. Celem uniknięcia dezaktualizacji pró-

operacyjnego związanego z tak dużym badaniem jak spis. Przedstawiony proces organizacji spisu tradycyjnego z coroczną aktualizacją jest drogi i technicznie bardzo trudny, wymagający kompleksowego wieloletniego planu badania, jego weryfikacji i doskonalenia. Jednak zastosowanie takiego podejścia może być niezbędne w krajach, w których regulacje prawne wymagają od instytucji statystyki publicznej dostarczania, z określoną częstotliwością, szacunków ludności na podstawie badania pełnego. Spis klasyczny z coroczną aktualizacją był przeprowadzany w Stanach Zjednoczonych (US Census Bureau, 2005) i od 2005 r. w Peru (INEI, 2005).

### 2.1.3. Spis kroczący

„Spis kroczący” stanowi alternatywę dla klasycznego modelu spisu powszechnego. Ma on charakter badania ciągłego, w którym kumulowane są wyniki wcześniejszych badań. Spis obejmuje obszar całego państwa i zazwyczaj dotyczy długiego okresu (całego roku, a nawet lat), w przeciwieństwie do klasycznego ujęcia, w którym jest przeprowadzany w krótkim czasie. Można wskazać dwa główne parametry spisu kroczącego: długość okresu gromadzenia danych (co się wiąże z wymaganą częstotliwością aktualizacji) oraz współczynnik określający wielkość próby (co zależy od dostępnego budżetu i szczegółowości podziału terytorialnego, dla którego mają być publikowane wyniki spisu). Przykładowo spis może być zaprojektowany jako cykl corocznych badań reprezentacyjnych, przy czym pojedyncze badanie umożliwia uogólnianie wyników na poziomie całego kraju. Z kolei kumulacja danych z trzech kolejnych badań pozwala na dokonanie uogólnień na poziomie regionalnym. Jednak przedstawienie wiarygodnych szacunków dla małych jednostek terytorialnych wymaga kumulacji danych z pięciu kolejnych lat. Organizacja corocznych badań może być różna, określająca czas gromadzenia danych jako cały rok, jeden kwartał, miesiąc lub nawet krótszy okres.

Od 2004 r. metodologia spisu kroczącego jest stosowana we Francji. Podkreślając ważną potrzebę ciągłego monitorowania zmian dotyczących ludności i struktury społeczeństwa, Francuzi powołują się na uregulowania prawne z 2002 i 2003 r.<sup>67</sup>, które umożliwiają coroczną publikację wyników spisu ludności dla każdej gminy.

We Francji wdrożenie procedury spisu kroczącego wymagało zastosowania złożonych technik losowania oraz modelowania i estymacji<sup>68</sup>. Ponadto niezmier-

---

by „matki” należało wówczas eliminować budynki wyburzone oraz dolosowywać budynki nowo wybudowane.

<sup>67</sup> Ustawa nr 2003–485 o spisie ludności opublikowana 8 czerwca 2003 r. (Official Journal) oraz ustawa nr 2002–276 z 27 lutego 2002 r. (artykuły 156 do 158) określająca zasady przeprowadzania spisu i upowszechniania oficjalnych wyników dla każdej gminy, <https://www.insee.fr/en/metadonnees/source/s1156>, dostęp do strony internetowej francuskiego urzędu statystycznego (The National Institute of Statistics and Economic Studies INSEE) 6 września 2017 r.

<sup>68</sup> Informacje zaczerpnięto ze strony internetowej INSEE: <https://www.insee.fr/en/metadonnees/source/s1156>, dostęp 6 września 2017 r.

nie ważne były konsultacje z wszystkimi interesariuszami, w tym użytkownikami danych, organami administracji rządowej i samorządowej. Za główną zaletę tego podejścia uznano większą częstotliwość aktualizacji: tradycyjny spis ludności był przeprowadzany mniej więcej co dziesięć lat, a spis kroczący umożliwił coroczną aktualizację danych. Sposób badania jest zróżnicowany w zależności od wielkości gminy, łączy badanie pełne (gminy poniżej 10 tys. mieszkańców – rocznie 1/5 obszaru) oraz reprezentacyjne dla 40-procentowej próby (gminy powyżej 10 tys. – próba ok. 8% mieszkań rocznie). Łącznie, kumulując coroczne wyniki, w okresie pięciu lat badaniu podlega ok. 70% populacji. Zebrane w ten sposób dane pozwalają na szacunki różnych charakterystyk według stanu na 1 stycznia środkowego roku pięcioletniego okresu. W ten sposób zdefiniowano okres referencyjny. Pierwsze pięcioletnie badanie spisowe przeprowadzono we Francji w latach 2004–2008. Umożliwiło ono opublikowanie wyników pełnego spisu według stanu ludności na 1 stycznia 2006 r. Od tego momentu każdego roku przedstawiane są szacunki stanu i struktury populacji na podstawie danych z ostatnich pięciu lat (wyniki sprzed pięciu lat są pomijane, a dane są aktualizowane wynikami bieżącymi).

Podobnie jak we wcześniejszych spisach, badanie jest przeprowadzane za pomocą krótkiego kwestionariusza określanego jako główny, oraz długiego dla badania uzupełniającego. Kwestionariusz krótki składa się z 25 pytań, które dotyczą danych osobowych, płci, wieku, narodowości, daty i miejsca urodzenia, miejsca zamieszkania przed rokiem, wykształcenia, zawodu. Natomiast kwestionariusz uzupełniający dotyczy zagadnień ekonomiczno-społecznych, sektora zatrudnienia, aktywności ekonomicznej, struktury gospodarstwa domowego. Wypełniają go wszystkie osoby zamieszkujące w mieszkaniach wylosowanych do próby (gminy liczące powyżej 10 tys. mieszkańców) oraz 25% mieszkań w gminach poniżej 10 tys. mieszkańców. Dla każdego budynku wypełniany jest także oddzielny kwestionariusz obejmujący 15 pytań.

Zaletą spisu kroczącego jest, do pewnego stopnia, łagodzenie obciążeń ludności udziałem w badaniu, obniżenie kosztów i pracochłonności. Co więcej, każdego roku można doskonalić proces przeprowadzania badania, począwszy od techniki losowania próby po upowszechnianie wyników i zastosowanie nowoczesnych technologii. Główną wadą jest to, że spis kroczący nie spełnia wymogu jednoczesności badania. Sytuacja komplikuje się w szczególności w zakresie możliwości porównań, przykładowo dla małych obszarów, z powodu różnego czasu obserwacji i aktualizacji danych.

#### **2.1.4. Spis oparty na rejestrach**

Pierwsze spisy wykorzystujące dane rejestrów administracyjnych (a więc co najmniej kilka źródeł informacji) przeprowadzono w Finlandii i Norwegii już w 1970 r. (Myrskylä, 2004; Tønder, 2008). Początkowo wykorzystano jedynie dane rejestru ludności, a następnie w kolejnych spisach stopniowo wprowadzano dalsze rejestry dotyczące zagadnień społecznych i gospodarczych. W rundzie spisów 1980 r.

w Danii, 1990 r. w Finlandii i 2000 r. w Norwegii w pełni „wygenerowano” dane spisu ludności z danych rejestrów administracyjnych (tab. 2.5). Główna idea spisów opartych na rejestrach sprowadza się do jak najbardziej efektywnego wykorzystania już istniejących źródeł informacji. Podstawowe znaczenie ma rejestr ludności i mieszkań, którego dane, na poziomie jednostek indywidualnych, są łączone z danymi rejestru ubezpieczeń, służby zdrowia, podatkowego, edukacyjnego itp., tworząc tzw. „złoty rekord”. Łączenie rejestrów jest przeprowadzane na podstawie klucza identyfikacyjnego. W Polsce takim identyfikatorem może być np. numer PESEL. W procesie łączenia danych z różnych źródeł mogą być także wykorzystywane różne metody statystycznej integracji danych (D’Orazio, Di Zio i Scanu, 2006; Little i Rubin, 2002; Rässler, 2002; Winkler, 2005). Coraz szersze wykorzystanie zasobów administracyjnych dla celów statystycznych, prowadzące do zastąpienia danych badań reprezentacyjnych, można określić transformacją do systemu statystyki opartego na rejestrach administracyjnych (Wallgren i Wallgren, 2014).

W latach 1964–1969 we wszystkich państwach skandynawskich powstały Centralne Rejestry Ludności, w których utworzono unikatowe numery identyfikacyjne dla wszystkich osób. W następnych latach utworzono kolejne rejestry, dotyczące m.in. przedsiębiorstw, mieszkań, wykształcenia, zatrudnienia (por. też tab. 2.5).

Tabela 2.5. Rok utworzenia rejestru i rok pierwszego wykorzystania w spisie ludności według typu rejestru, państwa skandynawskie

Rodzaj rejestru	Dania		Finlandia		Norwegia		Szwecja	
	utworzenie	zastosowanie w spisie	utworzenie	zastosowanie w spisie	utworzenie	zastosowanie w spisie	utworzenie	zastosowanie w spisie
Ludności	1968	1981	1969	1970	1964	1970	1967	1975
Przedsiębiorstw	1975	1981	1975	1980	1965	1980	1963	1975
Mieszkań	1977	1981	1980	1985	2001	2011	2008?	2011?
Warunków mieszkaniowych	1977	1981	1980	1985	2001	2011	2008?	2011?
Wykształcenia	1971	1981	1970	1975	1970	1980	1985	1990
Zatrudnienia	1979	1981	1987	1990	1978	2001	1985	1985
Rodzin	1968	1981	1978	1980	1964	1980	1960	1975
Gospodarstw domowych <sup>a</sup>	1968	1981	1970	1975	2001	2011	2011?	2011?
Dochodów	1970	1981	1969	1970	1967	1980	1968	1975
Spis całkowicie oparty na rejestrach		1981		1990		2011		2011?

<sup>a</sup>Gospodarstwo domowe zostało określone jako osoby wspólnie zamieszkujące w jednym mieszkaniu

Uwaga: Kursywa i znak zapytania pojawiające się w odniesieniu do utworzenia/wykorzystania niektórych rejestrów w Szwecji i Norwegii, występują bez wyjaśnienia w oryginalnym tekście.

Źródło: (Tønder, 2008, s. 3).

Z początkiem lat 70. zaczęto wykorzystywać dane rejestrów do celów statystycznych. Pierwsze zastosowania dotyczyły statystyki ludności i dochodu. Czas wprowadzenia rejestru i jego wykorzystania w statystyce różnił się w poszczególnych państwach i w zależności od zakresu merytorycznego. Jednak wszędzie stosowano identyczną zasadę – gdy uznano dane rejestru za wiarygodne, wykorzystywano je w spisie. Utworzenie rejestrów obejmujących cały zakres merytoryczny spisu oraz ich pozytywna ocena umożliwiły wygenerowanie danych spisu ludności w całości z rejestrów administracyjnych. Wśród podstawowych czynników warunkujących realizację spisu opartego na rejestrach Tønder (2008) wskazuje na odpowiednie regulacje prawne, akceptację społeczną, zintegrowany system identyfikacyjny oraz istnienie kompleksowego, wiarygodnego systemu rejestrów dla potrzeb administracyjnych.

Jednym z zasadniczych warunków zastosowania metodologii spisu opartego na rejestrach jest funkcjonowanie w danym państwie dobrej jakości rejestru ludności obejmującego całą badaną populację (ludność rezydującą). Rejestr ludności pełni w spisie rolę rejestru centralnego, który jest łączony w pierwszej kolejności z rejestrami budynków i mieszkań oraz z rejestrem podmiotów gospodarczych. Wszystkie jednostki indywidualne są łączone za pomocą unikatowych numerów identyfikacyjnych, tworząc spójny system. Pojedyncze osoby są łączone, tworząc rodziny i gospodarstwa domowe przypisywane do mieszkań i budynków, w których mieszkają, oraz do przedsiębiorstw, w których pracują. Z kolei wszystkie obiekty: domy, mieszkania i przedsiębiorstwa, są precyzyjnie zlokalizowane poprzez system współrzędnych geograficznych.

Fundamentalne znaczenie mają odpowiednie regulacje prawne umożliwiające korzystanie przez statystykę publiczną z danych rejestrów oraz dobra współpraca z gestorami, co zapewnia płynne zasilanie bazy danych spisowych. Do konstrukcji systemu statystyki opartej na źródłach administracyjnych niezbędna jest ustawa upoważniająca urząd statystyczny do korzystania z danych rejestru ludności na poziomie jednostek indywidualnych oraz ich integracji z danymi innych rejestrów. Akty prawne zapewniają także odpowiednią ochronę danych indywidualnych. Legislacja w krajach skandynawskich dopuszcza korzystanie przez instytucje statystyki państwowej z danych, które zostały zgromadzone dla innych celów, z ograniczeniem do celów statystycznych oraz badawczych. Obowiązuje w tym przypadku zasada „one-way traffic”, co oznacza, że dane administracyjne w posiadaniu urzędu statystycznego stają się poufne, nawet jeśli w posiadaniu organów administracyjnych mają charakter publiczny (Tønder, 2008). Warto w tym miejscu uzupełnić, że dane zebrane do celów statystycznych są zawsze poufne.

Bezwzględnie konieczna jest ocena jakości danych rejestrów, adekwatności ich zasobów do potrzeb spisu, możliwości integracji. Bez spełnienia tych warunków trudno byłoby sprostać oczekiwaniom formułowanym wobec danych spisu ludności.



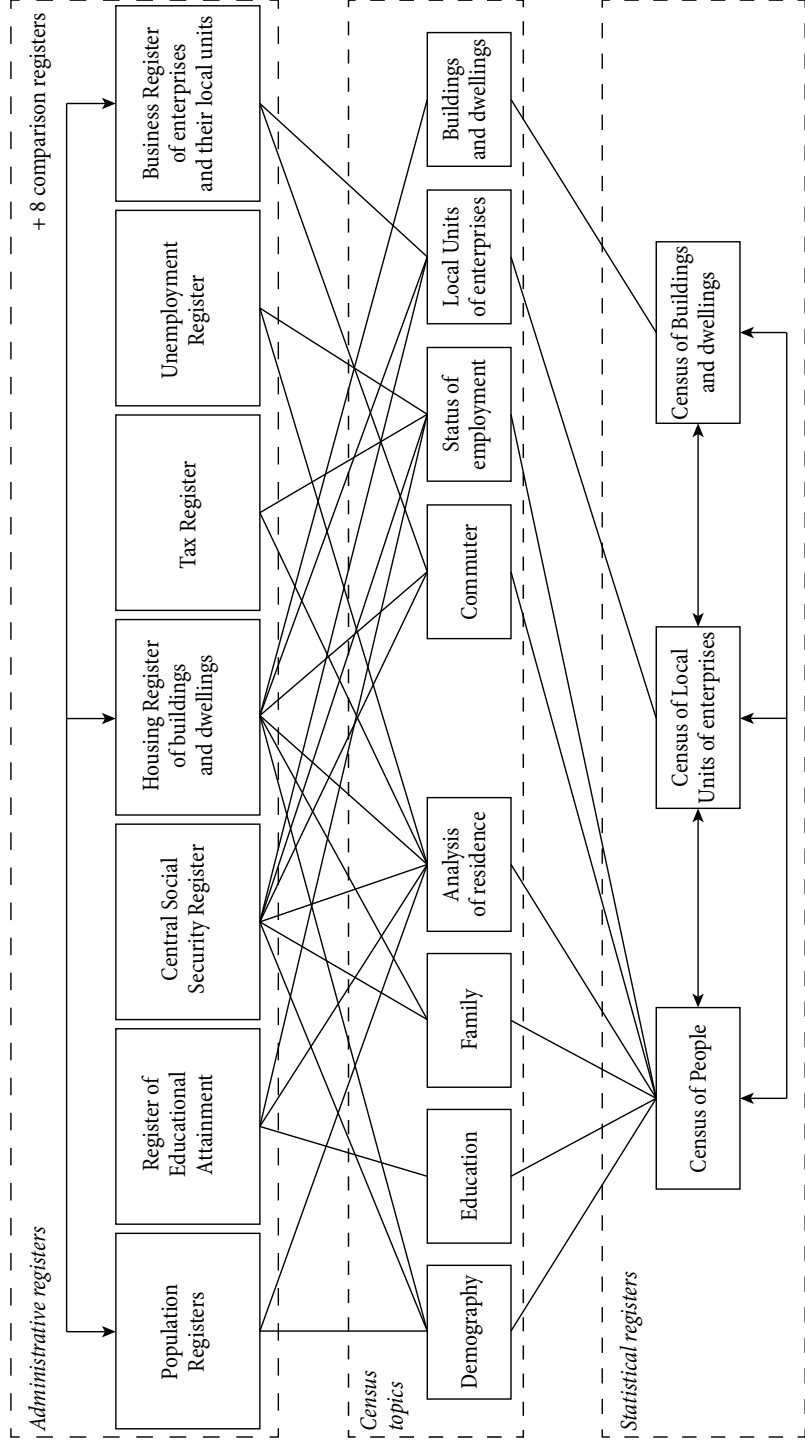
Spis oparty na rejestrach ma niewątpliwie wiele zalet, wśród których najważniejsza jest możliwość redukcji kosztów i obciążeń społecznych oraz zwiększenia częstości pozyskiwania danych spisowych. Nie bez znaczenia są także korzyści wynikające z dostępności danych rejestrów dla dowolnego poziomu jednostek terytorialnych. Do zalet tego podejścia należy zaliczyć poprawę spójności i wiarygodności danych oraz ich bezpieczeństwa. Natomiast zasadniczą wadą spisu opartego na rejestrach jest ograniczenie zakresu informacji spisowych jedynie do tych cech, które są zawarte w rejestrach. Dodatkowo, często stosowane w rejestrach definicje, odpowiadając potrzebom obowiązujących w danym państwie przepisów prawnych, odbiegają od unormowanych i porównywalnych w skali międzynarodowej definicji, pojęć i klasyfikacji statystycznych. Wymusza to nieustanne prace związane z harmonizacją i poprawą spójności. Ponadto zmiany regulacji prawnych skutkują zmianami definicji pozyskiwanych danych, co może utrudniać porównywalność w czasie.

Szczególny i wart omówienia jest przypadek Austrii. Austria jest jedynym dużym państwem, które w czasie jednej dekady zdecydowało się na kompleksową zmianę metodologii spisu z w pełni tradycyjnego na całkowicie oparty na rejestrach. Decyzja rządu o zmianie metodologii spisu została podjęta w 2000 r., co zostało zapisane w odpowiednich regulacjach prawnych (Federal Law Gazette, 2015). W okresie dwóch lat utworzono Centralny Rejestr Rezydentów (CRR) i podjęto odpowiednie ustawy umożliwiające Statistics Austria, austriackiemu urzędowi statystycznemu, utworzenie rejestru ludności, który zawierał dane historyczne aktualizowane na podstawie rejestru rezydentów CRR (por. rys. 2.1). Wszystkie osoby mieszkające w Austrii otrzymały unikalne numery identyfikacyjne powiązane z głównym miejscem zamieszkania. Na podstawie odpowiednich regulacji prawnych mieszkańcy zostali zobowiązani do zameldowania/wymeldowania w ciągu trzech dni roboczych od momentu zmiany miejsca zamieszkania (Asamer i in., 2014, 2016; Berka, Humer, Moser, Lenk, Rechta i Schwerer, 2012; Cetkovic, Humer, Lenk, Moser, Schnetzer i Schwerer, 2012; Lenk i in., 2014).

Spis w Austrii przeprowadzono, wykorzystując dane ośmiu administracyjnych rejestrów<sup>69</sup> „podstawowych” i ośmiu rejestrów „porównawczych” służących celom weryfikacji (Asamer i in., 2016) (rys. 2.1). Dane z rejestrów administracyjnych

<sup>69</sup> Statistics Austria jest gestorem czterech spośród ośmiu podstawowych rejestrów: (1) Business register of enterprises and their local units (BR), (2) Housing register of buildings and dwellings (HR), (3) Register of educational attainment (EAR); (4) Register of enrolled pupils i students (PSR). Pozostałe podstawowe rejestry administracyjne to: (5) Central population register (CPR), (6) Central social security register (CSSR), (7) Tax register (TR), (8) Unemployment register (UR).

Wykorzystano też siedem następujących rejestrów „porównawczych”: (1) Child allowance register (CAR), (2) Central foreigner register (CFR) – available as of 2008, (3) Registers of public servants of the federal state and the Länder (RSP), (4) Register of car owners (COR), (5) Register of social welfare recipients (SWR), (6) Conscription Register (CR), (7) Register of alternative civilian service (ACSR).



Uwaga: Rejestry: Register of educational attainment (EAR) oraz Register of enrolled pupils & students (PSR) zostały przedstawione na schemacie łącznie

Rysunek 2.1. Źródła danych w spisie opartym na rejestrach, Austria 2011

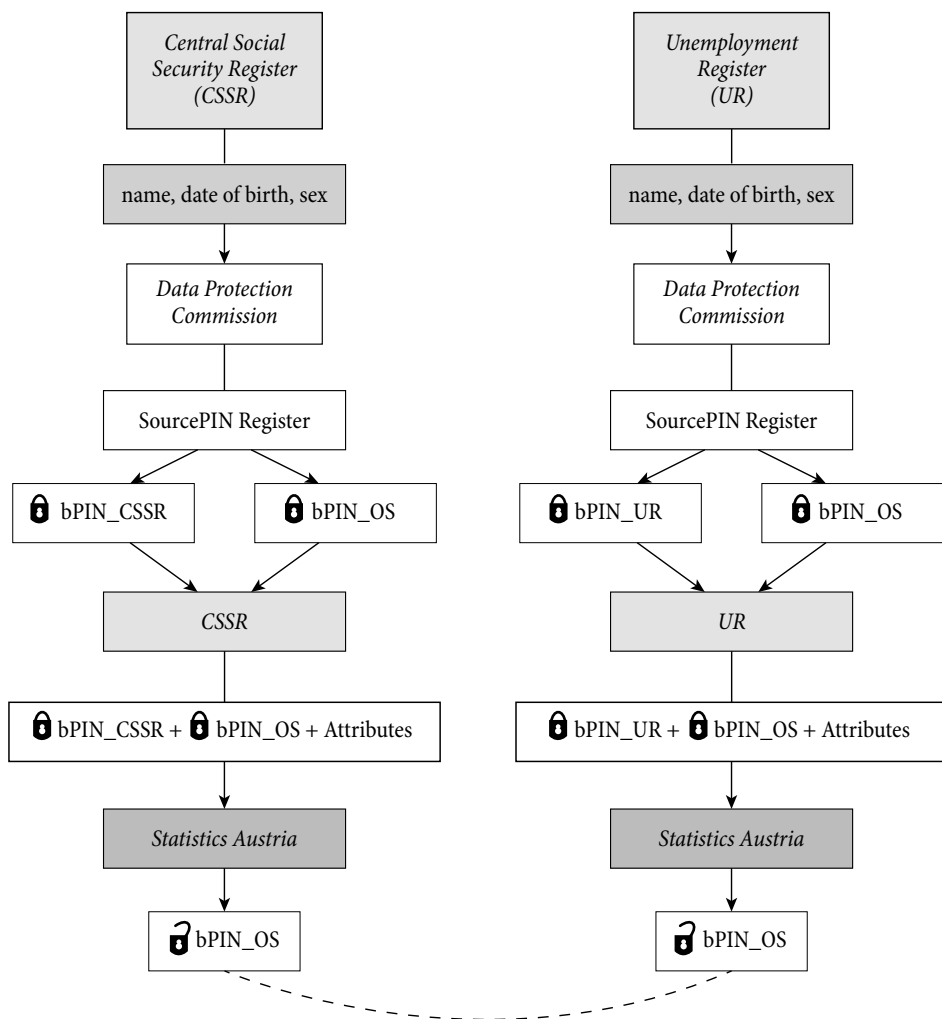
Źródło: (Asamer i in., 2016)

są gromadzone w tematycznych „kostkach danych” (data cubes) określonych na podstawie obszarów tematycznych spisu. Kostki danych umożliwiają utworzenie rejestrów statystycznych stanowiących realizację zadań w zakresie spisu ludności, mieszkań i przedsiębiorstw. Szkielet spisu stanowią rejestry umożliwiające zliczenie populacji w każdym z powyższych obszarów: ludności, mieszkań i budynków oraz przedsiębiorstw, a także zawierające podstawowe cechy charakteryzujące wyszczególnione populacje. Rejestry określone jako podstawowe są prowadzone przez Statistics Austria. W celu zgromadzenia odpowiadającego potrzebom zestawu zmiennych spisowych wykorzystano dane od ponad 50 gestorów. Organizacja spisu w Austrii wzbudza ogromne uznanie, ale kilka rozwiązań zastosowanych w ramach spisu w Austrii wartych jest szczególnego podkreślenia. Są to między innymi podejście gwarantujące poufność danych, dbałość o poprawność szacunku ludności rezydującej czy system zapewnienia jakości. Niektóre z nich zostaną nieco szerzej omówione.

### **Ochrona poufności danych osobowych**

Uregulowania prawne w Austrii dopuszczają łączenie rejestrów administracyjnych pod warunkiem zachowania pełnej poufności danych (Federal Law Gazette, 2015, s. 26). W Statistics Austria przyjęto rozwiązanie wykorzystujące dziedziczne numery identyfikacyjne (Branch-Specific Personal Identification Number for Official Statistics, bPIN OS), które tworzą pierwszy w Europie narodowy system identyfikacji dla potrzeb statystyki publicznej (Kukutai i in., 2015, s. 16). Numery te są swego rodzaju szyfrem generowanym niezależnie przez Komisję Ochrony Danych Osobowych (Data Protection Commission) dla danych rejestrów administracyjnych, z uwzględnieniem nazwiska, daty i miejsca urodzenia. Zaszifrowane, zanonimizowane dane, po usunięciu nazwisk, identyfikowane przez numer bPIN OS, są przekazywane do Statistics Austria. Numery bPIN OS są unikalne (rys. 2.2). Dane spisu są generowane w wyniku badania zgodności wewnątrz i pomiędzy trzema głównymi populacjami: ludności, budynków i mieszkań oraz przedsiębiorstw w wyniku łączenia odpowiednich rekordów. Rejestr budynków i mieszkań (poprzez numer identyfikacyjny budynku) łączy wszystkie trzy populacje i obszary tematyczne spisu. Numer identyfikacyjny b-PIN OS oraz numer identyfikacyjny Głównego Austriackiego Związku Ubezpieczycieli Społecznych (Main Association of Austrian Social Security Organisations, HV-ID) umożliwia z kolei połączenie pojedynczych osób z danymi rejestru przedsiębiorstw.

System ochrony danych, oprócz anonimowego gromadzenia danych według systemu bPIN OS, obejmuje także wyszukiwanie „ryzykownych rekordów”. Jako ryzykowne rekordy danych identyfikowane są te, które reprezentują unikalne lub rzadkie przypadki kombinacji kilku cech. Wówczas jest stosowany specjalny system kodowania uniemożliwiający identyfikację jednostek na podstawie szczególnej kombinacji zmiennych.



Rysunek 2.2. System zabezpieczenia poufności danych rejestrów w spisie opartym na rejestrach, Austria 2011

Źródło: (Lenk, 2014)

### Zasada redundancji

Podczas spisu w Austrii rejestry „porównawcze” odgrywały rolę wspierającą, pozwalającą na uzupełnienie informacji spisowych oraz na dodatkową weryfikację krzyżową. Wykorzystano tutaj tzw. zasadę redundancji – nadmiaru informacji. Zdefiniowano dwie koncepcje zmiennych: „wielokrotnych” (*multiple attributes*) i „centralnych” (*central attributes*) (Lenk, 2008). Podział na zmienne „wielokrotnie” występujące i „centralne” nie był wykluczający się. Wielokrotne występowanie

zmiennych w kilku rejestrach oznacza sytuację powtórzenia tej samej informacji, a więc jej „nadmiar”. Właśnie ów „nadmiar” informacji – redundancja – jest wykorzystywany do badania zgodności zmiennych powtarzających się w kilku rejestrach. Każdą z takich zmiennych przypisano do jednej kostki danych, obszaru tematycznego, w którym była możliwa jej ewentualna modyfikacja. Dla każdej ze zmiennych „wielokrotnych” badanie zgodności przeprowadzono, przyjmując odpowiednio zdefiniowane reguły. Do zmiennych takich należały przede wszystkim: data urodzenia, płeć, stan cywilny, kraj urodzenia, obywatelstwo, główne miejsce zamieszkania. Zastosowanie zasady redundancji przedstawiono na przykładzie zmiennej płeć. W tym przypadku dla określenia ostatecznie „zatwierdzonej” wartości zmiennej zdefiniowano cztery reguły (tab. 2.6).

Tabela 2.6. Ilustracja zasady redundancji w spisie opartym na rejestrach, Austria 2011

bPIN OS	Nazwa zmiennej i rejestru							Zatwierdzona wartość	Reguła
	SEX CPR	SEX CSSR	SEX TR	SEX UR	SEX COR	SEX SWR	SEX RSP	SEX_VALID	RULE
ID3457	1	1	1	0	0	0	1	1	1
ID3458	1	0	0	0	0	0	0	1	2
ID3459	0	2	2	0	0	2	0	2	3
ID3460	2	2	1	1	0	1	0	2	4

Uwagi:

Nazwy rejestrów wykorzystanych w przykładzie określenia płci są następujące:

CPR – Central Population Register; CSSR – Central Social Security Register, TR – Tax Register, UR – Unemployment Register, Register of car owners – COR, Register of social welfare recipients – SWR, Registers of public servants of the federal state and the Länder – RSP.

bPIN OS – statystyczny indywidualny numer identyfikacyjny.

**Reguły zastosowane w przykładzie są następujące:**

R1: Identyczna zawartość zmiennej płeć we wszystkich źródłach.

R2: Jeśli dostępna jest wartość zmiennej płeć jedynie w rejestrze CPR, to wartość ta jest akceptowana.

R3: Jeśli wartość zmiennej płeć nie jest określona w CPR, to przyjmowana jest wartość zmiennej z rejestru CSSR.

R4: Jeśli wartość zmiennej płeć w rejestrach CPR I CSSR jest identyczna, to jest ona akceptowana nawet, jeśli nie jest ona zgodna z wartością zmiennej płeć w innych rejestrach.

Źródło: (Asamer i in., 2016; Lenk, 2008).

Koncepcja zmiennych „centralnych” wskazywała te istotne zmienne, które wykorzystywano w różnych obszarach tematycznych. Podobnie jak w przypadku zmiennych wielokrotnie występujących każda ze zmiennych „centralnych” została przypisana do jednego obszaru tematycznego. Wartość ustalona dla tej zmiennej w określonej kostce danych była obowiązująca we wszystkich pozostałych obszarach. Przykładowe zmienne „centralne” to: płeć, rodzaj miejsca zamieszkania, obszar geograficzny.

### Analiza rezydencji

Jednym z najważniejszych zadań każdego spisu jest oszacowanie liczby ludności. Zgodnie z zaleceniami międzynarodowymi w Austrii przyjęto definicję ludności rezydującej, przy czym warunkiem uwzględnienia w spisie była (i) informacja o głównym miejscu zamieszkania w Austrii w CPR, (ii) zamieszkiwanie w Austrii co najmniej trzy miesiące przed spisem oraz (iii) pozytywny wynik procedury analizy rezydencji (Asamer i in., 2014; Lenk, 2014). Procedura określona jako analiza rezydencji miała na celu identyfikację osób, które według stanu na moment krytyczny spisu (31 października 2011 r.) miały status rezydenta w Austrii. Procedura ta wykorzystuje, w pewnym sensie, „nadmiar” informacji z różnych źródeł, co pozwala na weryfikację aktualności i zgodności rejestrów. Zastosowanie zasady redundancji sprowadzało się do sprawdzenia poprawności spisania każdej osoby poprzez potwierdzenie informacji z podstawowego rejestru ludności CPR w innym rejestrze (tab. 2.7). W przypadku braku informacji z co najmniej dwóch rejestrów stosowano procedurę wyjaśniającą. Walidacja na podstawie badania relacji pomiędzy rejestrami i – w konsekwencji – określania „aktywności” poszczególnych osób miała za zadanie ograniczenie liczby procedur wyjaśniających. Istota takiej procedury sprowadzała się do poszukiwania tzw. „znaku życia” (*sign of life*) według poniżej opisanego algorytmu. Zastosowanie procedury wyjaśniającej umożliwiało podjęcie decyzji o spisaniu danej osoby.

Tabela 2.7. Ilustracja analizy rezydencji w spisie opartym na rejestrach, Austria 2011

bPIN OS	Nazwa rejestru oraz fakt występowania / aktywności osoby o określonym bPIN OS							Zatwierdzona wartość
	RESID CPR	RESID CSSR	RESID TR	RESID UR	RESID COR	RESID SWR	RESID RSP	RESIDENCE_ VALID
ID3457	√	√	√	×	×	×	√	√
ID3458	√	×	×	×	√	×	×	√
ID3459	√	√	√	×	×	√	×	√
ID3460	×	×	√	×	×	×	×	×

Uwagi:

Nazwy rejestrów wykorzystanych w przykładzie określenia płci są następujące:

CPR – Central Population Register; CSSR – Central Social Security Register, TR – Tax Register, UR – Unemployment Register, Register of car owners – COR, Register of social welfare recipients – SWR, Registers of public servants of the federal state and the Länder – RSP.

bPIN OS – oznacza statystyczny indywidualny numer identyfikacyjny.

√ – występuje w rejestrze.

×

Źródło: (Asamer i in., 2016; Lenk, 2008).

Analiza rezydencji – procedura szacowania liczby ludności:

1. Osoba musi być uwzględniona w podstawowym rejestrze ludności Central Population Register (CPR) oraz w co najmniej jednym jeszcze rejestrze.

2. Osoby, które zostały zidentyfikowane tylko w rejestrze podstawowym CPR, podlegają badaniu wyjaśniającemu, które polega na wysłaniu oficjalnego listu z pytaniem o zamieszkiwanie w Austrii w momencie krytycznym spisu.
3. Na podstawie uzyskanych odpowiedzi następuje reidentyfikacja tych osób. Dla zachowania poufności procedurę prowadzi biuro ochrony danych osobowych.
4. W spisie są uwzględniane tylko te osoby, które potwierdziły zamieszkiwanie w Austrii poprzez odesłanie podpisanego formularza.
5. W przypadku braku potwierdzenia osoba „może być odrzucona”. W takim przypadku jest stosowana procedura dodatkowej weryfikacji.
6. Wszystkie rekomendacje „odrzucen” są przekazywane do urzędów lokalnych, które mają prawo zakwestionować propozycję „odrzucenia”. Dla „spisania” danej osoby jest jednak wymagane przedstawienie „dowodów” jej obecności.

Ponadto określone zostały dodatkowe procedury weryfikacji stosowane w celu uniknięcia przeszacowania liczby ludności. Są to procedury weryfikacji podwójnego spisowania oraz weryfikacji zgonów (przy wykorzystaniu rejestru ubezpieczeń społecznych – CSSR). Osoby zmarłe przed 1 listopada 2011 r., które były uwzględnione w spisie na podstawie rejestru CPR jako posiadające status rezydenta w Austrii, zostały wykreślone. Weryfikacja w tym zakresie pozwoliła wskazać 3763 osoby o takim statusie. W spisie uwzględniono także potrzebę retrospektywnej oceny aktualności rejestru, która dotyczyła m.in. opóźnienia rejestracji zdarzeń ruchu naturalnego i wędrownego ludności.

#### **Retrospektywna ocena aktualności rejestru CPR**

A. W celu uniknięcia niedoszacowania z powodu liczby ludności wynikającej z opóźnień w rejestracji zdarzeń dane CPR pobrano z uwzględnieniem sześciu miesięcy opóźnienia od momentu krytycznego. Korekta dotyczyła przede wszystkim:

- Opóźnienia zgłoszeń urodzeń w rejestrze CPR. Urodzenia są rejestrowane zazwyczaj z kilkudniowym opóźnieniem (w Polsce ustawa przewiduje 14 dni na zarejestrowanie urodzenia).
- Rejestracji w nowym miejscu zamieszkania. Przepisy przewidują dokonanie rejestracji w okresie trzech dni od momentu przeprowadzki.

Przyjęto zasadę 90 dni, zgodnie z którą osoby niezarejestrowane w momencie krytycznym spisu, ale zarejestrowane w CPR przed i po momencie krytycznym były uwzględniane w spisie, jeśli przerwa nie przekraczała 90 dni (art. 7 ust. 3 ustawy o spisie powszechnym).

B. W celu uniknięcia „turystyki” spisowej wprowadzono zasadę 180 dni (art. 7 ust. 2 ustawy o spisie powszechnym).

Zgodnie z tą zasadą osoby, które mieszkały w danej gminie (A) krócej niż 180 dni w okresie referencyjnym<sup>70</sup>, a następnie wróciły do gminy (B), z której wcześ-

---

<sup>70</sup> Okres referencyjny jest definiowany jako czas wokół momentu krytycznego, tj. 31.10.2011 r.

niej przybyły, nie były zliczone w gminie (A), w której rzeczywiście przebywały w momencie krytycznym spisu, ale w gminie (B), z której przybyły i do której powrócili. W spisie powszechnym zidentyfikowano 2925 osób (tj. 0,03% populacji), które nie zostały policzone w miejscu, w którym przebywały w momencie krytycznym spisu. Zasada ta nie wprowadzała zmian w szacunku liczby ludności, lecz jedynie zmiany dotyczącej miejsca zamieszkania.

### 2.1.5. „Spis wirtualny”

„Spis wirtualny”, jest szczególnym rodzajem spisu łączącym dane rejestrów administracyjnych oraz badań reprezentacyjnych. Doświadczenia ostatniego spisu holenderskiego przeprowadzonego w 1971 r. metodą tradycyjną dotyczyły znacznego wzrostu odsetka odmów, braku chęci uczestniczenia w badaniu oraz wysokich kosztów i wpłynęły na decyzję o rezygnacji. W kolejnych dekadach prowadzono intensywne badania metodologiczne dotyczące rozwoju rejestrów, poprawy ich jakości oraz statystycznej integracji z wynikami badań reprezentacyjnych. Dane administracyjne coraz szerzej wykorzystywano w statystyce publicznej oraz w spisach 1981 i 1991 r. Zintegrowane dane rejestrów oraz badań specjalnych pozwoliły utworzyć w Holandii Statystyczną Bazę Danych Społecznych (Social Statistical Database). Stanowi ona od 2001 podstawę holenderskiego „spisu wirtualnego”.

Gerards (2012) definiuje „spis wirtualny” jako złożony system zintegrowanych mikrobaz danych o spójnych definicjach, które zawierają szczegółowe dane demograficzne, ekonomiczne i społeczne dotyczące populacji osób, gospodarstw domowych, mieszkań i miejsc pracy<sup>71</sup>. Statystyczna Baza Danych Społecznych łączy rejestry ludności, zatrudnionych i samozatrudnionych, rejestr podatkowy, emerytur i ubezpieczenia społecznego, pomocy społecznej, rejestr budynków. Dodatkowo dla celów spisu dokonuje się integracji danych administracyjnych z wynikami badania siły roboczej przeprowadzonego w okresie około spisowym. Połączenie danych z różnych źródeł jest przeprowadzone na poziomie jednostkowym przy wykorzystaniu różnych metod statystycznej integracji danych, w tym za pomocą unikatowego klucza identyfikacyjnego. Mikrointegracja z kolei ma na celu poprawę jakości danych w połączonych źródłach poprzez wyszukiwanie i korygowanie błędów na poziomie jednostek. Różne metody ważenia i kalibracji na poziomie mikro są stosowane w celu poprawy estymacji i spójności oraz zgodności numerycznej uogólnianych wyników na różnych poziomach agregacji.

Wśród zalet „spisu wirtualnego” wskazuje się nie tylko redukcję kosztów i eliminację obciążeń społecznych, ale także szybkość publikacji wyników i porównywalność do wcześniejszych spisów holenderskich oraz spisów aktualnie przeprowadzanych w innych państwach (Nordholt, 2005; Nordholt, Hartgers i Gircour, 2004).

<sup>71</sup> W przyjazny sposób idea „spisu wirtualnego” jest prezentowana w internecie na stronie: <https://www.youtube.com/watch?v=SLpDkcyenf0>.



„Spis wirtualny” jest unikatowym badaniem w skali europejskiej i światowej, może dostarczać danych w sposób ciągły, oraz, co podkreślają autorzy, zyskał akceptację społeczeństwa holenderskiego. „Spis wirtualny” nie jest oczywiście badaniem pozbawionym wad. Można w tym miejscu wymienić identyczne właściwości ze wskazanymi w spisie opartym na rejestrach. Są to m.in. całkowita zależność od danych administracyjnych oraz ich gestorów, dla których priorytetem nie jest dbałość o spełnienie wymogów statystyki publicznej. Jedną z istotnych wad jest ograniczona dostępność zmiennych spisowych oraz opóźnienia w rejestracji.

### 2.1.6. Metody mieszane

Korzystanie ze źródeł administracyjnych staje się coraz powszechniejsze. Jednak nie wszystkie państwa dysponują rejestrem ludności, poza tym jego jakość czy zakres merytoryczny bywają niewystarczające. Sytuacja taka, w połączeniu z brakiem źródeł administracyjnych obejmujących zagadnienia społeczne i ekonomiczne, powoduje, że dane z rejestrów nie pozwalają na pozyskanie wszystkich wymaganych wyników spisu powszechnego. Jednak jeżeli rejestr ludności jest prowadzony i pozytywnie przejdzie proces ewaluacji, pojawia się możliwość wykorzystania zawartych w nim danych jako podstawy pełnego spisu powszechnego. Informacje z rejestru ludności mogą zostać uzupełnione danymi z innych istniejących rejestrów administracyjnych. W rundzie spisów około 2000 r. niektóre państwa przygotowały częściowo wypełnione formularze zawierające adres zamieszkania i nazwiska mieszkańców uzyskane z rejestru ludności, a także z innych źródeł administracyjnych. Dane te zostały zweryfikowane i poprawione podczas wywiadów. Inne wymagane dane, które nie były dostępne ze źródeł administracyjnych, zostały uzupełniane w czasie spisu lub w badaniu reprezentacyjnym<sup>72</sup>. W 2011 r. także Polska dołączyła do grona państw przeprowadzających spis przy zastosowaniu metody mieszanej. Rachmistrze spisowi odwiedzający<sup>73</sup> wylosowane do badania reprezentacyjnego mieszkania dysponowali częściowo wypełnionymi tzw. długimi formularzami dla osób w nich zamieszkujących. Również osoby, które wypełniały zeznanie spisowe przez internet<sup>74</sup>, po zalogowaniu się otrzymały już częściowo wypełniony krótki lub długi formularz, w zależności od tego, czy obejmował je spis pełen, czy znalazły się w próbie badania reprezentacyjnego.

<sup>72</sup> Strona internetowa ONZ Population and housing censuses: <https://unstats.un.org/unsd/demographic/sources/census/alternativeCensusDesigns.htm#classicregister>.

<sup>73</sup> W spisie 2011 r., obok danych pobranych z systemów informacyjnych, pozyskanych w samospisie internetowym, były one gromadzone w formie wywiadu telefonicznego wspomaganego programem komputerowym (CATI) oraz w formie wywiadu rejestrowanego na przenośnym urządzeniu elektronicznym prowadzonym przez rachmistrza spisowego (CAPI).

<sup>74</sup> Formularze internetowe dostępne były przez 14 dni od daty pierwszego zalogowania w specjalnej aplikacji, do której dostęp wymagał podania numeru Pesel oraz hasła, które należało zdefiniować, wypełniając formularz identyfikacyjny (<https://nsp2011.spis.gov.pl/nsp.form/FormularzID.aspx/>, obecnie strona ta nie jest aktywna).

Metoda mieszana może być realizowana w wielu wariantach (tab. 2.1–2.3). W klasycznym ujęciu będzie to połączenie danych rejestru administracyjnego ze spisem tradycyjnym. Jednak w praktyce takie rozwiązanie nie znalazło zastosowania. Na stronach internetowych ONZ znajdują się informacje o planowanym takim przedsięwzięciu w Hiszpanii<sup>75</sup>, jednak ostatecznie spis przeprowadzony w 2011 r. bazował na danych rejestru ludności i badania reprezentacyjnego (INEE, 2011). W zależności od tego, które ze źródeł zostanie uznane za podstawowe, możemy mówić o spisach bazujących na rejestrach (zintegrowanych źródłach administracyjnych lub tylko podstawowych rejestrach) i regularnych badaniach reprezentacyjnych bądź specjalnych badaniach reprezentacyjnych organizowanych w ramach spisu. Badania takie obejmują, podobnie jak spis, szeroki zakres merytoryczny zagadnień społecznych i ekonomicznych. Zazwyczaj są one przeprowadzane na tzw. dużej próbie, znacznie większej aniżeli standardowo w regularnych badaniach, takich jak BBGD czy BAEL. Badanie reprezentacyjne w NSP 2011 obejmowało próbę liczącą 20%, a w Niemczech 10%.

Wcześniejsze wypełnienie kwestionariuszy spisowych danymi z rejestrów oznacza wiele korzyści. Przede wszystkim skraca czas przeprowadzania wywiadu, przez co zmniejsza koszty badania oraz obniża obciążenie respondentów, a to z kolei wpływa na poprawę jakości danych. W przypadku zastosowania metody mieszanej wykorzystującej dane z rejestru oraz spis klasyczny dysponujemy w zasadzie dwoma źródłami informacji z badania pełnego: pochodzącymi z rejestru oraz pozyskanymi bezpośrednio od respondentów. Sytuacja ta stwarza doskonałą okazję do weryfikacji kompletności spisu oraz jakości źródeł administracyjnych. Porównanie danych rejestru z informacjami od respondentów umożliwia poznanie nieścisłości i opóźnień w rejestracji zdarzeń, ocenę ich kompletności. Dzięki temu możliwa jest bezpośrednia aktualizacja rejestrów, ale także wprowadzenie procedur poprawiających ich jakość w przyszłości. W ten sposób możliwe jest pozyskanie dla celów statystycznych źródeł administracyjnych oraz regularne publikowanie zawartych w nich danych, nie tylko w okresie spisu. Podejście takie stwarza perspektywę coraz szerszego wykorzystania rejestrów w przyszłości. Jest to szczególnie ważne w odniesieniu do rejestru ludności, który mógłby być systematycznie wykorzystywany przy sporządzaniu bilansu ludności.

Spis łączący elementy podejścia tradycyjnego oraz opartego na rejestrach miał mieć zalety obu podejść i coraz mniej wad ujęcia tradycyjnego, takich jak np. wysokie koszty i duże obciążenie respondentów czy czas przetwarzania danych. Przykładowo, według pierwszych, niezrealizowanych planów, spis 2011 r. w Hiszpanii miał wykorzystywać rejestry oraz dane z badania wyczerpującego, przy czym zwracano uwagę na dwie zasadnicze różnice w stosunku do spisu opartego

---

<sup>75</sup> W 2005 r. Hiszpania deklarowała zastosowanie ujęcia klasycznego oraz wykorzystanie rejestru ludności, jednak już wstępne plany dotyczące spisu zakładały również realizację badań reprezentacyjnych (INEE, 2005).

na rejestrach i badaniu reprezentacyjnym (INEE, 2005). Po pierwsze, dane niedostępne z rejestrów miały być pozyskiwane, podobnie jak w spisie tradycyjnym, z badania pełnego. Po drugie, szacunki liczby ludności wynikające z rejestrów nie miały być automatycznie akceptowane, lecz miały być weryfikowane i korygowane w badaniu pełnym. Ujęcie to miało na celu ewaluację rejestrów pod względem kompletności, by w przyszłości spis mógł całkowicie bazować na danych administracyjnych. Jednak albo ewaluacja rejestrów wypadła pozytywnie, albo koszty dostosowania rejestru do potrzeb spisu i badania tradycyjnego uznano za zbyt wysokie, gdyż ostatecznie wykorzystano dane rejestrów i badania reprezentacyjnego. Zalety podwójnego badania pełnego, jak zmniejszenie obciążeń ludności przez częściowe wypełnienie kwestionariuszy czy czasu trwania wywiadu, okazały się niewspółmierne do oczekiwań.

Jednym z wariantów metody mieszanej jest spis oparty na danych rejestrów, które są uzupełniane wynikami badania reprezentacyjnego. Rozwiązanie takie jest stosowane w sytuacji, gdy rejestry nie zawierają wszystkich wymaganych danych spisowych. W podejściu takim dane ze źródeł administracyjnych są łączone na poziomie jednostkowym z danymi badania reprezentacyjnego (tworzony jest tzw. złoty rekord). Wówczas uogólnianie danych może wymagać zastosowania odpowiednich metod ekstrapolacji lub systemu ważenia. Podejście takie wydaje się zbliżone do zastosowania krótkiego formularza i długiego formularza spisowego. Wówczas informacje dotyczące cech podstawowych są określane na podstawie badania pełnego. Natomiast dla tematów uzupełniających szacunki są przeprowadzane na podstawie próby losowej.

## 2.2. NSP 2011 – spis ludności metodą mieszaną

Wiele państw zdecydowało się na wdrożenie nowej metodologii wykorzystującej dane rejestrów administracyjnych oraz innych alternatywnych źródeł informacji. Polska była jednym z krajów, które wybrały metodę mieszaną. Dane dla potrzeb spisu 2011 r. były pobierane ze źródeł administracyjnych i pozadministracyjnych, samospisu internetowego oraz z badania reprezentacyjnego. Podstawę prawną przeprowadzenia spisu stanowiła Ustawa z dnia 4 marca 2010 r. o narodowym spisie powszechnym ludności i mieszkań w 2011 r. (Ustawa, 2010) wraz z aktami wykonawczymi oraz rozporządzenie (WE) Parlamentu Europejskiego i Rady Nr 763/2008 z dnia 9 lipca 2008 r. w sprawie spisów powszechnych ludności i mieszkań (UE, 2008).

**Spis przeprowadzono w formie badania pełnego, obejmującego całą populację, oraz badania reprezentacyjnego, które miało objąć 20-procentową próbę.** W badaniu pełnym posłużono się tzw. krótkim formularzem wypełnionym danymi rejestrów administracyjnych, które zostały uzupełnione informacjami z samospisu

internetowego. Natomiast w badaniu reprezentacyjnym (tzw. długi formularz) zastosowano nowoczesne, wspomagane komputerowo technologie gromadzenia danych, bez formularza papierowego.

Przeprowadzony w 2011 r. spis obejmował kategorię ludności faktycznie zamieszkałej i rezydującej<sup>76</sup>. Jednak w zasadzie wszystkie publikacje spisowe dotyczą ludności zdefiniowanej jako faktyczna. Przy ustalaniu zakresu tematycznego spisu uwzględniono analizę wykorzystania danych poprzedniego spisu oraz nowych potrzeb informacyjnych, a także wyniki konsultacji społecznych przeprowadzonych w 2008 r. Brano pod uwagę również możliwości pozyskania tych informacji z innych źródeł, poza tradycyjnym wywiadem, konieczność zachowania pełnej porównywalności w czasie oraz zobowiązania międzynarodowe określone w rozporządzeniu nr 763/2008. Ostatecznie ustalono 12 następujących obszarów badawczych (GUS, 2013a):

1. Geograficzna charakterystyka ludności – miejsce przebywania, miejsce zamieszkania w okresie międzypisowym i przyczyny jego zmiany.
2. Demograficzna charakterystyka osób – płeć, wiek, stan cywilny (formalno-prawny i faktyczny), kraj urodzenia, w tym rodziców, obywatelstwo.
3. Gospodarstwa domowe i rodziny: wielkość i skład gospodarstwa domowego i rodziny, rodziny biologiczne i zrekonstruowane, rodziny niepełne, pozycja osób w gospodarstwie domowym i rodzinie.
4. Wykształcenie – poziom wykształcenia, kontynuacja nauki; rodzaj szkoły, dziedzina i kierunek kształcenia.
5. Migracje wewnętrzne i zagraniczne, w tym badanie emigracji Polaków, emigracji zarobkowej, reemigracji oraz imigracji cudzoziemców do Polski.
6. Dzietność kobiet.
7. Narodowość i język.
8. Wyznanie (przynależność do kościoła lub związku wyznaniowego).
9. Niepełnosprawność prawna i biologiczna.
10. Ekonomiczna charakterystyka osób, w tym:
  - 10.1 Bieżąca aktywność ekonomiczna: pracujący w pracy głównej i dodatkowej, bezrobotni, bierni zawodowo, charakterystyka zawodowa pracujących.
  - 10.2 Stała aktywność ekonomiczna dla pracujących w indywidualnych gospodarstwach rolnych.
  - 10.3 Dojazdy do pracy.
11. Główne i dodatkowe źródło utrzymania osób.
12. Źródła utrzymania gospodarstwa domowego, samodzielność gospodarowania i zamieszkania.

Odpowiedzi w niektórych z wyżej wyszczególnionych obszarów były gromadzone na zasadzie dobrowolności. Były to m.in. kwestie dotyczące pozostawiania

---

<sup>76</sup> Szczegółowa dyskusja definicji kategorii ludności została przedstawiona w punkcie 1.2.3 niniejszego opracowania.

w związku nieformalnym, wyznania – przynależności do kościoła lub związku wyznaniowego, liczby urodzonych dzieci oraz planów prokreacyjnych. Powszechny spis ludności obejmował także tematykę mieszkaniową, w której uwzględniono cztery grupy zagadnień: rodzaj zamieszkaných pomieszczeń; charakterystykę mieszkań, charakterystykę budynków, w tym formę własności, oraz informacje o tytule prawnym do zajmowanego mieszkania.

W spisie 2011 r. realizowanym w Polsce metodą mieszaną wykorzystano dane rejestrów administracyjnych oraz systemów pozaadministracyjnych. Zebrano także dane w sposób bezpośredni od ludności w ramach badania reprezentacyjnego oraz tzw. badania pełnego. Wykorzystano tutaj następujące kanały pozyskiwania danych: CAII, CATI, CAPI. Ponadto zostały przeprowadzone dwa pełne badania, obejmujące osoby przebywające w obiektach zbiorowego zakwaterowania oraz osoby bezdomne. W wyniku szczegółowych badań zasobów informacyjnych około 300 rejestrów i systemów, dane do NSP 2011 pozyskano z:

1. Systemów i rejestrów administracyjnych:

– centralnych:

- Ministerstwa Finansów,
- Ministerstwa Spraw Wewnętrznych i Administracji (zbiory PESEL i Centralnej Ewidencji Pojazdów CEP)
- Ministerstwa Sprawiedliwości,
- Urzędu do Spraw Cudzoziemców,
- Zakładu Ubezpieczeń Społecznych,
- Kasy Rolniczego Ubezpieczenia Społecznego,
- Agencji Restrukturyzacji i Modernizacji Rolnictwa,
- Narodowego Funduszu Zdrowia (Centralnego Wykazu Ubezpieczonych),
- Głównego Urzędu Geodezji i Kartografii (Państwowego rejestru granic oraz powierzchni jednostek podziału terytorialnego kraju i ortofotomapy dla obszaru Polski),
- Państwowego Funduszu Rehabilitacji Osób Niepełnosprawnych,
- Inspekcji Jakości Handlowej Artykułów Rolno-Spożywczych;

– rozproszonych:

- urzędów marszałkowskich (Bazy Danych Obiektów Topograficznych),
- starostw powiatowych (ewidencji gruntów i budynków i systemów informacyjnych Powiatowych Urzędów Pracy – PUP),
- powiatowych/miejskich zespołów ds. orzekania o niepełnosprawności,
- urzędów gmin/miast (ewidencji podatkowej nieruchomości EPN i gminnych zbiorów meldunkowych).

2. Systemów pozaadministracyjnych

- zarządców i administratorów zasobów mieszkaniowych;
- przedsiębiorców wykonujących działalność gospodarczą w zakresie sprzedaży energii elektrycznej (systemów informacyjnych dotyczących odbior-

ców energii elektrycznej wykorzystujących energię elektryczną na potrzeby mieszkaniowe);

- dostawców publicznie dostępnych usług telekomunikacyjnych (systemów informacyjnych dotyczących abonentów).

Dużym przedsięwzięciem organizacyjnym i technologicznym było zebranie ponad 10 000 zbiorów danych z systemów informacyjnych administracji publicznej prowadzonych przez urzędy gmin i starostwa powiatowe. W tym celu przygotowano specjalną platformę informatyczną służącą do pozyskiwania danych z tych systemów za pomocą aplikacji internetowej umożliwiającej transfer zbiorów danych do serwera zasobowego GUS w bezpiecznym kanale szyfrowanym.

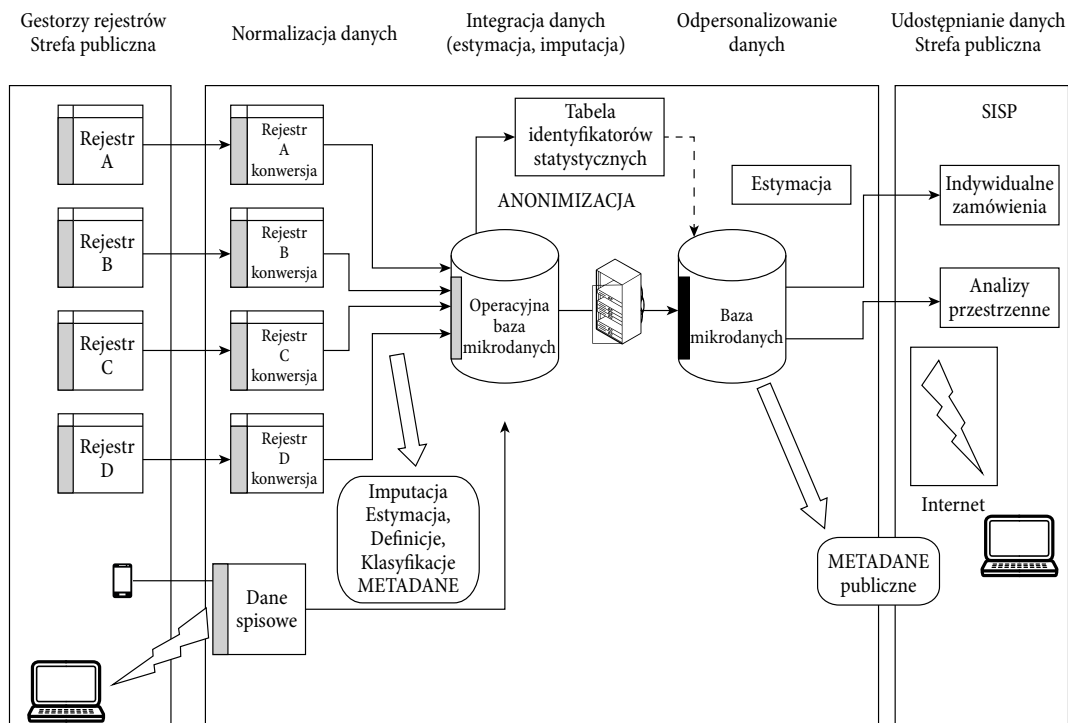
**Badanie pełne** przeprowadzono, zgodnie z ustawą o NSP 2011, poprzez jak najszersze wykorzystanie zasobów rejestrów administracyjnych oraz samospis internetowy. Oznaczało to, że w pierwszej kolejności formularz spisowy został zasilony informacjami z dostępnych źródeł administracyjnych<sup>77</sup>. Ankieta w badaniu pełnym zawierała 16 pytań tworzących tzw. „formularz krótki”. Każdy respondent miał możliwość weryfikacji oraz uzupełnienia danych wygenerowanych z rejestrów. Z tej możliwości skorzystało ok. 10% ludności. Dla osób, które nie spisały się przez internet, przyjęto formularz wypełniony na bazie informacji dostępnych w źródłach administracyjnych<sup>78</sup>. Korzystanie z danych administracyjnych było związane z koniecznością wdrożenia odpowiednich procedur normalizacji, synchronizacji oraz walidacji i weryfikacji ze specjalnie wyodrębnionymi danymi referencyjnymi oraz szczególną ochroną danych osobowych (rys. 2.3).

**Badanie reprezentacyjne** w NSP 2011 planowano jako obejmujące bardzo dużą, 20-procentową próbę losową (ponad 8 mln osób mieszkających lub przebywających w ok. 2744 tys. mieszkań spośród prawie 13,5 mln mieszkań znajdujących się a w operacie losowania) (GUS, 2012c, s. 3, 2013a). Losowanie próby przeprowadzono według schematu warstwowego. Podział na warstwy był bardzo głęboki i zróżnicowany w zależności od typu powiatu i gminy oraz obszaru miejskiego i wiejskiego. Ostatecznie mieszkania były losowane z każdej z prawie 70,5 tys. warstw, zaś wielkość próby w poszczególnych warstwach wahała się od niemal 6% do ponad 49%. Zakres tematyczny badania reprezentacyjnego w NSP 2011<sup>79</sup> uwzględniał szeroki zakres tematyczny, tzw. długi formularz zawierał ponad 120 pytań. Respondenci odpowiadali przeciętnie na 70–80 pytań, w zależności od

<sup>77</sup> Uzyskane dane zostały użyte do wygenerowania podpowiedzi w ankietach spisowych i przyspieszenia w ten sposób gromadzenia informacji od respondentów.

<sup>78</sup> Wyjątkiem były tutaj osoby podlegające badaniu reprezentacyjnemu, z którymi przeprowadzono wywiad metodą CATI bądź CAPI, oraz osoby, które podlegały jednemu z dwóch dodatkowych badań pełnych przeprowadzonych w ramach NSP 2011. Badaniami tymi objęto zbiorowości osób w obiektach zbiorowego zamieszkania oraz bezdomnych.

<sup>79</sup> Zakres obszarów tematycznych w spisie reprezentacyjnym, tzw. długim formularzu, został przedstawiony w aneksie (aneks II.2).



Objaśnienie :

Szare paski na elementach składowych procesu pobierania, normalizacji i integracji danych oznaczają dane osobowe, identyfikacyjne. Po utworzeniu operacyjnej bazy mikrodanych dane osobowe miały zostać zastąpione, przy użyciu kluczy przejścia, identyfikatorami statystycznymi (anonimizacja). Baza mikrodanych – z czarnym paskiem – używana do dalszego przetwarzania wyników, jest bazą w której niemożliwe jest zidentyfikowanie poszczególnych podmiotów bez znajomości kluczy przejścia. Operacyjna baza mikrodanych wraz z tabelą identyfikatorów statystycznych (zawierającą klucze przejścia) miała zostać zniszczona natychmiast po opracowaniu i zatwierdzeniu wyników spisu na poziomie krajowym oraz ich zatwierdzeniu przez Eurostat, nie później niż w dwa lata po spisie. Do tego czasu operacyjna baza mikrodanych i tabela identyfikatorów statystycznych miała podlegać szczególnej ochronie.

Czarny pasek oznacza identyfikatory statystyczne uniemożliwiające powiązanie danych z konkretną osobą.

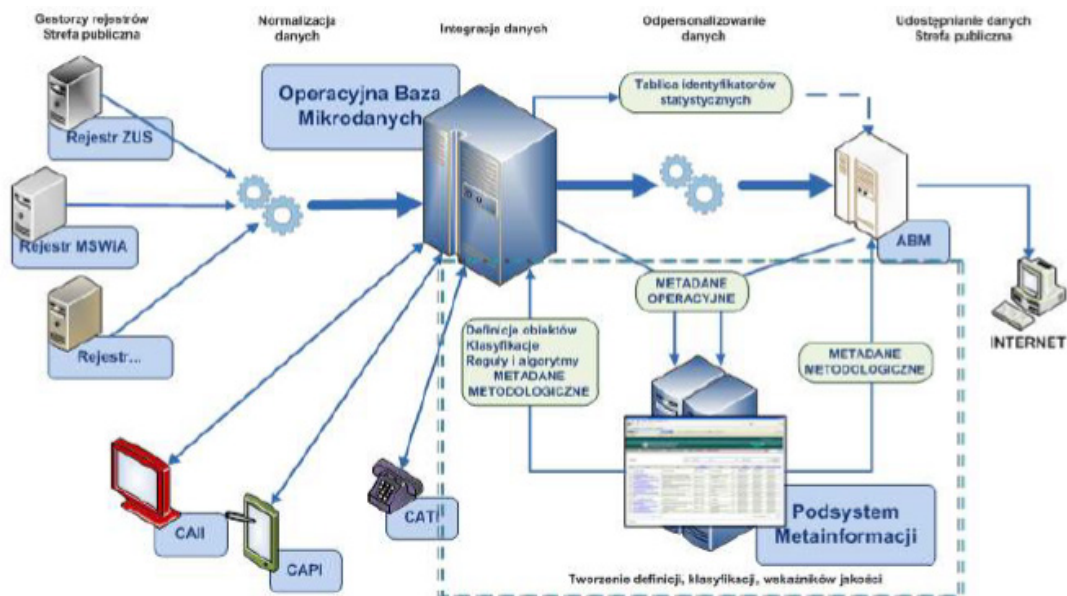
Rysunek 2.3. Schemat przetwarzania danych spisowych, NSP 2011

Źródło: (Dygaszewicz, 2009)

płci, wieku, mobilności i aktywności zawodowej. Badanie przeprowadzono głównie<sup>80</sup> metodą wywiadu bezpośredniego (CAPI) z wykorzystaniem elektronicznej wersji kwestionariusza spisowego, a takie urządzenia, jak palmtopy, hand-heldy i inne, posłużyły do bezpiecznego i sprawnego przesyłania pozyskanych informacji. Zastosowane rozwiązania miały przede wszystkim zmniejszyć koszty spisu oraz obciążenie osób objętych spisem, przy jednoczesnym zachowaniu dobrej jakości wyników spisu.

<sup>80</sup> Z możliwości samospisu przez internet skorzystało ok. 2% osób.

Dane pozyskiwane w trakcie realizacji spisu były umieszczane w Operacyjnej Bazie Mikrodanych (OBM). W zakres OBM wchodziło 11 podsystemów wykonujących wiele zadań<sup>81</sup>. Po weryfikacji poprawności danych w OBM były generowane dane odpersonalizowane (Złoty Rekord), które następnie były przekazywane do Analitycznej Bazy Mikrodanych (ABM, rys. 2.4). Baza ta stanowi zbiór odpersonalizowanych danych spisowych w ostatecznej wersji, które są podstawą analiz oraz przygotowania publikacji spisowych (GUS, 2014b, s. 141 i nast.).



Rysunek 2.4. Informatyczny system spisowy z wyróżnieniem OBM oraz ABM wraz z podsystemem Metainformacji

Źródło: (GUS, 2014b, s. 159)

Podobnie jak w przypadku OBM, również Analityczna Baza Mikrodanych składa się z wielu podsystemów umożliwiających realizację określonych zadań. Są to m.in. podsystemy: analiz informacji, metadanych systemu ABM, monitorowania jakości, przetwarzania danych, udostępniania danych, zarządzania bezpieczeństwem informacji, zarządzania i administracji, zarządzania potrzebami i produktami. Wyróżniony na rys. 2.3 podsystem metainformacji, składający się z trzech

<sup>81</sup> W skład Operacyjnej Bazy Mikrodanych wchodziły następujące podsystemy, których nazwy odzwierciedlają funkcje OBM: podsystem czyszczenia i importu danych z rejestrów, obsługi CAxI, korekty danych, integracji i weryfikacji danych, obsługi operatu spisowego, ładowania wykazu gospodarstw dla PSR 2010, ładowania wykazu adresowo-mieszkaniowego dla NSP 2011, administracji, zarządzania i konfiguracji, eksportu do ABM, raportów operacyjnych, przetwarzania danych.



modułów; metadanych metodologicznych, technicznych i operacyjnych, pozwala określić m.in., jakie dane znajdują się zasobach ABM, jakie jest ich źródło, jakie są ich definicje, klasyfikacje, w jaki sposób zostały przetworzone, gdzie się znajdują.

Proces przetwarzania danych w ABM miał na celu w szczególności realizację takich zadań, jak integracja danych, walidacja, imputacja, tworzenie danych pochodnych. Ważnym etapem był dwuetapowy proces uogólniania wyników badania reprezentacyjnego. Etap I sprowadzał się do ważenia przy zastosowaniu wag wynikających z zastosowanego schematu losowania dla jednostek badania, jakimi były mieszkania. Uwzględniono tutaj korektę wynikającą z braków odpowiedzi (13,7% odmów<sup>82</sup>). Na etapie II przeprowadzono ważenie przy zastosowaniu skalirowanych wag dla osób. Zastosowana w spisie kalibracja miała na celu dostosowanie struktur ludności według płci i wieku uogólnionych na podstawie badania reprezentacyjnego na poziomie powiatów do analogicznych struktur ludności ustalonych w spisie pełnym<sup>83</sup>.

Analityczna Baza Mikrodanych została zintegrowana z aplikacją Census Hub opracowaną przez Eurostat.

### 2.3. Alternatywne metody przeprowadzania spisów w rundzie 2010

Opisując alternatywne metody przeprowadzania spisów, odwołujemy się zazwyczaj do wykorzystania dodatkowych źródeł informacji. W ostatniej dekadzie są to przede wszystkim dane administracyjne, a w szczególności rejestr ludności. To właśnie istnienie centralnego, elektronicznego rejestru ludności obejmującego całą populację umożliwiło transformację klasycznej metodologii spisów ludności do metod alternatywnych. Jak zauważają Poulain i Herm (2013), historia rejestrów ludności jest mało znana, choć sięga starożytności<sup>84</sup>. Opracowanie Poulain i Herm należy do nielicznych przedstawiających historię rejestrów, ale przede wszystkim autorzy analizują ich zawartość, zasady funkcjonowania i sposób aktualizacji, współpracę z instytucjami statystyki publicznej pod kątem możliwości wykorzystania rejestrów do pogłębionej analizy demograficznej, szczególnie w zakresie migracji, oraz tworzenia systemu statystyki ludnościowej.

Warto podkreślić być może oczywistą kwestię, że w rejestrze ludności indywidualną jednostką jest człowiek (każdej osobie odpowiada jeden rekord). Rejestry

---

<sup>82</sup> W raporcie (GUS, 2012a, s. 31) GUS podaje informację o konieczności korekty pierwotnych wag z powodu 13,7% braków wypełnionych ankiet mieszkaniowych w badaniu reprezentacyjnym.

<sup>83</sup> W procesie kalibracji przyjęto, że populację referencyjną stanowiła ludność według kryterium faktycznego miejsca zamieszkania określona na podstawie badania pełnego.

<sup>84</sup> Poulain i Herm (2013) przytaczają tu przykład rejestru osób i gospodarstw domowych prowadzony za czasów dynastii Han w Chinach w II w. p.n.e. Ciekawym jest fakt, że większość rejestrów ludności powstała na bazie danych spisów ludności, począwszy od Belgii w 1846 r. po Austrię w 2001 r.

powstają w trybie ciągłym, co oznacza, że każde zaistniałe zdarzenie na bieżąco podlega rejestracji, zatem z natury rzeczy zasadniczą charakterystyką rejestrów jest ich aktualność<sup>85</sup>. Aktualizacja rejestru oznacza uwzględnienie wszystkich administracyjnych informacji dotyczących każdej osoby (tworzenie i likwidację rekordów, wprowadzanie danych, uaktualnienie, korektę, uzupełnienie). Rejestr ludności dostarcza więc aktualnych danych, które można agregować w dowolnych przekrojach terytorialnych w określonym momencie czasu, przedstawiać szacunki struktur badanej populacji według cech demograficznych i społecznych zawartych w rejestrze. Umożliwia to monitorowanie sytuacji ludnościowej, obserwowanie zmian, planowanie polityki społecznej.

Na podstawie danych rejestrów możliwe jest, przynajmniej potencjalnie, przeprowadzenie szczegółowej analizy demograficznej. Wiąże się to z faktem, że obok informacji o zdarzeniach demograficznych (wchodzących w skład ruchu naturalnego i wędrownego ludności) dotyczących indywidualnych jednostek w określonym przekroju terytorialnym, według czasu kalendarzowego jest jeszcze możliwość badania zdarzeń według wieku (czasu własnego jednostek) oraz określania historii zdarzeń, uwzględniania ich kolejności i odstępów między nimi, a także przeprowadzenia analizy kohortowej. Zazwyczaj są to informacje o urodzeniach, małżeństwach, zgonach i migracjach.

Jak wynika z zestawienia Poulain i Herm (2013) na rok 2010, spośród objętych badaniem 30 państw europejskich w czterech nie funkcjonował żaden rejestr ludności. Są to Wielka Brytania, Portugalia, Irlandia i Francja (tab. 2.8). W kolejnych trzech państwach (Cypr, Grecja i Malta), choć nie było rejestru jako takiego, na poziomie lokalnym funkcjonowały bazy danych zawierające dane dotyczące pojedynczych osób. Najwcześniej powstał rejestr ludności w Islandii (1953 r.), około 10 lat później powstały rejestry w państwach skandynawskich. W Polsce rejestr ludności powstał w 1979 r. Zaledwie w dwóch państwach, w Islandii i w Hiszpanii, rejestry ludności są prowadzone przez instytucje statystyki publicznej. W większości przypadków gestorem rejestru jest ministerstwo spraw wewnętrznych.

Większość państw prowadzących rejestr ludności wykorzystuje dane w nim zawarte w ramach spisu ludności. Stopień tego wykorzystania, w rundzie spisów 2010 r. był bardzo różny. W Austrii, Danii, Finlandii, Islandii, Norwegii i Szwecji spis był w całości oparty na danych rejestru. W innych państwach rejestr ludności był jednym ze źródeł wykorzystanych w spisie. Zaledwie w pięciu państwach (Bułgaria, Węgry, Luksemburg, Rumunia i Słowacja) nie powiązано danych rejestru z przeprowadzonym spisem (Poulain i Herm, 2013, s. 197). Sytuacja ta wskazuje na tendencję rozszerzania wykorzystania rejestrów ludności w spisie. Powszechna

---

<sup>85</sup> Wnikliwa ocena rejestru PESEL pod względem możliwości wykorzystania w spisie ludności wykazała około dwutygodniowe opóźnienia w rejestracji urodzeń (Józefowski i Rynarzewska-Pietrzak, 2010).

komputeryzacja oraz rozwój nowoczesnych technologii komunikacyjnych mogą sprzyjać centralizacji rejestrów oraz ich dalszemu wykorzystaniu dla celów statystycznych. Więcej informacji o wykorzystaniu danych rejestrów ludności przez statystykę publiczną zamieszczono w aneksie (por. aneks II.3).

Tabela 2.8. Nazwy centralnych rejestrów ludności i instytucji je prowadzących w trzydziestu państwach europejskich

Państwo		Nazwa rejestru	Typ rejestru	Gestor rejestru	Rok
Austria	AT	Zentrales Melderegister (ZMR)	Centralny Rejestr Ludności	Federal Interior Ministry	2002
Belgia	BE	Registre national des personnes physiques	Centralny Rejestr Ludności	Interior and Equal Opportunity Ministry	1985
Bułgaria	BG	Naselenie Esgraon	Centralny Rejestr Ludności	Regional Development Ministry	1977
Cypr	CY	Archeio Plithismou	Lokalny Rejestr Cywilny	Interior Ministry	–
Czechy	CZ	Centrální registr obyvatelstva	Centralny Rejestr Ludności	Interior Ministry	1980
Dania	DK	Det Centrale Personregister	System Informacji o Ludności	Interior Affairs Ministry	1968
Estonia	EE	Rahvastikuregister	Centralny System Informacji o Ludności	Interior Ministry	2002
Finlandia	FI	Väestötietojärjestelmä	Centralny System Informacji o Ludności	Interior Ministry	1970
Francja	FR	–	–	–	–
Grecja	GR	Dimotologio	Lokalny Rejestr Cywilny	Interior Ministry	–
Hiszpania	ES	Fichero de Coordinación de los Padrones municipales (Padrón Continuo)	Centralny Rejestr Ludności	National Statistical Institute (INE)	1996
Holandia	NL	Gemeentelijke basisadministratie persoonsgegevens	Lokalny Rejestr Ludności (online)	Municipalities and Interior and Kingdom Relations Ministry	1994
Irlandia	IE	–	–	–	–
Islandia	IS	Thjóðskrá	Centralny Rejestr Ludności	Statistics Iceland	1953
Litwa	LT	Gyventojų registro tarnyba prie Vidaus reikalų ministerijos	Centralny Rejestr Ludności	Interior Ministry	1992
Luksemburg	LU	Répertoire général des personnes physiques	Centralny Rejestr Ludności	State Information Technology Centre (CTIE)	1979
Łotwa	LV	Latvijas Republikas Iedzīvotāju reģistrs	Centralny Rejestr Ludności	Interior Ministry	1992
Malta	MT	Public Register	Odrębne Rejestry Cywilne na Malcie i Gozo	Home Affairs and National Security Ministry	–

Państwo		Nazwa rejestru	Typ rejestru	Gestor rejestru	Rok
Niemcy	DE	Melderegister	Lokalny / Centralny Rejestr Ludności (w zależności od Landu), Centralny Rejestr Obcokrajowców	Municipalities Länder Interior Ministries	–
Norwegia	NO	Det sentrale folkeregister	Centralny System Informacji o Ludności	Tax Administration of Ministry of Finance	1964
Polska	PL	CBD PESEL (Powszechny Elektroniczny System Ewidencji Ludności)	Centralny Rejestr Ludności	Voivodships and Interior Ministry	1979
Portugalia	PT	–	–	–	–
Rumunia	RO	Registrul permanent de evidenta populatiei	Centralny Rejestr Ludności	Interior Ministry	1996
Słowacja	SK	Register obyvateľov Slovenskej republiky	Centralny Rejestr Ludności	Interior Ministry	1980
Słowenia	SI	Centralni register prebivalstva	Centralny Rejestr Ludności	Interior and Public Administration Ministry	1971
Szwajcaria	CH	Registre des habitants, Zentrales Ausländerregister/ registre central des étrangers/ Registro centrale degli stranieri	Lokalny Rejestr Ludności i Centralny Rejestr Obcokrajowców	Municipalities and Federal Department of Justice and Police	–
Szwecja	SE	Registret över totalbefolkningen	Centralny System Informacji o Ludności	Tax Agency of Ministry of Finance	1968
Węgry	HU	Népszámlálás	Centralny Rejestr Ludności	Interior Ministry	1975
Włochy	IT	Anagrafe della popolazione residente	Lokalny Rejestr Ludności (w trakcie centralizacji)	Interior Ministry	–
Zjednoczone Królestwo	UK	–	–	–	–

Uwaga:

PR – Population Register – Rejestr Ludności.

CR – Civil Register – Rejestr Cywilny.

PIS – Population Information System – System Informacji o Ludności.

Źródło: (Poulain i Herm, 2013, s. 190–191).

Wykorzystanie danych rejestrów administracyjnych w spisie ludności odnotowano w Finlandii i Norwegii w 1970 r. W ostatniej rundzie spisów 2010 r. zaobserwowano ogromną zmianę (Gołata, 2014). Spośród 55 badanych przez ONZ państw (UNECE, 2012) większość (31 państw, tj. 56%) zastosowała tradycyjną metodę przeprowadzania spisu (tab. 2.9). Jednak w poprzedniej rundzie spisów 2000 r. 73% państw wybrało tradycyjną formę. Zatem odsetek ten zmniejszył się o 17 punktów procentowych. Ponad 25% państw, które w 2000 r. stosowały podejście tradycyjne, w 2010 r. wprowadziło zmiany metodologii, decydując się na wykorzystanie źródeł administracyjnych: w ośmiu państwach w połączeniu

z badaniami reprezentacyjnymi, w Austrii wyłącznie, a we Francji i w Stanach Zjednoczonych stosując inne podejścia. Ponadto żadne z państw, które w poprzednich spisach korzystały z zasobów administracyjnych, nie zrezygnowało z dotychczasowego podejścia. Wręcz przeciwnie, obserwowano raczej ograniczenie się wyłącznie do danych rejestrów (Belgia i Słowenia). Udział państw stosujących spis oparty na rejestrach zwiększył się o 100%, stanowiąc 14,5% wśród badanych 55 krajów. Jeśli jednak uwzględnić państwa, w których zastosowano także metodę złożoną, udział ten wzrósł z 18% w spisach rundy 2000 do 40% w rundzie 2010.

Tabela 2.9. Metody przeprowadzania spisów ludności w wybranych państwach w latach 2000–2010

Metoda przeprowadzania spisu w rundzie 2000	Metoda przeprowadzania spisu w rundzie 2010				Łącznie
	tradycyjna	złożona	spis oparty na rejestrach	inna	
	29	8	1	2	40
Tradycyjna	Albania, Armenia, Azerbejdżan, Białoruś, Bułgaria, Chorwacja, Czarnogóra, Cypr, Grecja, Gruzja, Irlandia, Kanada, Kazachstan, Kirgistan, Luksemburg, Macedonia, Malta, Mołdawia, Monako, Portugalia, Rosja, Rumunia, Serbia, Słowacja, Tadżykistan, Turkmenistan, Ukraina, Węgry, Wielka Brytania	Czechy, Estonia, Izrael, Lichtenstein, Litwa, Polska, Turcja, Włochy	Austria	Francja, USA	
Złożona	0	4	2	0	6
		Holandia, Hiszpania, Łotwa, Szwajcaria	Belgia, Słowenia		
Spis oparty na rejestrach	0	0	4	0	4
			Andora, Dania, Finlandia, Norwegia		
Brak spisu	2	2	1	0	5
	Bośnia-Hercegowina, San Marino	Islandia, Niemcy	Szwecja		
Łącznie	31	14	8	2	

Źródło: (UNECE, 2012).

Wyniki kolejnych badań, opublikowanych przez Europejską Komisję Ekonomiczną ONZ dotyczących źródeł informacji wykorzystanych w spisach 2010 r. (UN, 2013c), wskazują, że blisko 16% państw zastosowało inne aniżeli tradycyjne badania w terenie. Rozkład terytorialny państw według wykorzystywanych w spisie źródeł informacji nie był równomierny. Wśród państw europejskich ponad 43% wykorzystało głównie administracyjne źródła informacji<sup>86</sup> (tab. 2.10). Podobne podejście zastosowało 10% krajów Azji oraz 25% państw Ameryki Północnej. W Afryce i Ameryce Południowej wszystkie państwa pozostały przy tradycyjnej metodzie przeprowadzania spisu.

W Europie podczas spisów ludności zaplanowanych na lata 2010–2014 podejście tradycyjne wybrało tylko 21 państw, spis oparty na rejestrach przeprowadzono w pięciu państwach (rys. 2.5). Na wybór metody mieszanej wykorzystującej źródła administracyjne i badanie reprezentacyjne zdecydowało się 13 państw, w tym Polska (Valente, 2010). Metodę mieszaną wybrano także w Niemczech, gdzie spis ludności przeprowadzono po bardzo długiej przerwie<sup>87</sup>. Wykorzystano w tym celu dane rejestrów administracyjnych urzędów meldunkowych i Federalnego Urzędu Pracy, a badaniem reprezentacyjnym objęto próbę losową 10% mieszkańców (Lamla, 2010). Podejście spisu krocącego, oparte na próbie rotacyjnej (UN, 2008), zastosowano w jednym państwie europejskim – we Francji (rys. 2.5).

Wykorzystanie nowoczesnych technologii w procesie gromadzenia informacji w większości państw (ponad 42%) oznaczało wykorzystanie danych rejestrów administracyjnych oraz gromadzenie informacji za pośrednictwem internetu (tab. 2.10). Jednak w żadnym kraju nie zdecydowano się na wyłączenie zastosowanie samospisu internetowego. Również odsetek państw, w których zastosowano wyłącznie tradycyjną metodę wywiadu bezpośredniego, wynosił blisko 53%. Podkreślimy, że Polska należała do państw, w których całkowicie zrezygnowano z formularza papierowego.

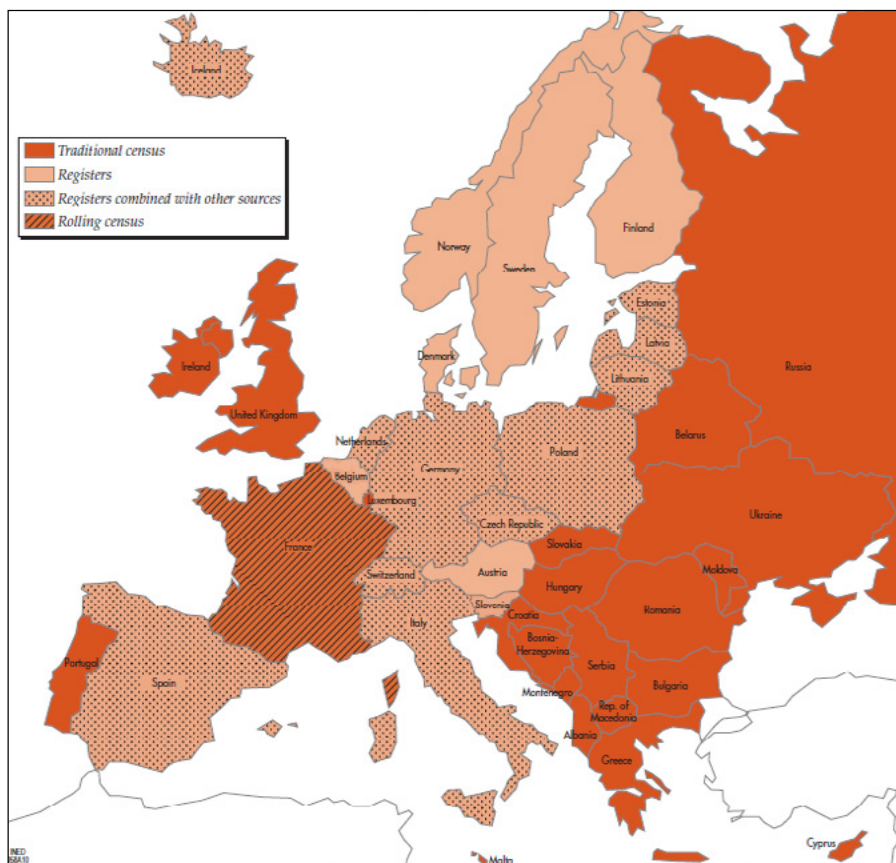
Internet czy samospis internetowy z pewnością nie jest jedyną formą gromadzenia danych, ale może być postrzegany jako metoda transformacji dotychczasowego doświadczenia spisu tradycyjnego z wykorzystaniem nowoczesnych technologii. Doświadczenia rundy 2010 są pod tym względem obiecujące, ale też wyraźnie wskazują, że w kolejnej rundzie 2020 może być jeszcze sporo osób, które będą w tym zakresie potrzebować wsparcia. Wnioski takie dotyczą nie tylko Polski.

Podsumowując powyższe informacje o metodach przeprowadzania spisu ludności w kolejnych rundach lat 1985–1995–2005–2015, można sformułować następujące wnioski (Kukutai i in., 2015, s. 13–14):

- wzrasta liczba państw przeprowadzających spis ludności;
- zwiększa się różnorodność stosowanych metod przeprowadzania spisu;

<sup>86</sup> Jako główne źródło informacji przyjęto to, które było podstawą szacunku łącznej liczby ludności.

<sup>87</sup> W Niemczech ostatnie spisy przeprowadzono jeszcze przed zjednoczeniem: w RFN w 1987 r., a w NRD w 1981 r.



Rysunek 2.5. Metody przeprowadzania spisów ludności w krajach europejskich, 2010–2011

Źródło: (Valente, 2010, s. 3)

Tabela 2.10. Metody gromadzenia danych w spisach rundy 2010 według regionów

Region	Metoda gromadzenia informacji							
	wywiad bezpośredni		samospis papierowy		samospis internetowy		rejestr administracyjny	
	łącznie	wyłączenie	łącznie	wyłączenie	łącznie	wyłączenie	łącznie	wyłączenie
Wszystkie państwa	117	73	39	4	32	0	25	10
Afryka	29	24	4	0	0	0	1	0
Azja	37	18	13	0	12	0	5	0
Europa	23	13	14	3	13	0	17	9
Ameryka Południowa	21	17	3	1	4	0	0	0
Ameryka Północna	3	0	2	0	1	0	1	1
Oceania	4	1	3	0	2	0	1	0

Źródło: (UNECE, 2012).

- częstsze jest wykorzystanie rejestrów administracyjnych w spisie ludności (zarówno jako podstawowego źródła informacji, jak i jako uzupełniającego źródła danych);
- coraz powszechniej obserwujemy dywersyfikację spisu na (i) część „obowiązkową”, w której są gromadzone podstawowe dane – tzw. krótki formularz, oraz (ii) część objętą badaniem reprezentacyjnym gromadzącą informacje uzupełniające – tzw. długi formularz; oznacza to powszechne wykorzystanie wyników badań reprezentacyjnych jako danych spisowych;
- częste jest odwoływanie się do dążenia do poprawy jakości spisu jako istotnej przyczyny zmiany metodologii.

Poprawa jakości danych spisu była też w badaniu ONZ wskazywana jako jedno z istotnych kryteriów sukcesu spisu. Porównanie oceny jakości spisów tradycyjnych i opartych na rejestrach będzie przedmiotem rozważań w kolejnym rozdziale.

## 2.4. Spisy tradycyjne i nowej generacji w świetle międzynarodowych kryteriów jakości

### 2.4.1. Jakość spisów według kryteriów Eurostatu

Porównanie przedstawionych metod przeprowadzania spisów nie jest zadaniem prostym, w szczególności pod względem jakości. Podejmując próbę odniesienia się do wskazanych przez Eurostat sześciu kryteriów jakości (rozdz. 1.4), trzeba być świadomym ich specyficznej oceny w zależności od konkretnych uwarunkowań. W nawiązaniu do sformułowanej tezy, że spisy **wykorzystujące więcej aniżeli jedno źródło danych są lepszej jakości, aniżeli bazujące wyłącznie na jednym źródle**, poniższe rozważania przedstawiają dylematy oceny z perspektywy kryteriów jakości według Eurostatu. Zostaną one omówione w kolejnych punktach tego rozdziału.

#### 2.4.1.1. Przydatność

Przydatność jest to stopień, w jakim są spełnione obecne i potencjalne wymagania użytkowników. Ocenie powinny podlegać także stosowane pojęcia, definicje i klasyfikacje pod kątem tego, czy spełniają potrzeby osób fizycznych, ministerstw i urzędów centralnych, przedsiębiorców, naukowców itp.

W rozdziale pierwszym omówiono tzw. minimum spisowe oraz wyniki konsultacji przeprowadzonych przez GUS w celu oceny potrzeb użytkowników oraz stopnia ich zaspokojenia w kontekście spisu planowanego na 2021 r. Zestawienie zamieszczone w aneksie I.6 pozwala zauważyć, że ocena dokonana przez samych użytkowników jest stosunkowo wysoka (poza dojazdami do pracy we wszystkich obszarach poziom zaspokojenia potrzeb został oceniony na powyżej 88%). Ocena



ta jest wyraźnie wyższa w przypadku organizacji i instytucji aniżeli indywidualnych użytkowników. W wyniku przeprowadzonych konsultacji zostały także zebrane informacje dotyczące oczekiwań respondentów w zakresie niezaspokojonych potrzeb, dodajmy, że wszystkie należą do tzw. tematów uzupełniających. Planując badanie statystyczne, można oczywiście określić wszystkie potrzeby interesariuszy oraz skonstruować odpowiedni kwestionariusz, który pozwoli zebrać wskazane dane. Pozostaje jednak pytanie, na ile respondenci okażą się karni, by nie odmówić udziału w badaniu, oraz cierpliwi, by udzielić odpowiedzi na szeroki zakres pytań, niejednokrotnie poruszających tematy osobiste i drażliwe.

Porównując tematy uwzględniane w polskich spisach ludności od 1921 r.<sup>88</sup>, zauważono sukcesywne poszerzanie ich zakresu. **Na tle wcześniejszych spisów przeprowadzonych metodą tradycyjną, a więc z założenia opartych na jednym źródle danych, spis metodą mieszaną z 2011 r. jest najbardziej wyczerpujący.**

Wśród podstawowych ograniczeń spisów opartych na rejestrach podaje się często argument niemożności pozyskania wszystkich potrzebnych informacji, gdyż nie są one dostępne w źródłach administracyjnych. Jednym z najczęściej przytaczanych jest brak informacji dotyczących rodzin. Pojawia się wówczas potrzeba rekonstrukcji rodzin, która prowadzi do opracowywania metod tzw. parentyzacji. Mogą one czerpać z metod demografii historycznej, w której rekonstrukcję rodzin przeprowadzano, wykorzystując różne źródła, np. na podstawie ksiąg metrykalnych (Kukło, 2009). Ciekawą alternatywą może być metoda własnych dzieci (*own children method*), która umożliwia szacunek cząstkowych współczynników płodności (Feeney, 1975; Krapf i Kreyenfeld, 2015)<sup>89</sup>.

Do zmiennych, jakie można pozyskać tylko w badaniach specjalnych, zalicza się np.:

- stan cywilny faktyczny – związki nieformalne;
- rodzina – struktura, typ biologiczny;
- wykształcenie – w niedalekiej perspektywie możliwe wydaje się jeszcze pełniejsze korzystanie z zasobów zintegrowanego systemu informacji o nauce i szkolnictwie wyższym POL-on w zakresie wyższego wykształcenia oraz z Systemu Informacji Oświatowej Ministerstwa Edukacji Narodowej; problemem może być dostępność danych dla starszych osób, które kończyły edukację przed utworzeniem tych rejestrów;
- narodowość;
- język ojczysty;

<sup>88</sup> Zestawienia przedstawione w aneksach I.8 i I.9 zaczerpnięto z pracy Krywułt-Albańskiej (2012), która skorzystała także z opracowań Holzera (2003).

<sup>89</sup> Wykorzystuje ona dane spisowe o liczbie dzieci, które klasyfikuje się według rocznych grup wieku i wieku matki. Nazwa metody wiąże się z tym, że wiek matki może być sprecyzowany tylko dla dzieci z tego samego gospodarstwa domowego, w którym spisana została matka, stąd „dzieci własne kobiety z danego gospodarstwa domowego”.

- wyznaczenie;
- środki komunikacji i czas dojazdu do pracy.

Dążenie do pełnej i kompleksowej informacji w każdym zakresie jest oczywiste i zrozumiałe, choć czasem trudne do zrealizowania w jednym badaniu, nawet o tak wszechstronnym zakresie jak spis ludności. Do planowania rozwoju aglomeracji miejskiej i gospodarki komunalnej niezbędne mogą być informacje dotyczące środków komunikacji czy infrastruktury. Warto jednak zwrócić uwagę na możliwość pozyskania przynajmniej niektórych takich danych z istniejących źródeł. Mogą one się okazać nie tylko bardzo przydatne, ale wręcz łatwiejsze i tańsze do pozyskania aniżeli w ujęciu tradycyjnego spisu, a także charakteryzować się większą częstotliwością.

W 2009 r. w Ośrodku Statystyki Miast Urzędu Statystycznego w Poznaniu zostało zrealizowane badanie *Przepływy ludności związane z zatrudnieniem w Polsce w 2006 r.* Celem tego badania było zapewnienie informacji o ludności dojeżdżającej do pracy oraz kierunkach i natężeniu dojazdów do pracy na podstawie danych z bazy POLTAX udostępnionej przez Ministerstwo Finansów (Akcelikli i in., 2010). Ze względu na dostępność informacji badanie dotyczyło pracowników najemnych. Skojarzenie informacji o miejscu zamieszkania oraz źródle dochodów osób pracujących dostępnych z formularzy PIT-11/8B z danymi dotyczącymi pracodawców z formularza PIT-40 pozwoliło na konstrukcję macierzy dojazdów do pracy<sup>90</sup>.

Przykład ten dobrano celowo, gdyż było to jedno z pierwszych w Polsce badań przeprowadzonych w skali całego kraju, w którym wykorzystano nowe źródło danych, rejestr administracyjny POLTAX. Było to badanie tanie i niepowodujące dodatkowych obciążeń ludności. Eksperymentalny charakter badania wynikał nie tylko z nowatorskiego ujęcia, ale również czasu potrzebnego na jego przygotowanie oraz zaangażowanej liczby osób (siedem), która wdrożyła nowe, nieznanie wcześniej rozwiązanie: przeprowadziła integrację baz danych, pozyskała dane, zabezpieczyła ich poufność<sup>91</sup>, w celu wizualizacji wykorzystwała narzędzia GIS itd. W związku z badaniem przeprowadzono wnikliwą ocenę jakości pozyskanych danych, konsultacje w gronie ekspertów oraz potencjalnych beneficjentów, przedstawicieli władz samorządowych.

W celu uzupełnienia analizy dojazdów do pracy o dane na temat czasu dojazdu, środków komunikacji oraz infrastruktury można pozyskać dane z nietypowych źródeł pozastatystycznych, takich jak np. czujniki natężenia ruchu, monitoring miejski, serwisy internetowe Google Maps, zasoby operatorów sieci komórkowych<sup>92</sup>,

<sup>90</sup> W badaniu przyjęto pewne założenia, które mogły wpłynąć na jakość wyników, jak np. tożsamość siedziby pracodawcy z miejscem świadczenia pracy. Innym problemem był ograniczony zakres udostępnionych przez Ministerstwo Finansów danych.

<sup>91</sup> Wykorzystanie danych podatkowych mogło budzić emocje społeczne, choć zupełnie nieuzasadnione.

<sup>92</sup> Ciekawe i zachęcające są wyniki badania Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics przeprowadzonego przez konsorcjum Eurostatu, Urzędu Statystycznego

rozkłady jazdy PKP S.A. oraz e-podróżnik.pl. (Beręsewicz i Szymkowiak, 2015; Bul, 2016). Przykładem pierwszego wdrożenia big data w statystyce publicznej jest projekt zrealizowany w Statistics Netherlands polegający na wykorzystaniu drogowych pętli indukcyjnych (*traffic loop detection*) do pomiaru natężenia ruchu samochodowego. Informacje z około 13 tys. punktów pomiarowych umożliwiły oszacowanie liczby samochodów poruszających się po drogach oraz określenie natężenia ruchu według wielkości pojazdów (Beręsewicz i Szymkowiak, 2015; Daas, Puts, Buelens i van den Hurk, 2014).

Spis korzystający z większej liczby źródeł spełnia kryterium przydatności. Podkreślimy też pełną zgodność z zasadami spisu: **powszechność, czyli objęcie całej populacji, oraz indywidualne spisywanie**. Uniwersalność przejawia się w fakcie, że większość cech (dostępnych z rejestrów) dotyczy całej populacji. Jest to niewiarygodnie pozytywna cecha. Dzięki pozyskaniu danych z rejestru bezpośrednio o każdej indywidualnej osobie niweluje się wielkość błędów danych z drugiej ręki.

**Większa liczba źródeł w sposób naturalny oznacza możliwość poszerzenia zakresu tematycznego spisu.** Doskonałym przykładem jest ujęcie wielu tematów dodatkowych poprzez czerpanie z zasobów kolejnych rejestrów, jak w praktyce wdrażano spis oparty na rejestrach w państwach skandynawskich (tab. 2.5). Do podstawowego rejestru ludności dołączono rejestr dochodów<sup>93</sup>, zatrudnienia, przedsiębiorstw<sup>94</sup>, wykształcenia, mieszkań, gospodarstw domowych, warunków mieszkaniowych. Obecnie rejestry są prowadzone przez instytucje w formie elektronicznej, często z automatyczną formą odczytu i przesyłu danych. Jednak wykorzystanie danych rejestru w spisie w każdym przypadku było poprzedzone wieloletnią współpracą statystyków z gestorem rejestru dotyczącą poprawy jakości, kompletności, wyjaśnienia pojęć. W efekcie jakość rejestrów ulega sukcesywnej poprawie, nie tylko w państwach skandynawskich (Tønder, 2008). Doskonałym przykładem jest także Austria (Asamer i in., 2014, 2016; Lenk, 2014) czy Hiszpania (Vega i Argüeso, 2016), a także Polska (Beręsewicz i in., 2018; GUS, 2014b). Również systemy informacyjne przedsiębiorstw działających w zakresie sprzedaży energii elektrycznej, wody, usług telekomunikacyjnych coraz częściej odpowiadają standardom jakości (Dehnel i Gołata, 2006, 2012). Wprowadzenie danych rejestru czy systemu pozaadministracyjnego do wykorzystania w spisie w każdym

---

Estonii oraz firmę Positium. Celem projektu była ocena możliwości wykorzystania danych pochodzących z sieci komórkowych do szacowania migracji wewnętrznych i zagranicznych w Estonii (Ahas i in., 2014).

<sup>93</sup> W Polsce nie ma tradycji szacowania wysokości dochodu w spisach ludności, jednak możliwość taką stwarza potencjalna integracja z systemem informacyjnym Ministerstwa Finansów/urzędów skarbowych prowadzonym przez organy podatkowe (POLTAX).

<sup>94</sup> Spis 2011 r. w Austrii obejmował spis ludności, spis przedsiębiorstw i ich lokalnych jednostek zatrudnienia oraz spis mieszkań (Asamer i in., 2014).

państwie jest uwarunkowane uznaniem ich za wiarygodne przez statystykę publiczną. Ostatecznie, gdy zasoby rejestrów i systemów zostają dołączone do spisu i są analizowane w wielowymiarowym kontekście, zyskują dodatkową przydatność w zarządzaniu państwem.

#### **2.4.1.2. Dokładność – bliskość do wartości rzeczywistych**

Odpowiedź na pytanie o dokładność wyników spisu nie jest łatwa, gdyż dotykamy tutaj kwestii „wartości rzeczywistej czy prawdziwej”, albo jeszcze szerzej – prawdy. Nie wnikając w filozoficzne niuanse określenia prawdy, można sformułować pytanie, co przemawia za uznaniem szacunków bazujących na jednym źródle za „bardziej prawdziwe” bądź nieprawdziwe czy bardziej różne od prawdziwych w porównaniu do szacunków bazujących na dwóch lub więcej źródłach. Logiczne rozumowanie zdaje się przemawiać na korzyść przypuszczenia, że jeśli dwa źródła, lub kilka źródeł, dostarczą zgodnych szacunków – bardziej nas utwierdzą w przekonaniu o ich bliskości do wartości prawdziwej aniżeli jedno źródło. Przyjmując powyższy tok rozumowania, konieczne jest uwzględnienie charakteru porównywanych badań pod względem pokrycia populacji celu. Dokonując oceny spisu, ograniczamy dyskusję do badań pełnych. Z tego powodu eliminujemy z dalszych rozważań błędy losowe, pozostając przy ocenie błędów kompletności i treści. Specyficznym problemom oceny jakości spisów jako badań pełnych, w tym ocenie kompletności, jest w całości poświęcony rozdział trzeci, dlatego w poniższych rozważaniach dokładności zamieszczono jedynie ogólny komentarz.

W spisie tradycyjnym ocena jakości w tym zakresie sprowadza się zazwyczaj do porównań z wynikami innych badań: spisu kontrolnego<sup>95</sup> lub wcześniejszych spisów. Również w spisie wykorzystującym wiele źródeł danych nie ma przeszkód do zastosowania analogicznej procedury.

W spisie opartym na rejestrach wykorzystanie poszczególnych źródeł jest uwarunkowane wcześniejszą pozytywną oceną ich jakości, w tym kompletności. Przykładem specjalnych metod stosowanych w tym celu jest austriacka analiza rezydencji (rozdz. 2.1.4). Jej podstawą jest walidacja rekordów rejestru ludności na podstawie badania relacji pomiędzy rejestrami (co najmniej dwoma) potwierdzającymi „aktywność” poszczególnych osób. Dodatkowo w sytuacjach wątpliwych są stosowane specjalne procedury wyjaśniające, które uwzględniają możliwość odwołań (Asamer i in., 2014). Jak przedstawiono we wcześniejszych rozważaniach, w procedurze spisu bazującego na wielu źródłach jest wbudowany mechanizm weryfikacji jakości, także w odniesieniu do treści. Decyzja dotycząca wartości każdej cechy dla każdej z badanych jednostek jest wynikiem analizy relacji między rejestrami. Wykorzystuje się tutaj zasadę redundancji oraz określony dla każdej z cech algorytm weryfikacji zgodności.

<sup>95</sup> Spis kontrolny zostanie szczegółowo omówiony w rozdz. 3.2.2.

### 2.4.1.3. Terminowość i punktualność

Punktualność jest związana z opóźnieniem czasowym pomiędzy rzeczywistą a planowaną datą dostarczenia danych wynikającą z oficjalnego kalendarza publikacji ustanowionego w zarządzeniach lub uzgodnionego wcześniej z klientami. Natomiast przez **terminowość**, zgodnie z zapisami Europejskiego Kodeksu Praktyk Statystycznych (Eurostat, 2011), należy rozumieć czas między wydarzeniem a publikacją informacji.

Ponieważ wyniki badań statystyki publicznej są ogłaszane według ściśle określonego harmonogramu<sup>96</sup>, skoncentrujemy się raczej na porównaniu terminowości. Wyniki spisu, podobnie jak i innych, badań są opracowywane i ogłaszane zgodnie z przyjętym porządkiem. Polskę obowiązują także regulacje międzynarodowe, które w przypadku NSP 2011 wyznaczały termin przekazania danych do Eurostatu w ciągu dwóch lat po zakończeniu spisu, tj. do końca 2013 r.

**Terminowość i aktualność danych** spisu tradycyjnego jest ograniczona ze względu na zazwyczaj długi czas między momentem realizacji badania i publikacji wyników. W Polsce pierwsze wyniki ostatniego spisu przeprowadzonego metodą tradycyjną (moment krytyczny: 20 maja 2002 r.) zostały opublikowane po upływie roku – 11 czerwca 2003 r., a kolejne raporty – 16 czerwca i 1 września 2003 r.<sup>97</sup>. Ostatnia publikacja danych NSP 2002 ukazała się 15 lipca 2008 r., a więc po upływie sześciu lat. Jeśli porównać czas publikacji wyników spisu przeprowadzonego metodą mieszaną (moment krytyczny: 1 kwietnia 2011 r.), pierwsze (wstępne) wyniki GUS opublikował w rekordowo krótkim czasie – 22 grudnia 2011 r. Kolejna publikacja również wyników wstępnych ukazała się 22 marca 2012 r., z okazji Kongresu Demograficznego, a więc przed upływem roku. Wyniki NSP 2011, które nie zostały określone jako wstępne, opublikowano 26 lipca 2012 r., a kolejne raporty dopiero w kwietniu i maju 2013 r., czyli po dwóch latach. Ostatnia publikacja wyników NSP 2011 dostępna na stronie GUS ukazała się z datą 10 listopada 2016 r., czyli z opóźnieniem pięciu i pół roku.

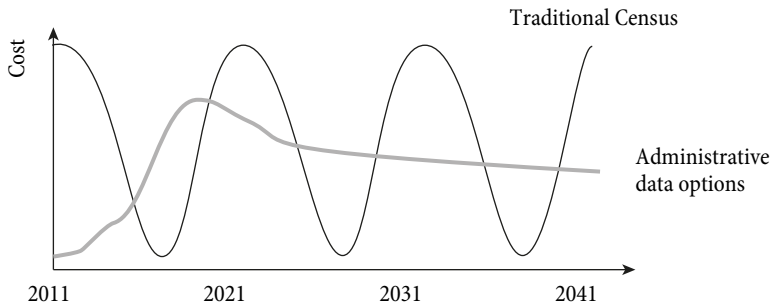
Przedstawiona wyżej praktyka publikacji wskazuje, że uwzględniając tempo zmian ludnościowych, wyniki spisu bardzo szybko tracą aktualność. Nie jest to tylko polski problem. Ze względu na rozmiar operacji spisowych oraz ich złożoność opóźnienie od momentu badania do publikacji wynosi 18 miesięcy (Calder i Teague, 2012, s. 5), a według innych szacunków nawet około dwóch lat (Vega i Argüeso, 2016). Tymczasem zgodnie z obowiązującymi zasadami wyniki spisu są

<sup>96</sup> Roczny harmonogram publikacji danych oraz kalendarium miesięczne są dostępne na stronie internetowej Eurostatu, <http://ec.europa.eu/eurostat/news/release-calendar>, w odniesieniu do statystyk europejskich oraz GUS-u <http://stat.gov.pl/dla-mediow/harmonogramy-publicacji-danych/> w przypadku statystyk krajowych. Wyznaczone jest również miejsce w GUS, w którym odbywają się konferencje i briefingi prasowe oraz określona jest standardowa godzina w ciągu dnia na publikację statystyk europejskich. Wszelkie odstępstwa od harmonogramu publikacji są z wyprzedzeniem ogłaszane i uzasadniane oraz ustalany jest nowy termin publikacji.

<sup>97</sup> Dokładne daty publikacji wyników spisu pochodzą ze strony internetowej GUS, dla NSP 2002: <http://stat.gov.pl/spisy-powszechne/narodowe-spisy-powszechne/narodowy-spis-powszechny-2002/>, a dla NSP 2011: <http://stat.gov.pl/spisy-powszechne/nsp-2011/>.

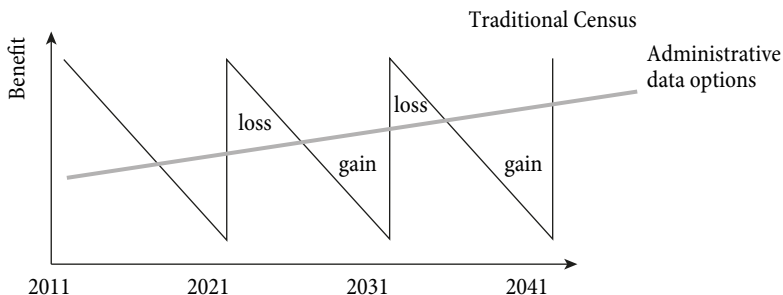
uznawane jako referencyjne w okresie do kolejnego spisu – czyli przez około 10 lat. Dotychczas obowiązujące standardy przewidują korektę danych wstecz, zgodnie z informacjami otrzymanymi w bieżącym spisie.

Szczególnie szybka jest utrata aktualności danych spisowych w populacji doświadczającej gwałtownych przemian procesów demograficznych, w tym intensywnych migracji. Ze względu na opóźnienia aktualność danych spisowych jest dyskusyjna także w chwili publikacji wyników. Natomiast dane administracyjne charakteryzuje stała aktualność. Podejście oparte na rejestrach może zapewnić częstsze aktualizacje, przez co zmniejszy się spadek dokładności w okresie międzyspisowym. Potencjalny zysk dokładności szacunków spisowych w ujęciu tradycyjnym i opartym na danych administracyjnych autorzy programu Beyond 2011<sup>98</sup> przedstawiają w ujęciu dziesięcioletniego cyklu badania, dążąc do zrozumienia profilu korzyści dla każdej opcji (rys. 2.6 i 2.7).



Rysunek 2.6. Potencjalne profile kosztów spisu tradycyjnego i wykorzystującego dane administracyjne

Źródło: (Calder i Teague, 2012, s. 4)



Rysunek 2.7. Potencjalne profile korzyści tradycyjnego spisu i wykorzystującego dane administracyjne

Źródło: (Calder i Teague, 2012, s. 5)

<sup>98</sup> Beyond 2011 to projekt realizowany w Wielkiej Brytanii po spisie 2011 r. w celu wypracowania rekomendacji dla kolejnego spisu.

Na podstawie przeprowadzonych analiz Calder i Teague (2012) przedstawili korzyści danych administracyjnych w stosunku do spisu tradycyjnego. Przygotowując badanie oparte na rejestrach, które z natury są na bieżąco aktualizowane, można tak zaplanować i zautomatyzować szereg zadań, by skrócić czas potrzebny na udostępnienie danych i opublikować je stosunkowo bliżej daty referencyjnej. Uwzględnienie możliwości automatyzacji procedur procesu przekazywania danych oznacza ich większą aktualność oraz skrócenie czasu od momentu referencyjnego. Dodatkowo, wykorzystanie danych ze źródeł administracyjnych umożliwia podejście zapewniające publikowanie danych o ludności z większą częstotliwością – np. roczną.

#### **2.4.1.4. Dostępność i przejrzystość – kanały dystrybucji, metadane, formy publikacji**

Kryterium dostępności i przejrzystości wynika z 15. zasady Europejskiego Kodeksu Praktyk Statystycznych. Zasada ta głosi, że „Statystyki europejskie należy prezentować w jasnej i zrozumiałej formie, upowszechniać w odpowiedni i dogodny sposób, udostępniać zgodnie z zasadą bezstronności wraz z właściwymi metadanymi i wytycznymi” (Eurostat, 2011, s. 8). W zakresie **dostępności** należałoby oceniać procedury składania zamówień na dane, czas potrzebny na realizację dostawy, politykę cenową, dostępność mikro- i makrodanych<sup>99</sup>, formę udostępniania (wydruk, pliki, CD-ROM, internet) itp. **Przejrzystość** natomiast odnosi się do odpowiednich metadanych (informacji, wyjaśnień, dokumentacji itp.).

Jakość w zakresie udostępniania i przejrzystości danych zależy od kultury i dojrzałości instytucji, regulacji prawnych, stosowania zaawansowanych metod badawczych oraz rozwiązań technologicznych. Jakość w tym zakresie nie musi być związana z wielością wykorzystanych źródeł. Jednak doświadczenie wskazuje, że właśnie przeprowadzenie badania na podstawie więcej aniżeli jednego źródła przyczynia się do lepszej jakości metadanych. Wynika to z konieczności zastosowania zaawansowanych metod, które wymagają nie tylko przedstawienia definicji i stosowanych klasyfikacji, ale przede wszystkim opisu i wyjaśnień w zakresie algorytmu, wyników porównań, badania spójności mikrodanych w trakcie procesu realizacji badania. Wykonane operacje często przekładają się na sposób przedstawienia tych informacji w materiałach źródłowych.

#### **2.4.1.5. Porównywalność – czasowa, geograficzna i porównywalność dziedzin**

Kryterium **porównywalności** wiąże się z zapewnieniem zgodności stosowanych pojęć, definicji i klasyfikacji w czasie, przestrzeni oraz w przekroju różnych dziedzin. Jest to podstawowy wymóg wszelkich analiz, którego przestrzeganie jest istotnym problemem zarówno w spisach tradycyjnych, jak i wykorzystujących wiele różnych źródeł. Zapewnienie porównywalności w wymiarze przestrzennym, w tym pomiędzy krajami i (lub) regionami, jest między innymi przedmiotem prac

<sup>99</sup> Dostęp do mikrodanych regulują ścisłe protokoły.

organizacji międzynarodowych, których szczególnym wyrazem są rekomendacje ONZ (por. rozdz. 1.2.1).

Państwa starają się stosować do tych zaleceń, przynajmniej w zakresie minimum spisowego. W spisach tradycyjnych pytania kwestionariusza spisowego oraz odpowiednie instrukcje dla rachmistrzów spisowych uwzględniają określone zalecenia. Uzyskane odpowiedzi przyjmuje się jako prawdziwe i nie podlegają one weryfikacji. W ten sposób uzyskuje się zgodność deklaratywną. W spisach korzystających ze źródeł administracyjnych uzyskuje się informacje odpowiadające określonym regulacjom prawnym, często zmieniającym się w czasie oraz nieporównywalnym pomiędzy państwami. W celu uzyskania pożądanej informacji wymagane jest stosowanie określonych przez metodologów algorytmów. Zamieszczony w uwagach metodycznych szczegółowy opis sposobu pozyskania danej zmiennej umożliwia wnikliwą analizę zgodności oraz rozbieżności, tak istotną w analizach porównawczych.

#### **2.4.1.6. Spójność – zgodność informacji z różnych źródeł w różnych klasyfikacjach i ujęciach**

Spójność statystyki oznacza możliwość zgodnego przedstawienia danych w różnych konfiguracjach, dla różnych zastosowań. Powszechnie dane z jednego źródła uważa się za spójne, podczas gdy w przypadku korzystania z wielu źródeł mogą wystąpić różne definicje, klasyfikacje, metody szacunku. Jednak również przy jednym, wyczerpującym badaniu nie zawsze uzyskiwano spójne szacunki w przekrojach regionalnych z uwzględnieniem różnych domen. Zapewnienie spójności wymaga z pewnością bardziej kompleksowego podejścia w sytuacji wielu źródeł, które uwzględniają zaawansowane metody edycji danych, integracji, estymacji, kalibracji, oceny precyzji itd. Stanowią one w rzeczywistości istotę tego podejścia w badaniach statystycznych, dążąc do wykazania wiarygodności szacunku opartego na wielu źródłach.

#### **2.4.2. Ekonomiczno-społeczne problemy jakości spisów**

W przedstawionej dyskusji różnych metod przeprowadzania spisów ludności według kryteriów jakości Eurostatu starano się wyłowić argumenty za i przeciw, ukazując korzyści z zastosowania określonej metody i jej ograniczenia. Podstawowe znaczenie ma w tym zakresie identyfikacja czynników warunkujących wprowadzenie i praktyczne zastosowanie alternatywnych metod spisu łącznie z wyszczególnieniem kosztów przedsięwzięcia oraz uzyskanych wyników. Po przedstawieniu opisu różnych metod oraz ich dyskusji, w świetle przyjętych kryteriów jakości, podjęto także próbę syntezy z perspektywy takich dodatkowych, ale ważnych uwarunkowań. Kukutai i in. (2015), przeprowadzając ocenę spisu tradycyjnego, wskazują cztery najczęściej przedstawiane czynniki: (i) koszty spisu, (ii) zaufanie społeczne do



statystyki publicznej i zmniejszający się odsetek respondentów uczestniczących w badaniu, (iii) zachowanie poufności danych oraz (iv) potrzebę terminowego i częstego dostarczania danych. Niezbędne wydaje się jeszcze ujęcie w tym miejscu piątego czynnika – (v) regulacji prawnych.

#### 2.4.2.1. Koszt spisu

W pierwszej kolejności Kukutai i in. (2015) wskazują niepohamowany wzrost kosztów tradycyjnych spisów. Odwołując się do przykładu Stanów Zjednoczonych, podaje, że koszty spisu w 2010 r. sięgające 13 miliardów USD oznaczały podwojenie kosztów spisu przeprowadzonego w 2000 r., w którym z kolei podwojono koszt poprzedniego spisu.

Ten wzrost kosztów autorzy tłumaczą zmianami społecznymi skutkującymi wzrostem odsetka populacji „trudnej do zbadania” (*hard to reach population*). Wskazują w tym przypadku na kilka kwestii. Po pierwsze, rodziny stanowiące dotąd podstawową jednostkę badania coraz częściej mieszkają w kilku miejscach, a dzieci, wskutek upowszechnienia modelu rodziny patchworkowej, współzamieszkują w kilku gospodarstwach domowych. Wzrost kosztów spisu wynika także z większego zróżnicowania społecznego w wymiarze narodowościowym i z konieczności pokonania bariery językowej, właściwego przygotowania rachmistrzów w celu umożliwienia poprawnego spisu.

Z drugiej strony podawane są przykłady szerszego wykorzystania nowoczesnych technologii redukujących koszty. Znakomitym i często podawanym przykładem jest tutaj Estonia, gdzie 67% ludności spisało się przez internet, korzystając z kwestionariusza online (GUS, 2017d; Kukutai i in., 2015). Również Polska jest przykładem państwa, w którym z powodzeniem wykorzystano elektroniczną formę formularza oraz zastosowano metodę CAPI, CATI oraz CAII w celu zgromadzenia informacji, a rachmistrze spisowi korzystali z handheldów wyposażonych nie tylko w system internetowego przekazywania wypełnionych formularzy, ale także w mapy numeryczne i fotografie ułatwiające lokalizację (GUS, 2014b). Kukutai i in. (2015) przytaczają informacje o obniżeniu kosztów spisu przez likwidację formularza papierowego<sup>100</sup> oraz zmniejszenie liczby rachmistrzów spisowych ze 170 tys. w 2002 r. do 17 tys. w 2011 r. (GUS, 2014b, s. 353 i nast.).

Komentując wysokie koszty spisów tradycyjnych, warto spojrzeć na nie z jeszcze jednej perspektywy. Mianowicie Coleman (2013) podkreśla marnotrawstwo kapitału ludzkiego. W jego opinii większy byłby pożytek dla statystyki publicznej z trwałego zatrudnienia nielicznej grupy osób regularnie zdobywających doświadczenie i wiedzę ekspercką, niż z intensywnego szkolenia ogromnej rzeszy osób, które są zatrudnione przez krótki okres realizacji spisu.

<sup>100</sup> W spisie tradycyjnym w Polsce drukowano zwykle ok. 80 milionów arkuszy A4 papieru wysokiej jakości, głównie ze względu na wymagania kosztownej technologii OCR (GUS, 2014b, s. 68).

#### 2.4.2.2. Zaufanie społeczne do statystyki publicznej i zmniejszający się odsetek respondentów uczestniczących w badaniu

Ważnym problemem jest zaufanie społeczne do statystyki oficjalnej oraz gwarancja bezpieczeństwa i poufności danych spisowych. Zdaniem Colemana (2013), ludzie coraz rzadziej wierzą w zapewnienia, że dane spisowe zostaną zachowane w tajemnicy, szczególnie przed fiskusem. Wyniki badań opinii społecznej potwierdzają kryzys zaufania na całym świecie. Barometr zaufania Edelmana w 2017 r. ujawnił, że zaufanie ogółu społeczeństwa do wszystkich czterech kluczowych instytucji – biznesu, rządu, organizacji pozarządowych i mediów – znacznie się zmniejszyło, odkąd w 2012 r. Edelman objął badaniem ten segment<sup>101</sup>. Sytuację tę potęgują takie wydarzenia jak kryzys statystyki greckiej czy decyzje władz kanadyjskich dotyczące spisu (długi formularz). Z poglądem tym polemizuje Kukutai, który przyznaje, że ludzie coraz więcej uwagi poświęcają bezpieczeństwu danych i gwarancjom poufności, jednak w wielu państwach Europy zaufanie do rządów w ostatnich 30 latach nie tylko nie spadło, ale wręcz wzrosło (Kukutai i in., 2015, s. 7). Coleman (2013) z kolei uważa, że zaufanie to jest kwestią kultury politycznej, a nie reakcją ludności na sposób gromadzenia danych czy przeprowadzania spisu. W państwach skandynawskich, w których od lat 70. ubiegłego wieku są wykorzystywane rejestry administracyjne, podejście to cieszy się powszechną aprobatą. Społeczeństwo pozytywnie ocenia wykorzystanie już istniejących zasobów danych z uwagi na efektywność i racjonalność takiego podejścia. Jednak trudno nie przyznać, że w państwach o doświadczeniach totalitarnych, w których pamięta się strach przed inwigilacją i represjami, trudno o zgodę w kwestii wykorzystania zasobów administracyjnych. Przykładem mogą być Niemcy, gdzie dwukrotnie odkładano spis ludności, raz z powodów finansowych, a drugi raz z uwagi na zbyt daleko idące zmiany w metodologii spisu (Kukutai i in., 2015), oraz Francja (Poulain i Herm, 2013).

Coleman (2013) zauważa prawidłowość, że długi formularz spisowy w spisie tradycyjnym, ale także w spisie metodą mieszaną, wpływa na obniżenie uczestnictwa w badaniu. Potwierdzają to także wyniki innych badań specjalnych, w których dramatycznie zmniejsza się odsetek respondentów biorących udział w badaniu. W NSP 2011 r., pomimo obowiązkowego charakteru badania, udział braków odpowiedzi wyniósł 13,7%. Choć stosuje się zaawansowane metody estymacji, odmowy udziału w badaniu wpływają na jakość szacunków. Coleman (2013, s. 338) przytacza wyniki badań ONS, według których w tradycyjnym spisie 2001 r. w Wielkiej Brytanii udało się zbadać 90% populacji celu, a w przypadku populacji trudnej do zbadania („*hard to reach*”) było to zaledwie 62%. Autor podkreśla, że mimo iż błędy tradycyjnych spisów mogą być znaczne, to traktowane są jako „*gold standard*”.

W tym miejscu ważne wydaje się zwrócenie uwagi na problemy związane z integracją danych w spisach korzystających z wielu źródeł. Wyraźnie różna jest jakość

<sup>101</sup> Barometr zaufania Edelmana jest publikowany od 1952 r. na stronie internetowej: <https://www.edelman.com/global-results/>.

danych spisowych pochodzących z rejestru o zazwyczaj blisko 100-procentowym pokryciu badanej populacji oraz z długiego formularza otrzymanego w badaniu specjalnym, w którym, pomimo wszelkiej staranności, było 30% odmów. Jedne i drugie dane zostały zgromadzone w ramach spisu, a poddawane integracji i kalibracji tworzą całość. Jednak wszystkie powinny podlegać także niezależnej ocenie. W publikacjach spisowych powinna być publikacja o jakości każdej cechy niezależnie, nie tylko o źródle, z jakiego pochodzi, ale także wielkości błędu szacunku, odsetku braków odpowiedzi itp.

#### 2.4.2.3. Zachowanie poufności danych

Niebezpieczeństwo szczególnie wiąże się z wykorzystaniem indywidualnych numerów identyfikacyjnych (np. numeru Pesel) jako klucza umożliwiającego bezpośrednią integrację danych różnych rejestrów. Numer identyfikacyjny jest związany z każdym człowiekiem „od kołyski po grób”, a zgromadzenie danych z różnych rejestrów rodzi możliwość utworzenia i ciągłej aktualizacji niewiarygodnej bazy danych o każdym z nas, jakiej nie zgromadził dotąd żaden spis ludności. Istnienie takiej bazy rodzi z kolei ryzyko niewłaściwego jej wykorzystania. Niebezpieczeństwo związane z brakiem ochrony czy niedostatecznym zabezpieczeniem danych doprowadziło do wdrożenia zasady „odcinania” numerów identyfikacyjnych, jednostronnego przekazywania danych od gestorów do urzędów statystycznych oraz zakazu bezpośredniego łączenia baz danych według rzeczywistych identyfikatorów. Przykładowy system zabezpieczenia poufności przedstawiono w podrozdziale 2.1.4, opisując austriacki spis oparty na rejestrach, oraz w podrozdziale 2.2 w odniesieniu do NSP 2011. Przyjęte rozwiązania obejmowały zapewnienie bezpieczeństwa w zakresie infrastruktury informatycznej i telekomunikacyjnej, sprzętu, oprogramowania, baz danych, fizycznej ochrony danych, kwestii prawnych, organizacji badania w odniesieniu do wszystkich faz przeprowadzania spisu (Dygaszewicz, 2009; GUS, 2014b). W Europie nie odnotowano przypadku naruszenia prywatności danych<sup>102</sup>.

Na marginesie dyskusji o zachowaniu poufności danych warto zwrócić uwagę na odmienny pogląd, że obowiązujące regulacje są przejawem swoistej „poprawności politycznej” w badaniu niektórych zagadnień życia rodzinnego czy niepełnosprawności. Przykładowo w NSP 2011 informacje o osobach niepełnosprawnych w badaniu reprezentacyjnym były zbierane na zasadzie dobrowolności (Slany, 2014). Podejście takie było konsekwencją zapisu Konstytucji RP (art. 51), który nie dopuszcza pozyskiwania informacji o stanie zdrowia (w tym o niepełnosprawności) na zasadzie obowiązku. W rezultacie jednak szacunek liczby osób niepełnosprawnych biologicznie w NSP 2011 jest zaniżony<sup>103</sup>.

<sup>102</sup> Jednak incydenty takie zdarzały się, a ostrzeżeniem może być zdarzenie, jakie miało miejsce w Izraelu. W 2011 r. aresztowano tam sześć osób za kradzież i dystrybucję całego rejestru ludności Izraela zawierającego dane dotyczące dziewięciu milionów osób (Coleman, 2013, s. 346).

<sup>103</sup> W efekcie wyniki NSP 2011 r. wykazały, że liczba osób, które zadeklarowały ograniczenie zdolności do wykonywania zwykłych czynności podstawowych dla swojego wieku i (lub) posiadały

#### 2.4.2.4. Częstotliwość spisu

Jak wskazano wcześniej (2.4.1.3), spis bazujący na źródłach administracyjnych umożliwi publikację danych o ludności z większą częstotliwością. Podejście takie umożliwi likwidację jednego z istotnych ograniczeń spisu przeprowadzanego metodą tradycyjną, jakim jest trudność ukazania dynamiki zmian w populacji. W szczególności jest to ważne w państwach doświadczających intensywnych migracji wewnętrznych i zagranicznych. Według Colemana (2013) sięgające dwóch lat opóźnienie w dostępie do danych spisowych ogranicza zdolność polityki społecznej do nadążenia za zmianami demograficznymi i społecznymi. Brak aktualnych danych w ujęciu regionalnym ogranicza możliwość właściwego wykonywania zadań przez samorządy lokalne. Wskazując powody, dla których należy zaniechać tradycyjnych spisów ludności, Coleman (2013) argumentuje, że spis przedstawia statystyczny obraz populacji raz na 10 lat, podczas gdy świat zmienia się bardzo szybko<sup>104</sup>. W dyskusji tej Coleman przytacza słowa innego autorytetu, Leslie Kisha<sup>105</sup>, który miał zauważyć, że spis „was great for its time, but not for today. Having censuses for local data once every 10 years is just not good enough anymore, and it hasn't been for some time” (Coleman, 2013, s. 336)<sup>106</sup>.

Zmiany społeczne wynikające z mobilności edukacyjnej, zawodowej, związane z migracjami wewnętrznymi i zagranicznymi powodują, że badana ludność przeobraża się w populację trudną do zbadania (Coleman, 2013, s. 337). Coleman wskazuje raczej na przyczyny określane jako mobilność społeczna i przestrzenna aniżeli zmiany norm i zachowań dotyczących formowania rodziny oraz typów gospodarstw domowych. Podkreśla wzrost liczby osób mieszkających jednocześnie w dwóch miejscach, znaczenie imigrantów, zwłaszcza nielegalnych, osób nadmiernie przedłużających pobyt, które mogą unikać kontaktu z urzędnikami. Coleman zauważa, że w spisie powszechnym mogą być powielane informacje z innych oficjalnych źródeł. Przypadek Polski potwierdza duże znaczenie wszystkich wskazanych tutaj czynników obejmujących wzrost mobilności społecznej i przestrzennej oraz zmiany norm i stylu życia, które mogą skutkować także niedoszacowaniem ludności spisowej.

---

ważne orzeczenie kwalifikujące je do zaliczenia do zbiorowości osób niepełnosprawnych, wynosiła 4697,5 tys., co stanowiło 12,2% ludności kraju, wobec 14,3% w 2002 r.

<sup>104</sup> Zasadne wydaje się w tym miejscu wyjaśnienie, że przedstawiony przy okazji omawiania zapotrzebowania na dane spisowe argument za potrzebą spisu nie wskazywał na konieczność tradycyjnego badania, ale generalnie świadczył o niezbędności danych spisu do ukazania dynamiki przemian demograficznych.

<sup>105</sup> Kish (27.07.1910–7.10.2000), urodzony w Popradzie statystyk, autorytet w zakresie metody reprezentacyjnej, w 1975 r. jeden z założycieli international Association of Survey Statisticians (IASS). W latach 1977–79 wiceprezydent, a w latach 1983–85 prezydent IASS. Założyciel Institute for Social Research na Uniwersytecie Michigan. Współpracował z U.S. Bureau of the Census.

<sup>106</sup> Stwierdzenie Kisha można przedstawić następująco: w swoim czasie spis był wielkim osiągnięciem, ale nie dzisiaj. Dysponowanie danymi w przekrojach lokalnych raz na 10 lat już nie wystarcza.

#### 2.4.2.5. Regulacje prawne

Spis powszechny jest badaniem przeprowadzanym na podstawie odpowiednich aktów prawnych. W Polsce jest to ustawa o narodowym spisie powszechnym ludności i mieszkań, którą uchwała Sejm przed każdą edycją badania (Ustawa, 2010). Ustawie sejmowej towarzyszą odpowiednie akty wykonawcze w formie rozporządzeń. Ponadto państwa członkowskie Unii Europejskiej są zobowiązane do przestrzegania ustaleń organizacji międzynarodowych, w tym przypadku jest to rozporządzenie Parlamentu Europejskiego i Rady Nr 763/2008 w sprawie spisów powszechnych ludności i mieszkań (UE, 2008). Członkostwo w Unii wiąże się z obowiązkiem przekazywania odpowiednich danych do Eurostatu, co wymaga niezbędnych wyjaśnień prawnych i metodologicznych (EU, 2011).

W zasadzie można by poprzestać na powyższym stwierdzeniu, ponieważ wszelkie szczegółowe uregulowania prawne muszą respektować akty nadrzędne, ale także z nich wynikać. Jednak inaczej kształtuje się sytuacja, gdy regulacje prawne dotyczą jednego badania, nawet bardzo wszechstronnego i o zasięgu ogólnokrajowym, i inaczej, jeśli w realizację badania jest dodatkowo włączonych wiele niezależnych organizacji i instytucji. Powodzenie badania zależy wówczas od zdefiniowania zasad współpracy, określenia zadań poszczególnych jednostek, sumiennej i terminowej ich realizacji, wzajemnego poszanowania. Organizacja spisu wykorzystującego dane rejestrów administracyjnych wymaga budowy systemu, w którym wszystkie organy administracji państwowej i samorządowej świadomie i odpowiedzialnie współuczestniczą w tworzeniu statystyki oficjalnej – gromadzeniu danych niezbędnych do sprawnego funkcjonowania państwa.

Przede wszystkim urząd statystyczny musi mieć prawo korzystać z zasobów danych tworzonych przez inne organy administracji państwowej. Zasady przekazywania tych danych, ich zakres, częstotliwość aktualizacji, forma przygotowania i gwarancje bezpieczeństwa muszą być szczegółowo doprecyzowane. Niewywiązanie się z ustawowo określonych obowiązków może być porównane do odmowy udziału w badaniu, co według ustawy pozwala na nałożenie kary, a w przypadku instytucji państwowej powinno być niedopuszczalne. Jeśli przy organizacji NSP 2011 w Polsce korzystano ogółem z zasobów 27 rejestrów administracyjnych oraz z trzech systemów pozaadministracyjnych, słuszny wydaje się postulat powołania przez prezesa Rady Ministrów Grupy Międzyministerialnej ds. Spisu, w której skład wchodziłoby obok prezesa GUS ministrowie kierujący resortami będącymi gestorami wykorzystywanych w spisie rejestrów. Ustalenia na najwyższym szczeblu umożliwiają sprawną realizację działań na szczeblu operacyjnym.

Może warto dodatkowo podkreślić, że ustawa winna gwarantować nieodpłatne przekazywanie danych do urzędu statystycznego i że dane powinny być w formie elektronicznej uzgodnionej z urzędem. Ponieważ łączenie danych rejestrów jest zabronione, w pierwszej kolejności urząd jest zobowiązany do gwarantującego poufność przekodowania danych przed rozpoczęciem badań, analiz. Ważnym

elementem współpracy urzędu statystycznego i gestora rejestrów jest wypracowanie systemu poprawy jakości. Chodzi o to, że dochowując zasady poufności, GUS nie może wskazać, że przykładowo adres miejsca zamieszkania konkretnej osoby jest niewłaściwie zapisany, ale może zasugerować gestorowi rejestru weryfikację zgodności z systemem TERYT.

Korzystanie z danych systemów informacyjnych administracji publicznej przez GUS jest coraz powszechniejsze. W Polsce zagadnienie to zostało formalnie uregulowane w Ustawie z dnia 29 czerwca 1995 r. o statystyce publicznej z późniejszymi zmianami (Obwieszczenie, 2016, art. 13). Określa ona, że organy administracji rządowej i jednostki samorządu terytorialnego oraz inne instytucje, w tym Narodowy Bank Polski (NBP), nieodpłatnie przekazują służbom statystyki publicznej zgromadzone dane administracyjne w zakresie, formie i terminach określanych każdorazowo w Programie Badań Statystycznych Statystyki Publicznej (PBSSP). Zasilanie statystyki publicznej danymi administracyjnymi jest traktowane jako priorytet w przepisach krajowych (art. 13.5)<sup>107</sup> oraz w międzynarodowych zaleceniach i standardach (GUS, 2014b, s. 54 i nast.).

---

<sup>107</sup> Artykuł 13:

1. Organy administracji publicznej, Zakład Ubezpieczeń Społecznych, Narodowy Fundusz Zdrowia, Komisja Nadzoru Finansowego, a także inne państwowe lub samorządowe osoby prawne, organy rejestrowe oraz inne podmioty prowadzące rejestry urzędowe przekazują lub udostępniają nieodpłatnie służbom statystyki publicznej zgromadzone dane administracyjne w szczegółowym zakresie, postaci i terminach, określonych w programie badań statystycznych statystyki publicznej, w szczególności w postaci zbiorów danych z systemów teleinformatycznych, w tym wyników pomiarów, danych monitoringu środowiska, a w przypadku braku systemu teleinformatycznego – w innej utrwalonej postaci.
2. Organy i podmioty, o których mowa w ust. 1, przekazują dane administracyjne łącznie z informacjami objaśniającymi zawartość informacyjną tych danych (metadane).
3. Organy i podmioty, o których mowa w ust. 1, są obowiązane do:
  - 1) stosowania standardów klasyfikacyjnych, numeru identyfikacyjnego i oznaczeń kodowych przyjętych w rejestrze terytorialnym, o których mowa w rozdziale 6;
  - 2) przekazywania, na wniosek Prezesa Głównego Urzędu Statystycznego, informacji o zawartości informacyjnej eksploatowanych systemów informacyjnych administracji publicznej oraz rejestrów urzędowych;
  - 3) informowania Prezesa Głównego Urzędu Statystycznego o projektowanym zakresie informacyjnym systemów informacyjnych administracji publicznej oraz rejestrów urzędowych na etapie ich tworzenia lub modernizacji.
4. Prezes Głównego Urzędu Statystycznego jest uprawniony do zgłaszania organom i podmiotom, o których mowa w ust. 1, wniosków dotyczących zawartości informacyjnej i wymogów jakości danych administracyjnych w celu umożliwienia wykorzystania tych danych na potrzeby statystyki publicznej.
5. Przy wyborze źródeł danych na potrzeby statystyki publicznej w pierwszej kolejności uwzględnia się dane administracyjne.
6. Prezes Głównego Urzędu Statystycznego jest obowiązany do udostępniania organom administracji rządowej i jednostkom samorządu terytorialnego danych dotyczących realizowanych przez nie zadań, w ujęciu odpowiednim do ustawowego usytuowania poszczególnych zadań publicznych.

---

## 3. JAKOŚĆ SPISÓW<sup>108</sup>

---

### 3.1. Metody oceny jakości spisów ludności

#### 3.1.1. Klasyfikacje błędów i metod oceny jakości spisów

W klasyfikacji błędów statystycznych wyróżnia się dwa zasadnicze rodzaje: błędy losowe i nielosowe<sup>109</sup>. Błędy losowe wynikają z korzystania z próby losowej zamiast z danych dotyczących całej populacji i doczekały się wielu opracowań. W literaturze przedmiotu poświęca się im dużo uwagi, kosztem niedoceny znaczenia błędów nielosowych (Płatek i Särndal, 2001). Jak zauważa Szreder (2009, s. 11), jednym z wyzwań współczesnej statystyki jest konieczność opracowania w badaniach próbkowych dobrych technik pomiaru innych rodzajów błędów, poza błędem losowania. Opinia ta wydaje się słuszna w odniesieniu do wszystkich rodzajów badań statystycznych. Warto więc powtórzyć za Szrederem (2009, s. 11) szczególnie aktualne stwierdzenie Mahalanobisa: „Ogólna postawa jest taka, że błąd nielosowy postrzega się jako coś, co nie dotyczy statystyka, a w każdym razie uważany jest za niezbyt jasną kwestię, którą statystyk nie musi się przejmować”<sup>110</sup>.

---

<sup>108</sup> Rozdział 3 powstał w ramach projektu: Estymacja pośrednia w zakresie badania niepełnosprawności na podstawie NSP 2011. Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/11/B/HS4/01472.

<sup>109</sup> Na temat błędów statystycznych w polskiej literaturze przedmiotu dostępne są dwie monografie Kordosa (1987, 1988) oraz inne opracowania tego autora, np. (Kordos, 2003, 2007, 2016, 2017). Problem ten poruszany jest również w pracach z zakresu metody reprezentacyjnej np. przez Zasępę (1972) i Szredera (2010) czy w licznych artykułach publicystycznych tego autora, np. *Jak rodzi się błąd* (2013). W światowej literaturze i praktyce badań statystycznych jest to problem szczególnej troski. Jej wyrazem są *Deklaracja Jakości Europejskiego Systemu Statystycznego* przyjęta przez Eurostat we wrześniu 2001 r. (EU, 2001), a także regularne konferencje europejskie poświęcone tej problematyce (European Conference on Quality and Methodology in Official Statistics, np. Wiesbaden (2004), Cardiff (2006), Rzym (2008), Helsinki (2010), Ateny (2012), Wiedeń (2014), Madryt (2016), której kolejna edycja w 2018 r. planowana jest w Krakowie), oraz konferencje ONZ.

<sup>110</sup> Cytat z artykułu Szredera (2009), który przytacza także oryginalny tekst Mahalanobisa: „In fact, the general attitude is to look upon the non-sampling error as something, which does not concern the statistician, or in any case is a kind of dirty job, which a highbrow statistician need not

Z kolei błędów losowych zazwyczaj się nie rozważa w przypadku badań wyczerpujących, są jednak opracowania, które wskazują zasadność ich uwzględnienia również w takim przypadku<sup>111</sup>. Mimo że spis ludności jest badaniem pełnym, w dobie transformacji trzeba uwzględnić także występowanie błędów losowych, których znaczenie wzrasta przy zastosowaniu alternatywnych metod. Błędy losowe wymagają uwagi w spisach przeprowadzanych metodą mieszaną, gdy uwzględnia się możliwość integracji danych z różnych źródeł: rejestrów administracyjnych i badań reprezentacyjnych, a także w spisach kroczących. Znaczenie błędów losowych wzrasta w spisach przeprowadzanych przy pomocy innych metod aniżeli tradycyjne.

W spisach przeprowadzanych metodą tradycyjną wśród błędów nielosowych wyróżnia się zasadniczo dwa rodzaje błędów<sup>112</sup>. Są to błędy kompletności (pokrycia<sup>113</sup>) i treści. Wśród błędów kompletności rozważa się trzy typy, a mianowicie: pominięcia, powtórzenia oraz błędne zakwalifikowania jednostek (rys. 3.1). Pominięcia skutkują zazwyczaj niedoszacowaniem (*undercoverage*), a powtórzenia (duplikacje, podwójne spisanie) – przeszacowaniem (*overcoverage*) spisywanej populacji. W większości przypadków błędy pokrycia są ujemne, ponieważ niedoszacowanie zazwyczaj przewyższa przeszacowanie. W analizie błędów pokrycia szczególnie ważne jest uwzględnienie ich źródeł i rozkładu przestrzennego. Jeszcze inna klasyfikacja<sup>114</sup> wprowadza kategorię błędów netto i brutto<sup>115</sup> (Baldrige, Brown, Jones i Keane, 1985; UN, 2010). Błędy treści z kolei dotyczą charakterystyk poszczególnych jednostek i mogą wynikać z niewłaściwych odpowiedzi, opracowania, analizy i prezentacji czy niewłaściwego zakwalifikowania wariantów badanych cech. W przytoczonym zestawieniu klasyfikacji błędów nielosowych, w spisie korzystającym z wielu źródeł danych, uwzględniono także błędy dotyczące danych z badań reprezentacyjnych oraz rejestrów administracyjnych i źródeł pozaadministracyj-

---

bother about” (Mahalanobis, 1951). Professional Training in Statistics, „Bulletin of the International Statistical Institute”, vol. 33, No. 5).

<sup>111</sup> Doskonałym przykładem jest w tej kwestii praca Paradysza (2009) ukazująca rolę błędów losowych w populacji generalnej dla małego obszaru na przykładzie cząstkowych współczynników płodności kobiet i cząstkowych współczynników zgonów.

<sup>112</sup> Szreder wyszczególnia, poza błędem losowania, cztery inne kategorie: błędy pokrycia jednostek badanej zbiorowości przez operat losowania; spowodowany brakiem odpowiedzi respondentów; pomiaru, związany z zarejestrowaniem nieprawdziwych informacji; przetwarzania zebranych danych statystycznych (2009, s. 11).

<sup>113</sup> Określenia błędów kompletności i pokrycia będą stosowane zamiennie.

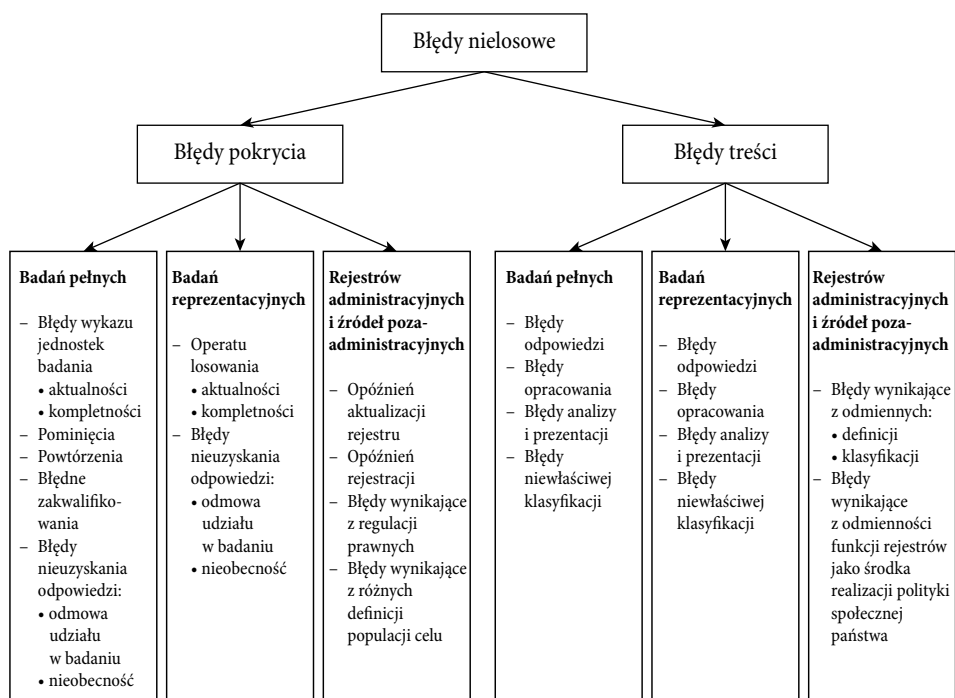
<sup>114</sup> W pracy Zasępy można znaleźć rozróżnienie znacznie większej liczby przyczyn błędów (Zasępa, 1972, s. 271). Dokonuje on też grupowania błędów na (A) spowodowane niedostatecznym przygotowaniem badania, (B) popełnione podczas zbierania danych i (C) błędy w opracowaniu zebranego materiału.

<sup>115</sup> Błędy pokrycia brutto definiuje się jako sumę trzech składników: powtórzeń, błędnych włączeń (klasyfikacji) oraz pominięć. Natomiast błędy pokrycia netto są różnicą pomiędzy wielkością przeszacowania i niedoszacowania. W pracy (Baldrige i in., 1985), na przykładzie analizy struktury wieku ludności, przedstawiono również ujęcie netto i brutto w odniesieniu do błędów treści.



nych. W tym ostatnim przypadku szczególne znaczenie mają definicje i klasyfikacje, które uwzględniają raczej określone cele polityki społecznej danego państwa czy obowiązujące regulacje aniżeli względy porównywalności międzynarodowej, w tym zalecenia ONZ. Wszystkie te błędy mogą rzutować na jakość spisu.

W ocenie jakości spisów podstawowym zagadnieniem jest ocena kompletności. W opublikowanej przez Dywizję Statystyczną ONZ instrukcji dotyczącej oceny jakości spisów (UN, 2010) stwierdza się wprost, że uwzględnia się w niej różne rodzaje błędów i metody ich korekty, lecz największą uwagę poświęca się badaniu kompletności. Również Kordos (2007, s. 564) podkreśla, że kompletność jest podstawowym elementem oceny. Do innych ważnych zagadnień jakości spisów zalicza się także kwestie schematu doboru próby w części badania reprezentacyjnego (długi formularz), projektowania kwestionariusza, planowania i wykorzystania wyników badania kontrolnego, oceny integracji statystycznej, badania zgodności pól, definicji zmiennych, systemu szacunków dualnych, klasyfikacji, tabulacji, oceny błędów treści. W niniejszym opracowaniu skoncentrowano się na kilku zaledwie problemach ewaluacji spisu związanych przede wszystkim z oceną pokrycia, ale także z obserwowaną transformacją metod przeprowadzania spisów wynikającą z korzystania z wielu źródeł informacji oraz stosowania nowoczesnych technologii ICT.



Rysunek 3.1. Klasyfikacja błędów nielosowych w spisach wykorzystujących różne źródła danych

Rozważając metody oceny jakości tradycyjnych spisów ludności, należy zauważyć ich mnogość oraz różnorodność pod względem mierników, rodzajów analiz, stopnia ich złożoności czy dostępności danych wykorzystanych w procesie ewaluacji. Kordos (2007, s. 566) podaje bardzo wiele różnych metod oceny spisów. Można je sklasyfikować według źródeł informacji wykorzystywanych w procesie ewaluacji. Wyróżnilibyśmy wówczas techniki oceny: (i) oparte jedynie na danych przeprowadzonego właśnie spisu, oparte na badaniu wewnętrznej spójności; (ii) oparte na porównaniach z danymi innych istniejących źródeł; a także (iii) wymagające przeprowadzenia dodatkowych badań (tab. 3.1). Konstrukcja tabeli 3.1 jest oparta na propozycji zestawienia metod ewaluacji spisów zaprezentowanej w (Baldrige i in., 1985). Klasyfikacja ta cieszy się dużym uznaniem i z niewielkimi tylko uzupełnieniami nadal jest ujmowana w standardach międzynarodowych (UN, 2010). W tabeli 3.1 wprowadzono dalsze rozszerzenie zarówno w zakresie wykorzystanych w ocenie spisu źródeł danych, jak i metod.

Tabela 3.1. Metody oceny błędów spisu w zależności od liczby źródeł danych wykorzystywanych w procesie ewaluacji

Metody oceny jakości spisów w zależności od wykorzystanych źródeł danych	Rodzaj błędów	
	błędy pokrycia	błędy treści
<b>Jedno źródło</b>		
– analiza demograficzna danych spisowych	×	×
– techniki oceny poprzez badanie wewnętrznej spójności	×	×
– studia interpenetracyjne w ramach spisu powszechnego	×	×
<b>Metody ewaluacji wykorzystujące wiele już istniejących źródeł danych</b>		
(i) Badania wykorzystujące metody statystycznej integracji danych		
– badanie rekordów	×	×
– porównania z wynikami bieżących badań gospodarstw domowych	×	×
(i) Badania niewykorzystujące metod statystycznej integracji danych		
– analizy demograficzne wykorzystujące dane poprzednich spisów	×	×
– porównania z danymi rejestrów administracyjnych	×	×
– porównania z wynikami istniejących badań gospodarstw domowych	×	×
<b>Metody ewaluacji uwzględniające dodatkowe badania</b>		
– spisy kontrolne dla oszacowania błędów pokrycia i treści	×	×
– pospisowa integracja danych	×	×
– powtórne wywiady		×
– dodatkowe badania specyficznych populacji np. mniejszości narodowych i wybranych grup etnicznych	×	×
– badania fokusowe zadowolenia respondentów ze sposobu zbierania informacji w różnych zakresach tematycznych, by poznać wiedzę o ich reakcji i odpowiedziach		×

Źródło: Na podstawie (Baldrige i in., 1985, s. 9; Gołata, 2012; UN, 2010, s. 12 i nast.).

Podstawowa metoda weryfikacji jakości spisu to analiza zgodności szacunków ludności pomiędzy spisami. Niezmiernie ważne jest porównanie wyników spisu z danymi z innych źródeł, w tym poprzednich spisów, bieżących badań gospodarstw domowych oraz danych rejestrów administracyjnych. Gdy dysponujemy odpowiednimi danymi jednostkowymi, istnieje możliwość porównania rekordów. Wówczas dane spisu mogą być łączone (parowane) z danymi z alternatywnych źródeł, co pozwala na specyficzną ocenę dokładności. Ewaluacja danych spisu to także studia penetrujące<sup>116</sup>, analiza interpretacyjna, badanie efektywności spisu, ocena kompletności. W tym ostatnim przypadku szczególne zastosowanie znajduje spis kontrolny określany jako badanie pospisowe. Badanie to może być także wykorzystane do oceny błędów treści. W ocenie jakości spisu w zakresie grup ludności trudnych do zbadania można stosować metody etnograficzne, sieci społecznych, które umożliwiają określenie wpływu mobilności na stopień pokrycia tych specyficznych populacji. Ponadto warto wspomnieć o badaniach satysfakcji wśród użytkowników danych czy badaniach fokusowych przeprowadzanych w celu poznania przyczyn określonych typów odpowiedzi i zachowań respondentów.

W rekomendacjach międzynarodowych (UN, 2015a) zaleca się kompleksowe podejście do oceny jakości spisu, poczynając od określenia celu ewaluacji, następnie poprzez opracowanie programu oceny spisu, wskazanie stosowanych metod, przeprowadzenie oceny, aż po przedstawienie informacji o akceptacji wyników spisu. W państwach o niedoskonałym systemie rejestracji, gdzie liczba ludności podlega szacunkom, jest prawdopodobne, że liczba ludności określona na podstawie spisu będzie różna. W takich sytuacjach wskazane jest wyjaśnienie różnic i ich przyczyn. Wśród metod oceny jakości spisu wyróżnione są dwie grupy: metody analizy demograficznej oraz spis kontrolny<sup>117</sup>. ONZ zaleca stosowanie metod analizy demograficznej począwszy od najprostszycy porównań struktur według płci i wieku (piramidy wieku), poprzez uwzględnienie różnych współczynników demograficznych umożliwiających szacowanie liczby ludności, po skomplikowane modele ludności zastojowej, które mogą być źródłem informacji o zgodności i poprawności wyników spisu. Wyróżnia się tu cztery grupy metod. Pierwsza sprowadza się do międzyspisowych szacunków stanu i struktury populacji na podstawie danych bieżącej rejestracji ruchu naturalnego i wędrownego ludności (urodzenia, zgony i migracje). Druga metoda to zastosowanie projekcji demograficznej opartej na strukturze ludności z poprzedniego spisu z uwzględnieniem szacunków

---

<sup>116</sup> Studia penetrujące mogą być stosowane w celu oceny wpływu określonych operacji spisowych na ogólną wielkość błędu. Metoda ta ma charakter eksperymentalny, służący do oceny efektów różnej organizacji poszczególnych, przenikających się operacji spisowych w działaniach związanych z gromadzeniem danych. Wyniki studiów penetrujących mogą być także wykorzystywane w ocenie dokładności szacunków.

<sup>117</sup> Ze względu na znaczenie spisu kontrolnego jako szczególnie zalecanej metody ewaluacji spisu zostanie on nieco bardziej szczegółowo omówiony w kolejnym podrozdziale.

plodności, umieralności i migracji na podstawie różnych źródeł oraz porównanie uzyskanych szacunków z wynikami ocenianego spisu. Trzecia metoda sprowadza się do porównania wyników analizy kohortowej na podstawie danych poprzedniego spisu z wynikami ocenianego spisu (kohortowe współczynniki przeżycia). Czwarta z wyróżnionych w zaleceniach ONZ metod demograficznych to szacunek korygujących współczynników pokrycia z zastosowaniem metod analizy regresji. Współczynniki te winny zapewnić spójność struktury ludności według dwóch kolejnych spisów (*cohort survival regression method*).

Wiele z wyszczególnionych wyżej metod zachowuje swoją aktualność i przydatność zarówno w przypadku spisu tradycyjnego, jak i przeprowadzonego według podejścia alternatywnego. Należy jednak wyraźnie wyróżnić dwie kwestie:

1. Wykorzystanie innych źródeł, poza ocenianym spisem, znacznie zwiększa ofertę metod badań i analiz. W przypadku dysponowania jedynie danymi ocenianego spisu musimy się w zasadzie ograniczyć do badania wewnętrznej zgodności wyników z uwzględnieniem zasad analizy demograficznej, studiów interpretacyjnych. Możliwości rzetelnej oceny wyników spisu są wówczas bardzo ograniczone, zgodnie z zasadą *eine Quelle – keine Quelle* (Paradysz, 2007a). Uwzględnienie wielu źródeł, chociażby wyników poprzedniego spisu ludności, diametralnie zmienia sytuację.
2. Zmiany w metodologii spisów powodują, że samo jego przeprowadzenie oznacza jednoczesną pracę z wieloma źródłami danych. Automatycznie więc rozszerzają się możliwości ewaluacji. Z uwagi na obowiązkową harmonizację, badanie zgodności pojęć, definicji, konstrukcję tzw. złotego rekordu<sup>118</sup>, statystyczną integrację danych itp. proces ewaluacji jest obecny na wszystkich etapach badania spisowego<sup>119</sup>. Ewaluacja uwzględniająca dodatkowe źródła, a w szczególności badania, które umożliwiają statystyczną integrację danych, dostarcza także nowych możliwości oceny. Korzystanie z niezależnych źródeł danych rozszerza zakres oceny poprzez empiryczną ewaluację w kontekście zgodności rozkładów badanych zmiennych z różnych źródeł, ich wzajemnych relacji, spójności szacunków. Zastosowanie metod statystycznej integracji wykorzystującej dane rejestrów umożliwia nie tylko bardziej wnikliwą i wszechstronną analizę, ale przede wszystkim analizę wolną od powielania błędów ocenianego badania<sup>120</sup>.

<sup>118</sup> Złoty rekord w bazie danych spisowych to rekord zawierający zmienne z różnych źródeł. Są to dane zarówno ze spisu pełnego (czyli pozyskiwane z rejestrów i systemów administracyjnych), jak i z badania reprezentacyjnego. Wszystkie zmienne są poddane szczegółowemu badaniu jakości uwzględniającemu synchronizację, harmonizację, w tym weryfikację pojęć, definicji, kategorii itp.

<sup>119</sup> Nie oznacza to, że w metodzie tradycyjnej badanie jakości nie mogło być prowadzone na każdym etapie spisu. Teraz jednak stanowi ono integralny element procedury statystycznej integracji danych. Jak zauważa Kordos (2007, s. 568), w niektórych państwach przyjmuje się, że zagwarantowanie odpowiednich procedur badania jakości na poszczególnych etapach eliminuje nawet potrzebę przeprowadzania spisu kontrolnego.

<sup>120</sup> Różne przykłady zastosowania analizy demograficznej do ewaluacji wyników NSP 2002 dla Polski ogółem, w odniesieniu do struktur ludności według płci i wieku oraz w przekroju terytorial-

Wallgren i Wallgren (2013, 2015) omawiają ocenę jakości systemów statystycznych opartych na rejestrach jako proces składający się z dwóch części, z których każda ma dwa poziomy. Pierwszy poziom to analiza samego źródła. Zawiera ona omówienie metadanych oraz analizę mikrodanych dotyczących analizowanego źródła w celu ustalenia jego trafności i dokładności. Druga część to analiza porównawcza źródła z rejestrem podstawowym oraz z innymi źródłami w systemie zawierającymi podobne zmienne. Systematyczne porównywanie badań i rejestrów w systemie generuje nową wiedzę o jakości badań, stwarza możliwości ich zmian w celu poprawy jakości.

### 3.1.2. Spis kontrolny

Wśród zalecanych przez organizacje międzynarodowe metod oceny jakości spisu szczególne miejsce zajmuje spis kontrolny<sup>121</sup> (Chipperfield, Brown, i Bell, 2017; Kordos, 2017; Newell i Smallwood, 2010; UN, 2009, 2010, 2013a). Jest to specjalne badanie reprezentacyjne – replika spisu – przeprowadzane w sposób niezależny bezpośrednio po zakończeniu spisu. Zasadniczym celem spisu kontrolnego jest oszacowanie błędu pokrycia zarówno w skali całego kraju, jak i w przekroju terytorialnym (por. też Kordos, 2007). Z tego powodu istotne jest takie zaprojektowanie badania, aby możliwe były również szacunki dla domen wyodrębnionych terytorialnie, a także opartych na podstawowych cechach demograficznych, takich jak wiek i płeć. Dzięki porównaniu precyzyjnych szacunków z wynikami spisu dla wskazanych domen możliwe jest uzyskanie informacji o kompletności, a tym samym oszacowanie wielkości błędów. Równie istotnym zadaniem badania pospisowego jest dostarczenie informacji o błędach treści w stosunku do wybranych charakterystyk, choć zgodnie z sugestią zawartą w brytyjskiej nazwie badania, służy ono jedynie weryfikacji kompletności spisu.

Idea spisu kontrolnego sprowadzająca się do niezależnej repliki spisu właściwego oznacza jednocześnie szereg wymagań odnośnie do organizacji tego badania. Przede wszystkim niezależność badania pospisowego sprowadza się do organizacji badania przez odrębną instytucję, niezależną od tej, która przeprowadzała spis właściwy. Badanie powinno być powierzane najbardziej doświadczonym, najlepszym ankietom, jednak powinna być zapewniona rozdzielność obwodów spisowych, w których przeprowadzali oni spis właściwy, z tymi, w których przeprowadzają badanie kontrolne. Wywiady winny się odbywać ze szczególną starannością,

---

nym, na podstawie różnych źródeł danych (rejestrów administracyjnych, poprzednich spisów czy ewidencji ruchu naturalnego ludności) przedstawiono w raportach podgrupy ds. metod statystyczno-matematycznych na rzecz spisów (powołanej przez prezesa GUS w ramach Projektu PSR 2010 i NSP 2011, która pracowała pod kierunkiem Paradysza) np. (CSR, 2008, 2009a, 2009b) oraz pracy Paradysza (2010).

<sup>121</sup> W Polsce badanie to jest nazywane *spisem kontrolnym* czy *badaniem pospisowym*, w Wielkiej Brytanii tradycyjnie używa się nazwy *census coverage survey* (CCS), a zalecenia ONZ mówią o *post enumeration survey* (PES).

z zachowaniem weryfikacji list adresowych oraz indywidualnym spisywaniem. Należy tutaj uwzględnić wielokrotne podejmowanie próby skontaktowania się z potencjalnymi respondentami (tak często, jak to konieczne, o różnych porach dnia i w różnych dniach tygodnia), natomiast w przypadku nieobecności w spisie właściwym można pozostawić formularz samospisu do przesłania pocztą. Spis kontrolny jest badaniem reprezentacyjnym, a to wiąże się z wymogiem wysokiej jakości operatu losowania, zastosowaniem schematu doboru próby i takiej jej liczebności, by uogólnienie wyników badania pospisowego w przekroju określonych domen było precyzyjne i uwzględniało różne metody estymacji, w tym estymacji dla małych domen. Często praktykowanym rozwiązaniem jest też zastosowanie systemu estymacji dualnej, co z kolei wymaga uwzględnienia właściwych metod integracji danych, harmonizacji definicji i pojęć, zapewnienia, że parowane są odpowiednie jednostki. Zastosowana metodologia oraz wyniki badania pospisowego oczywiście powinny być opublikowane (Gołata, 2016b; Hogan, Cantwell, Devine, Mule i Velkoff, 2013; Kordos, 2007; Plewis, Simpson i Williamson, 2011; UN, 2009).

Przegląd różnych badań poprawności szacunku stanu ludności zawiera praca Newell i Smallwooda (2010), którzy omawiają m.in. takie badania, jak: Post Enumeration Survey (PES), Census Coverage Survey (CCS) i Reverse Record Check (RRC), Dual System Estimation (DSE), Residents Temporarily Overseas (RTO). Są to w zasadzie badania pospisowe czy spisy kontrolne, oraz metody sprawdzania rekordów i podwójnych estymacji, które umożliwiają weryfikację kompletności (pokrycia) szacunków spisu właściwego. Badania takie są prowadzone w wielu państwach, np. w Australii, Kanadzie, Japonii, Nowej Zelandii, Wielkiej Brytanii czy USA (Newell i Smallwood, 2010). Metoda podwójnego systemu szacunków jest oparta na założeniu niezależności między danymi spisu właściwego i spisu kontrolnego (różne metody gromadzenia danych, inny personel i inny, niezależnie określony operat losowania). W metodzie tej stosuje się procedury dopasowujące poszczególne rekordy oraz szczegółową weryfikację poprawności klasyfikacji według wieku, płci i jednostki terytorialnej. Następnie jest konstruowana dwuwymiarowa tablica kontyngencji dla umożliwienia porównania oszacowań. Dysponując danymi z rejestru ludności i badania pospisowego (CCS), można zastosować podobne podejście (tab. 3.2) w celu oceny kompletności spisu (Brown, Abbott i Diamond, 2006; Plewis i in., 2011).

Tabela 3.2. Tablica podstawowa – schemat podwójnego systemu estymacji

Źródło danych	Rekordy	Spis kontrolny		
		obserwowane	nieobserwowane	razem
PESEL – podstawowy Rejestr ludności	Obserwowane	$n_{11}$	$n_{12}$	$n_{1.}$
	Nieobserwowane	$n_{21}$	$n_{22}$	$n_{2.}$
	Razem	$n_{.1}$	$n_{.2}$	$n_{..}$

Źródło: Na podstawie (Brown i in., 2006; Plewis i in., 2011).

Porównując liczebności poszczególnych komórek tablicy kontyngencji, można stwierdzić zgodność bądź rozbieżność szacunków. Przykładowo  $n_{12}$  oznacza liczbę osób zaobserwowanych w rejestrze ludności, ale niewystępującą w spisie kontrolnym. Różnice w tym wymiarze wydają się możliwe i znaczące, ponieważ Polska jest krajem intensywnej migracji, a osoby migrujące (niezależnie od czasu ich pobytu za granicą) są włączone do rejestru, chyba że powiadomią o wyjeździe z kraju, dokonując wymeldowania na pobyt stały. Kolejny rzeczywisty problem może dotyczyć osób, które wyjechały za granicę i tam zmarły. Jeśli nikt nie poinformuje polskiego urzędu o śmierci, osoba taka może „żyć w rejestrze” nawet do 200 lat. W rezultacie, zgodnie z rejestrem, najstarszy człowiek świata żyje w Polsce (Kuc, 2014). Kolejną istotną niespójnością jest liczba  $n_{21}$ , reprezentująca osoby obserwowane w spisie kontrolnym, ale niezarejestrowane w PESEL. Liczba ta może się odnosić do wszystkich cudzoziemców przebywających w Polsce (nawet przez ponad 12 miesięcy), którzy nie mają statusu legalnego rezydenta. Status ten jest związany z zameldowaniem na pobyt stały i wymaga złożenia dokumentu potwierdzającego prawo stałego pobytu. W przypadku niespełnienia wymogów formalnych dotyczących stałego pobytu nawet imigranci mieszkający w Polsce przez ponad 12 miesięcy nie byli liczeni w spisie jako rezydenci.

Przy założeniu niezależności ( $\theta = \frac{n_{11} n_{22}}{n_{21} n_{12}} = 1$ ), w ramach danej grupy płci i wieku,

w obserwowanej jednostce przestrzennej można oszacować kompletność spisu oraz liczbę nieobserwowanych osób  $n_{22}$ . Jednak gdy założenie niezależności nie będzie spełnione, wówczas oszacowanie w systemie dualnym będzie obciążone  $\theta > 1$ , DSE jest niedoszacowane bądź  $\theta < 1$ , DSE jest przeszacowane (Brown i in., 2006). Założenia homogeniczności i niezależności są bardzo mocne i z wielu względów mogą nie być spełnione (Plewis i in., 2011). Różnice prawdopodobieństwa odpowiadającego sytuacji „bycia włączonym” do rejestru są możliwe, szczególnie według płci i wieku oraz w przekroju terytorialnym. Niektóre osoby mogą też być pozbawione możliwości bycia uwzględnionym w rejestrze, np. z powodu określonych regulacji prawnych. Proces łączenia rekordów może też być utrudniony z innych powodów, takich jak migracje, opóźnienia w rejestracji nowo narodzonych dzieci czy po prostu inercji w przekazywaniu aktualizacji rejestru.

Przydatność wyników spisu kontrolnego zależy od poprawności jego przeprowadzenia, co w znacznej mierze jest uzależnione od właściwego schematu losowania próby. W wielu krajach przygotowanie badań kontrolnych jest jednym z elementów całego systemu badania jakości spisu, których przykładami mogą być projekty One Number Census (ONC) po spisie 2001 r. czy Beyond 2011 po spisie 2011 r. w Wielkiej Brytanii albo kanadyjski Census Program Transformation Project (CPTP) czy amerykański CPEX – 2010, Census Program for Evaluations

and Experiments<sup>122</sup>. Celem tych badań, obok oceny jakości spisu i zastosowanych procedur, jest przygotowanie rekomendacji dla przyszłych spisów.

Dobrze zaplanowany i przeprowadzony spis kontrolny jest często podstawą korekty szacunków wyników spisu właściwego. W tym celu są wykorzystywane wskaźniki pokrycia na poziomie kraju oraz w przekroju domen. W rekomendacjach ONZ znajdujemy jednak stwierdzenie, że również w wielu państwach nie dokonuje się korekty spisów na podstawie badań kontrolnych. Podkreśla się, że dla świadomych użytkowników danych społeczno-gospodarczych, planistów i decydentów szacunek błędu pokrycia jest pomocny w rozsądnej interpretacji i wykorzystywaniu wyników spisu. Dodatkowo pomiar błędu treści jest przydatny dla planujących przyszłe spisy w celu przeprowadzenia rewizji zastosowanego kwestionariusza, procedur i sposobu realizacji badania.

W Polsce w dotychczasowej praktyce spisów ludności nie przeprowadzono rzetelnego, spełniającego podstawowe kryteria badania kontrolnego. Nigdy w dotychczasowej praktyce wyniki badań kontrolnych nie zostały wykorzystane do korekty danych spisu<sup>123</sup>. Przynajmniej GUS takiej informacji nie opublikował (Kordos, 2007).

Z dostępnych do tej pory opracowań wyników badań pospisowych w Polsce na uwagę zasługują opracowania Zasępy (1988, 1993). Autor podkreśla, że nawet przy najlepszej organizacji prac spisowych oraz stosowaniu różnych zabezpieczeń kontrolnych należy się liczyć z popełnieniem pewnych błędów. W trosce o jakość danych spisowych będących podstawą ważnych decyzji władz administracyjnych autor podnosi wybrane kwestie metodologiczne, analizuje rozwiązania przyjęte w innych państwach, poddaje pod rozważenie różne metody oceny dokładności wyników spisów, by, jak pisze, „zebrane doświadczenia mogły być wykorzystane do poprawy jakości przyszłego spisu” (Zasępa, 1988).

Zasępa (1988) podkreśla, że pierwszy spis kontrolny przeprowadzony w Polsce w 1978 r. objął stosunkowo małą próbę losową, a porównanie jego wyników z wynikami spisu właściwego dostarczyło informacji o różnicy wynoszącej 0,3%<sup>124</sup>. Zasępa

<sup>122</sup> Projektom oceny jakości badań spisowych poświęca się wiele czasu i uwagi. Są to bardzo kosztowne przedsięwzięcia. Przykładowo w Wielkiej Brytanii rezultaty prac oraz informacje metodologiczne są dostępne na stronie Office for National Statistics (ONS) <https://www.ons.gov.uk/census> jak również współpracujących z ONS ekspertów z Southampton Statistical Sciences Research Institute (S3RI) z Uniwersytetu w Southampton: <http://www.energy.soton.ac.uk/category/research/energy-behaviour/census-2022/>. Informacje nt. Kanadyjskiego programu są dostępne na stronie Statistics Canada: <http://www12.statcan.gc.ca/census-recensement/fc-rf/98-506-x/98-506-x2017001-eng.cfm>. Natomiast raporty dotyczące programów amerykańskich są dostępne na stronie US Census Bureau <https://www.census.gov/2020census>.

<sup>123</sup> Jednym z powszechnych błędów kompletności spisów są braki niemowląt. Prawdopodobnie już w latach 70. GUS dokonywał korekty, zastępując spisane liczby niemowląt liczbą dzieci urodzonych w roku spisu. Jednak informacji takiej nie zamieszczano w publikacjach pospisowych.

<sup>124</sup> Jak podaje Zasępa (1988), dwie trzecie pominięć stwierdzono w Warszawie i kilku innych większych miastach w Polsce.



wynik ten uznaje za pozytywny, ale dodaje, że spis kontrolny był przeprowadzony tą samą metodą co spis podstawowy (kontrolowany), co oznacza, że przyczyny pominięć mogły się powtarzać. Uważa, że zbieżności wyników obu spisów można było upatrywać również w fakcie przeprowadzenia spisu kontrolnego przez ten sam personel, osoby współodpowiedzialne za spis na danym terenie, a więc zainteresowane wykazaniem braku błędów pokrycia.

Z kolei w opracowaniu Szablowskiego, Wieczorkowskiego i Wesołowskiego (1996) znajdujemy propozycję zastosowania indeksu zgodności w ocenie jakości mikrospisu 1995 r. W opinii Nowak (1998) w spisach kontrolnych 1988 i 1995 r. zarówno wielkość próby, jak i sposób jej doboru były „odpowiednie”. Spis kontrolny po mikrospisie 1995 r. przeprowadzono metodą reprezentacyjną. Z próby wylosowanej do mikrospisu badaniem objęto 1176 wylosowanych obwodów spisowych. W wylosowanych obwodach uwzględniono wszystkie nieprzebadane mieszkania, których łącznie było 12 600. Niestety nie są dostępne oceny tych spisów ani wyniki badań pospisowych.

Spis kontrolny do NSP 2002 przeprowadzono w czerwcu 2002 r. Badanie objęło 903 spośród 177 591 obwodów wybranych w losowaniu proporcjonalnym warstwowym. Zbadano 60 029 mieszkań oraz 192 tys. osób. W spisie kontrolnym uwzględniono 27 cech. W publikacjach, które się ukazały bezpośrednio po NSP 2002, znajdujemy deklaracje, że wyniki spisu kontrolnego posłużą do analizy i oceny jakości danych spisowych. Obiecano, że efekty takiej analizy będą prezentowane w kolejnych ogólnopolskich publikacjach tematycznych, a także zostaną udostępnione na stronie internetowej GUS. Szczegółowe zasady metodologiczne spisu kontrolnego miały zostać przedstawione w publikacji poświęconej metodologii spisu ludności i mieszkań 2002 r. Wyniki spisu kontrolnego nie zostały jednak opublikowane. Także w przygotowanym przez ONZ zestawieniu państw przeprowadzających spisy kontrolne brakuje informacji o spisie kontrolnym dla NSP 2002 (UN, 2010, s. 22).

Nie przedstawiono również wniosków z tego badania dla kolejnego spisu w 2011 r. W opracowaniu prezentującym metodologię NSP 2011 (GUS, 2014b, s. 306) znajdujemy informację, że „celem spisu kontrolnego było sprawdzenie kompletności przeprowadzonego spisu właściwego w zakresie badania reprezentacyjnego, a także dostarczenie informacji służącej do oceny błędów nielosowych, a przede wszystkim oceny błędów pokrycia popełnianych z powodu podwójnego spisywania osób, ewentualnego dopisywania fikcyjnych osób lub opuszczania osób oraz błędów treści (wynikających z braków odpowiedzi lub błędów w odpowiedziach)”. Spis kontrolny do NSP 2011 przeprowadzono w dniach od 1 do 11 lipca 2011 r. Próba losowa do spisu kontrolnego wynosiła 3,0% (ok. 80 tys.) ogólnej liczby 2744 tys. mieszkań wylosowanych wcześniej do badania reprezentacyjnego. Były to mieszkania, w których przebywała przynajmniej jedna osoba z przypisanym numerem telefonu stacjonarnego lub komórkowego oraz wszystkie mieszkania, których

mieszkańcy spisali się przez internet. Formularz zawierał 14 pytań. Ponadto spis kontrolny został przeprowadzony przez ankierów przez telefon (metodą CATI) z wykorzystaniem elektronicznej aplikacji. Sposób realizacji spisu kontrolnego był zgodny z ustawą (Ustawa, 2010, s. art. 22 ust. 2), w której mowa o tym, że spis jest przeprowadzany drogą wywiadu telefonicznego. Zatem w żadnym wymiarze spis kontrolny nie spełniał warunków określonych dla starannie zaprojektowanego badania reprezentacyjnego, przeprowadzonego przez najlepiej przeszkolony personel, przy zastosowaniu najbardziej precyzyjnej metody wywiadu. Nie mógł więc zapewnić osiągnięcia celu, jakim było sprawdzenie kompletności przeprowadzonego spisu właściwego, poprawności uzyskanych danych oraz zgodności tych danych ze stanem faktycznym.

### 3.2. Problemy jakości w spisach nowej generacji

W spisach powszechnych rundy 2010 coraz powszechniej korzystano z wielu różnych źródeł danych, rejestrów administracyjnych, badań reprezentacyjnych i źródeł pozastatystycznych. Alternatywna metoda przeprowadzania wymaga jednocześnie zastosowania wielu metod statystycznych, m.in.: imputacji danych, łączenia rekordów, kalibracji, estymacji pośredniej przy użyciu zmiennych pomocniczych ze źródeł zewnętrznych. Realizacja poszczególnych metod może stanowić jedno z kryteriów oceny spisu, np.: poprawa jakości rejestrów, dokładność szacunków, obniżenie kosztów, wykorzystanie nowoczesnych technologii informacyjno-komunikacyjnych w gromadzeniu i rozpowszechnianiu danych. Z uwagi na nową metodologię, w tym nowe źródła danych, ocenę jakości można rozpatrywać w sposób strukturalny. W spisie austriackim (Berka i in., 2012) zaproponowano trójstopniowe podejście do oceny i poprawy jakości surowych, połączonych i imputowanych danych w trzech hiper-wymiarach: (dokumentacja ( $HD_D$ ), przetwarzanie wstępne ( $HD_P$ ) i zewnętrzne źródło ( $HD_E$ )). Podejście takie pomogło spełnić wymagania transparentności, dokładności i odpowiedniości danych (Lenk i in., 2014). Omawiając zmianę metodologii spisu, niektórzy autorzy wskazują na potrzebę konceptualizacji pomiaru jakości (Zhang, 2011).

Zagadnieniu jakości spisu w Statistics Austria poświęcano wiele uwagi już na etapie przygotowań (Lenk i in., 2014). Metoda gromadzenia danych z istniejących rejestrów administracyjnych wymagała zupełnie nowej koncepcji oceny jakości. W tym celu wypracowano podejście oceniające jakość danych w poszczególnych rejestrach i w powiązanej z nimi spisowej bazie danych na poziomie indywidualnych charakterystyk dla różnych wymiarów jakości: ( $HD_D$ ), ( $HD_P$ ), ( $HD_E$ ). Ideą tego podejścia było stworzenie wskaźnika jakości dla każdej charakterystyki spisu. Jednocześnie proces poprawy jakości prowadzono w każdym rejestrze, poddając ocenie dokumentację, gromadzenie i przechowywanie danych oraz błędy formalne. Ponadto badano spójność danych rejestru z innymi źródłami zewnętrznymi. Proces

oceny jakości ściśle powiązано z procesem przetwarzania danych, zapewniając jednak jego niezależność w tym sensie, by ocena pojedynczego rejestru oraz spisu opartego na rejestrach były odrębne.

Jak zauważają Berka i in. (2012) oraz Wallgren i Wallgren (2013), podczas budowy złotego rekordu w spisie austriackim 2011 r. przeprowadzono bardzo szczegółowe badania i analizę porównawczą. Zebrano i przeanalizowano metadane dotyczące około 300 różnych rejestrów administracyjnych. Wszystkie zmienne w tych systemach zostały ocenione pod kątem możliwości uzyskania informacji na temat populacji mieszkań i budynków, zgodnie z zaleceniami i klasyfikacjami Europejskiej Komisji Gospodarczej Organizacji Narodów Zjednoczonych (UNECE) i Eurostatu (UNECE, 2006a). Także w ramach przygotowań do spisu ludności z 2011 r. w Polsce Główny Urząd Statystyczny zgromadził dane wielu rejestrów administracyjnych i przeprowadził na dużą skalę badania ich przydatności do celów spisu w zakresie zgodności definicji i klasyfikacji (Gołata, 2009; GUS, 2014b; Józefowski i Rynarzewska-Pietrzak, 2010; Kobus, Smolka i Nowakowska, 2009). W wyniku szczegółowej analizy wielu rejestrów ostatecznie do spisu wybrano 28. Wśród nich należy wymienić następujące systemy: Powszechny Elektroniczny System Ewidencji Ludności (PESEL), Centralny Rejestr Zakładu Ubezpieczeń Społecznych (ZUS), Kasy Rolniczego Ubezpieczenia Społecznego (KRUS), rejestr Krajowej Ewidencji Podatników (KEP), rejestr Narodowego Funduszu Zdrowia (NFZ), Krajowy Rejestr Urzędowy Podziału Terytorialnego Kraju (TERYT), dane z Państwowego Funduszu Rehabilitacji Osób Niepełnosprawnych (EKSMoON). Informacje zebrane ze źródeł administracyjnych, które zostały odpowiednio uporządkowane i podzielone na warstwy, zostały również wykorzystane do stworzenia operatu losowania w badaniu reprezentacyjnym.

Przy zmianie metodologii spisu niezmiernie ważna jest społeczna akceptacja i pozytywna ocena. Tymczasem opinie i uwagi społeczności naukowej były bardzo zróżnicowane: niektóre opinie wyrażały pełne poparcie i uznanie, inne były negatywne. Tradycyjny spis ludności był postrzegany przez społeczeństwo, samorząd, a także przez wielu naukowców jako niepodważalne źródło „pewnych” i niekwestionowanych informacji (Barwiński, 2014; *Raport*, 2011). Istnieją jednak również wyraźne oceny wskazujące, że poprzednie ustalenia nie były idealne ze względu na błędy kompletności (Sakson, 2002; Slany, 2014; Śleszyński, 2004a, 2005) oraz ze względu na fakt, że dane nie zostały zebrane bezpośrednio od respondentów (Paradysz, 2002b). Paradysz (2010) podkreśla potrzebę krytycznej oceny poprzednich spisów i sugeruje wykorzystanie wszystkich dostępnych źródeł danych w celu poprawy szacunków spisu. Z drugiej strony brak informacji z rejestrów dla rodzin i gospodarstw domowych, a także niedostępność tych danych dla szczegółowego podziału terytorialnego<sup>125</sup> często się uważa za wadę podejścia mieszanego (Gołata, 2013b).

<sup>125</sup> Informacje takie były dostępne jedynie z badania reprezentacyjnego.

Błędy pokrycia odnoszą się zarówno do zaniżenia, jak i zawyżenia liczby ludności z powodu pominięć, podwójnego spisowania lub błędnego włączenia. W tradycyjnym spisie powszechnym typowa sytuacja dotyczyła niedoszacowania liczby ludności (Paradysz, 2002b, 2010). Wytyczne operacyjne dotyczące spisów kontrolnych (UN, 2010) prezentują zastosowanie różnych procedur wraz z wynikami uzyskanymi dla wybranych krajów, w większości niedoszacowanie (*undercoverage*). Jeśli chodzi o spis oparty na rejestrach, podkreśla się znaczenie skrupulatnej analizy rezydencji w celu wyeliminowania wszystkich osób (rekordów), dla których nie uzyskano „znaku życia” (Lenk, 2008, 2014). W Statistics Austria wszystkie osoby ujęte tylko w Centralnym Rejestrze Ludności, a nieaktywne w innych źródłach zostały poproszone w formie pisemnej o potwierdzenie głównego miejsca zamieszkania. W wyniku tej procedury początkowa populacja została umniejszona o około 0,5%. Niestety, GUS nie dostarczył żadnych informacji na temat stosowania podobnej procedury, aby uniknąć przeszacowania liczby ludności w spisie z 2011 r. w Polsce.

Może warto w tym miejscu podkreślić, że analiza jakości spisów w 2011 r. w wielu państwach UE wykazała istotne problemy dotyczące danych bądź metadanych (EU, 2014, s. 5). W celu zidentyfikowania wspólnych problemów podjęto inicjatywę wymiany raportów jakości pomiędzy urzędami statystycznymi oraz przygotowania wspólnego raportu przez Eurostat. Wyniki analiz mają służyć poprawie jakości spisów planowanych w państwach UE na 2021 r.

### 3.3. Jakość spisów tradycyjnych w Polsce

Kwestia jakości tradycyjnych spisów jest stosunkowo rzadko podejmowana i mało popularna. Wręcz przeciwnie, jak podkreślono omawiając kontrowersje dotyczące podstawowych właściwości i charakterystyk oraz zapotrzebowanie na dane spisowe, przekonanie społeczne o nieomyślności tradycyjnego spisowania jest głęboko zakorzenione, przynajmniej w polskim społeczeństwie.

W Polsce okresu międzywojennego przeprowadzono dwa spisy ludności<sup>126</sup>: 30.09.1921 r. i 9.12.1931 r., których doniosłość i znaczenie są nie do przecenienia, aczkolwiek ocena wartości ich wyników nie jest jednoznaczna. Jak szacuje Krzywicki (2002), odchylenie danych dotyczących ludności w spisie 1921 r. nie przekraczało 3%. Autor podkreśla jednak znaczne niedoszacowanie liczby ludności w miastach, np. w Krakowie o 2,2–2,7%, w Przemyśle o blisko 8%, w Tarnowie nawet o 10% (Krzywicki, 2002, s. 135–139), oraz wskazuje powiaty i gminy, w których ludność wiejska uchyliła się od zeznań statystycznych (Krzywicki, 2002, s. 133–134). Ocenę kompletności szacunku ludności w miastach Krzywicki przeprowadził na

<sup>126</sup> Kolejny, trzeci spis ludności był planowany na 1941 r.

podstawie porównań ze statystyką aprowizacyjną, którą dodatkowo poddał rewizji. Z kolei jeśli chodzi o dane spisu w 1931 r., poważny problem wiąże się z pogłoską o sfałszowaniu jego wyników, szczególnie w zakresie struktury narodowościowej. Informacje takie przedstawia Szturm de Szterm (2002) w artykule opublikowanym przez *Tygodnik Polski – The Polish Weekly* nr 1, 16 marca 1946 r., który opatrzył wnikliwym komentarzem Tomaszewski. Szerzej o kwestiach narodowościowych w spisie 1931 r. pisze także Buławski (2002a, 2002b).

Jak zauważa Kordos (2007), jakiś sposób weryfikacji spisów ludności był stosowany w każdym z polskich spisów. W przypadku pierwszych spisów, przeprowadzonych w latach 1921, 1931, 1950, 1960 i 1970, stosowano raczej niewielkie próby wybrane w sposób nielosowy (Kordos, 1988). Zadania oceny kompletności w polskich spisach ludności okresu powojennego podejmowali się m.in.: Jończy, (2010), Kędelski (1990), Kordos (2007), Sakson (2002), Śleszyński (2004a, 2004b, 2005), ale nie sposób pominąć opracowań Paradysza (1992, 1999, 2002a, 2002b, 2009, 2010). Paradysz (2010) stwierdza, że błędy pokrycia o różnym charakterze występowały we wszystkich spisach powojennych i zapewne także wcześniej (w okresie międzywojennym). Zauważa też, że w ocenie kompletności wszystkich tych spisów dostępnych jest co najmniej pięć rodzajów źródeł danych:

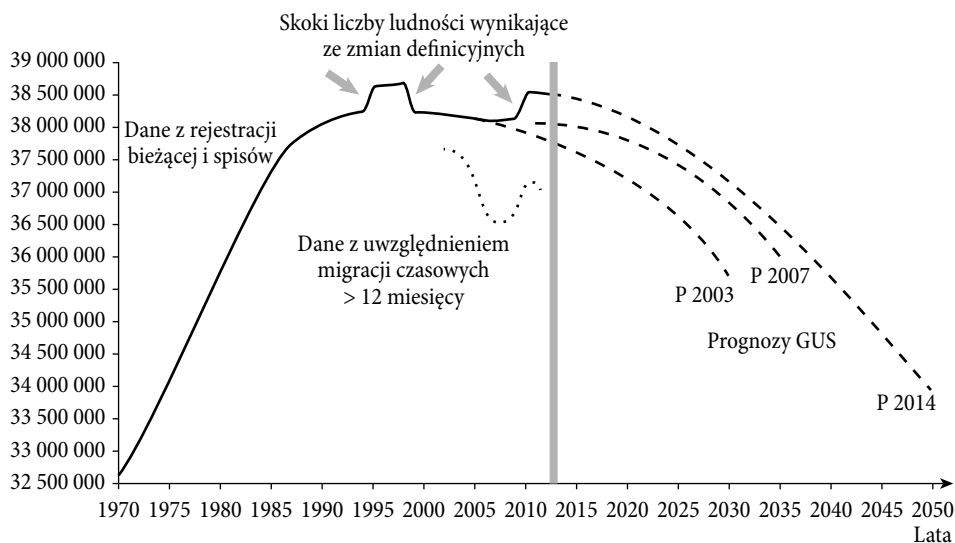
- bieżąca rejestracja zdarzeń demograficznych,
- dane poprzedniego spisu ludności,
- bieżące struktury ludności (z bilansu ludności),
- struktury oszacowane na podstawie regularnie przeprowadzanych badań specjalnych (Badanie Aktywności Ekonomicznej Ludności – BAEL, Badanie Budżetów Gospodarstw Domowych – BBGD)<sup>127</sup>,
- dane rejestru ludności – PESEL.

Z kolei Śleszyński (2014) analizuje szacunki liczby ludności na podstawie bilansów i spisów z perspektywy prognoz. Zauważa on, że wyniki kilku ostatnich prognoz GUS wskazują na „ciągłe przesuwanie depopulacji w przyszłość” (rys. 3.2).

W opinii Śleszyńskiego (2014) wartości stanu ludności w prognozie z 2014 r. (GUS, 2014a) dla roczników powyżej 2020 r. są przesunięte o około 10 lat w stosunku do wcześniejszych projekcji. Zdaniem autora, wynika to z faktu, że stanem wyjściowym dla przedstawionej przez GUS prognozy jest ludność według kryterium formalnego<sup>128</sup> opartego na zameldowaniu. Uważa on, że to przesuwanie nie jest uzasadnione merytorycznie niedoskonałościami bieżącej ewidencji ludności. Silne skoki szacunków liczby ludności o amplitudzie 300–400 tys. osób są spowodowane zmianami definicyjnymi (w latach: 1994/1995, 1998/1999, 2009/2010, por. rys. 3.2).

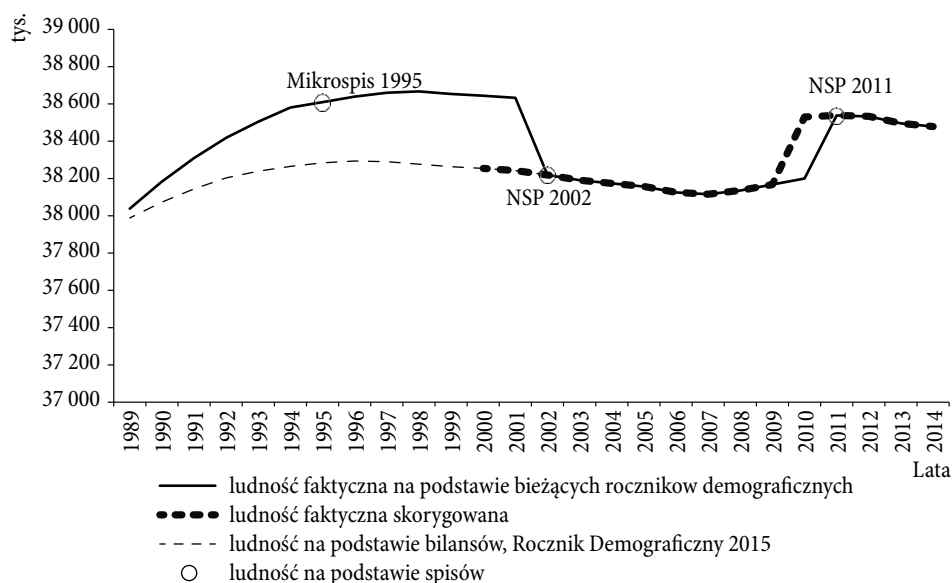
<sup>127</sup> Badanie budżetów gospodarstw domowych również byłoby możliwe. Należy jednak zachować daleko idącą ostrożność z uwagi na duże odsetki odmów.

<sup>128</sup> Śleszyński wprowadza tutaj określenie kryterium formalnego, które odpowiada ludności faktycznej (por. rozdz. 1.2.3).



Rysunek 3.2. Zmiana liczby ludności Polski w latach 1970–2050 według rejestrowanych i szacowanych przez GUS stanów liczby ludności oraz trzech ostatnich prognoz GUS: P 2003, P 2007, P 2014

Źródło: (Śleszyński, 2014, s. 153)



Rysunek 3.3. Liczba ludności Polski w latach 1989–2014: rozbieżności szacunków według spisów i bilansów

Źródło: Na podstawie danych GUS

Analiza szacunku stanów ludności opartych na wynikach bieżącej ewidencji (na podstawie danych bieżących roczników demograficznych) oraz bilansów uwzględniających korekty stanów sporządzone na podstawie danych spisu ludności (tab. 3.3, rys. 3.3) potwierdza wnioski Śleszyńskiego. Zauważmy sięgającą ponad 390 tys. osób różnicę między stanem ludności na 31.12.2001 r. przed korektą i po korekcie danymi NSP 2002. Natomiast dla 2011 r. obserwujemy dokładnie przeciwny kierunek korekty, ale jej wielkość jest zbliżona – 330 tys. osób. Zmiana relacji pomiędzy szacunkiem stanu ludności wg spisu i metody bilansowej jest prawdopodobnie wynikiem zmiany metodologii przeprowadzania spisu z tradycyjnej, uwzględniającej kategorię ludności deklaratywnej, na metodę mieszaną, która jest oparta na ludności ujętej w rejestrze.

Tabela 3.3. Rozbieżności szacunków liczby ludności według spisów oraz bilansów (w tys.)

Spis/Bilans	Stan wg bilansu na dzień		Stan wg spisu
	31.12.2001	31.12.2002	
NSP 20 maja 2002			20.05.2002
Stan przed spisem	38 632,5	38 218,5	38 230,1
Stan po korekcie	38 242,2	38 218,5	
Różnica: bilans – spis	390,3		
NSP 1 kwietnia 2011			31.04.2011
Stan przed spisem	38 200,0	38 538,3	38 538,5
Stan po korekcie	38 529,9	38 538,5	
Różnica: bilans – spis	-329,9		

Źródło: Na podstawie danych GUS.

Śleszyński (2014) uważa, że projekcje powielają niedoszacowania i przeszacowania stanów ludności wynikające z trwałego pozostawania poza granicami kraju mimo formalnego zameldowania. Na podstawie danych NSP 2011 poza granicami Polski na okres powyżej 12 miesięcy pozostawało 1 564 580 osób. Dlatego, zdaniem autora, dla celów praktycznych bardziej wartościowa będzie prognoza oparta na definicji ludności rezydującej (GUS, 2015). Niestety, okazuje się, że definicja osoby rezydującej w ujęciu polskiej statystyki publicznej jest oparta na kryterium „formalnym” (por. rozdz. 1.2.3 i 1.2.4). Są podstawy do obaw, że wielu odbiorców zostało wprowadzonych w błąd, przyjmując, że stosowane pojęcie ludności rezydującej nie uwzględnia migracji długookresowych. Jest to bardzo poważny problem, który bezwzględnie wymaga wyjaśnienia, gdyż obciąża nie tylko prognozę, ale także wszystkie bieżące szacunki, nie tylko demograficzne, ale także ekonomiczne i społeczne<sup>129</sup>.

<sup>129</sup> Wróćmy do tego problemu na początku kolejnego podrozdziału.

Zastosowanie metod analizy demograficznej stwarza w zasadzie nieograniczone możliwości porównań i oceny wyników spisu. W zależności od przedmiotu oceny i zastosowanej procedury Paradysz (2010) wyróżnia m.in. „metodę zmartwychwstałych trzydziestolatków”, „zgonów duchów”, „oficjalnych struktur”. W wyniku oceny przeprowadzonej na podstawie Badania Dzietności kobiet NSP 1970 Paradysz (2002c) stwierdził bardzo duże braki kobiet o najkrótszym stażu małżeńskim, które sięgały nawet 30%<sup>130</sup>. Autor zauważył także występowanie błędów systematycznych polegających na nadreprezentatywności „okrągłych” lat zawarcia związku małżeńskiego (1960, 1965, 1970) przy jednoczesnych „wyrwach” w latach sąsiednich (1959 i 1961, 1964 i 1966, 1969), które ocenia jako efekt pozyskiwania danych „z drugiej ręki” (Paradysz, 2002b, 2010).

Natomiast w odniesieniu do NSP 2002 Paradysz (2010) zauważa znaczne rozbieżności między poszczególnymi kategoriami ludności spisowej: faktycznej i stałej oraz deklaratywnej ludności faktycznej i deklaratywnej ludności stałej, które określa mianem „szarej strefy”. Szczególnie istotne rozbieżności dotyczą mieszkańców dużych miast będących ośrodkami akademickimi, gdzie okresowo na kwaterach prywatnych zamieszkuje duża liczba młodzieży. Zauważone przez Paradysza rozbieżności potwierdzają wyniki analiz Śleszyńskiego (por. 2004a, 2004b, 2005), który badał różnice w liczbie ludności z szacunków spisowych i z szacunków na podstawie bieżącej ewidencji. Różnice te są zazwyczaj przypisywane stosowaniu odmiennej metodologii w obu badaniach. Autor zwraca także uwagę na znaczenie nierejestrowanej migracji. Szczególnie duże różnice w ujęciu terytorialnym dotyczą wielkich aglomeracji miejskich Warszawy (ok. 100 tys.) i Katowic (70 tys.). Przyczyn ujawnionych rozbieżności Śleszyński poszukuje także w zróżnicowanym terytorialnie udziale w badaniu – odsetek braków odpowiedzi sięga 20% (przy średniej dla Polski wynoszącej 2%) i jest zróżnicowany według miejsca zamieszkania, poziomu wykształcenia i źródła utrzymania (por. Śleszyński 2004a, 2004b, 2005).

Najbardziej kompleksową ocenę kompletności w odniesieniu do dotychczasowych spisów ludności przedstawił Paradysz w wielokrotnie cytowanym tu opracowaniu (2010). Objęła ona ocenę:

- kompletności małych dzieci i młodzieży z wykorzystaniem informacji o urodzeniach z ostatnich lat przed spisem oraz rejestracji ruchu ludności;
- kompletności z wykorzystaniem informacji z poprzedniego spisu oraz rejestracji ruchu ludności;
- pominięć ludzi starych w spisach.

Wnioski z przedstawionego przez Paradysza badania można w sposób syntetyczny ująć następująco:

<sup>130</sup> Paradysz sformułował hipotezę, że przyczyną tych błędów może być mobilność postmatrymonialna młodych małżeństw.



- NSP 1988 – dochodzące do 30% braki kobiet o najkrótszym stażu małżeńskim;
- NSP 2002 – brak ponad 10% najmłodszych niemowląt w wieku do 6 miesięcy;
- NSP 2002 – pominięcia ludności o podwyższonej mobilności (20–29 lat);
- NSP 2002 – brak ludzi sędziwych powyżej 90. roku życia.

W kolejnym podrozdziale podjęta zostanie próba podobnej oceny kompletności w odniesieniu do ostatniego spisu z 2011 r., który został przeprowadzony przy zastosowaniu metody mieszanej. Badanie to nie stanowi porównania jakości danych w zależności od metody spisu, gdyż wyniki każdego spisu są efektem oddziaływania wielu czynników, również o charakterze społecznym i gospodarczym, np. takich jak intensywność migracji po wstąpieniu Polski do Unii Europejskiej. Niemniej jednak przedstawiona analiza może być przydatna w ocenie jakości badań i danych spisów ludności w Polsce.

### 3.4. Ocena jakości NSP 2011

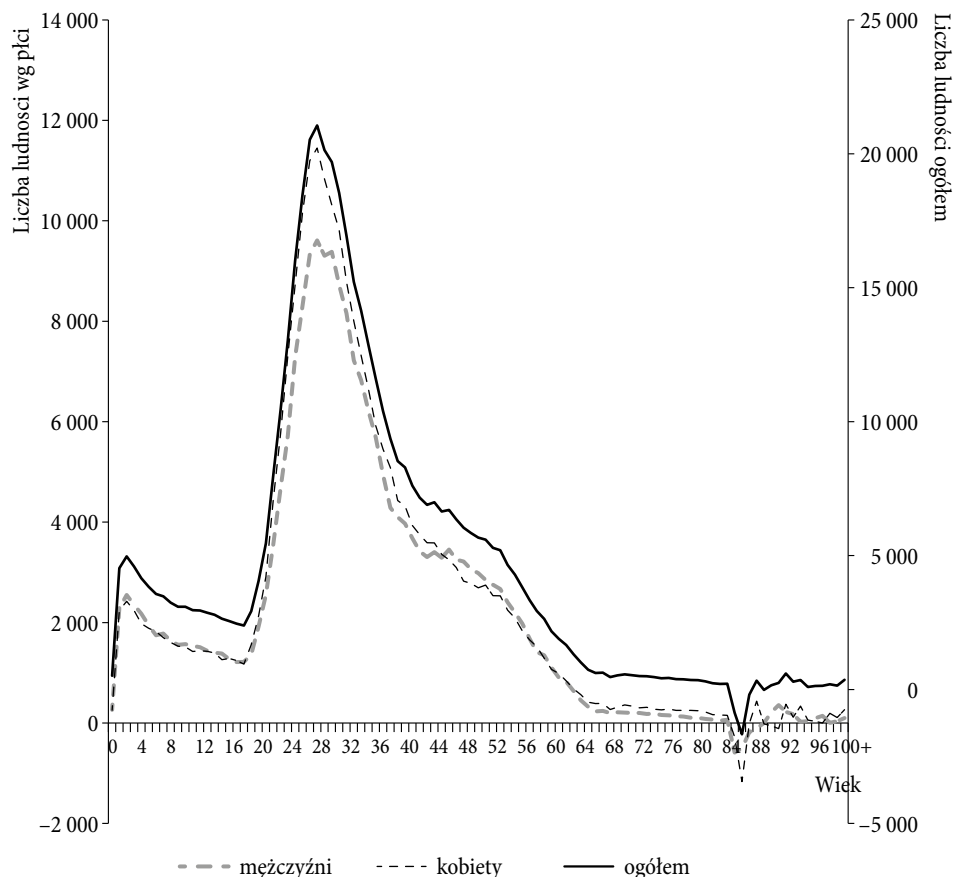
#### 3.4.1. Ocena kompletności NSP 2011 z wykorzystaniem metod analizy demograficznej

Poniższej podjęto próbę syntetycznej, acz w miarę kompleksowej oceny jakości spisu ludności przeprowadzonego w 2011 r. Poniższe badanie zostało przeprowadzone na podstawie ogólnie dostępnych informacji z publikacji spisowych oraz innych komunikatów zamieszczanych na stronie GUS.

Zasadniczą kwestią w ocenie jakości NSP 2011 jest odniesienie do oszacowania liczby ludności. Jak zauważono wcześniej omawiając uwagi Śleszyńskiego (2014), szacowanie i prognozowanie liczby ludności jest niezmiernie złożonym i trudnym problemem. Jakość prognoz zależy od wielu czynników, przyjętych założeń dotyczących kształtowania się podstawowych procesów demograficznych, stosowanych metod, ale przede wszystkim od przyjętego stanu wyjściowego. Ponieważ dane spisu ludności są przyjmowane jako najbardziej wiarygodne, stanowiące podstawę korekty stanu ludności wynikającego z bilansu, to właśnie one są często podstawą przygotowywanych prognoz. Tak też było w przypadku prognoz sporządzonych przez GUS w 2014 i 2015 r. (GUS, 2014a, 2015). Badania Śleszyńskiego kończy konkluzja o większej przydatności prognoz opartych na stanie wyjściowym określonym dla ludności rezydującej. Podobne były też konkluzje rozważań w podrozdziale 1.2.4 oraz wyniki badań Bijaka i Kupiszewskiego wyraźnie wyartykułowane w ocenie założeń prognozy i wpływu migracji zagranicznej na bilans demograficzny Polski (Kupiszewski i Bijak, 2012).

Przedstawiając liczbę ludności według NSP 2011, znajdujemy ludność faktyczną wynoszącą 38 511 824 osób oraz szacunek ludności rezydującej w wysokości 38 044 565 osób (tab. 1.4, rys. 3.4). Uwzględniając jednak diskutowane kategorie

i definicje ludności (podrozdział 1.2.3) oraz szacunki migracji w NSP 2011, naturalne jest oczekiwanie innych wielkości (rys. 3.5).



Rysunek 3.4. Różnica między liczbą ludności rezydującej a faktycznej według płci i wieku, NSP 2011

Źródło: Na podstawie (Gołata, 2016a)

Przedstawione przez GUS ostateczne szacunki ludności rezydującej na podstawie NSP 2011 wskazują, że wielkość ta jest niższa od liczby ludności faktycznie zamieszkałej o 467 259 osób (223 481 mężczyzn i 243 778 kobiet). Z kolei rozkład według wieku dostarcza informacji, że największe rozbieżności pomiędzy obydwo- ma kategoriami dotyczą ludności w wieku od 20 do 50 lat (rys. 3.4). Przy czym w wieku 20 lat następuje gwałtowny wzrost różnicy, która osiąga maksimum dla wieku 28 lat. W odniesieniu do osób w wieku od 23 do 37 lat liczba rezydentów jest o ponad 10 tys. niższa od liczby osób według kryterium faktycznego zamieszkania, dla pojedynczego rocznika wieku. Jednak wobec ponad półtora miliona

osób, które przebywają za granicą dłużej aniżeli 12 miesięcy, nieprzekraczająca pół miliona osób różnica pomiędzy ludnością faktyczną a rezydującą jest zaskoczeniem (tab. 3.4).

Tabela 3.4. Szacunek liczby ludności rezydującej ogółem, NSP 2011

Kategoria ludności	Szacunek NSP 2011
Ludność faktyczna – LF Stali mieszkańcy Polski, w tym osoby przebywające czasowo za granicą (bez względu na czas), ale zameldowane na stałe	38 511 824
Emigranci przebywający za granicą powyżej 3 miesięcy – E(+3)	2 017 501
Emigranci przebywający za granicą powyżej 12 miesięcy – E(+12)	1 564 580
Imigranci przebywający w Polsce powyżej 3 miesięcy – I(+3)	40 097
Imigranci przebywający w Polsce powyżej 12 miesięcy – I(+12)	27 215
Ludność rezydująca – szacunek GUS – LR	38 044 565
Szacunek ludności rezydującej RAF według wynikającego z definicji następującego toku rozumowania: LF – E(+12) + I(+12) = 38 511 824 – 1 564 580 + 27 215 = 36 974 459 = RAF	36 974 459

Źródło: Na podstawie (Gołata, 2016a).

Oczekiwanie innych wielkości ma swoje źródło w różnicy między definicją ludności faktycznej oraz rezydującej w NSP 2011. Przypomnijmy, że kwestia dotyczy sposobu ujmowania osób przebywających za granicą. W odniesieniu do kategorii ludności faktycznie zamieszkałej długość czasu pobytu za granicą nie ma znaczenia. Natomiast w ujęciu ludności rezydującej pojawia się kryterium co najmniej 12 miesięcy. Można więc wnioskować, że różnica między ludnością faktyczną a rezydującą w NSP 2011 wynosi ponad 1,5 mln, a jej szacunek jest mniej więcej zgodny z wynikającą z definicji logiką rozumowania (tab. 3.4). Stosując podejście przedstawione w tab. 3.4, otrzymujemy „surowy szacunek liczby ludności rezydującej” (RAF) wynoszący 36 974 459 osób (Gołata, 2016a).

Poszukiwanie źródła zaobserwowanych rozbieżności wiązało się z wnikliwą analizą metod szacunku emigrantów i imigrantów w NSP 2011 oraz indywidualnymi konsultacjami z pracownikami GUS. Stwierdzono, że zgodnie z polskimi przepisami cudzoziemcy przebywający w Polsce dłużej niż 12 miesięcy nie są zaliczani do kategorii rezydentów, o ile nie mają statusu rezydenta. Posiadanie statusu rezydenta wiąże się przede wszystkim z rejestracją na pobyt stały, który wymaga złożenia dokumentu potwierdzającego prawo stałego pobytu. Ta dokumentacja różni się w zależności od kraju pochodzenia imigranta i może wymagać skomplikowanych procedur<sup>131</sup>. W przypadku niespełnienia wymogów formalnych

<sup>131</sup> Dla osoby spoza Unii Europejskiej dokumentacja ta zawiera m.in. zezwolenie na długoterminowe miejsce pobytu w UE, decyzję o przyznaniu statusu uchodźcy w Rzeczypospolitej Polskiej, przyznanie ochrony uzupełniającej lub zezwolenie na pobyt tolerowany w Polsce.

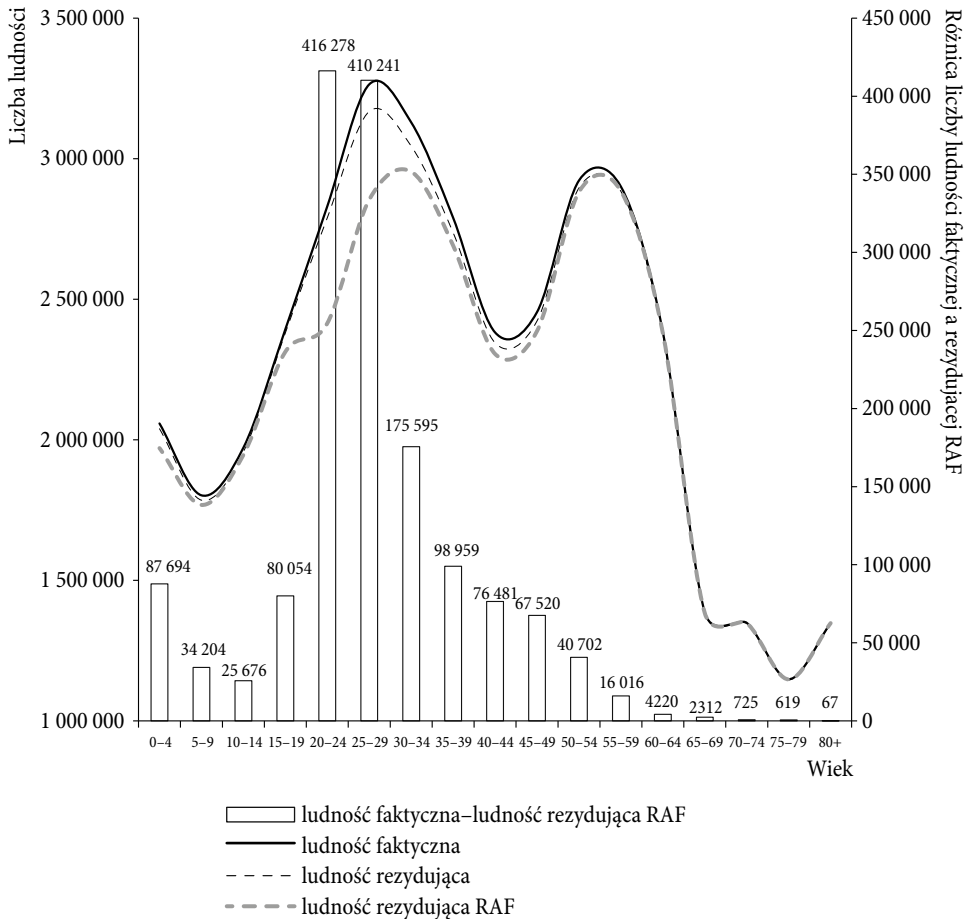
dotyczących stałego pobytu nawet imigranci mieszkający w Polsce przez ponad 12 miesięcy nie byli liczeni jako rezydenci w spisie; i odwrotnie, emigranci z Polski, którzy nie wywiązali się z obowiązku powiadomienia władz o wyjeździe z terytorium Rzeczypospolitej Polskiej (nie dokonali wymeldowania na pobyt stały), zostali włączeni do ludności rezydującej, niezależnie od czasu ich pobytu za granicą (Gołata, 2016a; GUS, 2013b).

Pomijając ograniczenia wynikające z obowiązujących regulacji prawnych, w opracowaniu Gołaty (2016a) przedstawiono szacunek ludności rezydującej według płci i wieku. Wykorzystano w tym celu dostępne na podstawie spisu informacje o strukturze demograficznej ludności faktycznie zamieszkałej oraz o analogicznych strukturach długookresowych imigrantów i emigrantów. Mając świadomość niekompletnych danych o osobach migrujących, także według wieku<sup>132</sup>, w celu weryfikacji jakości szacunków wykorzystano elementy statystyki lustrzanej, porównując struktury emigrantów na podstawie NSP 2011 oraz na podstawie wyników spisu ludności w Anglii i Walii dla ludności polskiej narodowości i polskiego obywatelstwa (Gołata, 2016a). Jednak wobec dostępności struktur demograficznych dla migrantów oraz Polaków w Anglii i Walii przedstawione szacunki dotyczą pięcioletnich grup wieku (rys. 3.5). Najwyraźniejsze rozbieżności dotyczą w szczególności osób w wieku od 20 do 29 lat. Abstrahując od dopełnienia obowiązku wymeldowania, wielkość różnicy między ludnością faktyczną a rezydującą (RAF) tylko dla dwóch grup wieku wynosi ponad 825 tys. osób. Rozszerzając analizę o kolejną grupę – do 34 lat – otrzymujemy różnicę wynoszącą ponad milion osób. Są to osoby w wieku największej aktywności zawodowej, matrymonialnej i prokreacyjnej, zazwyczaj bezpośrednio po zakończeniu edukacji. Ubytek ten, wyraźnie widoczny na rys. 3.5, odzwierciedla bezpośredni efekt migracji. Na rysunku 3.5 jest również uwidoczniony pośredni efekt migracji, w postaci blisko 88 tys. dzieci w wieku 0–4 lat, które bądź wyemigrowały ze swoimi rodzicami, bądź wręcz urodziły się poza granicami kraju<sup>133</sup>. Ekonomiczne konsekwencje emigracji ludności w najbardziej mobilnych grupach wieku nie są przedmiotem niniejszego opracowania, jednak ich znaczenie dla rozwoju demograficznego i gospodarczego państwa ma charakter zasadniczy.

Nie wnikając w szczegóły przedstawionej przez GUS prognozy ludności (GUS, 2014a) oraz wspomnianych badań Śleszyńskiego (2014), ważne wydaje się w tym miejscu przedstawienie potencjalnych konsekwencji dla analiz prospektywnych (tab. 3.5). Wynikają one z przyjęcia jako stanu wyjściowego prognozy liczby ludności faktycznej zamiast rezydującej, a w przypadku prognozy ludności rezydującej określonych definicji i regulacji (GUS, 2014a, 2015).

<sup>132</sup> W wynikach NSP 2011 nie podano informacji o wieku 83% emigrantów.

<sup>133</sup> Problem szacunku dzieci w NSP 2011, również z uwzględnieniem elementów statystyki lustrzanej, będzie przedmiotem dyskusji w kolejnym podrozdziale.



Rysunek 3.5. Liczba ludności rezydującej i faktycznej według płci i wieku, NSP 2011 oraz szacunki własne

Źródło: Na podstawie (Gołata, 2016a)

Zauważmy przede wszystkim, że zgodnie z prognozą GUS można oczekiwać spadku liczby ludności Polski o ponad 4,5 miliona w perspektywie 2050 r. Oznacza to, że z wyjściowego stanu 38, 5 miliona osób oszacowanego na podstawie NSP 2011, w 2050 r. można oczekiwać, że liczba ludności wyniesie 33,9 miliona. Należy wyraźnie podkreślić, że przyjęcie jako wyjściowy stanu ludności faktycznej NSP 2011 skutkuje szacunkiem 36,5 miliona osób dla 2035 r., podczas gdy wielkość ta mogłaby być stanem wyjściowym, gdyby przyjąć jako podstawę ludność rezydującą z uwzględnieniem migracji długookresowych (RAF) (36,97 miliona, tab. 3.4). Z kolei Fihel (2015) przedstawiła modyfikację prognozy GUS, nie zmieniając założeń dotyczących płodności i umieralności, ale uwzględniając migracje długookresowe,

*Tabela 3.5. Stan ludności w świetle prognozy GUS 2013–2050 oraz jej modyfikacji*

Płeć	2013	2014	2015	2020	2025	2030	2035	2040	2045	2050
Stan ludności według prognozy GUS 2014–2050										
Ogółem	<b>38 495 659</b>	38 461 750	38 419 004	38 137 795	37 741 455	37 185 061	<b>36 476 780</b>	35 668 224	34 817 352	<b>33 950 563</b>
Mężczyźni	18 629 535	18 613 099	18 592 593	18 460 535	18 272 421	18 001 713	17 663 027	17 293 917	16 926 402	16 563 540
Kobiety	19 866 124	19 848 651	19 826 411	19 677 260	19 469 034	19 183 348	18 813 753	18 374 307	17 890 950	17 387 023
Zmiana liczby ludności w stosunku do okresu bazowego 2013										
Ogółem		-33 909	-76 655	-357 864	-754 204	-1 310 598	-2 018 879	-2 827 435	-3 678 307	<b>-4 545 096</b>
Mężczyźni		-16 436	-36 942	-169 000	-357 114	-627 822	-966 508	-1 335 618	-1 703 133	-2 065 995
Kobiety		-17 473	-39 713	-188 864	-397 090	-682 776	-1 052 371	-1 491 817	-1 975 174	-2 479 101
Zmiana liczby ludności w stosunku do okresu bazowego 2013 = 100%										
Ogółem		99,91	99,80	99,07	98,04	96,60	94,76	92,66	90,44	<b>88,19</b>
Mężczyźni		99,91	99,80	99,09	98,08	96,63	94,81	92,83	90,86	88,91
Kobiety		99,91	99,80	99,05	98,00	96,56	94,70	92,49	90,06	87,52
Stan ludności według modyfikacji prognozy z uwzględnieniem migrantów w tys.										
Modyfikacja Fihel: emigracja i imigracja czasowa				36 466		35 459		33 916		32 122
Modyfikacja Fihel: 50% powrotów				37 287		36 369		34 812		32 942

Źródło: Na podstawie (Fihel, 2015; GUS, 2014a).

a także przyjmując trzy scenariusze powrotu emigrantów: 25%, 50% i 75%, które rozłożyła równo na dwie fale w latach 2015 i 2020. Uzyskane szacunki prognozy sugerują około 2-milionowe zawyżenie prognozy GUS, będące konsekwencją stanu wyjściowego abstrahującego od migracji. Fihel (2015) szacuje, że 36,5 miliona wynosić będzie liczba ludności w Polsce w 2020 r., a nie – jak prognozuje GUS – 15 lat później – w 2035 r.

Przedstawiona w dalszej części rozdziału ocena kompletności danych NSP 2011 dotyczy trzech grup ludności: niemowląt i dzieci, ludności w wieku produkcyjnym oraz osób starszych. Nie zastosowano tutaj podziału ludności według ekonomicznych grup wieku<sup>134</sup>, jednak gdy tylko było to możliwe, porównywano wskaźniki jakości dla różnych spisów. W ocenie kompletności spisu przeprowadzonego metodą mieszaną wykorzystano metody analizy demograficznej oraz dostępne źródła informacji. Poza danymi NSP 2011 oraz poprzedniego spisu z 2002 r. odwołano się do danych ewidencji ruchu naturalnego ludności w zakresie urodzeń i zgonów. W odniesieniu do ludności w wieku produkcyjnym wykorzystano wyniki Badania Aktywności Ekonomicznej Ludności (BAEL). Wszystkie dane pochodzą z publikacji GUS, w większości udostępnionych na stronie internetowej w bazie danych Demografia, w Banku Danych Lokalnych czy zamieszczonych w zakładce „obszary tematyczne” oraz wcześniejszych wyników badań specjalnych publikowanych w wersji papierowej. Dokonując porównań, odwołano się do wcześniejszych analiz dotyczących spisów przeprowadzonych metodą tradycyjną, w szczególności opracowania Paradyśza (2010) dla NSP 2002.

### **3.4.2. Jakość szacunków dzieci i niemowląt<sup>135</sup> – analiza kompletności urodzeń**

W ocenie kompletności szacunku liczby małych dzieci i młodzieży wykorzystano informacje o urodzeniach z ostatnich lat przed spisem z ewidencji ruchu naturalnego ludności, a także z poprzedniego spisu z 2002 r. Paradyśz (2010, s. 55) zaproponował ewaluację przy pomocy metody „przetrwania między urodzeniem a spisem”, którą stosuje w dwóch wariantach, bez uwzględniania umieralności niemowląt i migracji zagranicznych oraz z uwzględnieniem umieralności. Metoda ta sprowadza się do porównania liczby spisanych niemowląt z liczbą urodzeń żywych z ewidencji ruchu naturalnego oraz oszacowania odsetka braków – tj. niespisanych niemowląt (bez uwzględnienia umieralności i z uwzględnieniem). Sprawdzenie wpływu migracji zagranicznych na odsetek niespisanych niemowląt jest utrudnione z uwagi na brak dostatecznie szczegółowych danych statystycznych. W ocenie Paradyśza dla NSP

<sup>134</sup> Ujęcie takie wiązałoby się z koniecznością kilkukrotnego przeszacowywania stanu ludności na podstawie wyników odpowiednich spisów, co uniemożliwiłoby porównania.

<sup>135</sup> Niemowlęta to dzieci do ukończenia pierwszego roku życia, w analizie demograficznej są to osoby w wieku zero ukończonych lat.

2002 takich przypadków nie było więcej niż 100, co zdecydowało o uznaniu wpływu tego czynnika za nieznaczny. Z kolei w odniesieniu do NSP 2011 autorka niniejszej monografii uważa, że migracje zagraniczne mogą mieć istotne znaczenie tak dla oceny stanu ludności, jak i dla ruchu naturalnego<sup>136</sup>. Jednak brak wiarygodnych danych GUS i trudność ich oszacowania spowodowały pominięcie wpływu migracji w ocenie kompletności dzieci i niemowląt w NSP 2011, aczkolwiek wskazano rozbieżności, sugerując możliwe ich przyczyny oraz konsekwencje.

W pierwszej kolejności zostaną przedstawione szacunki rozbieżności między wielkościami wykazanymi w spisie 2002 i 2011 r. oraz wynikającymi z rejestru urodzeń. Porównując dane dla dwóch spisów, odwołano się do szacunków Paradyśa (2010, s. 55). Wykonując analogiczne obliczenia, wyznaczono odsetek braków bez uwzględnienia zgonów (OB) oraz z uwzględnieniem zgonów (OBz) według następujących formuł:

$$OB(0, 2002, NSP 2002) = \frac{R(0, 2002, NSP 2002)}{U(2002, m = 1 \div 5)} \cdot 100\%, \quad (3.1)$$

$$OBz(0, 2002, NSP 2002) = \frac{Rz(0, 2002, NSP 2002)}{U(2002, m = 1 \div 5)} \cdot 100\%, \quad (3.2)$$

$$R(0, 2002, NSP 2002) = U(2002, m = 1 \div 5) - L(0, 2002, NSP 2002), \quad (3.3)$$

$$Rz(0, 2002, NSP 2002) = U(2002, m = 1 \div 5) - Z(0, 2002, m = 1 \div 5) + \\ - L(0, 2002, NSP 2002). \quad (3.4)$$

W celu szczegółowego wyjaśnienia sposobu szacunku wszystkich różnic absolutnych i względnych poniżej przedstawiono przyjęte oznaczenia, w których wskazano jednocześnie sposób szacowania poszczególnych wielkości:

OB(0, 2002, NSP 2002) – odsetek braków niemowląt, urodzonych w 2002 r., w NSP 2002 (różnica względna),

OB(0, 2011, NSP 2011) – odsetek braków niemowląt, urodzonych w 2011 r., w NSP 2011 (różnica względna),

OB(0, NSP 2002) – odsetek braków niemowląt w NSP 2002 (różnica względna),

OB(0, NSP 2011) – odsetek braków niemowląt w NSP 2011 (różnica względna),

R(0, 2002, NSP 2002) – braki niemowląt urodzonych w 5 pierwszych miesiącach 2002 r. w NSP 2002 (różnica w liczbach absolutnych),

R(0, 2011, NSP 2011) – braki niemowląt urodzonych w 3 pierwszych miesiącach 2011 r. w NSP 2011 (różnica w liczbach absolutnych),

<sup>136</sup> Próby dociekań autorki w zakresie oceny skutków migracji zagranicznych dla szacunku stanu i ruchu naturalnego ludności są przedstawione m.in. w pracach (Gołata, 2016a, 2016b).



- R(0, NSP 2011) – braki niemowląt, w NSP 2011 (różnica w liczbach absolutnych),
- U(2002, m = 1 ÷ 5) – liczba urodzeń żywych w 5 pierwszych miesiącach 2002 r., od 1.01 do 31.05.2002 r.,
- U(2011, m = 1 ÷ 3) – liczba urodzeń żywych w 3 pierwszych miesiącach 2011 r., od 1.01 do 31.03.2011 r.,
- U(2001, m = 6 ÷ 12) + U(2002, m = 1 ÷ 5) – liczba urodzeń żywych od 1.06.2001 r. do 31.05.2002 r.,
- U(2010, m = 4 ÷ 12) + U(2011, m = 1 ÷ 3) – liczba urodzeń żywych od 1.04.2010 r. do 31.03.2011 r.,
- L(0, 2002, NSP 2002) – liczba dzieci urodzonych w 2002 r. (1.01.2002–31.05.2002) i wykazanych w NSP 2002 na podstawie opracowania Paradyśza (2010),
- L(0, 2011, NSP 2011) – liczba dzieci urodzonych w 2011 r. (1.01.2011–31.03.2011) i wykazanych w NSP 2011, na podstawie opracowanego przez GUS bilansu ruchu naturalnego w I kwartale 2011 r.,
- L(0, NSP 2002) – liczba niemowląt (dzieci w wieku 0 ukończonych lat) wykazanych w NSP 2002,
- L(0, NSP 2011) – liczba niemowląt (dzieci w wieku 0 ukończonych lat) wykazanych w NSP 2011,
- Z(0, 2002, m = 1 ÷ 5) – zgony niemowląt urodzonych w 5 pierwszych miesiącach 2002 r. przed spisem, tj. w okresie od 1.01 do 31.05.2002 r. (na podstawie podwójnej klasyfikacji zgonów: według miesięcy i wieku niemowląt),
- Z(2001, m = 6 ÷ 12) + Z(2002, m = 1 ÷ 5) – zgony niemowląt urodzonych w okresie od 1.06.2001 r. do 31.05.2002 r. i zmarłych przed spisem, tj. w okresie od 1.06.2001 r. do 31.05.2002 r. (na podstawie podwójnej klasyfikacji zgonów: według miesięcy i wieku niemowląt),
- Z(0, 2011, m = 1 ÷ 3) – zgony niemowląt urodzonych w 3 pierwszych miesiącach 2011 r. przed spisem, tj. w okresie od 1.01 do 31.03.2011 r. (na podstawie podwójnej klasyfikacji zgonów: według miesięcy i wieku niemowląt),
- Z(2010, m = 4 ÷ 12) + Z(2011, m = 1 ÷ 3) – zgony niemowląt urodzonych w okresie od 1.04.2010 r. do 31.03.2011 r. i zmarłych przed spisem, tj. w okresie od 1.04.2010–31.03.2011 r. (na podstawie podwójnej klasyfikacji zgonów: według miesięcy i wieku niemowląt),
- P(0, 2002, NSP 2002) – liczba żywo urodzonych dzieci w 5 pierwszych miesiącach 2002 r., które powinien objąć NSP 2002,
- P(0, NSP 2002) – liczba niemowląt, które powinien objąć NSP 2002,
- P(0, 2011, NSP 2011) – liczba żywo urodzonych dzieci w 3 pierwszych miesiącach 2002 r., które powinien objąć NSP 2011,
- P(0, NSP 2011) – liczba niemowląt, które powinien objąć NSP 2011,
- Rz(0, 2002, NSP 2002) – braki niemowląt urodzonych w 5 pierwszych miesiącach 2002 r. w NSP 2002 z uwzględnieniem zgonów (różnica w liczbach absolutnych),
- Rz(0, NSP 2002) – braki niemowląt w NSP 2002 z uwzględnieniem zgonów (różnica w liczbach absolutnych),
- Rz(0, 2011, NSP 2011) – braki niemowląt urodzonych w 3 pierwszych miesiącach 2011 r. w NSP 2011 z uwzględnieniem zgonów (różnica w liczbach absolutnych),
- Rz(0, NSP 2011) – braki niemowląt w NSP 2011 z uwzględnieniem zgonów (różnica w liczbach absolutnych),
- OBz(0, 2002, NSP 2002) – odsetek braków niemowląt urodzonych w 5 pierwszych miesiącach 2002 r. w NSP 2002 z uwzględnieniem zgonów (różnica względna),
- OBz(0, NSP 2002) – odsetek braków niemowląt w NSP 2002 z uwzględnieniem zgonów (różnica względna),
- OBz(0, 2011, NSP 2011) – odsetek braków niemowląt urodzonych w 3 pierwszych miesiącach 2011 r. w NSP 2011 z uwzględnieniem zgonów (różnica względna),
- OBz(0, NSP 2011) – odsetek braków niemowląt w NSP 2011 z uwzględnieniem zgonów (różnica względna),
- R(0, 2002, NSP 2002) = U(2002, m = 1 ÷ 5) – L(0, 2002, NSP 2002),

$$\begin{aligned}
R(0, 2011, NSP 2011) &= U(2011, m = 1 \div 3) - L(0, 2011, NSP 2011), \\
R(0, 2002, NSP 2002) &= (U(2001, m = 6 \div 12) + U(2002, m = 1 \div 5)) - L(0, NSP 2002), \\
R(0, 2011, NSP 2011) &= (U(2010, m = 4 \div 12) + U(2011, m = 1 \div 3)) - L(0, NSP 2011), \\
Rz(0, 2002, NSP 2002) &= U(2002, m = 1 \div 5) - Z(0, 2002, m = 1 \div 5) - L(0, 2002, NSP 2002), \\
Rz(0, 2011, NSP 2011) &= U(2011, m = 1 \div 3) - Z(0, 2011, m = 1 \div 3) - L(0, 2011, NSP 2011), \\
Rz(0, NSP 2002) &= (U(2001, m = 6 \div 12) + U(2002, m = 1 \div 5)) - (Z(2001, m = 6 \div 12) + Z(2002, \\
& m = 1 \div 5)) - L(0, NSP 2002), \\
Rz(0, NSP 2011) &= (U(2010, m = 4 \div 12) + U(2011, m = 1 \div 3)) - (Z(2010, m = 4 \div 12) + Z(2011, \\
& m = 1 \div 3)) - L(0, NSP 2011), \\
OBz(0, 2002, NSP 2002) &= (Rz(0, 2002, NSP 2002)/P(0, 2002, NSP 2002)) \cdot 100\%, \\
OBz(0, NSP 2002) &= (Rz(0, NSP 2002)/P(0, NSP 2002)) \cdot 100\%, \\
OBz(0, 2011, NSP 2011) &= (Rz(0, 2011, NSP 2011)/P(0, 2011, NSP 2011)) \cdot 100\%, \\
OBz(0, NSP 2011) &= (Rz(0, NSP 2011)/P(0, NSP 2011)) \cdot 100\%.
\end{aligned}$$

Przed wszystkim zauważmy, że bez względu na metodę przeprowadzania spisu, tj. zarówno w NSP 2002, jak i w NSP 2011, porównanie danych z ewidencji ruchu naturalnego ludności z liczbą spisanych niemowląt wskazuje na brak kompletności – brakujące w spisie dzieci (tab. 3.6). Należy także przyznać, że wielkość tych braków w spisie 2002 r. wynosząca 8183 była wyraźnie niższa od różnicy wykazanej dla 2011 r., wynoszącej 14745 niemowląt. Porównanie pomijające możliwe zgony niemowląt jest oczywiście nieprecyzyjne, nawet w sytuacji zmniejszającej się umieralności niemowląt. Nadal największa jest umieralność niemowląt w pierwszym miesiącu życia. Uwzględnienie zgonów niemowląt pozwala tę niecisłość skorygować. Stosując metodę przetrwania między urodzeniem a spisem z uwzględnieniem umieralności (tab. 3.7), otrzymujemy oczywiście nieco odmienne wartości różnic absolutnych i względnych, jednak zaobserwowane wcześniej relacje pozostają aktualne. Potwierdza się brak dzieci, które wykazała ewidencja urodzeń (tab. 3.6). Jest on nadal wyższy w NSP 2011 aniżeli w NSP 2002, tak w wymiarze absolutnym, jak i względnym. W spisie 2002 r. brakuje ok. 6,3 tys. niemowląt, a w spisie 2011 r. – 12,9 tys. Udział brakujących niemowląt stanowi odpowiednio 1,8% i 3,2% liczby niemowląt oszacowanej jako właściwa na podstawie ewidencji ruchu naturalnego na moment krytyczny spisów 2002 i 2011. Oznacza to, że odsetek brakujących niemowląt w NSP 2011 był o 80% wyższy niż w NSP 2002. W każdym ze spisów udział brakujących chłopców był o około 40–57% wyższy aniżeli dziewczynek. Oczywiście uwzględnienie zgonów niemowląt wpłynęło na zmniejszenie wskazanych odsetków braków (w porównaniu do danych z tab. 3.6), jednak nie jest ono jednakowe. Odsetek brakujących niemowląt w NSP 2002 zmniejszył się o 25%, natomiast w NSP 2011 zaledwie o 12%. Możliwe że o połowę niższa korekta rozbieżności szacunków w przypadku danych NSP 2011 jest związana z faktem, że to właśnie ewidencja ruchu naturalnego, będąca podstawą rejestru PESEL, stanowiła zasadnicze źródło danych w spisie przeprowadzonym metoda mieszaną.

Analizując kompletność spisu w odniesieniu do najmłodszych niemowląt, ze względu na moment krytyczny spisu, Paradysz uwzględnił niemowlęta urodzone

w 5 pierwszych miesiącach 2002 r., natomiast dla NSP 2011 przyjęto dzieci urodzone w pierwszym kwartale 2011 r. Braki wskazane przez Paradysza dla NSP 2002 są bardzo duże i sięgają  $R(0, 2002, NSP 2002) = 15\ 829$  dzieci, co w przypadku dziewczynek stanowi blisko 11% niemowląt urodzonych w 2002 r. (tab. 3.6). Uwzględnienie zgonów, mimo że rozważania dotyczą najmłodszych niemowląt, nie wpływa na istotne zmniejszenie tego odsetka:  $OBz(0, 2002, NSP 2002) = 9,96$  (tab. 3.7).

Tabela 3.6. Ocena dokładności szacunku liczby niemowląt w NSP 2002 i 2011 bez uwzględniania ich zgonów i migracji zagranicznych

Płeć niemowląt	Liczba dzieci urodzonych w 2002 r. i wykazanych w NSP 2002	Liczba urodzeń żywych od 1.01 do 31.05.2002 r.	Brak niemowląt w czasie spisu urodzonych od 1.01 do 31.05.2002 r.	
			w liczbach bezwzględnych	w %
	L(0, 2002, NSP 2002)	U(2002, m = 1 ÷ 5)	R(0, 2002, NSP 2002)	OB(0, 2002, NSP 2002)
Ogółem	133 268	149 097	15829	10,62
Chłopcy	68 189	76 582	8393	10,96
Dziewczęta	65 079	72 515	7436	10,25
Płeć niemowląt	Liczba dzieci w wieku zero lat wykazanych w NSP 2002	Liczba urodzeń żywych od 1.06.2001 do 31.05.2002 r.	Brak niemowląt w NSP 2002	
			w liczbach bezwzględnych	w %
	L(0, NSP 2002)	U(2001, m = 6 ÷ 12) + U(2002, m = 1 ÷ 5)	R(0, NSP 2002)	OB(0, NSP 2002)
Ogółem	351 662	359 845	8183	2,33
Chłopcy	180 116	184 975	4859	2,70
Dziewczęta	171 546	174 870	3324	1,94
Płeć niemowląt	Liczba dzieci urodzonych w 2011 r. i wykazanych w NSP 2011*	Liczba urodzeń żywych od 1.01 do 31.03.2011 r.	Brak niemowląt w czasie spisu urodzonych od 1.01 do 31.03.2011 r.	
			w liczbach bezwzględnych	w %
	L(0, 2011, NSP 2011)	U(2011, m = 1 ÷ 3)	R(0, 2011, NSP 2011)	OB(0, 2011, NSP 2011)
Ogółem	95 231	95 572	341	0,36
Chłopcy	49 239	49 431	192	0,39
Dziewczęta	45 992	46 141	149	0,32
Płeć niemowląt	Liczba dzieci w wieku zero lat wykazanych w NSP 2011	Liczba urodzeń żywych od 1.04.2010 do 31.03.2011 r.	Brak niemowląt w NSP 2011	
			w liczbach bezwzględnych	w %
	L(0, NSP 2011)	U(2010, m = 4 ÷ 12) + U(2011, m = 1 ÷ 3)	R(0, NSP 2011)	OB(0, NSP 2011)
Ogółem	389 903	404 648	14745	3,64
Chłopcy	200 592	209 699	9107	4,34
Dziewczęta	189 311	194 949	5638	2,89

Uwaga: \* Na podstawie tablic bilansowych GUS opublikowanych dla 2011 r. w bazie danych Demografia.

Źródło: (Paradysz, 2010) oraz szacunki własne na podstawie danych GUS z NSP 2002, NSP 2011, roczników demograficznych i bazy danych Demografia.

Zupełnie odmiennie przedstawia się analiza dla NSP 2011. Odsetek brakujących najmłodszych niemowląt (urodzonych od 1 stycznia do końca marca 2011 r.) jest stosunkowo niewielki:  $OB(0, 2011, NSP 2011) = 0,36$ . Prawdopodobnie jednak uzyskany rezultat nie jest porównywalny z informacją przedstawioną przez Paradysza, który powołuje się na niepublikowane dane GUS, na których podstawie oszacował liczbę dzieci urodzonych w 2002 r. i wykazanych przez NSP 2002. Ponieważ informacje takie nie są dostępne dla NSP 2011, zdecydowano się zastąpić tę wielkość liczbą urodzeń przedstawioną przez GUS w tablicy bilansowej dla I kwartału 2011 r. Założono, że bilans ten przygotowano z wykorzystaniem danych NSP. Zatem liczba 95 231 oznacza urodzenia w I kwartale 2011 r. na podstawie bilansu, a 95 572 to liczba urodzeń na podstawie ewidencji. Niewielka rozbieżność wynosząca 341 dzieci (tab. 3.7) została prawie całkowicie zniwelowana liczbą 272 niemowląt urodzonych i zmarłych w I kwartale 2011 r. (tab. 3.6). Wynika stąd, że niedoszacowanie najmłodszych niemowląt wyniosło zaledwie 0,07%, co w porównaniu do 9,96% dla NSP 2002 mogło sugerować znaczną poprawę kompletności w spisie wykorzystującym dane rejestrów. Jednak badanie kompletności wszystkich niemowląt przeprowadzone na podstawie danych spisu i ewidencji ruchu naturalnego nie potwierdza tych optymistycznych wniosków. Jak zauważono wcześniej, rozbieżności rosną do  $Rz(0, NSP 2011) = 12\ 892$ , co daje odsetek braków wynoszący  $OBz(0, NSP 2011) = 3,2$ . Oznacza to, że prawdopodobnie dane bilansowe nie odzwierciedlały wielkości ujętych w spisie. Niemniej jednak informacja ta wskazuje na poprawność metodologii stosowanej przy szacowaniu bilansu<sup>137</sup>. Ponieważ nie dysponujemy informacjami, które pozwoliłyby na wyjaśnienie rozbieżności szacunku liczby najmłodszych niemowląt w wieku do 6 miesięcy w NSP 2002, przedstawionych przez Paradysza oraz w wieku do 3 miesięcy w NSP 2011 analizowanych w niniejszym opracowaniu, na tym poprzestaniemy. Analizowane rozbieżności należy uznać za bardzo szczegółowe i ważne do rozważenia w kontekście oceny jakości rejestrów oraz metody sporządzania bilansu.

Zauważmy, że spis w 2011 r. został przeprowadzony metodą mieszaną z uwzględnieniem danych rejestrów administracyjnych, w tym przede wszystkim rejestru PESEL, co tłumaczy oczekiwanie większej zgodności niż w spisie przeprowadzonym metodą tradycyjną. Pewnym wyjaśnieniem rozbieżności w tym zakresie może być zauważona właściwość rejestru PESEL, a mianowicie opóźnienie rejestracji urodzeń (Józefowski i Rynarzewska-Pietrzak, 2010). Fakt ten wymaga szczególnej uwagi oraz nie tylko poprawy w odniesieniu do jakości rejestru, ale także jego aktualizacji w czasie i synchronizacji rejestracji urodzeń oraz zgonów z momentem krytycznym.

<sup>137</sup> Oczywiście bilans powinien dać w pełni zgodne rezultaty, jednak liczba brakujących 69 niemowląt może być wytłumaczona opóźnieniem rejestracji.

Tabela 3.7. Ocena dokładności szacunku liczby niemowląt w NSP 2002 i 2011 z uwzględnieniem zgonów

Płeć niemowląt	Liczba żywo urodzonych dzieci w 5 pierwszych miesiącach 2002 r.				Brak niemowląt	
	ogółem	spisanych w NSP 2002	zmarłych przed spisem	które powinien objąć spis	w liczbach bezwzględnych	w %
	U(2002, m = 1 ÷ 5)	L(0, 2002, NSP 2002)	Z(0, 2002, m = 1 ÷ 5)	P(0, 2002 NSP 2002)	Rz(0, 2002, NSP 2002)	OBz(0, 2002, NSP 2002)
Ogółem	149 097	133 268	1 087	148 010	14 742	9,96
Chłopcy	76 582	68 189	604	75 978	7 789	10,25
Dziewczęta	72 515	65 079	483	72 032	6 953	9,65
Płeć niemowląt	Liczba dzieci w wieku zero lat wykazanych przez NSP 2002				Brak niemowląt	
	ogółem	spisanych w NSP 2002	zmarłych przed spisem	które powinien objąć spis	w liczbach bezwzględnych	w %
	U(2001, m = 6 ÷ 12) + U(2002, m = 1 ÷ 5)	L(0, NSP 2002)	Z(2001, m = 6 ÷ 12) + Z(2002, m = 1 ÷ 5)	P(0, NSP 2002)	Rz(0, NSP 2002)	OBz(0, NSP 2002)
Ogółem	359 845	351 662	1 837	358 008	6 346	1,77
Chłopcy	184 975	180 116	1 010	183 965	3 849	2,09
Dziewczęta	174 870	171 546	827	174 043	2 497	1,43
Płeć niemowląt	Liczba żywo urodzonych dzieci w 3 pierwszych miesiącach 2011 r.				Brak niemowląt	
	ogółem	spisanych w NSP 2011*	zmarłych przed spisem	które powinien objąć spis	w liczbach bezwzględnych	w %
	U(2011, m = 1 ÷ 3)	L(0, 2011, NSP 2011)	Z(0, 2011, m = 1 ÷ 3)	P(0, 2011, NSP 2011)	Rz(0, 2011, NSP 2011)	OBz(0, 2011, NSP 2011)
Ogółem	95 572	95 231	272	95 300	69	0,07
Chłopcy	49 431	49 239	135	49 296	57	0,12
Dziewczęta	46 141	45 992	137	46 004	12	0,03
Płeć niemowląt	Liczba dzieci w wieku zero lat wykazanych przez NSP 2011				Brak niemowląt	
	ogółem	spisanych w NSP 2011	zmarłych przed spisem	które powinien objąć spis	w liczbach bezwzględnych	w %
	U(2010, m = 4 ÷ 12) + U(2011, m = 1 ÷ 3)	L(0, NSP 2011)	Z(2010, m = 4 ÷ 12) + Z(2011, m = 1 ÷ 3)	P(0, NSP 2011)	Rz(0, NSP 2011)	OBz(0, NSP 2011)
Ogółem	404 648	389 903	1 853	402 795	12 892	3,20
Chłopcy	209 699	200 592	1 009	208 690	8 098	3,88
Dziewczęta	194 949	189 311	844	194 105	4 794	2,47

Uwaga: \* Na podstawie tablic bilansowych GUS opublikowanych dla 2011 r. w bazie danych Demografia.

Źródło: (Paradysz, 2010) oraz szacunki własne na podstawie danych GUS z NSP 2002, NSP 2011, roczników demograficznych i bazy danych Demografia.

Oceniając kompletność szacunku liczby osób najmłodszych wiekiem, zwrócono uwagę na niedoszacowanie liczby niemowląt. Stosując analogiczną metodę w odniesieniu do starszych dzieci w wieku 1–9 lat, otrzymano jednak odmienną informację – o przeszacowaniu ich liczby (tab. 3.8). Jest to sytuacja wyjątkowa, gdyż pomimo możliwości „podwójnego spisania” ewentualna liczba takich przypadków jest zazwyczaj niższa od liczby „pominięć”. Natomiast dane przedstawione w tabeli 3.8 wskazują wyraźnie, że w NSP 2011 ujęto dzieci, które nie zostały objęte ewidencją urodzeń. Dane spisu nie tylko przewyższają liczby urodzonych dzieci (zarejestrowanych jako urodzone w Polsce), ale są także wyższe od szacunków po uwzględnieniu zgonów niemowląt oraz postarzeniu dzieci<sup>138</sup>. Różnice są znaczne, dla dzieci w wieku 2 i 3 lat sięgają po 20 tys., a nawet tę wielkość przekraczają. W wymiarze względnym zaobserwowana w spisie nadwyżka liczby dzieci stanowi ponad 5% liczebności danej grupy wieku. Można przypuszczać, że jest to efekt nieuwzględnionych migracji zagranicznych. Jeśli chodzi o polską statystykę urodzeń, znajdujemy rozróżnienie pomiędzy urodzeniami w kraju i poza jego granicami, ale według danych GUS różnica ta nie jest znaczna i wskazuje, na przykład dla 2011 r., na 446 urodzeń za granicą.

Tabela 3.8. Ocena dokładności szacunku liczby dzieci w NSP 2011 z uwzględnieniem zgonów

Wiek*	NSP 2011			Ewidencja po postarzeniu**			Różnica absolutna			Różnica (w %)		
	Og	M	K	Og	M	K	Og	M	K	Og	M	K
0	389 903	200 592	189 311	402 795	208 690	194 105	-12 892	-8 098	-4 794	-3,2	-3,9	-2,5
1	432 349	221 916	210 433	416 192	214 764	201 428	16 157	7 152	9 005	3,9	3,3	4,5
2	435 505	223 003	212 502	415 826	213 547	202 279	19 679	9 456	10 223	4,7	4,4	5,1
3	409 794	210 089	199 705	389 246	199 970	189 276	20 548	10 119	10 429	5,3	5,1	5,5
4	390 447	200 302	190 145	376 450	193 515	182 935	13 997	6 787	7 210	3,7	3,5	3,9
5	370 695	190 057	180 638	361 750	185 863	175 887	8 945	4 194	4 751	2,5	2,3	2,7
6	360 629	184 683	175 946	355 427	182 357	173 070	5 202	2 326	2 876	1,5	1,3	1,7
7	355 800	182 810	172 990	351 755	181 213	170 542	4 045	1 597	2 448	1,1	0,9	1,4
8	351 311	180 379	170 932	346 440	178 385	168 055	4 871	1 994	2 877	1,4	1,1	1,7
9	363 960	186 676	177 284	360 867	185 526	175 341	3 093	1 150	1 943	0,9	0,6	1,1

Og – ogółem, M – mężczyźni, K – kobiety.

\* Wiek określono w latach ukończonych na moment krytyczny NSP 2011.

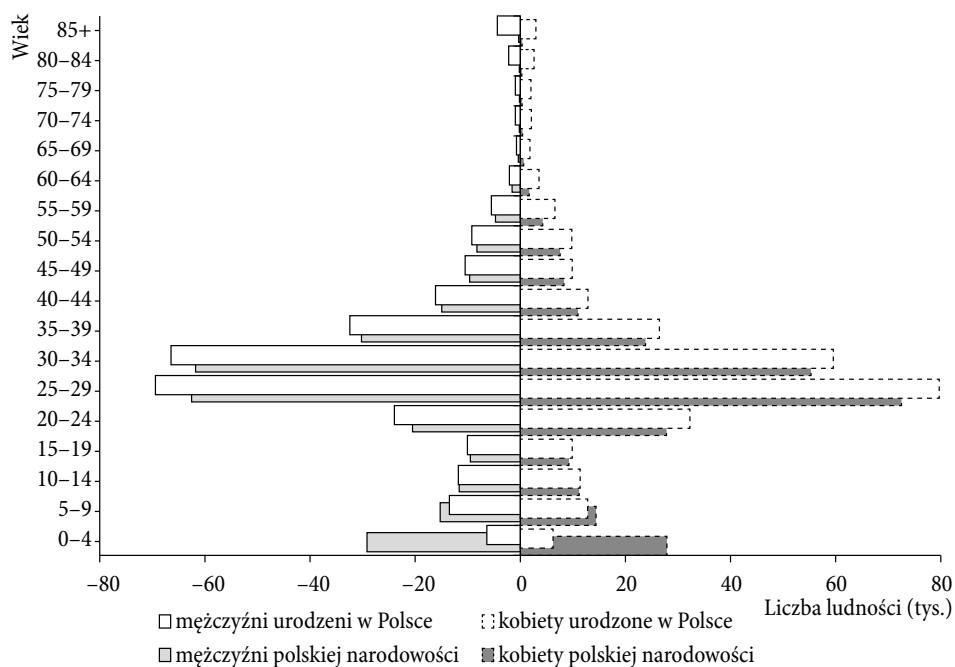
\*\*Dane o urodzeniach z ewidencji ruchu naturalnego ludności zostały w pierwszej kolejności skorygowane przez uwzględnienie rzeczywistej liczby zgonów niemowląt według podwójnej klasyfikacji, tj. według wieku i miesiący czasu kalendarzowego. Dla pozostałych lat wieku w obliczeniach przyjęto prawdopodobieństwo przeżycia roku według płci z tablic trwania życia odpowiednich dla danego roku kalendarzowego.

Źródło: Na podstawie bazy danych Demografia GUS oraz roczników demograficznych.

<sup>138</sup> Liczby niemowląt określone według momentu krytycznego spisu pomniejszono o zarejestrowane liczby zgonów niemowląt, a dla kolejnych lat wieku dokonano „postarzenia” na podstawie prawdopodobieństwa przeżycia z odpowiednich dla wieku i roku kalendarzowego tablic trwania życia.

Uwzględnienie łącznie pięciu lat w żadnej mierze nie pokrywa brakującej liczby dzieci. Sprawy nie wyjaśnia także dołączenie do rozważań dzieci urodzonych w Polsce przez cudzoziemki (w 2010 r. było to 529 urodzeń). Dodatkowo należy pamiętać, że spisowa kategoria ludności faktycznej nie obejmuje cudzoziemców mieszkających w Polsce.

Wykazana przez spis nadwyżka dzieci w wieku 1–4 lat w stosunku do szacunków wynikających z ewidencji urodzeń może wynikać z uwzględnienia w spisie ludności faktycznej obejmującej wszystkie osoby przebywające za granicą, bez względu na czas nieobecności. Dotyczy to szczególnie licznych grup emigrantów w wieku największej aktywności prokreacyjnej, rejestrowania dzieci urodzonych za granicą i zgłaszania ich w rejestrze PESEL, których to faktów nie ujmują statystyki. Podjęte badania przy wykorzystaniu tzw. statystyki lustrzanej (Gołata, 2016b) pozwoliły zauważyć, że wynosząca około 57,5 tys. nadwyżka spisanej liczby dzieci w wieku 0–4 lat (tab. 3.8) w porównaniu do szacunków na podstawie ewidencji urodzeń może być wyjaśniona urodzeniami przez osoby, które wyemigrowały z Polski. W szczególności dotyczy to różnicy między liczbą osób polskiej narodowości i liczbą obywateli polskich w wieku 0–4 lat poza granicami kraju – przykładowo w Anglii i Walii różnica ta wynosi 50 tys. (rys. 3.6) (ONS, 2013, s. 23). Dane Urzędu Statystycznego Wielkiej Brytanii (Office for National Statistics ONS) odnotowały



Rysunek 3.6. Piramida wieku ludności polskiej narodowości oraz urodzonej w Polsce ludności rezydentów w Anglii i Walii, 2011

Źródło: (ONS 2013, Figure 10a, s. 25)

74 456 urodzeń przez Polki w Anglii i Walii w latach 2007–2010 (Jo, Dormon i Jefferies, 2012, s. 24).

Przedstawiona powyżej dyskusja nie jest gotowym rozwiązaniem ukazującym kompletny szacunek liczby ludności, stanowi jedynie propozycję określenia brakujących elementów. Przy czym należy mieć świadomość wielu upraszczających założeń i wynikających z nich konsekwencji. Szacunki bazujące na porównaniach statystyk różnych państw wymagają szczególnej wnikliwości i jednocześnie są trudne do sprawdzenia i oceny. Warto zauważyć, że Wielka Brytania jest najpopularniejszym krajem docelowym dla polskich emigrantów. Fakt ten jest argumentem uzasadniającym porównania struktury wieku emigrantów z Polski z NSP 2011 ze strukturą wieku imigrantów z Polski na podstawie spisu w Wielkiej Brytanii, w szczególności w Anglii i Walii. W okresie od grudnia 2003 r. do grudnia 2013 r. populacja osób urodzonych w Polsce rezydujących w Wielkiej Brytanii wzrosła z 75 000 do 697 000 (ONS, 2013).

### 3.4.3. Jakość szacunków – analiza kompletności na podstawie wskaźników „przetrwania”

Powyższe spostrzeżenia zostały potwierdzone także w wyniku analizy demograficznej przeprowadzonej z uwzględnieniem danych poprzedniego spisu z 2002 r. (tab. 3.9). Zastosowano w tym przypadku tzw. „współczynniki przetrwania”. Stosując ideę badania wzdłużnego, śledzimy losy poszczególnych generacji na podstawie danych kolejnych spisów. Analizę taką na podstawie danych powojennych spisów: NSP 1960, 1970, 1978, 1988 i 2002, a więc w szerszym horyzoncie czasowym, przeprowadził Paradysz (2010), otrzymując czasem zaskakujące rezultaty związane z pomijaniem przez wszystkie spisy dzieci i ich rodziców. W niniejszym opracowaniu przedstawiono wyniki badań jedynie dwóch ostatnich spisów.

$$WP_{\frac{2011}{2002}}(s, x) = \frac{P_{2011}(s, (x+9))}{P_{2002}(s, x)}, \quad (3.5)$$

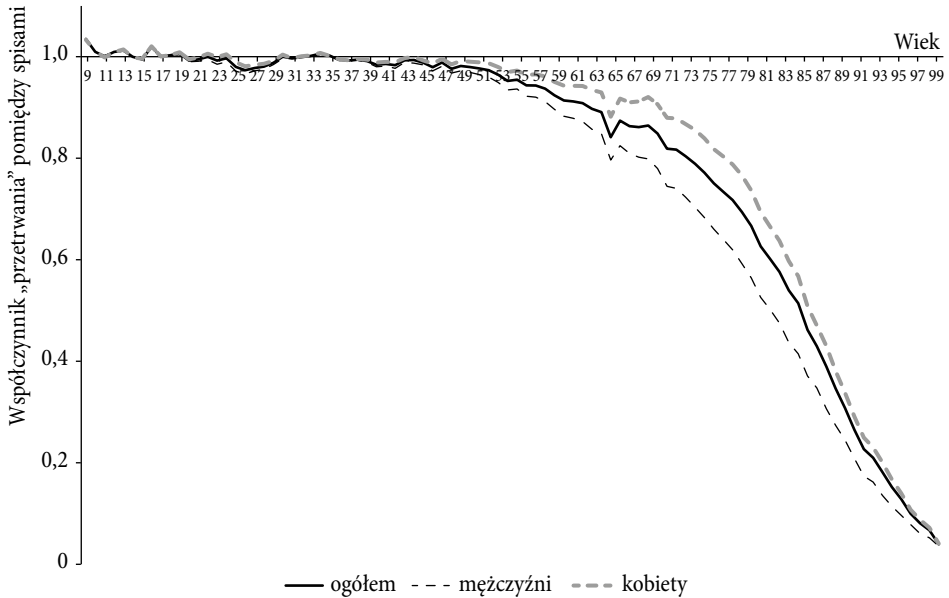
gdzie:

- $WP_{\frac{2011}{2002}}(s, x)$  – współczynnik „przetrwania generacji” pomiędzy spisami 2002 i 2011 według płci i wieku,  
 $P_t$  – populacja według NSP 2002 lub NSP 2011 ( $t = 2002$  lub  $t = 2011$ ),  
 $s$  – płeć:  $T$  – ogółem, tzn. obie płcie razem;  $F$  – kobiety;  $M$  – mężczyźni,  
 $x$  – wiek.

W pierwszej kolejności zaprezentujemy współczynniki „przetrwania generacji” dla okresu między dwoma kolejnymi spisami bez uwzględnienia migracji. Zostały one wyznaczone jako stosunek liczby osób w wieku  $x$  według NSP 2011



do liczby osób o 9 lat młodszych według NSP 2002 (rys. 3.7). Zauważmy, że dane spisu 2011 r. wskazują wyższą liczbę 9-latków oraz młodzieży kolejnych roczników wieku aniżeli dzieci w wieku zero, jeden i kolejnych lat wieku w spisie 2002 r. Wyższe od jednościi współczynniki przetrwania potwierdzają zauważone wcześniej przeszacowanie liczby dzieci w wieku 1–10 lat. Stosowane tutaj określenie „przeszacowanie” należy interpretować w powiązaniu z dyskusją kategorii ludności faktycznej i rezydującej. Bezwzględnie jednak osoby dziś żyjące powinny być ujęte we wcześniejszych statystykach. Dlatego wyższa od jednościi wartość współczynników przetrwania może być interpretowana w kategoriach „wskaźników zmartwychwstania” i sugerować niedoszacowanie niemowląt w spisie 2002 r., na co zwrócił uwagę także Paradysz (2010). Każda wyższa od jednościi wartość współczynnika przetrwania oznacza bezwzględną konieczność refleksji, gdyż oznacza błędy szacunku.



Rysunek 3.7. Współczynniki „przetrwania dla osób danej generacji” pomiędzy NSP 2002 i 2011, bez uwzględniania migracji

Źródło: Na podstawie danych NSP 2002 i 2011

Na uwagę zasługują również wyraźnie niższe od jednościi współczynniki przetrwania dla młodzieży w wieku 25–30 lat (rys. 3.7). Są one także niższe od współczynników przetrwania dla sąsiednich grup wieku. Kolejna anomalia wyższych od jednościi współczynników przetrwania generacji pomiędzy spisami pojawia się z kolei dla wieku 30, 33–35 lat. Paradysz szczególnie niskie współczynniki przetrwania w spisie 2002 r. dla tej grupy generacji tłumaczył zwiększoną intensywnością

migracji po 1989 r. Ponowne pojawienie się („zmartwychwstanie”) tych osób w spisie 2011 r. wskazuje kolejne błędy pokrycia (niedoszacowania) w NSP 2002, kiedy nie spisano osób nieobecnych. Natomiast ich uwzględnienie w spisie 2011 r. wynikało z oparcia szacunków na danych administracyjnych oraz przyjętej definicji ludności faktycznej<sup>139</sup>. Wiązanie owych, „wyższych od jedności” wskaźników z powrotami z emigracji wydaje się bezpodstawne, szczególnie w świetle dostępnych danych, w tym statystyki lustrzanej ze spisu w Anglii i Walii w 2011 r. Od 40. roku życia wartości współczynników „przetrwania” zaczynają maleć, szybciej dla mężczyzn i wolniej dla kobiet, co jest uzasadnione zróżnicowanym oczekiwanym trwaniem życia. Szczególnie niskie wartości współczynników dotyczą wieku 65 lat (0,8 dla mężczyzn i 0,88 dla kobiet), gdy obserwujemy nagłe obniżenie, a następnie wzrost wartości. Zaobserwowany „uskok” oznacza trudne do wytłumaczenia, szczególnie niedoszacowanie dla jednej rocznej grupy wieku.

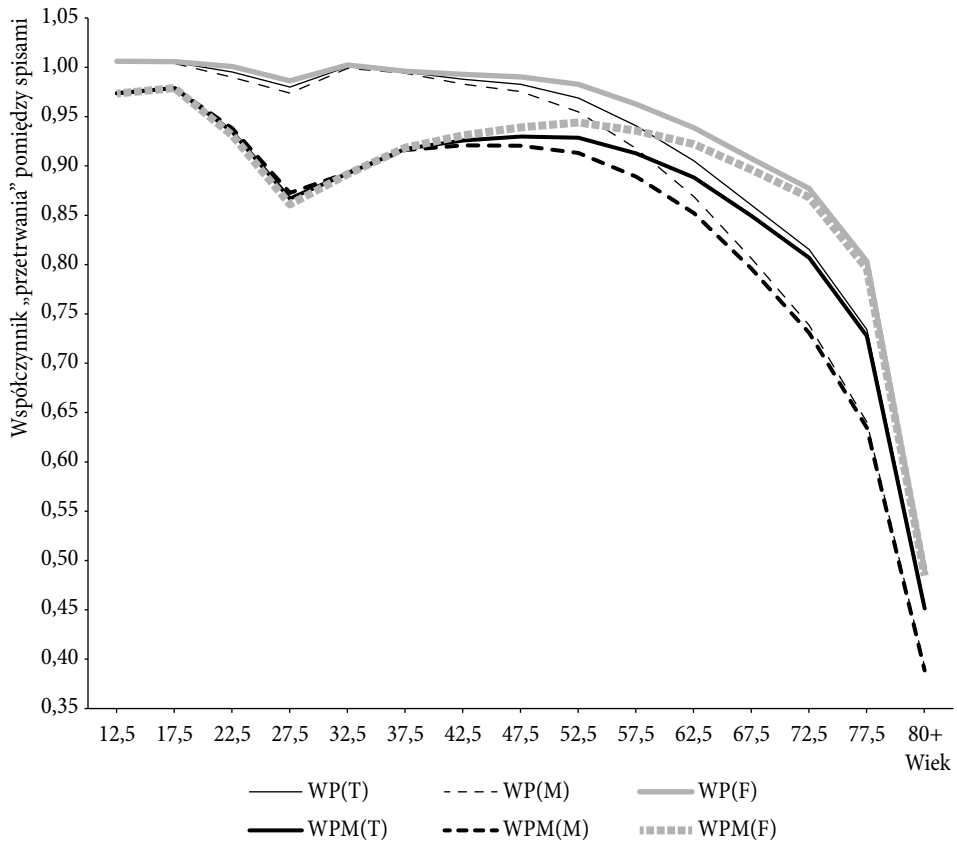
W dalszej analizie zostaną uwzględnione także migracje (rys. 3.8, tab. 3.9). Wydaje się, że właśnie intensywność migracji po akcesji do Unii Europejskiej, która została odzwierciedlona w stanie ludności faktycznej spisanej w NSP 2002, a nie została uwzględniona w 2011 r., jest jedną z zasadniczych przyczyn przekraczających jedność wskaźników przetrwania generacji pomiędzy spisami. W tym przypadku analiza zostanie przeprowadzona dla pięcioletnich grup wieku, co wynika z dostępności danych dotyczących migracji. Wiąże się to z odpowiednią modyfikacją współczynnika przetrwania generacji pomiędzy spisami, w którym nie tylko uwzględniono (zamiast rocznych) pięcioletnie grupy wieku, ale dodatkowo w liczniku wprowadzono saldo migracji zagranicznych. W ten sposób, przynajmniej częściowo skorygowano różnice wynikające z odmiennego ujęcia tej samej kategorii ludności w zależności od źródła danych.

$$\frac{WPM_{2011}}{2002}(s, x+4) = \frac{P_{2011}(s, ((x \div 4) + 9)) + I_{2011}(s, ((x \div 4) + 9)) - E_{2011}(s, ((x \div 4) + 9))}{P_{2002}(s, (x \div 4))}, \quad (3.6)$$

gdzie:

$\frac{WPM_{2011}}{2002}(s, x)$  – współczynnik „przetrwania generacji” pomiędzy spisami 2002 i 2011 r. według płci i wieku z uwzględnieniem migracji,  
 $P_t$  – populacja według NSP 2002 lub NSP 2011 ( $t = 2002$  lub  $t = 2011$ ),  
 $I_{2011}$  – populacja imigrantów według NSP 2011,  
 $E_{2011}$  – populacja emigrantów według NSP 2011,  
 $s$  – płeć:  $T$  – ogółem, tzn. obie płcie razem;  $F$  – kobiety;  $M$  – mężczyźni,  
 $x$  – wiek.

<sup>139</sup> Osoby przebywające za granicą, które nie wymeldowały się z pobytu stałego, zostały spisane jako obecne, bez względu na czas trwania nieobecności.



Rysunek 3.8. Współczynniki „przetrwania dla osób danej generacji” pomiędzy NSP 2002 i 2011, z uwzględnieniem migracji

Źródło: Na podstawie danych NSP 2002 i 2011

Tabela 3.9. Współczynniki „przetrwania dla osób danej generacji” pomiędzy NSP 2002 i 2011, z uwzględnieniem migracji i bez niej

Grupa wieku wg NSP 2011	Współczynniki „przetrwania generacji” pomiędzy NSP 2002 i 2011					
	bez uwzględnienia migracji			z uwzględnieniem migracji		
	ogółem	mężczyźni	kobiety	ogółem	mężczyźni	kobiety
10–14	1,0062	1,0063	1,0062	0,9737	0,9738	0,9735
15–19	1,0049	1,0040	1,0058	0,9790	0,9790	0,9790
20–24	0,9953	0,9900	1,0008	0,9346	0,9381	0,9310
25–29	0,9799	0,9739	0,9862	0,8670	0,8727	0,8610
30–34	1,0007	0,9991	1,0024	0,8924	0,8931	0,8918
35–39	0,9951	0,9941	0,9960	0,9173	0,9161	0,9185
40–44	0,9879	0,9830	0,9929	0,9259	0,9208	0,9311
45–49	0,9828	0,9754	0,9903	0,9298	0,9206	0,9391

Grupa wieku wg NSP 2011	Współczynniki „przetrwania generacji” pomiędzy NSP 2002 i 2011					
	bez uwzględnienia migracji			z uwzględnieniem migracji		
	ogółem	mężczyźni	kobiety	ogółem	mężczyźni	kobiety
50–54	0,9689	0,9549	0,9828	0,9287	0,9131	0,9441
55–59	0,9407	0,9178	0,9628	0,9129	0,8892	0,9357
60–64	0,9051	0,8689	0,9388	0,8885	0,8523	0,9221
65–69	0,8602	0,8060	0,9075	0,8494	0,7957	0,8962
70–74	0,8153	0,7385	0,8771	0,8069	0,7309	0,8681
75–79	0,7342	0,6404	0,8033	0,7276	0,6346	0,7961
80–84	0,6068	0,5075	0,6700	0,4513	0,3887	0,4848
85–89	0,4368	0,3511	0,4799	×	×	×
90–94	0,2533	0,1993	0,2775	×	×	×
95–100	0,1101	0,0845	0,1193	×	×	×

Źródło: Na podstawie danych NSP 2002 i 2011.

Uwzględnienie migracji w określeniu stanu ludności według płci i wieku przede wszystkim wyeliminowało nierealne wartości współczynników przetrwania dla osób określonych generacji. Sytuację tę można interpretować jako korektę szacunków spisowych. Dodatkowo porównanie współczynników przetrwania z uwzględnieniem i bez uwzględniania migracji ukazuje ogromny wpływ tego czynnika na szacunek stanu ludności. Jest on bardzo duży wśród dzieci i młodzieży i w najbardziej mobilnych grupach wieku, które jednocześnie charakteryzują się największą aktywnością ekonomiczną i prokreacyjną. Fakt ten ma ogromne znaczenie w wymiarze społecznym i gospodarczym.

#### 3.4.4. Jakość szacunków osób starszych – analiza stanu i porządku wymierania stulatków

Podjętą próbę oceny jakości szacunków spisowych w odniesieniu do osób starszych, porównano przede wszystkim szacunek stanu ludności oraz dane ewidencji ruchu naturalnego w zakresie zgonów. W każdym przypadku korzystano wyłącznie z danych zagregowanych opublikowanych przez GUS w rocznikach demograficznych bądź na stronie internetowej Urzędu w bazie danych Demografia. Ponieważ moment krytyczny spisu 2011 r. został określony jako północ z 31 marca na 1 kwietnia i dodatkowo najstarsza grupa ludności w opublikowanej strukturze według wieku obejmuje łącznie osoby liczące 100 i więcej lat, zdecydowano się odwołać do opracowanej na podstawie danych spisowych struktury ludności według płci i wieku oraz rocznika urodzenia (stan na 31.12.2011 r., a w dalszych analizach także do stanu na 31 grudnia kolejnych lat).

W końcu grudnia 2011 r. żyło w Polsce 25 757 osób urodzonych w 1916 r. lub wcześniej. Na podstawie danych o zgonach według roczników urodzenia w 2012 r. skumulowano zgony osób urodzonych w latach 1916 i wcześniej (1901–1916),

otrzymując informację, że zmarły 7353 osoby: 1611 mężczyzn i 5742 kobiet (tab. 3.10). Następnie, śledząc porządek wymierania generacji 1916 oraz wcześniejszych w kolejnych latach kalendarzowych do 2016 r., odnotowano, że w tym okresie zmarło łącznie 20 788 osób urodzonych w 1916 r. lub wcześniej. Bilansując stan ludności generacji 1916 r. i wcześniejszych z 31.12.2011 r. oraz łączną liczbę zgonów w latach 2012–2016 wśród osób tych generacji, otrzymano liczbę 4969 osób, w tym 969 mężczyzn i 4000 kobiet na dzień 31.12.2016 r. Z danych zawartych w tablicy 4: Ludność według płci i roku urodzenia, stan w dniu 31.12.2016 r. (w bazie Demografia GUS), odczytano, że w końcu 2016 r. żyły 4933 osoby badanych generacji, w tym 961 mężczyzn i 3972 kobiety. Przeprowadzony bilans jest więc bardzo zbliżony, a różnica między szacunkiem GUS a bilansem wynikającym z przeprowadzonej analizy zgonów dla ludności ogółem to – 36 osób (brakuje 8 mężczyzn i 28 kobiet urodzonych w 1916 r. i wcześniej). Wielkość ta stanowi 0,73% stanu podanego przez GUS.

Tabela 3.10. Analiza zgonów osób generacji 1916 r. i wcześniejszych w latach 2012–2016

Lp.	Wyszczególnienie	Ogółem	Mężczyźni	Kobiety
1	Liczba osób urodzonych w 1916 r. i wcześniej wg GUS, stan na 31.12.2011 r.	25 757	5 243	20 514
2	Liczba zgonów w 2012 r. wśród osób urodzonych w 1916 r. lub wcześniej	7 353	1 611	5 742
3	Liczba zgonów w 2013 r. wśród osób urodzonych w 1916 r. lub wcześniej	5 311	1 059	4 252
4	Liczba zgonów w 2014 r. wśród osób urodzonych w 1916 r. lub wcześniej	3 641	762	2 879
5	Liczba zgonów w 2015 r. wśród osób urodzonych w 1916 r. lub wcześniej	2 759	515	2 244
6	Liczba zgonów w 2016 r. wśród osób urodzonych w 1916 r. lub wcześniej	1 724	327	1 397
7	Łączna liczba zgonów w latach 2011–2016 wśród osób urodzonych w 1916 r. lub wcześniej (suma wierszy 2–6 bieżącej tabeli)	20 788	4 274	16 514
8	Bilans jako wynik analizy zgonów: różnica między stanem osób generacji 1916 r. i wcześniejszych wg GUS, na 31.12.2011 r. (wiersz 1) i liczbą zgonów osób tych generacji w latach 2012–2016 (wiersz 7), czyli bilans stanu stulatków i osób starszych na 31.12.2016 r.	<b>4 969</b>	<b>969</b>	<b>4 000</b>
9	Liczba osób urodzonych w 1916 r. i wcześniej wg GUS, stan na 31.12.2016	<b>4 933</b>	<b>961</b>	<b>3 972</b>
10	Różnica między liczbą osób urodzonych w 1916 r. i wcześniej wg GUS, stan na 31.12.2016 r. a szacunkiem tej liczby	<b>-36</b> <b>(0,72%)</b>	<b>-8</b> <b>(0,83%)</b>	<b>-28</b> <b>(0,70%)</b>

Uwaga: Ze względu na brak szczegółowych danych według wieku w analizie pominięto migracje zewnętrzne.

Źródło: Na podstawie danych GUS.

Oczywiście można bardzo wnikliwie dociekać źródeł rozbieżności. Analiza takich dociekań, poczynawszy od danych spisu z 2002 r., jest przedstawiona w aneksach III.1 i III.2. W tym miejscu zostanie pominięte szczegółowe rozważanie wykazanych rozbieżności. Zwrócona zostanie jedynie uwaga na dwie znaczące różnice. Obie dotyczą lat bezpośrednio poprzedzających spis.

Pierwsza znaczna różnica przedstawia szacunek liczby osób urodzonych w 1916 r. i wcześniej na 31.12.2002 r. na podstawie stanu ludności według wieku na 31.12.2001 r. i liczby zgonów według wieku i roku urodzenia w 2002 r. (tab. 3.11). Wyjściowy stan ludności na koniec grudnia 2001 r. określono na podstawie danych Rocznika Demograficznego 2002, a więc przed korektą opracowaną na podstawie NSP 2002. Pozwalają one na określenie stanu według pojedynczych roczników wieku tylko w ograniczonym zakresie, tj. do 85 lat oraz osób w wieku 85 i więcej lat łącznie. Uwzględnienie stanu wyjściowego na koniec roku kalendarzowego upoważnia do kontynuowania szacunków według roczników urodzenia. Dodatkowo dane o liczbie zgonów w 2002 r. są przedstawione w bazie danych Demografia według podwójnej klasyfikacji wieku i roku urodzenia, co pozwala na szacunek bilansu według roczników urodzenia na 31.12.2002 (tab. 3.11).

Tabela 3.11. Liczba osób urodzonych w latach 1921 i wcześniej według bazy danych Demografia GUS oraz bilansu liczby osób urodzonych w latach 1921 i wcześniej na podstawie stanu ludności według wieku na 31.12.2001 r. i liczby zgonów według wieku i roku urodzenia w 2002 r.

Rok urodzenia	Stan ludności wg roku urodzenia na 31.12.2002 r., na podst. bazy danych Demografia (NSP 2002)			Szacunek stanu ludności wg roku urodzenia na 31.12.2002 r., na podst. stanu ludności na 31.12.2001 r. <sup>a</sup> wg wieku i liczby zgonów w 2002 r. wg wieku i roku urodzenia			Różnica między stanem 31.12.2002 r., na podst. bazy danych Demografia (NSP 2002) a szacunkiem stanu z analizy zgonów		
	Og	M	K	Og	M	K	Og	M	K
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8) = (2) - (5)	(9) = (3) - (6)	(10) = (4) - (7)
1921	122 676	38 775	83 901	123 007	40 263	82 744	-331	-1 488	1 157
1920	109 535	34 912	74 623	114 355	37 561	76 794	-4 820	-2 649	-2171
1919	86 917	26 300	60 617	85 587	27 076	58 511	1 330	-776	2 106
1918	55 069	16 019	39 050	59 566	17 714	41 852	-4497	-1 695	-2 802
1917	48 036	14 241	33 795	48 079	14 752	33 327	-43	-511	468
1916 i wcześniej	281 489	71 970	209 519	<b>280 330</b>	<b>78 109</b>	<b>202 221</b>	<b>1 159</b>	<b>-6 139</b>	<b>7 298</b>

<sup>a</sup>Na podstawie Rocznika Demograficznego 2002.

Og – ogółem, M – mężczyźni, K – kobiety.

Źródło: Na podstawie danych GUS.

Przed wszystkim oparte na spisie szacunki GUS dla większości roczników urodzenia były niższe od szacunków stanu otrzymanych na podstawie analizy zgonów. Zarówno absolutna, jak i względna różnica jest duża, sięgająca blisko 5 tys. dla osób urodzonych w 1920 r. oraz ponad 10,5% dla urodzonych w 1918 r. Z kolei dla osób urodzonych w 1916 r. i wcześniej niezgodność szacunków wynosząca 1159 osób rozkłada się nierównomiernie według płci, gdzie zaobserwowano ujemną różnicę liczby mężczyzn między szacunkami GUS i własnymi oraz dodatnią dla kobiet. Przy czym różnica dla ludności ogółem stanowiła około 0,4% szacunku GUS, ale dla mężczyzn było to już ponad 8,5%, a dla kobiet prawie 3,5%.

Druga rozbieżność dotyczy 2010 r.. Przeprowadzono analizę zgonów stulatków w latach 2002–2016, według algorytmu szacowania bilansu ludności opisanego na początku tego punktu (tab. 3.10 wraz z opisem), w odniesieniu do osób urodzonych równo 100 lat wcześniej. W analizie uwzględniono zgony według roczników urodzenia, otrzymując szacunek bilansu 100-latków na 31 grudnia kolejnego roku kalendarzowego (tab. 3.12 oraz aneks III.1 i III.2). Różnice przeprowadzonego szacunku w porównaniu do danych GUS dla badanego okresu są niewielkie: od zera dla stanu na 31.12.2016 r. do 10 dla stanu na 31.12.2003 r. Przy czym zero dla 2016 r. nie oznacza pełnej zgodności bilansu, gdyż odnotowano różnice według płci, otrzymując o jeden większą liczbę mężczyzn i o jeden mniejszą liczbę kobiet. Jediną znaczącą różnicę odnotowano w 2010 r., dla którego różnica między przeprowadzonym szacunkiem a danymi GUS wyniosła 1331 osób. Okazuje się jednak, że nie popełniono błędu w obliczeniach, a zaobserwowana rozbieżność wynika z przeprowadzonej przez GUS korekty stanu i struktury ludności na podstawie wyników NSP 2011. Zestawiając obydwa rodzaje danych dla 2010 r. (tab. 3.12), łatwo zauważyć, że gdyby nie skorygowano struktury, to przeprowadzony bilans przewyższa dane GUS o dwie osoby: jednego mężczyznę i jedną kobietę<sup>140</sup>.

Wyniki spisów są wykorzystywane między innymi do korekty bieżących szacunków stanu i struktury ludności nie tylko w roku przeprowadzania spisu, ale także w latach poprzednich<sup>141</sup>. Przyjmując to wyjaśnienie dla danych z 2010 r. i następných lat, można potwierdzić dużą zgodność prezentowanego w niniejszym opracowaniu szacunku bilansu liczby osób starszych z szacunkami przedstawionymi przez GUS. Fakt spójności danych ewidencji zgonów osób starszych oraz szacunków spisowych i bazujących na tych danych bilansów dla kolejnych lat może stanowić podstawę pozytywnej oceny jakości szacunków spisowych.

<sup>140</sup> Autorka nie ma niestety możliwości dalszej weryfikacji szacunków, tak jak przedstawiono w odniesieniu do 2002 r. (tab. 3.13).

<sup>141</sup> Jak podaje GUS, dane dotyczące liczby i struktury ludności prezentowane w bazie Demografia dla lat 2002–2009 zostały opracowane przy uwzględnieniu wyników NSP 2002; rok 2010 jest prezentowany w dwóch wersjach, tj. jako opracowanie uwzględniające wyniki NSP 2002 oraz przy uwzględnieniu wyników NSP 2011, natomiast od 2011 r. dane są opracowywane przy uwzględnieniu wyników NSP 2011 (<http://demografia.stat.gov.pl/bazademografia/Tables.aspx>).

Tabela 3.12. Analiza zgonów w najstarszych grupach wieku

	Wyszczególnienie	Ogółem	Mężczyźni	Kobiety
2001	Liczba osób w wieku 85 i więcej lat (urodzonych w 1916 r. i wcześniej) na 31.12.2001 r.	339 778	95 265	244 513
2002	Liczba zgonów w 2002 r. wśród osób urodzonych w 1916 r. i wcześniej wg roczników urodzenia	59 448	17 156	42 292
	Szacunek bilansu liczby osób urodzonych w 1916 r. i wcześniej na 31.12.2002 r.	280 330	78 109	202 221
	Liczba osób urodzonych w 1916 r. i wcześniej na 31.12.2002 r.	281 489	71 970	209 519
	<b>Różnica</b>	<b>1 159</b>	<b>-6 139</b>	<b>7 298</b>
⋮	⋮	⋮	⋮	⋮
2009	Liczba zgonów w 2009 r. wśród osób urodzonych w 1909 r. i wcześniej wg roczników urodzenia	1 488	305	1 183
	Szacunek bilansu liczby osób urodzonych w 1909 r. i wcześniej na 31.12.2009 r.	3 766	923	2 843
	Liczba osób urodzonych w 1909 r. i wcześniej na 31.12.2009 r.	3 765	923	2 842
	Różnica	1	0	1
2010	Liczba zgonów w 2010 r. wśród osób urodzonych w 1910 r. i wcześniej wg roczników urodzenia	1 606	326	1 280
	Szacunek bilansu liczby osób urodzonych w 1910 r. i wcześniej na 31.12.2010 r.	4 419	1 115	3 304
	Liczba osób urodzonych w 1910 r. i wcześniej na 31.12.2010 r.	4 417	1 114	3 303
	Liczba osób urodzonych w 1910 r. i wcześniej na 31.12.2010 r. korekta wg NSP 2011	3 088	678	2 410
	<b>Różnica</b>	<b>1 331</b>	<b>437</b>	<b>894</b>
⋮	⋮	⋮	⋮	⋮
	Szacunek bilansu liczby osób urodzonych w 1915 r. i wcześniej na 31.12.2015 r.	4 763	945	3 818
2015	Liczba osób urodzonych w 1915 r. i wcześniej na 31.12.2015 r.	4 755	944	3 811
	Różnica	8	1	7
	Liczba zgonów w 2016 r. wśród osób urodzonych w 1916 r. i wcześniej wg roczników urodzenia	1 724	327	1 397
	Szacunek bilansu liczby osób urodzonych w 1916 r. i wcześniej na 31.12.2016 r.	4 933	962	3 971
2016	Liczba osób urodzonych w 1916 r. i wcześniej na 31.12.2016 r.	<b>4 933</b>	<b>961</b>	<b>3 972</b>
	Różnica	0	1	-1

Źródło: Szacunek własny na podstawie danych Rocznika Demograficznego 2002, baz danych GUS: Demografia, Bank Danych Lokalnych i zasobów danych obszarów tematycznych: Spisy Powszechne.

Od 2007 r. prezentowane dane dotyczące zgonów według roczników urodzenia są bardziej szczegółowe i uwzględniają liczby zgonów według pojedynczych roczników nawet dla osób urodzonych w XIX w. Stąd możliwe jest oszacowanie bilansu – różnicy liczby osób według roczników urodzenia i liczby zgonów według roczników urodzenia w poszczególnych latach. Wyniki szacunków przedstawiono



w aneksie III.1. Dane o zgonach według roczników urodzenia można zestawić bezpośrednio z danymi określającymi stan ludności według roczników urodzenia (tab. 3.13), co pozwala na weryfikację tezy Paradysza (2010) o brakach w spisie osób starszych wskazującą na „zgony duchów” w kolejnych latach po spisie. Rzeczywiście sytuacje takie się zdarzają. Przyjmując stan ludności na 31.12.2006, określony na podstawie danych NSP 2002 i, śledząc porządek wymierania kobiet urodzonych w 1909 r., w 2016 r. otrzymujemy ujemne liczby, tzn. zgony kobiet, których nie spisano. Zauważmy jednak, że to korekta stanu na podstawie NSP 2011 spowodowała obniżenie liczby osób wynikającej z bilansu opartego na danych NSP 2002 i ewidencji ruchu naturalnego. Przeprowadzając w 2011 r. spis wykorzystujący dane rejestru PESEL, GUS starał się uporządkować wiele kwestii. Jedną z nich były nadmierne liczby osób starszych w rejestrze, osób, które wyemigrowały i prawdopodobnie zmarły na obczyźnie, ale przez długi czas były uwzględniane w statystyce ludności.

Na marginesie analizy stanu i porządku zgonów stulatków warto zauważyć stały wzrost liczby osób w wieku 100 i więcej lat, od 1890 w 2003 r. do 4933 w 2016 r.

Tabela 3.13. Liczba osób urodzonych w 1920 r. i wcześniej oraz zgony według roczników urodzenia

Rok urodzenia	Stan na 31.12.2006 r.			Suma zgonów w latach 2007–2016 wg roczników urodzenia			Różnica między stanem na 31.12.2006 r. i sumą zgonów w latach 2007–2016		
	ogółem	mężczyźni	kobiety	ogółem	mężczyźni	kobiety	ogółem	mężczyźni	kobiety
1920	70 543	20 308	50 235	60 534	17 905	42 629	10 009	2 403	7 606
1919	52 627	14 448	38 179	48 153	13 396	34 757	4 474	1 052	3 422
1918	31 837	8 335	23 502	29 111	7 710	21 401	2 726	625	2 101
1917	26 445	7 030	19 415	24 824	6 558	18 266	1 621	472	1 149
1916	22 953	5 752	17 201	21 694	5 465	16 229	1 259	287	972
1915	19 875	4 809	15 066	19 281	4 588	14 693	594	221	373
1914	19 962	4 717	15 245	19 080	4 464	14 616	882	253	629
1913	15 191	3 621	11 570	14 727	3 448	11 279	464	173	291
1912	11 910	2 599	9 311	11 155	2 372	8 783	755	227	528
1911	7 927	1 776	6 151	7 582	1 643	5 939	345	133	212
1910	6 023	1 327	4 696	5 524	1 183	4 341	499	144	355
1909	3 539	763	2 776	3 507	723	2 784	32	40	-8
1908	2 381	538	1 843	2 227	486	1 741	154	52	102
1907	1 673	402	1 271	1 459	332	1 127	214	70	144
1906 i wcześniej	2919	680	2239	2022	434	1 588	897	246	651

Źródło: Szacunek własny na podstawie danych Rocznika Demograficznego 2002, baz danych GUS: Demografia, Bank Danych Lokalnych i zasobów danych obszarów tematycznych: Spisy Powszechne.

### 3.4.5. Ocena długiego formularza NSP 2011

Opisując metodę przeprowadzenia NSP 2011, wskazano na podział badania na część pełną oraz reprezentacyjną (długi formularz). W punkcie 2.2 krótko opisano schemat doboru próby, podkreślając głębokie warstwowanie. Koncepcja warstwowania została tak przemyślana, by zapewnić efektywność losowania dla zróżnicowanych jednostek, przy czym uwzględniono 12 dodatkowych cech przypisanych w operacie losowania poszczególnym mieszkańcom. Próba wylosowana do spisu liczyła ponad 2,7 mln. mieszkań (z blisko 13,5 mln) w utworzonych 70,5 tys. warstw (GUS, 2012a, s. 22). W powiatach przeprowadzono losowanie jednostopniowe, a w alokacji próby pomiędzy warstwy zastosowano metodę pierwiastkową, w której liczba mieszkań losowanych w każdym powiecie była proporcjonalna do pierwiastka kwadratowego z liczby mieszkań w populacji.

$$\Delta_p = \frac{\hat{y}_p - Y_p}{Y_p} \cdot 100, \quad (3.7)$$

gdzie:

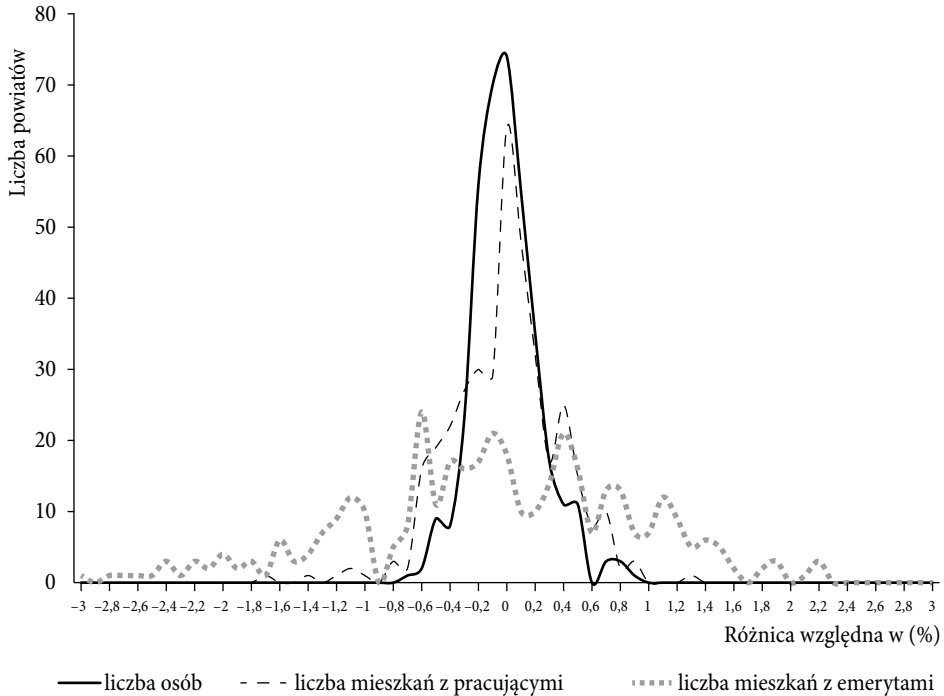
$\hat{y}_p$  – oszacowanie sumy wartości zmiennej  $Y$  w powiecie  $p$ -tym na podstawie wylosowanej próby,

$Y_p$  – suma wartości zmiennej  $Y$  w powiecie  $p$ -tym obliczona na podstawie danych z operatu losowania.

Wielkość próby, schemat doboru, a przede wszystkim głębokie warstwowanie oraz alokacja, zostały tak dobrane, by zapewnić pozyskanie precyzyjnych szacunków na poziomie powiatów z rozróżnieniem obszarów miejskich i wiejskich oraz dzielnic i delegatur w dużych miastach. Efektywność schematu losowania można ocenić na podstawie załączonych do opracowania (GUS, 2011a) informacji o wartości różnic względnych dla wybranych charakterystyk (wzór 3.7 oraz rys. 3.9).

Okazuje się, że otrzymane na podstawie próby oceny estymatorów bezpośrednich wartości globalnej wyodrębnionych charakterystyk są bardzo bliskie odpowiadającym im wartościom w populacji. Dla trzech spośród dwunastu cech<sup>142</sup>: (i) liczba osób, (ii) liczba mieszkań z osobą pracującą i (iii) liczba mieszkań z emerytem lub rencistą, ale bez osoby pracującej, średnie z różnic względnych w przekroju powiatów są tylko nieco niższe od zera, co sugeruje nieznaczne zaniżenie szacunków (-0,005; -0,017; -0,088, odpowiednio).

<sup>142</sup> W załączniku przedstawiono dane dotyczące 5 spośród 12 charakterystyk. Oprócz przedstawionych w opracowaniu są to jeszcze: liczba mieszkań z rolnikami oraz powierzchnia użytków rolnych. Dla obu tych zmiennych różnice względne między bezpośrednimi ocenami wartości globalnych oraz wartościami w populacji są zdecydowanie wyraźniejsze. W przypadku mieszkań z rolnikami sięgają 13%, ale dla powierzchni użytków rolnych dla niektórych powiatów są znaczne. Można przypuszczać, że mniejsza efektywność szacunku tych charakterystyk wynika z ich związku z rolnictwem.

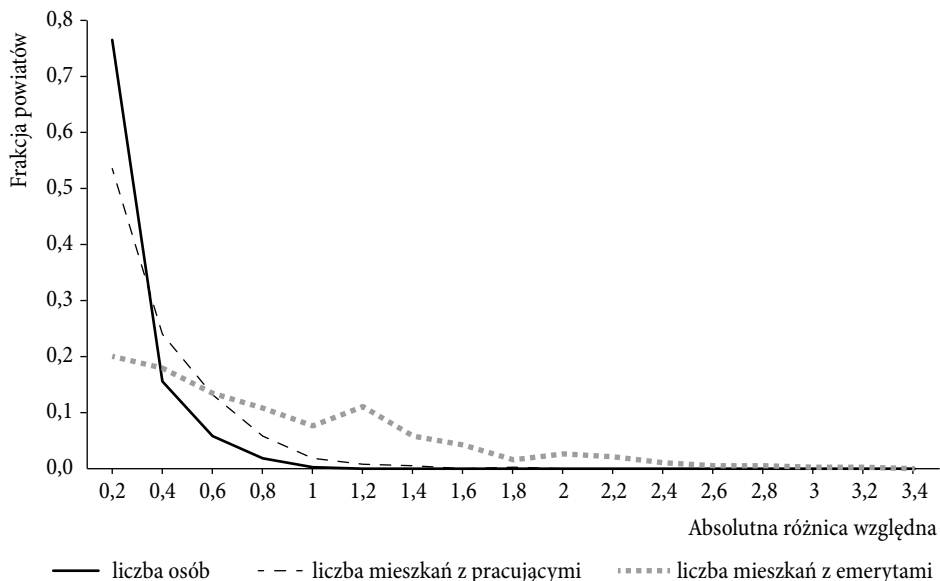


Uwaga: Efektywność schematu losowania była mierzona różnicą względną (w %) w przekroju powiatów dla zmiennych: (i) liczba osób, (ii) liczba mieszkań z osobą pracującą i (iii) liczba mieszkań z emerytem lub rencistą, ale bez osoby pracującej

*Rysunek 3.9. Rozkład powiatów według względnych różnic między oceną bezpośrednią wartości globalnej oraz wartością w operacji losowania dla wybranych charakterystyk, badanie reprezentacyjne NSP 2011*

Źródło: Na podstawie (GUS, 2011a)

Z kolei rozkład oszacowanych absolutnych różnic względnych jest skrajnie asymetryczny prawostronnie (GUS, 2012c) (por. rys. 3.10, tab. 3.14), choć występują istotne różnice w zależności od cechy. Największą asymetrię obserwujemy dla cechy „liczba osób”. W ponad 93% powiatów różnica względna między oceną estymatora bezpośredniego wartości globalnej liczby osób a wartością w populacji nie przekracza 0,2%. W żadnym z powiatów różnica ta nie jest większa aniżeli 1%. W odniesieniu do pozostałych zmiennych efektywność szacunków jest nieco słabsza. Dla liczby mieszkań z osobami pracującymi różnice względne dla ponad 90% powiatów nie przekraczają 0,4%, dla jednego zaledwie powiatu różnica wynosi 1,5%. Dla liczby mieszkań z emerytami i rencistami rozkład względnych różnic między ocenami bezpośrednimi i w populacji jest nieco bardziej płaski. Dla połowy powiatów różnice względne nie przekraczały 0,6%, ale dla żadnego nie były większe aniżeli 3%.



Uwaga: Efektywność mierzona była różnicą względną (w %) w przekroju powiatów dla zmiennych: (i) liczba osób, (ii) liczba mieszkań z osobą pracującą i (iii) liczba mieszkań z emerytem lub rencistą, ale bez osoby pracującej

Rysunek 3.10. Rozkład powiatów według absolutnych różnic względnych między oceną bezpośrednią wartości globalnej oraz wartością w operacji losowania dla wybranych charakterystyk, badanie reprezentacyjne NSP 2011

Źródło: Na podstawie (GUS, 2011a)

Tabela 3.14. Ocena efektywności schematu losowania: rozkład powiatów według absolutnej różnicy względnej między bezpośrednią oceną estymatora a wartością w operacji losowania dla wybranych charakterystyk, badanie reprezentacyjne NSP 2011

Absolutna różnica względna (w %)	Udział powiatów o wartości absolutnej różnicy względnej dla zmiennej		
	liczba osób	liczba mieszkań z pracującymi	liczba mieszkań z emerytami
[0,0–0,2]	0,765	0,536	0,201
[0,2–0,4]	0,156	0,240	0,179
[0,4–0,6]	0,058	0,132	0,135
[0,6–0,8]	0,018	0,058	0,108
[0,8–1,0]	0,003	0,018	0,077
[1,0–1,5]	0,000	0,013	0,190
[1,5–2,0]	0,000	0,003	0,063
[2,0–2,5]	0,000	0,000	0,034
[2,5–3,0]	0,000	0,000	0,011
[3,0–3,5]	0,000	0,000	0,003
Razem	1	1	1

Źródło: Na podstawie (GUS, 2011a).

Na podstawie przedstawionego przez GUS opisu losowania próby w spisie metodą reprezentacyjną (GUS, 2012c) można stwierdzić, że operat losowania mieszkań przygotowano starannie, tworząc zintegrowany wykaz adresowo-mieszkaniowy, wykorzystano różne źródła, informacje uzupełniono dodatkowymi zmiennymi. Jak wspomniano w podrozdziale 2.2, wyniki badania reprezentacyjnego zostały uogólnione na poziomie powiatów z uwzględnieniem przekroju miasto-wieś. Przy czym uogólnienie wyników przeprowadzono dwuetapowo: w pierwszej kolejności dla mieszkań, a następnie dla osób. Ponadto przeprowadzono korektę wag oraz ich kalibrację. Szczególnie istotna jest kwestia związana z licznymi brakami odpowiedzi wynoszącymi 13,7% mieszkań. Niestety, w opisie metodologicznym znajdujemy jedynie bardzo syntetyczną informację dotyczącą owej korekty, mianowicie że uwzględniono wskaźnik korygujący wyznaczany jako iloraz liczby mieszkań, które powinny być zbadane, do liczby mieszkań, dla których udało się zrealizować badanie. Dla wylosowanej próby liczącej 2684,2 tys. mieszkań zdołano zebrać wypełnione formularze dla 2317,7 tys. mieszkań. Stwierdzono brak 140,4 tys. mieszkań (5,2% próby), które potraktowano jako błąd operatu losowania. Dodatkowo w trakcie obwodu przedspisowego (już po etapie losowania) dopisano 105,4 tys. mieszkań, z których w 45,3 tys. przeprowadzono badanie (poza próbą było 43,1 tys. mieszkań). Inaczej wyznaczono wagi dla mieszkań i dla budynków jednorodzinnych, odrębnie też potraktowano część próby z mieszkaniem dopisanymi. Waga skorygowana była wyliczana jako iloczyn wyznaczonej korekty i wagi z losowania (GUS, 2011a, 2013a, s. 47 i nast.). Natomiast kalibracja w NSP 2011 miała na celu dostosowanie struktur według płci i wieku uzyskanych w wyniku badania reprezentacyjnego (długi formularz) do wyników spisu pełnego (krótki formularz).

Opis badania reprezentacyjnego w spisie nie informuje precyzyjnie o procedurze eliminacji wpływu odmów udziału w badaniu. Nie znamy rozkładu odmów w przekroju terytorialnym ani według żadnej innej cechy. Natomiast bardzo pozytywnie należy ocenić sposób prezentacji wyników spisu dla części reprezentacyjnej, gdzie zamieszczone zostały tablice precyzji szacunków. Zostały one opatrzone przykładową interpretacją oraz wyjaśnieniami sugerującymi zachowanie ostrożności w szczególnych przypadkach.

Jeśli chodzi o badanie reprezentacyjne w spisie, warto jeszcze podkreślić, że mimo że zaprojektowano je dla uogólnień na poziomie powiatów, projektując dobór próby uwzględniono przewidziane prawem wymagania danych dla niższych poziomów agregacji. Przykładowo, szczególne podejście zastosowano w NSP 2011 w celu szerszego uwzględnienia problemów mniejszościowych<sup>143</sup>. GUS zdecydował,

---

<sup>143</sup> Przy badaniu tych problemów konieczny jest samospis. Według ocen działaczy łemkowskich NSP obejmuje zaledwie niewielką ich część. Petro Trochanowski szacuje ich na 60–200 tysięcy. Przy nastrojach antyukraińskich w Polsce tylko niewielka część wyjawia swoją etniczność przed rachmistrzem spisowym. Podobnie jest zresztą w przypadku Żydów, których NSP wykazuje jeszcze mniej niż Łemków.

by w ramach NSP 2011 przeprowadzić pełne bezpośrednie badanie w 86 gminach wyróżnionych według kryterium „mniejszościowego”<sup>144</sup>. Zatem specjalnemu badaniu poddano gminy, co do których były przesłanki, na podstawie wcześniejszych informacji z NSP 2002, że występuje w nich znaczna frakcja ludności należącej do mniejszości (GUS, 2012a, s. 24). Przyjęto także założenie o dużym prawdopodobieństwie zmiany struktury narodowo-etnicznej (w porównaniu do ustalonej w NSP 2002). Dlatego w badaniu pełnym uwzględniono gminy, w których odsetek mniejszości był (w NSP 2002) dwukrotnie niższy, niż wskazuje kryterium mniejszościowe (co najmniej 10% ogółu mieszkańców)<sup>145</sup>.

Na koniec nasuwa się jeszcze jedna uwaga. W spisie nie przedstawiono informacji o sposobie szacunku charakterystyk w przekroju gmin, oczywiście poza tymi, które zostały objęte spisem pełnym (formularz krótki). Do tej pory opublikowano informacje o liczbie ludności w przekroju gmin oraz dane dotyczące ludności w rejonach statystycznych i obwodach spisowych według płci, ekonomicznych i 10-letnich grup wieku. Wszystkie te informacje były zawarte w krótkim formularzu. Możliwe, że w celu innych charakterystyk na poziomie gmin wykorzystano metody estymacji pośredniej. Powinny one być szczegółowo opisane, łącznie z wyborem i analizą zmiennych wspomagających, ich relacji z wielkościami szacowanymi, metodą oceny precyzji.

---

<sup>144</sup> Realizacja wynikających z ustawy uprawnień przyjmuje specyficzną formę w gminach, w których liczba mieszkańców należących do mniejszości narodowej, etnicznej czy posługujących się językiem regionalnym jest nie mniejsza niż 20% ogólnej liczby mieszkańców gminy. Gminy takie mają zagwarantowane prawo m.in. do swobodnego posługiwania się językiem, jego nauki bądź naukę w języku mniejszości oraz różnorodną działalność mającą na celu ochronę, zachowanie i rozwój tożsamości kulturowej mniejszości, z czym się wiążą określone dotacje z budżetu gminy i państwa.

<sup>145</sup> Jak wskazuje Łodziński (2012, s. 15), badanie to podjęto w wyniku ustaleń spotkania przedstawicieli GUS i MSWiA w lipcu 2010 r. Początkowo rozważano objęcie nim gmin, w których odsetek mniejszości był wyższy aniżeli 1% (według szacunku na podstawie NSP 2002).

---

## 4. PERSPEKTYWY SPISÓW LUDNOŚCI<sup>146</sup>

---

### 4.1. Perspektywy spisu w świetle wytycznych ONZ i Eurostatu

W końcu sierpnia 2014 r. odbyło się spotkanie Dyrektorów Statystyki Społecznej urzędów statystycznych państw Unii Europejskiej (DSS UE). Tematem spotkania był program spisów ludności i mieszkań w państwach UE w kolejnej rundzie – 2020 – oraz w dalszej perspektywie. Dyskutowano trzy grupy zagadnień: (i) jakość spisów rundy 2010, (ii) sposób przeprowadzenia spisów w 2021 r. oraz (iii) modernizację europejskiej statystyki ludności i mieszkań po 2021 r. (Beyond 2021). Echo tej dyskusji oraz kolejnych spotkań powtarza wśród statystyków wieść o końcu spisów. W rzeczywistości sprawa dotyczy nie tyle przysłowiowego „końca spisów”, co modernizacji całego systemu statystyki demograficznej i społecznej (EU, 2014, 2016).

Przede wszystkim podczas spotkania DSS UE ustalono, że w kolejnej rundzie spisy zostaną przeprowadzone zgodnie z dotychczas obowiązującymi w UE regulacjami prawnymi (UE, 2008). Przyjęto jednak, że konieczne są nowe regulacje dotyczące definicji, przekrojów, klasyfikacji, tabel krzyżowych (cross-tabulacji), metadanych oraz jakości raportowania. Podkreślono, że kwestie zasadniczych zmian dotyczących podejścia, stosowanych metod i wykorzystywanych źródeł danych są podejmowane na poziomie krajowym. Eurostat z kolei kontynuuje prace w zakresie modernizacji systemu statystyki społecznej. Zwrócono uwagę na potrzebę równoważenia istotnych i zmieniających się potrzeb związanych ze spisem w państwach UE w kontekście aspektów praktycznych i obciążeń wynikających z tworzenia statystyk. Jako istotny podjęto temat integracji danych statystycznych o ludności i danych przestrzennych w celu dostarczenia danych demograficznych według miejsca zamieszkania w przekroju terytorialnym dla siatki gridowej (1 km<sup>2</sup>, geo-grid: dyrektywa INSPIRE). Na marginesie warto zaznaczyć, że takie badania

---

<sup>146</sup> Rozdział 4 powstał w ramach projektu: Estymacja pośrednia w zakresie badania niepełnosprawności na podstawie NSP 2011. Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/11/B/HS4/01472.

dotyczące rozwoju demograficznego Poznania w kontekście starzenia się ludności są prowadzone w Ośrodku Statystyki Miast Urzędu Statystycznego w Poznaniu<sup>147</sup>.

Dyrektorzy Statystyki Społecznej UE (EU, 2014) uzgodnili, że fundamentalne zmiany w zakresie przebudowy europejskiej statystyki ludności i mieszkań planuje się przeprowadzić po spisie 2021 r. Będą one uwzględniać główne zmiany zachodzące w wielu państwach członkowskich, a mianowicie szersze wykorzystanie rejestrów administracyjnych oraz informacji z regularnie prowadzonych badań specjalnych. Rozważano także wykorzystanie wielu form gromadzenia danych, w szczególności badań prowadzonych przez urzędy statystyczne za pośrednictwem internetu, w tym pełnych. Duże nadzieje wiąże się z możliwością częstszego przeprowadzania spisów, jaką stwarza korzystanie z danych rejestrów administracyjnych. Wprost określono możliwość „corocznego raportowania” danych. Podejście takie stwarza możliwość bardziej równomiernego rozłożenia kosztów. Oczywiście nakreślona strategia jest bardzo elastyczna, uwzględniając prawdopodobny rozwój metod statystycznych oraz dostępność nowych źródeł danych, jak również pojawienie się nowych potrzeb względem danych statystycznych. W perspektywie kolejnych 10 lat należy wziąć pod uwagę możliwe zmiany polityczne, społeczne, uwarunkowania technologiczne i finansowe. Będąc świadomym powszechnego dostępu do danych z różnych źródeł oraz wyzwań i konkurencyjności wobec statystyki publicznej, w dyskusji podczas DSS UE (EU, 2014) zwrócono uwagę na dostępność bieżących danych o aktualnych problemach w różnych obszarach tematycznych spisu, kontynuowanie gromadzenia danych zharmonizowanych i dalszy postęp w tym zakresie, dostępność danych dla bardziej szczegółowych podziałów, w tym dane geo-grid (1 km<sup>2</sup>), możliwość badania relacji między zmiennymi w różnych ujęciach krzyżowych (cross-tabulacji), zwiększenie częstości gromadzenia i upowszechniania danych.

W 2015 r. Konferencja Statystyków Europejskich ONZ (UNECE CES<sup>148</sup>) opublikowała Rekomendacje dotyczące kolejnej rundy spisów ludności w 2020 r. dla regionu państw europejskich (UNECE, 2015). Ogólnie rzecz ujmując, rekomendacje ONZ (UN, 2015a) oraz Rekomendacje Konferencji Statystyków Europejskich ONZ (UNECE, 2015) są spójne. Oczywiście zalecenia ogólne ONZ mają szerszy charakter, a zalecenia CES są bardziej szczegółowe w odniesieniu do niektórych definicji i klasyfikacji. Wyróżniono tylko kilka tematów spisowych, np. „lokalizacja miejsca pracy”, które w zaleceniach globalnych określono jako uzupełniające, podczas gdy w zaleceniach Konferencji Statystyków Europejskich uznano te tematy za podstawowe (i podobnie w odwrotnym kierunku). W Rekomendacjach Konferencji Statystyków Europejskich uwagę zwraca jednak kilka kwestii sugerujących odrębność spisów w Europie. Jest to geostatystyczne ujęcie spisu, uwzględnienie nieformalnego wymiaru stanu cywilnego –związki nieformalne.

<sup>147</sup> Można tutaj wymienić następujące prace (Basarbowicz i in., 2015; Filas-Przybył, 2017).

<sup>148</sup> United Nations Economic Commission for Europe Conference for European Statisticians.



W opublikowanych przez Konferencję Statystyków Europejskich ONZ Rekomendacjach (UNECE, 2015) rozdział VI jest w całości poświęcony wprowadzeniu siatki gridowej (1 km<sup>2</sup>). Temat ten jest prezentowany jako dodatkowy, jednak dla państw UE jest to podejście o charakterze podstawowym. „Przypisanie” każdej osoby, dla której wypełniany jest formularz spisowy, oraz każdego budynku, dla którego gromadzone są dane, do konkretnej lokalizacji – określonej przez współrzędne geograficzne – oznacza powiązanie danych statystycznych z określonym miejscem w przestrzeni. Pojedyncze gridy mogą być dowolnie łączone, tak by odpowiadały rzeczywistym jednostkom terytorialnym, jak również w dowolny inny sposób, w zależności od potrzeb. Każdy kwadrat o boku 1 km (grid) może zostać opisany w wymiarze demograficznym, statystycznym, społecznym i ekonomicznym. Integracja danych statystycznych o ludności z przestrzenią wprowadza nowy wymiar analizy.

Gwałtowne zmiany społeczne spowodowały konieczność uwzględnienia w danych spisowych struktury ludności według stanu cywilnego w ujęciu rzeczywistym, a nie tylko formalnoprawnym. Informacje o związkach konsensualnych nie są oczywiście dostępne na podstawie źródeł administracyjnych. Zagadnienie „rzeczywistego stanu cywilnego” jako cechy podstawowej w państwach europejskich zostało przedstawione w rozdziale VII Rekomendacji poświęconym charakterystykom demograficznym. Wśród charakterystyk aktywności ekonomicznej (rozdział VIII) wyróżniono „wytwarzanie towarów na użytek własny”. Zagadnienie to jest tematem dodatkowym, jednak zostało uznane za bardzo ważny w krajach, w których taka praca ma szczególne znaczenie wśród aktywności na rynku pracy. W części III (rozdział XV) poświęconej charakterystykom zasobów mieszkaniowych duży nacisk położono na kwestie wyposażenia mieszkań i budynków, uznając, że w niektórych państwach mogą one być postrzegane jako czynniki dyskryminujące standard warunków mieszkaniowych.

## **4.2. Perspektywy spisów rundy 2020 w wybranych państwach**

### **4.2.1. Stany Zjednoczone Ameryki**

Zmiany metodologii przeprowadzania spisu proponowane w Stanach Zjednoczonych wynikają między innymi ze wzrastającej niechęci ludności do udziału w spisie. W 2000 r. Senat Stanów Zjednoczonych podjął niewiążącą rezolucję, aby nie karać osób, które odmówiły odpowiedzi na pytania „długiego formularza” spisowego, co w opinii Prewitt (2004, s. 1452) stanowiło skuteczne przyzwolenie społeczne na lekceważenie prawa. W rezultacie nastąpił wzrost odsetka odmów odpowiedzi na takie pytania „długiego formularza” jak wysokość czynszu: z 1% w spisie 1990 r. do 16% w spisie 2000 r. czy wysokość dochodu: z 10% do 20% odpowiednio. Jednym

z przedsięwzięć podjętych w celu zmniejszenia obciążeń społeczeństwa udziałem w badaniach statystycznych oraz uzyskania poprawy jakości było zastąpienie długiego formularza corocznym badaniem American Community Survey (ACS). Ponieważ każdego roku obejmuje ono ponad 3 miliony osób, zaplanowano, że będzie odgrywać rolę spisu kroczącego. Jednak również ACS jest przedmiotem kontrowersji politycznych i społecznych, czego wyrazem jest nieudana próba przegłosowania obowiązkowego statusu badania w Izbie Reprezentantów (Kukutai i in., 2015).

Oczekuje się, że wprowadzenie innowacyjnych zmian umożliwi redukcję kosztów o około 5,2 miliarda dolarów w porównaniu do sytuacji, gdyby w 2020 r. powtórzyć metodę przeprowadzania spisu z 2010 r. Planowane zmiany dotyczą czterech kluczowych kwestii: (1) rejestru adresów, (2) samospisu, (3) wykorzystania rejestrów administracyjnych oraz (4) monitorowania realizacji spisu. U.S. Census Bureau<sup>149</sup> wnikliwie ocenia procedury proponowane dla spisu 2020 r. w celu zapewnienia, że innowacje prowadzące do obniżenia kosztów nie spowodują obniżenia jakości. Analiza wpływu czterech kluczowych obszarów innowacji na jakość dostarcza następujących spostrzeżeń:

1. **Rejestr adresów** – planuje się wykorzystanie GIS i zdjęć lotniczych dla ułatwienia prac w terenie i zwiększenia ich efektywności. W celu aktualizacji rejestru adresów, obok systemów geograficznych, planuje się wykorzystać także rejestry administracyjne oraz zasoby danych biznesowych. Od 2015 r. trwają intensywne prace uwzględniające następujące zadania:
  - pomiar pokrycia i aktualizacja rejestru adresów (*Address Canvassing*),
  - lokalna weryfikacja i aktualizacja list adresowych w terenie (*Local Update of Census Addresses* – LUCA),
  - wykorzystanie programów geograficznych (GIS) dla ułatwienia pracy rachmistrzów spisowych w terenie,
  - aktualizacja instrukcji i procedur dotyczących gromadzenia danych.
 Zasadniczym celem tych działań ma być poprawa jakości głównego rejestru adresów – *Master Address File* (MAF). Zadanie to ma być zrealizowane w połowie 2019 r., tak by zaktualizowany MAF był gotowy do wykorzystania w spisie 2020 r.
2. **Samospis** – planuje się wprowadzenie samospisu internetowego oraz innych nowoczesnych metod gromadzenia informacji. W zasadzie wszelkie formy gromadzenia danych zaplanowano w taki sposób, by maksymalnie ułatwić udzielenie odpowiedzi: w każdym miejscu, w każdym czasie, głównie za pośrednictwem internetu. Przygotowywana jest także intensywna kampania promująca wszelkie działania związane ze spisem, a przede wszystkim zachęcająca do udziału w badaniu *American Community Survey* (ACS). Kampania ta ma

<sup>149</sup> U.S. Census Bureau to nazwa urzędu statystycznego w USA, Biuro Spisów.

charakter motywujący ludność do udziału w spisie, przede wszystkim poprzez zapewnienie o bezpieczeństwie danych i zachowaniu poufności. Wykorzystuje się różne kanały informacji, specjalne programy współpracy, dostosowane do różnych grup ludności. Przygotowanie samospisu i upowszechnianie informacji o tej formie gromadzenia danych jest ukierunkowane na realizację konkretnych celów, w tym:

- ograniczenie zbierania danych za pośrednictwem kwestionariusza papierowego – *Paper Data Capture*,
  - osiągnięcie jak największego udziału ludności w Samospisie internetowym – *Internet Self-Response*,
  - umożliwienie wprowadzania danych bez wykorzystania numerów identyfikacyjnych – *Non-ID Processing*,
  - zapewnienie systemu pomocy w wypełnieniu kwestionariuszy spisowych – *Census Questionnaire Assistance (CQA)*.
3. **Wykorzystanie rejestrów administracyjnych** oraz innych alternatywnych źródeł niestatystycznych jest planowane przede wszystkim w celu redukcji braków odpowiedzi i odmów udziału w badaniu, poprawy jakości oraz efektywności badania. Formułowane są tutaj następujące zadania:
- poprawa jakości list adresowych,
  - wzrost efektywności kampanii promującej spis i udział w badaniu ACS,
  - weryfikacja odpowiedzi respondentów, szczególnie tych, które nie zawierają spisowego numeru identyfikacyjnego w celu uniknięcia dostarczania nieprawdziwych danych,
  - aktualizacja rejestrów ,
  - redukcja braków odpowiedzi, próba ponownego dotarcia do respondenta (*Nonresponse Followup* – NRFU), optymalizacja liczby prób kontaktu z respondentem,
  - redukcja bezpośredniego gromadzenia danych przez ankieterów.
4. **Monitorowanie realizacji** spisu z wykorzystaniem nowoczesnych technologii – poprawa efektywności pracy rachmistrzów spisowych. Poprawa efektywności spisu dzięki nowoczesnym technologiom uwzględnia nie tylko automatyzację aktualizacji list adresowych, wykorzystanie rejestrów czy redukcję braków odpowiedzi (NRFU). Przedstawiono schemat organizacji prac związanych z gromadzeniem danych, który na każdym szczeblu uwzględnia intensywne wykorzystanie ICT umożliwiające wzrost efektywności zarządzania oraz pracy w terenie. Planowana hierarchiczna struktura organizacji pracy *Group Quarters (GQ)* uwzględnia:
- naczelników okręgowych biur spisowych (*aera census office managers*), którzy mają wykorzystywać nowoczesne metody w celu automatycznego i zoptymalizowanego podziału zadań; w efekcie uzyskana ma być większa widoczność statusu sprawy dla lepszego zarządzania personelem;

- zarządców obszarów spisowych (*census field managers*) wykorzystujących nowoczesne technologie w celu automatycznego raportowania w zakresie rekrutacji, szkoleń, płac i wydatków; w konsekwencji nastąpi przeprojektowanie systemu zapewnienia jakości;
- doradców w obszarach spisowych (*census field supervisors*), którzy mają mieć możliwość aktualizacji adresów i przeprowadzenia spisu korzystając z jednego urządzenia; podejście to zapewni poprawę komunikowania się w czasie spisu;
- rachmistrzów spisowych (*listers and enumerators*), którym nowoczesne technologie pozwolą na ograniczenie zużycia papieru oraz prac wykonywanych ręcznie.

W celu zapewnienia jakości gromadzonych danych planuje się taką organizację pracy, by usprawnić komunikowanie się osób zaangażowanych w prace spisowe na każdym etapie, a także zwiększyć przejrzystość działań i poprawę efektywności zarządzania. Planowane modyfikacje spisu w USA mają bardzo operacyjny charakter, jednak są bardzo głębokie i nie sprowadzają się jedynie do wykorzystania nowoczesnych technologii komunikacyjnych, ale także dotyczą innych dostępnych źródeł danych administracyjnych i pozaadministracyjnych.

#### 4.2.2. Wielka Brytania

Rada Statystyczna Wielkiej Brytanii (The UK Statistics Authority) zaleciła urzędowi statystycznemu (Office for National Statistics, ONS) dokonanie analizy różnych form przeprowadzania spisów i przedstawienie rekomendacji dla przyszłego spisu w Anglii i Walii<sup>150</sup>. W tym celu pod koniec 2013 r. ONS rozpoczął realizację programu Beyond 2011 uwzględniającego analizy i oceny niezależnych ekspertów, doświadczenia innych państw i różnych grup użytkowników oraz szerokie konsultacje społeczne (Calder i Teague, 2012; Leventhal, 2013). W wyniku tego programu zidentyfikowano dwa szeroko definiowane podejścia określane jako opcja tradycyjnego spisu w cyklu dziesięcioletnim oraz opcja wykorzystująca dane administracyjne i wyniki badania gospodarstw domowych. Zgodnie z harmonogramem badań w marcu 2014 r. Jil Matheson, narodowy statystyk i dyrektor naczelny Brytyjskiego Urzędu Statystycznego (*National Statistician and Chief Executive of the UK Statistics Authority*), przedstawił Radzie Statystycznej Wielkiej Brytanii rekomendacje (Matheson, 2014) dotyczące organizacji kolejnego spisu powszechnego w 2021 r. poparte niezależną opinią eksperta, prof. Chrisa Skinnera<sup>151</sup> (Skinner, Hollis i Murphy, 2013). W rekomendacjach tych znajdujemy opinię, w której Matheson stwierdza,

<sup>150</sup> Szkocja i Irlandia Północna przeprowadzają ocenę niezależnie.

<sup>151</sup> Osoba prof. Chrisa Skinnera jako niezależnego eksperta została rekomendowana przez Królewskie Towarzystwo Statystyczne (Royal Statistical Society), natomiast profesor Skinner zaprosił do współpracy dwóch kolejnych ekspertów, demografów prof. Mike'a Murphy'ego i Johna Hollisa.

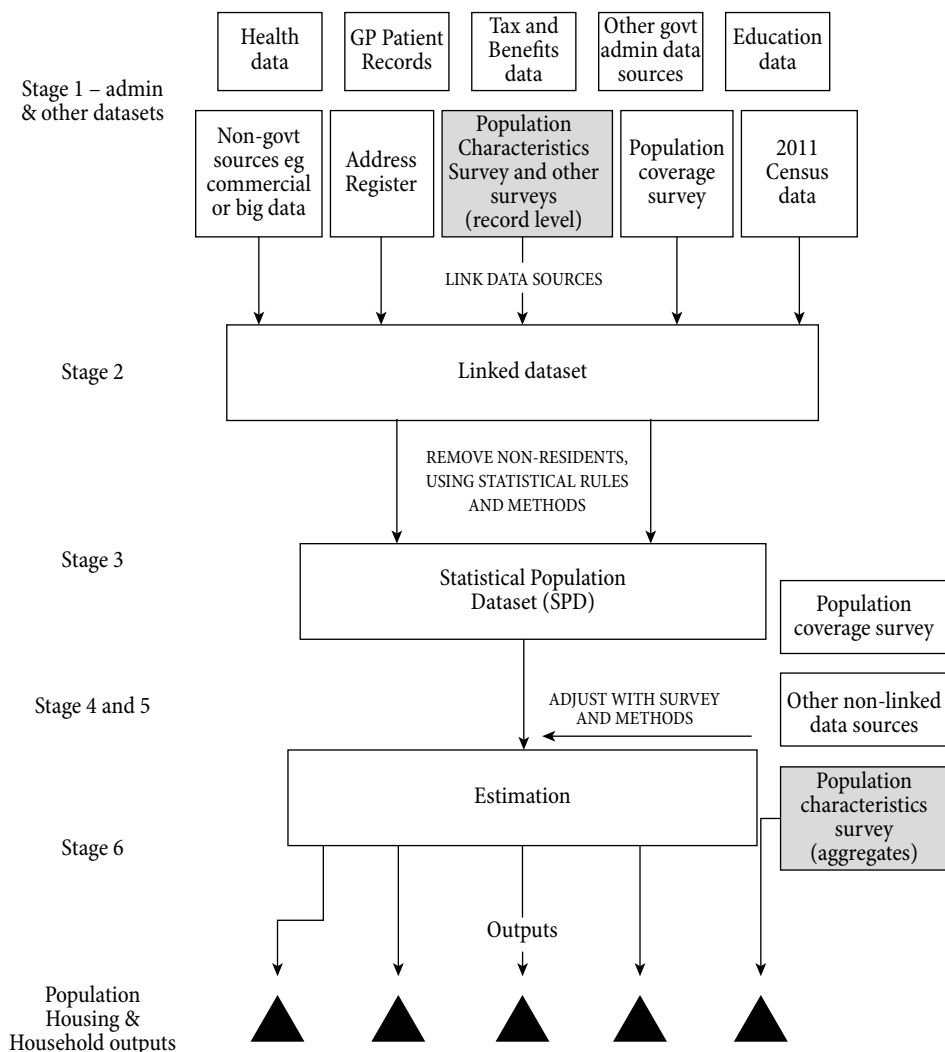
że ONS powinien jak najlepiej wykorzystać wszystkie źródła, łącząc dane spisu online z 2021 r. oraz rejestrów administracyjnych i badań reprezentacyjnych. Skinner w szczególności wyróżnił dwa zagadnienia. Po pierwsze, wskazał internetowy spis wszystkich gospodarstw domowych i zakładów zbiorowego zamieszkania w Anglii i Walii w 2021 r. jako nowoczesnego spadkobiercę tradycyjnego, organizowanego co dziesięć lat spisu ludności, korzystającego z formularzy papierowych<sup>152</sup>. Drugim elementem przedstawionym w rekomendacjach jest zwiększenie wykorzystania danych administracyjnych oraz badań reprezentacyjnych w celu wzmocnienia statystyk spisu internetowego 2021 r. oraz zwiększenia aktualności i poprawy statystyk między spisami.

Autorzy niezależnej recenzji jednoznacznie poparli spis online, pomimo pewnych wątpliwości dotyczących możliwości pokrycia populacji. Uznają, że rozwiązanie takie stanowi naturalną ewolucję tradycyjnego spisu ludności, czerpiąc z dotychczasowych doświadczeń oraz wdrażając innowacje technologiczne i dorobek najlepszych praktyk w zakresie przeprowadzania spisów powszechnych na całym świecie. Z jednej strony podkreślają, że w spisie online 2011 r. w Anglii i Walii uzyskano około 16% zwrotów. Zauważają jednak, że stopa zwrotu już w 2011 r. wyniosła 54% w spisie kanadyjskim, co ilustruje możliwy do osiągnięcia wynik<sup>153</sup> z przeszłości i perspektywę na przyszłość.

Bardziej sceptyczna opinia recenzentów dotyczy drugiego rekomendowanego elementu spisu 2021 r. – wykorzystania źródeł administracyjnych – jak zaznaczają – bardziej radykalnej zmiany metodologii spisowej. Proponowana przez ONS zmiana oznacza, zdaniem ekspertów, określenie populacji na podstawie dwóch odrębnych źródeł (rys. 4.1). Jednym ma być statystyczna populacja – złoty rekord (*statistical population dataset*) utworzony w wyniku połączenia różnych źródeł administracyjnych w celu corocznych szacunków ludności (*annual population estimates*). Drugim źródłem informacji o wszelkich innych charakterystykach tradycyjnie ujmowanych w formularzu spisowym mają być wyniki szacunków na podstawie specjalnego badania kroczącego obejmującego 4% próbę (*attribute estimates*). Z czasem przewidziana jest integracja obu tych źródeł danych. Eksperti stwierdzają, że dane administracyjne były skutecznie wykorzystywane w wielu krajach w spisach powszechnych. Jednak głównym warunkiem takiego rozwiązania jest istnienie odpowiedniego ustawodawstwa dotyczącego wymiany danych między urzędem statystycznym a instytucjami odpowiedzialnymi za poszczególne systemy administracyjne. Dotyczy to także odpowiedniości danych na potrzeby statystyki i terminowego ich udostępniania. Skinner, Hollis i Murphy (2013) uważają, że odpowiednie nowe prawodawstwo jest niezbędnym wymogiem kontynuowania prac w tym zakresie.

<sup>152</sup> Podobnie jak w 2011 r. ONS uznał potrzebę szczególnego wsparcia osób, które nie są w stanie samodzielnie przeprowadzić spisu online.

<sup>153</sup> Największy odsetek pokrycia w spisie internetowym w 2011 r. osiągnęła Estonia – 67%.



Rysunek 4.1. Schemat opracowywania wyników spisu ludności opartego na rejestrach administracyjnych

Źródło: Office for National Statistics: <https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/methodology/producingcensustypeoutputsfromanadministrativedatacensusadiscussionpaper>

Na stronie internetowej ONS, w specjalnej zakładce dedykowanej spisom ludności<sup>154</sup>, zamieszczane są na bieżąco informacje nt. przygotowań do spisu, o realizacji

<sup>154</sup> Link do tej strony jest następujący: <https://www.ons.gov.uk/census/censustransformationprogramme/consultations/the2021censusinitialviewoncontentforenglandandwales/updateon2021censusforenglandandwales>.

kolejnych zadań programu, uzupełnieniach i modyfikacjach. Przykładowo w dniu 5 listopada 2017 r. znajdujemy informację, że przez cały czas prowadzone są prace w celu umożliwienia realizacji spisu bazującego na źródłach administracyjnych po 2021 r. W celu zaspokojenia potrzeb użytkowników ONS bada możliwość replikacji informacji zebranych w drodze spisu na podstawie danych administracyjnych już posiadanych przez rząd uzupełnionych wynikami badań reprezentacyjnych. Na stronie ONS znajdziemy także informację o przedstawieniu wyników ostatnich analiz oraz konsultacjach na Forum Polityki Ludnościowej 13.12.2017 r. (ONS *Population and Public Policy Forum*), które pomogą sfinalizować prace nad przygotowaniem dokumentacji spisowej. Wiosną 2018 r. przewidziana jest publikacja *Census White Paper*<sup>155</sup>, dokumentu zawierającego dokładny opis konsultacji społecznych, zakres informacyjny, zapewnienie bezpieczeństwa, sposób przeprowadzenia spisu, oczekiwane wyniki oraz proces legislacyjny.

### 4.3. NSP 2021 w perspektywie planów GUS

W wyniku analizy prac metodologicznych i rezultatów przeprowadzonych konsultacji społecznych w końcu czerwca 2017 r. GUS opublikował dokument przedstawiający zakres informacyjny oraz planowaną metodę realizacji spisu ludności w 2021 r. (GUS, 2017d). W opracowaniu znajdujemy koncepcję realizacji spisu ludności i mieszkań NSP 2021, jaka została przedstawiona na posiedzeniu Komisji Metodologicznej GUS 30 czerwca 2017 r.<sup>156</sup>. Opracowując koncepcję kolejnego spisu, uwzględniono doświadczenia z rundy spisów 2010 r. oraz tendencje rozwoju międzynarodowej statystyki, w których jest wyraźne dążenie do coraz szerszego wykorzystywania źródeł administracyjnych, szczególnie w spisach ludności i mieszkań. Koncepcja GUS rozważa także założenia strategii Eurostatu, która przewiduje, że runda 2020 będzie ostatnią realizowaną w Europie według schematu dziesięcioletniego (GUS, 2017d, s. 3).

Podstawą przedstawionych rozwiązań była analiza potrzeb informacyjnych wobec spisu ludności, które wynikają z regulacji prawnych obowiązujących w Polsce oraz zobowiązań międzynarodowych (UE, 2008), a także potrzeb zgłaszanych przez organy administracji państwowej, różne instytucje i organizacje, naukowców,

---

<sup>155</sup> Dokument taki jest publikowany w Wielkiej Brytanii mniej więcej dwa–trzy lata przed spisem ludności. W przypadku poprzednich spisów były one opublikowane odpowiednio w marcu 1999 r. i w grudniu 2008 r. (ONS, 1999, 2008).

<sup>156</sup> Dokument ten jest wynikiem prac grupy ds. metodologii spisu ludności i mieszkań oraz grupy ds. oceny jakości danych, które działają w ramach Zespołu ds. opracowania podstaw prawnych i koncepcji realizacji powszechnego spisu rolnego 2020 r. oraz narodowego spisu powszechnego ludności i mieszkań 2021 r., powołanego Zarządzeniem wewnętrznym nr 14 prezesa GUS z dnia 30 czerwca 2015 r.

przedsiębiorców i indywidualnych użytkowników uczestniczących w przeprowadzonych przez GUS konsultacjach społecznych (GUS, 2017b). Wyniki te zostały krótko omówione w rozdziale 1.3 oraz w aneksach I.6 – I.9, gdzie przedstawiono obszary tematyczne dotychczasowych spisów, proponowane tematy przyszłego spisu oraz ocenę stopnia ich zaspokojenia i najczęściej zgłaszane dodatkowe potrzeby informacyjne. Określenie obszarów tematycznych spisu było związane z oceną możliwości pozyskania odpowiednich danych na podstawie dokonanego wcześniej przeglądu dostępnych źródeł informacji, przede wszystkim systemów administracyjnych oraz innych pozastatystycznych. Analizie poddano 35 systemów administracyjnych, w tym wcześniej niewykorzystywane przez statystykę publiczną takie bazy jak np. Centralna Baza Danych Ksiąg Wieczystych. Przeprowadzona analiza metadanych umożliwiła ocenę odpowiedniości informacji, ich zgodności i spójności dla potrzeb statystycznych oraz stopienia pokrycia w odniesieniu do badanych populacji (ludności, mieszkań, gospodarstw domowych i rodzin). Uwzględniono między innymi nazwy cech, dla których są gromadzone dane w rejestrach, ich definicje, a także stosowane identyfikatory i klasyfikacje. W wyniku analizy dokonano podziału cech na takie, dla których jest możliwe pełne pokrycie danymi z rejestrów i systemów administracyjnych (22 zmienne), oraz takie, które na podstawie rejestrów można tylko częściowo opracować według wymaganych klasyfikacji (44 zmienne). Wyodrębniono także 20 cech, dla których nie były dostępne informacje w rejestrach (całkowicie).

Przeprowadzona przez GUS analiza potrzeb informacyjnych spisu oraz możliwości ich pokrycia wykazała, że możliwe byłoby opracowanie wyników spisu wyłącznie na podstawie źródeł administracyjnych jedynie dla minimalnego zakresu informacyjnego spisu (GUS, 2017d, s. 10). Główną zaletą rejestrów jest dostępność danych pozwalająca na lokalizację jednostek według współrzędnych geograficznych ( $x, y$ ), np. dla osób według miejsca zamieszkania. Tak szczegółowa informacja umożliwia opracowanie danych dla dowolnego poziomu jednostek przestrzennych oraz dla dowolnie zdefiniowanych obszarów. Stwarza to możliwość realizacji zaleceń Eurostatu dotyczących opracowywania i przekazywania wybranych informacji na poziomie siatki kilometrowej (grid). GUS podkreśla także istotny niedostatek podejścia korzystającego wyłącznie z rejestrów administracyjnych, gdyż w sytuacji ograniczonego dostępu do danych, niektórych tematów nie można byłoby zrealizować. Istnieje wówczas niebezpieczeństwo niewypełnienia przez Polskę niektórych zaleceń międzynarodowych.

Wyrażona w opracowaniu GUS opinia, że dane pozyskiwane ze źródeł administracyjnych „odzwierciedlają stan formalnoprawny, a nie rzeczywistą skalę badanego zjawiska” (GUS, 2017d, s. 11), budzi pewną wątpliwość. Rzeczywiście, dane z rejestrów powinny przedstawiać stan formalnoprawny, jednak nie jest zrozumiałe, dlaczego autorzy zakładają, że nie jest on rzeczywisty. Prezentując w niniejszym opracowaniu dyskusję dotyczącą kategorii ludności oraz wynikające z tego tytułu



problemy szacunku (rozdz. 1.2.2), przedstawiono wątpliwości odnośnie do możliwości uzyskania w spisie tradycyjnym informacji o stanie „rzeczywistym”. Informacja uzyskana podczas wywiadu przeprowadzanego przez rachmistrza spisowego ma charakter deklaracyjny, ale czy przedstawia stan „rzeczywisty”? W każdej sytuacji pojawia się szereg wątpliwości. Można przypuszczać, że najbliższa prawdzie będzie informacja pozyskana podczas samospisu, jako bezpośrednia i nieskrępowana, niezależna od wpływu otoczenia (najbliższych członków rodziny, gospodarstwa domowego, sąsiadów, jak również rachmistrza spisowego). Wobec powyższego warto sformułować pytanie, czy prezentowanie danych według stanu formalnoprawnego nie stanowi pewnego uporządkowania. Wydaje się, że pomieszano tutaj dwie kwestie – możliwości pozyskania wszystkich danych z rejestru i ich prawdziwości.

Poszukując optymalnej metody realizacji spisu, tak by jak najpełniej zaspokoić potrzeby informacyjne na jak najniższym poziomie podziału terytorialnego, GUS odwołuje się do strategii przejścia na roczny system przekazywania informacji do Eurostatu (EU, 2016)<sup>157</sup>. Prawdopodobnie spisy po 2020 r. będą w jeszcze większym stopniu oparte na źródłach administracyjnych, a coroczne aktualizacje spowodują łatwiejszą ich realizację i bardziej równomierne rozłożenie kosztów. Planuje się utworzenie „nowej bazy spisowej/(operatu spisowego)” (GUS, 2017d, s. 11 i nast.), która będzie gromadzić dane pochodzące z różnych źródeł, statystycznych (spisy 2002, 2011, BAEL, EU-SILC, inne) i pozastatystycznych (rejestry administracyjne, systemy informatyczne).

Bardzo ważna jest deklaracja GUS, że idea konceptualizacji bazy ściśle wynika z potrzeby opracowania nowej metodologii bilansowania ludności, która „nie byłaby oparta jak dotychczas wyłącznie na adresie zameldowania osoby, a na jej faktycznym miejscu zamieszkania”. Warto w tym miejscu rozważyć używanie określenia „rezydencji” i stosowanie definicji rezydenta uwzględniającej określony w przepisach unijnych czas zamieszkiwania. Bezwzględnie jednak należy precyzyjnie zdefiniować kategorię spisywanej ludności. Odwoływanie się do „faktycznego miejsca zamieszkania” kojarzy się z dotychczas stosowanym pojęciem „ludności faktycznej” wraz z zastrzeżeniami dotyczącymi uwzględnienia osób przebywających za granicą bez względu na czas pobytu (rozdz. 1.2.2 oraz aneks I.5). Zapisy koncepcji oraz stosowane techniczne sformułowania nie zawsze jasno oddają intencje autorów i budzą uzasadnione dotychczasowym doświadczeniem wątpliwości.

Używane w koncepcji pojęcia bazy i operatu spisowego wymagają precyzyjnego określenia. Sposób ich zapisu w omawianym dokumencie: „baza/operat spisowy” (GUS, 2017d, s. 11 i nast.), sugeruje tożsamość i zamiennie traktowanie. Określenie operat spisowy kojarzy się także z operatem losowania, pojawia się więc pytanie o wzajemną relację, mianowicie czy operat spisowy można utożsamiać z określanym

<sup>157</sup> Zgodnie z informacją zawartą w opracowaniu GUS zadeklarował gotowość corocznego przekazywania danych w określonym zakresie statystyki społecznej do Eurostatu (GUS, 2017d, s. 11).

na podstawie spisu operatem losowania wykorzystywanym w badaniach specjalnych. Proces budowy bazy (operatu spisowego) zakłada dwa typy rejestrów:

1. Systemy referencyjne („rdzenie”), które cechują się najwyższą jakością danych, kompletnością i aktualnością oraz wysokim pokryciem podmiotowym. Mają one stanowić rdzeń, z którym będą integrowane pozostałe rejestry. Do prac projektowych przyjęto: PESEL<sup>158</sup>, KEP, ZUS, KRUS, NFZ, ARiMR.
2. Systemy zasilające, czyli pozostałe systemy, rejestry, bazy danych, wykazy, ewidencje, których zakres tematyczny odpowiada potrzebom statystyki.

Można przypuszczać, że głównym rejestrem będzie PESEL. Założono ponadto: „wykorzystanie wielu rejestrów do budowy osób poprzez wyszukanie unikalnych numerów PESEL, które przejdą kontrolę poprawności polegającej na sprawdzeniu wymaganej liczby znaków identyfikatora, zgodności cyfry kontrolnej oraz możliwości wystąpienia zakodowanej w numerze daty urodzenia w kalendarzu” (GUS, 2017d, s. 14). Sformułowanie tego założenia budzi nie tylko wątpliwości ze względu na techniczny język i „budowę osób”, ale także z powodu pominięcia kwestii merytorycznych. Budowa bazy ma być prowadzona w trzech etapach:

- I etap – podnoszenia jakości (przekształcanie systemów zewnętrznych w zbiory danych statystycznych),
- II etap – produkcji cech statystycznych (podejście „wielorzędowe”, integracja i generowanie wyniku, walidacja i korekta danych, imputacja),
- III etap – analizy jakości mikrodanych operacyjnych (ocena jakości wybranych procesów).

Zakłada się, że w konstruowanej bazie najważniejszy będzie adres zamieszkania, który ma być „dopisany” do rejestru PESEL, choć nie określono, w jaki sposób i na jakiej podstawie. W badaniu mają być uwzględnione „trzy rodzaje adresów: stały, czasowy oraz zamieszkania”, ale nie sprecyzowano, co oznaczają te trzy określenia (GUS, 2017d, s. 14). O hierarchii wyboru adresu z konkretnego zbioru ma decydować aktualność danych i referencyjność w zakresie danych adresowych.

Niestaranność sformułowań i – jak się wydaje – nie w pełni przemyślana koncepcja pod względem merytorycznym uwidaczniają się także w sformułowaniu dotyczącym niepewności odnośnie do pozyskania odpowiedzi. GUS podkreśla niepewność uzyskania odpowiedzi przez internet w związku z utrzymującą się w Polsce sytuacją, „która może skutkować nieobecnością w trakcie spisu istotnej części ludności – nierzadko całych rodzin – spowodowanej czasowym przebywaniem poza granicami kraju” (GUS, 2017d, s. 22). W opracowaniu stwierdzono, że w związku z powyższym istnieje poważne ryzyko uzyskania „dalece niewystarczającej (niesatysfakcjonującej) kompletności badania spisowego”. Stwierdzenie to nasuwa wątpliwość, że nadal nie uporządkowano definicji ludności i nie wiadomo,

<sup>158</sup> W tym zbiór osób wymeldowanych donikąd.

jaka kategoria ludności będzie spisywana. Czy nadal rezydentem będzie osoba, która posiada zameldowanie na pobyt stały, chociaż od lat przebywa za granicą?

Koncepcja GUS zakłada, że utworzona baza spisowa (operat spisowy) będzie odgrywać kluczową rolę, gdyż umożliwi określenie populacji objętej spisem według miejsca zamieszkania przed momentem krytycznym spisu. Dane zawarte w bazie zostaną wykorzystane w celu wstępnego wypełnienia kwestionariusza spisowego. Tak wypełniony formularz będzie podlegać aktualizacji przez respondentów (w czasie spisu). Podejście, w którym respondent dokonuje przede wszystkim weryfikacji danych, umożliwi skrócenie czasu realizacji spisu, a więc zmniejszenie obciążeń respondentów. Można także mieć nadzieję, że przyjęta forma będzie zachętą do samospisu internetowego. GUS deklaruje również, że elektroniczny kwestionariusz spisowy będzie zawierać logiczną kontrolę i walidację danych, co powinno wpłynąć na poprawę jakości danych.

Przyjęto, że podstawową (preferowaną) techniką gromadzenia danych będzie samospis internetowy, a uzupełniającymi będą wywiad telefoniczny i wywiad bezpośredni (*face-to-face*). Rzetelna wydaje się analiza dostępu do internetu oraz ostrożny szacunek zwrotu odpowiedzi w granicach 35–50%, przy intensywnej popularyzacji i promocji spisu. Pozytywnie też należy ocenić inicjatywę skorzystania z doświadczeń Ministerstwa Finansów wspierającego mieszkańców przy rozliczaniu deklaracji podatkowych (dyżury w urzędach skarbowych, specjalnie wydzielone miejsca publiczne z punktami obsługi). Przy realizacji spisu planuje się, aby takimi miejscami „wsparcia” dla ludności (poza urzędami statystycznymi) były urzędy gmin, szkoły, poczta, biblioteki.

W NSP 2021 planuje się także zastosowanie badań reprezentacyjnych, realizowanych „w tle” badania pełnego. Niestety, oprócz informacji o „nacisku na uzyskanie maksymalnej kompletności”, w koncepcji nie przedstawiono bardziej szczegółowych informacji w tym zakresie.

Z dużą nadzieją i optymizmem należy też ocenić wszystkie już podjęte przez GUS starania w zakresie poprawy jakości danych. Bardzo ważna jest świadomość i chęć przeprowadzenia dobrze zaplanowanego spisu kontrolnego. Przy czym GUS deklaruje wprost, że najbardziej „optymalnym rozwiązaniem (choć niewątpliwie kosztownym) powinno być przeprowadzenie badania kontrolnego przez niezależnych ekspertów np. firmę outsourcingową lub w innej formie (np. audyt zewnętrzny)” (GUS, 2017d, s. 22).

Ostatecznie przedstawiona przez GUS koncepcja realizacji spisu ludności mieszkańców 2021 jest następująca:

1. ustalony zakres informacyjny spisu zostanie zebrany od całej populacji, podlegającej badaniu, **poprzez kwestionariusz spisowy i przy wykorzystaniu następujących technik: samospisu internetowego (CAII), wywiadu telefonicznego (CATI) oraz wywiadu z udziałem rachmistrza spisowego (CAPI),**

2. zaproponowany **zakres informacyjny** spisu uwzględnia potrzeby użytkowników wyrażone w konsultacjach społecznych do spisu, jak również wynikające z krajowych i międzynarodowych aktów prawnych w odniesieniu do ludności rezydującej oraz wg definicji krajowej,
3. utworzona zostanie **Baza Spisowa/operat spisowy**, która realizować będzie szereg funkcji, od inicjalnego zasilenia kwestionariusza spisowego, kierowanego do respondentów w celu zweryfikowania danych, poprzez opracowanie i upowszechnienie danych wynikowych, do prowadzenia interaktywnej analizy danych od prostych programów raportujących, do złożonych systemów wyszukujących wzorce i regularności w dużych zbiorach danych z zastosowaniem zaawansowanych metod statystycznych lub data mining,
4. nacisk zostanie położony na pozyskanie **większego zakresu informacji** poprzez integrację różnych źródeł danych: **statystycznych i pozastatystycznych**, w tym administracyjnych<sup>159</sup>,
5. każdy etap prac (pozyskiwania, gromadzenia i opracowywania danych) będzie monitorowany poprzez **system opracowanych wskaźników jakości**,
6. informacje wynikowe będą opracowane na możliwie najniższym poziomie dostępności (dla wybranych zmiennych – **poziom siatki km<sup>2</sup>**),
7. przy budowie kwestionariusza – mając na uwadze jakość gromadzonych danych – szczególny nacisk zostanie położony na **konstrukcję pytań i wyjaśnień/instrukcji** (zrozumiały, nieformalny język przekazu) oraz komputerowe techniki wspomagające wypełnianie kwestionariusza (m.in. udoskonalone słowniki, odpowiednia wizualizacja/grafika pytania) dla każdej z zastosowanych technik realizacji spisu,
8. w szerszym zakresie będą wykorzystane **metody matematyczno-statystyczne** pozwalające na opracowanie danych dla możliwie wyczerpujących zestawień danych,
9. w celu zminimalizowania ewentualnego ryzyka uzyskania niewystarczającej (niesatysfakcjonującej) kompletności badania spisowego przewiduje się zastosowanie podejścia z **częściowym użyciem metody reprezentacyjnej**, realizowanej w ramach – niejako w tle – badania pełnego,
10. wyznaczenie dogodnego **terminu i daty krytycznej** badania spisowego z punktu widzenia dotarcia do respondentów, jak również synchronizacji danych pochodzących z różnych źródeł, co w istotny sposób wpływa na jakość wyników spisu,
11. **zintensyfikowana będzie współpraca z gestorami administracyjnych źródeł danych oraz jednostkami samorządu lokalnego**” (GUS, 2017d, s. 24–25).

---

<sup>159</sup> W aneksie IV.1 przedstawiono wykaz pozastatystycznych systemów informacyjnych, rejestrów i baz danych stanowiących potencjalne źródło danych dla NSP 2021.

Reasumując, w opracowaniu przedstawiono bardzo ciekawą propozycję przeprowadzenia NSP 2021 z uwzględnieniem badania pełnego oraz wykorzystaniem wielu źródeł informacji. W sytuacji gdy rejestry administracyjne, w ocenie statystyki publicznej, nadal nie spełniają wymagań umożliwiających zgromadzenie adekwatnych danych we wszystkich obszarach tematycznych, rozwiązanie takie jest uzasadnione. Kilka kwestii wydaje się jednak na tyle istotnych, że niezbędny jest dodatkowy komentarz.

### 1. Definicja podstawowych kategorii spisowych zgodnie ze standardami międzynarodowymi

Jednym z zasadniczych problemów jest nadal nieuporządkowanie definicji ludności. W koncepcji NSP 2021 znajdujemy informację, że już na „wejściu” uzyskamy tzw. populację odniesienia, tj. ludność, która podlega spisowi wraz z zestawem cech demograficznych według miejsca zamieszkania (w tym według współrzędnych geograficznych  $x$ ,  $y$ ). Jednak w przedstawionej przez GUS koncepcji nie znajdujemy definicji owej ludności. Być może założono, że zgodnie z wytycznymi Eurostatu będzie to ludność rezydująca. Jeśli jednak GUS pozostanie przy osobliwej interpretacji ludności rezydującej, tłumacząc ten fakt obowiązującymi regulacjami prawnymi, niezbędne wydaje się publikowanie dwóch stanów (i struktur!) ludności: rezydującej oraz rezydującej z uwzględnieniem migracji długookresowych<sup>160</sup>. Określenie to brzmi absurdalnie, ale jeśli rezydentem jest osoba, która spełnia wymogi formalnego zameldowania, a nie zamieszkania, to w efekcie otrzymujemy abstrakcyjne struktury ludności (spełniające wymogi formalne, ale nieodpowiadające rzeczywistości). Należy wyrazić obawę wynikającą z braku koncepcji rzetelnego rozwiązania definicji ludności rezydującej, o czym świadczy propozycja wątpliwego powrotu do wielokrotnie krytykowanych rozwiązań (również w tym opracowaniu, rozdz. 1.2.3., 3.3, 3.4.1). Swoistym precedensem w zakresie podwójnej publikacji danych o ważnym zjawisku gospodarczym jest publikacja dwóch wielkości bezrobocia, według BAEL oraz na podstawie rejestracji.

### 2. Zasada bezpośredniego udzielania informacji

Bezpośrednie udzielenie informacji pozwala uniknąć błędów odpowiedzi „z drugiej ręki”. Zarówno system oparty na rejestrach, jak i samospis internetowy jest swoistym gwarantem zachowania zasady bezpośredniości. Określenie „swoistej gwarancji” wynika z obawy o wypełnianie czy wręcz zachętę do wypełniania internetowych kwestionariuszy spisowych dla osób o ograniczonej znajomości internetu przez innych członków gospodarstwa domowego. Pewnym rozwiązaniem może być uwzględnienie w formularzu pola identyfikującego osobę wypełniającą formularz oraz pola umożliwiającego potwierdzenie odpowiedzi przez osobą spisywaną. Oczywiście rozwiązanie to nie jest pozbawione wad,

<sup>160</sup> Zasade jest przyjęcie czasu trwania migracji odpowiadającego definicji ludności rezydującej.

ma jednak tę zaletę, że uświadamia problem oraz umożliwia szacunek liczby odpowiedzi uzyskanych w sposób pośredni.

Niezbędna wydaje się odpowiednia akcja promująca konieczność bezpośredniego udzielania informacji bądź ich potwierdzania w sytuacji korzystania z pomocy. Obecnie obowiązuje społeczne przyzwolenie, niestety także w środowisku naukowym, na gromadzenie informacji od jednej osoby w gospodarstwie domowym, również w zakresie pytań drażliwych (wyznanie, narodowość, niepełnosprawność). Ponadto wyraźny jest powszechny brak świadomości konsekwencji błędów odpowiedzi wynikających z takiego podejścia, nie wspominając o braku zachowania zasad etycznych. Z omawianej koncepcji nie wynika, by w GUS podjęto temat świadomych działań zmierzających do minimalizacji błędów respondenta grupowego.

### 3. **Formy gromadzenia danych**

Obecnie proponuje się wiele kanałów przekazywania informacji poza bezpośrednim zasilaniem z rejestrów administracyjnych i źródeł pozastatystycznych. Podobnie jak w NSP 2011 są to:

- wywiad telefoniczny (CATI),
- wywiad z udziałem rachmistrza spisowego (CAPI),
- samospis internetowy (CAII).

Mankamentem tego rozwiązania jest brak formy umożliwiającej samospis dla osób „wykluczonych cyfrowo”. W kolejnym spisie po 2021 r. być może wystarczy samospis przez internet. Jednak obecnie niezbędne jest uzupełnienie kanałów gromadzenia danych o samospis w formie tradycyjnej, np. formularza papierowego pobieranego w urzędzie pocztowym, który po wypełnieniu można byłoby w tym urzędzie złożyć (w zamkniętej kopercie). Poczta byłaby zobowiązana wydrukować formularz „na żądanie” oraz przekazać zebrane formularze do urzędu statystycznego. Prawdopodobnie potrzeba taka dotyczyłaby bardzo nielicznej grupy ludności, często „trudnej do spisania”, są to mniejszości narodowe, bezdomni, osoby o specyficznych cechach, obawiające się udzielenia odpowiedzi za pośrednictwem innych osób, w tym rachmistrza spisowego.

### 4. **Wykorzystanie zasobów rejestrów administracyjnych oraz systemów pozaadministracyjnych**

- Dane spisu tradycyjnego a opartego na rejestrach

W koncepcji NSP 2021 wyraźnie określono, że w chwili obecnej zasoby rejestrów administracyjnych i pozaadministracyjnych systemów pozwalają na przeprowadzenie w Polsce spisu opartego na tych źródłach jedynie w minimalnym zakresie informacyjnym. „Wyprowadzenie„ zmiennych spisowych na podstawie dostępnych źródeł wymaga algorytmu wypracowanego przez metodologów, często z zastosowaniem zaawansowanych metod statystycznych i nowoczesnych technologii. Przedsięwzięcie to jest trudne i pracochłonne (Dygaszewicz, 2009, s. 26), niemniej jednak wykorzystanie alternatywnych

źródeł danych, ograniczenie obciążenia respondentów i minimalizacja kosztów, to coraz głośniej i powszechniej formułowane względem statystyki oczekiwania (Domański, 2014; Szreder, 2009, 2016, 2017, Witkowski, 2010, 2014).

Wśród zalet spisów tradycyjnych określa się możliwość pozyskania wszelkich potrzebnych informacji, co do których się zakłada, że są zgodne z prawdą i pozbawione błędów. Z kolei jako zasadnicze ograniczenie spisów opartych na rejestrach wskazuje się konieczność skorzystania z dostępnych danych, które nie zawsze odpowiadają oczekiwaniom. Paradysz (2007b, s. 197), na podstawie dociekliwych analiz, stwierdza, że „należy unikać [...] obciążania społeczeństwa badaniami, szczególnie tam, gdzie możliwe jest wykorzystanie tych informacji, które administracja państwowa wyegzekwowała od obywateli znacznie lepiej i skuteczniej niż by to mogła zrobić statystyka publiczna [...]”. Najbardziej jaskrawym przykładem są zeznania podatkowe. Statystyka publiczna nie jest w stanie uzyskać wyższej jakości informacji od tych, które wyegzekwowały od obywateli Urzędy Skarbowe”. Zatem w rejestrach znajdujemy zwykle „prawdziwe” dane, które jednak nie zawsze odpowiadają oczekiwaniom. Prawdopodobnie też nie zawsze zastosowanie zaawansowanych metod statystycznych (integracji, modelowania, estymacji, kalibracji) pozwoli pokonać te ograniczenia. Z drugiej strony dane administracyjne często są niedostatecznie wykorzystywane oraz, jak podkreśla Szreder (2017, s. 13), niedoceniana jest ich struktura panelowa umożliwiająca badanie zmienności w czasie.

Użytkownikom danych powinna towarzyszyć świadomość ograniczeń zarówno spisów opartych na rejestrach, jak i tradycyjnych, wynikających między innymi z kompletności czy kwestii podniesionych w poprzednich punktach.

#### – Rejestry w zakresie oświaty

Ciekawą możliwością zdaje się stwarzać wykorzystanie Systemu Informacji Oświatowej (SIO) oraz POLON dla pozyskania informacji nt. poziomu wykształcenia.

W chwili obecnej zasięg systemów może nie być wystarczający, szczególnie w odniesieniu do osób wcześniej urodzonych. Jednak wprowadzenie obowiązku elektronicznej ewidencji świadectw i dyplomów wraz z identyfikatorem PESEL już w tej chwili umożliwi coraz pełniejsze wykorzystanie tego źródła informacji. Warto także uwzględnić regulacje obejmujące nie tylko świadectwa w zakresie szkolnictwa podstawowego i średniego, ale także studiów I, II i III stopnia, łącznie z rejestrem dyplomów doktorskich, habilitacyjnych i tytułów naukowych.

### 5. Wykorzystanie niestandardowych źródeł informacji

#### – Niestandardowe źródła danych

Koncepcja spisu zakłada zbieranie informacji z różnych dostępnych źródeł, wskazuje przede wszystkim na rejestry administracyjne oraz pozastatystyczne systemy informacyjne i bazy danych (aneks IV.1). Podjęte działania są jak

najbardziej właściwe, jednak zasadne wydaje się wskazanie szerszej, być może bardzo perspektywicznej wizji. Chodzi o wykorzystanie takich źródeł, jak zasoby operatorów sieci komórkowych, monitoringu miejskiego, serwisów internetowych, portali społecznościowych i innych określanych mianem big data. Dotychczasowe doświadczenia statystyki publicznej w tym zakresie są ograniczone, ale przykład Holandii czy Estonii (rozdz. 2.4) wskazuje duże możliwości oraz potrzebę intensywnych badań. Nieświadome<sup>161</sup> generowanie spersonalizowanych danych w cyberprzestrzeni (Beręsewicz i Szymkowiak, 2015; Szreder, 2015) pozwala przedstawić hipotetyczną wizję przyszłości, że powszechne korzystanie z nowoczesnych technologii ułatwi samospis, czy wręcz „zautomatyzuje” badanie.

- Możliwości wykorzystania niestandardowych źródeł danych a potrzeba badań metodologicznych

Wykorzystanie niestandardowych, przede wszystkim internetowych źródeł danych wymaga jeszcze wielu prac metodologicznych, przede wszystkim pod względem oceny ich jakości oraz reprezentatywności, prawnej ochrony danych osobowych. Jak zauważa Szreder (2015, s. 9–10), rewolucja elektroniczna umożliwiła gromadzenie, przetwarzanie i przesyłanie wielkich zbiorów danych, w stosunku do których klasyczne narzędzia statystyki okazują się nieadekwatne. Prawdopodobnie powszechne wykorzystanie big data w statystyce publicznej jest kwestią odległej przyszłości. Jeśli jednak naukowcy i instytucje statystyki państwowej zarzucą prowadzenie eksperymentalnych badań, istnieje obawa, że pozbawione rzetelnej oceny źródła zyskają dużą popularność. Już w tej chwili bezkrytyczne podejście do prób pobranych na zasadzie autoselekcji respondentów, czy wręcz przypisywanie im waloru reprezentatywności ze względu na dużą liczebność, budzi uzasadniony niepokój i poważne zagrożenia (Beręsewicz, 2017; Beręsewicz i Szymkowiak, 2015; Szreder, 2017).

## 6. Współpraca między gestorami rejestrów:

- Formy współpracy

Zasady współpracy i przekazywania informacji regulują odpowiednie przepisy prawa. Okazuje się jednak, że nie zawsze tak zdefiniowana współpraca jest efektywna. Decydując się na wykorzystanie w spisie danych z rejestrów administracyjnych będących w gestii różnych resortów, sensowne wydaje się powołanie międzyresortowego zespołu ds. realizacji spisu, z formalnym określeniem jego kompetencji i odpowiedzialności w celu właściwego podziału

<sup>161</sup> Jest to odwołanie do klasyfikacji big data według sposobu powstawania danych w wyniku korzystania z usług sieci komórkowych oraz pasywnego zbierania danych wskutek korzystania z różnych „inteligentnych” urządzeń określanych mianem internetu rzeczy (Beręsewicz i Szymkowiak, 2015, s. 13).



zadań, a przede wszystkim ułatwienia współpracy i porozumienia. Chodzi tutaj nie tyle o operacyjny zespół do realizacji konkretnych zadań, bo takie z pewnością istnieją, ile o zespół decydentów na najwyższym szczeblu.

– Poprawa jakości rejestrów

Lepsza organizacja współpracy między gestorami rejestrów, a także wspólna odpowiedzialność będą też prowadzić do międzyresortowych uzgodnień sprzyjających poprawie jakości rejestrów. Zagadnienie to jest szczególnie trudne z uwagi na przepisy prawa chroniące dane osobowe. W większości państw europejskich istnieją przepisy dopuszczające korzystanie przez instytucje statystyki państwowej z danych administracyjnych, choć zakres tych regulacji jest zróżnicowany (rozdz. 2). Obowiązująca zasada jednostronnego przepływu informacji określa poufny charakter danych administracyjnych także dla organów statystycznych. Zasada ta oznacza, że GUS może wprowadzać korekty jedynie w bazie danych statystycznych, ale nie ma ani prawnych, ani technicznych możliwości dokonywania jakichkolwiek korekt danych rejestrów administracyjnych. Jak zauważa Tønder (2008), lata kontaktów pomiędzy instytucjami, wzajemne zrozumienie potrzeb, a nade wszystko świadomość społeczna celów wykorzystania danych administracyjnych pozwoliły w państwach skandynawskich na wypracowanie metod współpracy sprzyjających poprawie jakości rejestrów. Opinię taką potwierdzają również Austriacy (Cetkovic i in., 2012; Lenk i in., 2014) i Hiszpanie (Vega i Argüeso, 2016) oraz doświadczenia polskie (Beręsewicz i in., 2018; Dygaszewicz, 2009; GUS, 2014b, 2017b, 2017d; Paradysz, 2007b, 2010; Szymkowiak, 2011). Warto w tym miejscu uzupełnić, że dane zebrane dla celów statystycznych są zawsze poufne.

– Bezpieczeństwo danych

Ustalenie zasad współpracy między resortami winno respektować wszystkie określone prawem zasady poufności. Powołanie zespołu międzyresortowego wydaje się sprzyjać także budowie systemu ochrony danych osobowych umożliwiającego bezpieczną integrację danych rejestrów administracyjnych. Chodzi tu o wypracowanie rozwiązania na wzór austriackiego narodowego systemu identyfikacji dla potrzeb statystyki publicznej nadającego poszczególnym osobom szyfrowane i właściwie zabezpieczone identyfikatory statystyczne, np. bPIN (Szymkowiak, 2018).

**7. Wybrane kwestie metodologiczne. Szacowanie danych spisowych w ujęciu małych domen, zbiorowości „trudnych do zbadania” oraz cech szczególnie wrażliwych**

– Promocja samospisu w badaniu pełnym

GUS proponuje jeden formularz bez podziału na krótki i długi formularz. Samospis pozwala na weryfikację danych pobranych z rejestrów. W koncepcji podkreśla się, że jest to doskonała forma gromadzenia danych „wrażliwych”

(kwestie etniczne), gdyż zwiększa odsetek osób bezpośrednio uczestniczących w spisie oraz pozwala zmniejszyć odsetek braków odpowiedzi. Dodatkowo niekrępująca forma badania zachęca do udzielenia informacji.

Pomimo to GUS planuje zastosowanie podejścia z „częściowym użyciem metody reprezentacyjnej” (GUS, 2017d, s. 24). Nie zostało co prawda wyjaśnione, jak należy rozumieć „częściowe” użycie metody reprezentacyjnej. Gdyby jednak dokonano podziału na badanie pełne i reprezentacyjne, warte byłoby rozważenia, w niezbędnym zakresie, rozszerzenie krótkiego formularza o pytania drażliwe dotyczące niepełnosprawności. Zagadnienia te mogą być szczególnie istotne ze względu na intensywnie starzejącą się populację.

- Charakterystyka zbiorowości osób, które odmówiły udziału w spisie, oraz zastosowanie specjalnych metod niwelujących skutki odmów

W NSP 2011 nie przeprowadzono analizy braków odpowiedzi i odmów udziału w badaniu. Brakuje podstawowych charakterystyk zbiorowości osób, które odmówiły udziału w spisie. Nie wiemy, jakie jest ich rozmieszczenie terytorialne, zróżnicowanie według płci, wieku itp.

Przeprowadzając spis metodą mieszaną, gdy rekordy dla poszczególnych osób są inicjalnie zasilone danymi administracyjnymi, można przedstawić całkiem bogatą charakterystykę tej części populacji, która nie uczestniczyła w spisie. Pozyskana w ten sposób wiedza pozwoliłaby na konstrukcję modeli umożliwiających szacunek obciążenia. Prawdopodobnie w chwili obecnej analiza taka jest już niemożliwa, gdyż jest to uzależnione od zakresu informacji w Analitycznej Bazie Mikro danych. Duże oczekiwania w zakresie badania odmów i analizy braków odpowiedzi w NSP 2021 można wyrazić w związku z proponowaną konstrukcją Bazy Operatu Spisowego. Utworzona w GUS statystyczna baza spisowa umożliwiłaby, przykładowo poprzez łączenie danych z badań reprezentacyjnych z danymi ze statystycznej bazy spisowej, dokonywanie oceny charakterystyk osób (gospodarstw domowych), które nie biorą udziału w badaniach.

Zgodnie z informacjami zawartymi w publikacjach, w NSP 2011 nie przeprowadzono kalibracji w celu niwelacji skutków odmów udziału w badaniu (przypomnijmy, że odsetek ten wyniósł 13,7%). Zastosowano co prawda korektę wag. Jednak informacje w uwagach metodologicznych publikacji spisowych nie są szczegółowe. Wiadomo jedynie, że celem zastosowanej korekty wag było uzyskanie zgodnych szacunków liczby ludności na podstawie różnych źródeł oraz że korekta ta uwzględniała charakter budynku.

- Zastosowanie specjalnych metod badania w celu identyfikacji mniejszości narodowych i etnicznych oraz innych zbiorowości „trudnych do zbadania” z wykorzystaniem dodatkowych źródeł danych

Mogą to być odpowiednio dobrane, w zależności od dostępności zmieniających pomocniczych, estymatory statystyki małych obszarów, wykorzystujące

zmiennie na poziomie jednostek indywidualnych, na poziomie domen, albo typu SPREE czy kalibracyjne. W przypadku szacunku wartości globalnej każdej z charakterystyk konieczne jest także określenie możliwych do wykorzystania źródeł zmiennych pomocniczych. W przypadku niepełnosprawności mogą to być np. dane rejestru NFZ, ZUS, European Health Interview Survey (EHIS) czy jeszcze inne.

Ważne jest wykorzystanie metod estymacji pośredniej z uwzględnieniem korelacji przestrzennej oraz systemu informacji geograficznej w celu poprawy precyzji szacunków w przekroju małych domen (jednostek przestrzennych bądź innych małych grup, np. osób bezdomnych). Należy podkreślić potrzebę wypracowania metod estymacji umożliwiających publikację danych spisowych w zakresie wybranych charakterystyk nie tylko na poziomie gmin. Chodzi o szacunek wybranych charakterystyk spisowych w przekroju domen niestandardowych określonych na podstawie siatki gridowej. Zagadnienie to jest szczególnie istotne w analizie dojazdów do pracy i związanej z nią delimitacji lokalnych rynków pracy oraz obszarów metropolitalnych, których granice nie pokrywają się z granicami jednostek administracyjnych.

#### 8. Zastosowanie zaawansowanych metod badawczych

Wykorzystanie wielu źródeł w spisie nowej generacji wynika z ich dostępności oraz pozytywnej weryfikacji jakości. Nie byłoby to możliwe, gdyby nie rozwój nowoczesnych technologii i elektroniczna forma tych źródeł. Ale uwzględnienie wielości źródeł wymaga też zastosowania odpowiedniej metodologii. Statystycy muszą odpowiedzieć na wiele problemów związanych z integracją źródeł i oceną ich wiarygodności, co motywuje bezwzględnie konieczny, intensywny rozwój metod statystycznych. W większości zgodnej opinii wielu statystyków to rozwój metod statystycznych nie nadaża za rozwojem technologii informatycznych (Szreder, 2009, 2016, 2017).

Niemniej prace metodologiczne, także prowadzone w Polsce, są coraz bardziej zaawansowane. Odwołując się jedynie do prac prowadzonych w ośrodku poznańskim (Beręsewicz i in., 2018), warto zwrócić uwagę na następujące kwestie:

- Integracja danych z różnych źródeł – identyfikator statystyczny  
Ważnym zagadnieniem jest zabezpieczenie poufności danych, szczególnie w kontekście spisu wieloźródłowego wykorzystującego dane rejestrów administracyjnych oraz innych pozaadministracyjnych systemów informacyjnych. W spisie 2011 r. „naturalnym” identyfikatorem osób był numer PESEL, choć bezpośrednio integracja zbiorów została przeprowadzona za pomocą unikalnego, sztucznego klucza połączeniowego<sup>162</sup> (GUS, 2014b, s. 203). Za-

<sup>162</sup> Oprócz numeru PESEL w identyfikacji rekordów i łączeniu ze zbiorem referencyjnym, jakim był wykaz adresowo-mieszkaniowy, wykorzystano dane adresowe, system TERYT, numer

sadniczym celem konstrukcji tego specjalnego klucza identyfikacyjnego bazującego na identyfikatorach z rejestrów było umożliwienie łączenia zasobów administracyjnych z zapewnieniem ochrony danych wynikającej z ustawy o statystyce publicznej oraz z ustawy o ochronie danych osobowych.

Korzystając z numeru PESEL, można łączyć dane rejestrów administracyjnych<sup>163</sup>. Powstaje jednak pytanie o możliwość integracji danych z rejestru z danymi badania reprezentacyjnego, szczególnie wówczas, gdy w badaniu nie wprowadzono numeru PESEL. W NSP 2011 informacje tzw. „długiego formularza” zebrano w specjalnym badaniu reprezentacyjnym. Zamiast tego można postulować wykorzystanie istniejących, regularnie realizowanych badań specjalnych (np. BBGD, BAEL). Obecnie badania reprezentacyjne nie są przystosowane do integracji z rejestrami administracyjnymi. Natomiast zadanie utworzenia identyfikatora statystycznego, będącego kombinacją wartości wybranych zmiennych, jest w tym przypadku trudne z powodu licznych błędów. W szczególności efektywne zastosowanie probabilistycznego łączenia rekordów uniemożliwiają błędnie bądź w różny sposób zapisane adresy w badaniach reprezentacyjnych i w rejestrach, błędne daty urodzeń, które są kluczowe w konstrukcji identyfikatora statystycznego.

Warte rozważenia jest utworzenie identyfikatora statystycznego na wzór austriackiego bPIN, który umożliwiłby integrację wielu różnych źródeł danych, jednocześnie nie będąc w sprzeczności z zapisami ustawy o statystyce publicznej. Taki identyfikator nie pozwalałby identyfikować osób, umożliwiłby jedynie proces łączenia, co jest niezmiernie istotne ze względu na możliwości zastosowania metod estymacji pośredniej wykorzystujących zmienne pomocnicze na poziomie jednostek indywidualnych.

– Kalibracja, imputacja, *repeated weighting* i inne korekty wag

Jednym z zadań w spisach opartych na wielu źródłach jest konieczność odpowiedniego zintegrowania i zachowania spójności pomiędzy wynikami badania reprezentacyjnego z danymi pochodzącymi z rejestrów, które zgromadzono w „złotym rekordzie”. W tym celu stosuje się różne metody ważenia (Gerards, 2012; Stachurski i Żądło, 2017; Statistics Netherlands, 2011; Tanton, Vidyattama, Nepal i McNamara, 2011). W NSP 2011 zastosowano kalibrację wag, metodę polegającą na skorygowaniu wag wynikających ze schematu losowania próby z wykorzystaniem informacji dodatkowych (spoza próby) tak, aby spełnione były odpowiednie równania kalibracyjne<sup>164</sup>. W efekcie uzyskuje się równowagę rozumianą w sensie zgodności struktur w różnych przekrojach

---

NIP, REGON. Klucz identyfikacyjny został wygenerowany za pomocą transformacji CUST\_OBM\_SURRKEY.

<sup>163</sup> Również poprzez konstrukcję identyfikatora statystycznego.

<sup>164</sup> Podobne procedury stosują statystycy także w innych państwach. Zastosowana w NSP 2011 metoda kalibracji wykorzystywała odpowiednio zaadaptowane i zmodyfikowane oprogramowanie

(wagi sumują się nie tylko do liczby jednostek populacji, ale także według różnych przekrojów). Wagi kalibracyjne dostarczają nieobciążonych oszacowań, estymatory kalibracyjne cechują się większą precyzją aniżeli estymatory tradycyjne (Szymkowiak, 2009, 2011, 2012, 2014; Szymkowiak i Beręsewicz, 2011). Zastosowanie kalibracji w NSP 2011 umożliwiło dopasowanie struktur z badania reprezentacyjnego do znanych wartości globalnych z rejestrów administracyjnych w odniesieniu do wszystkich badanych przekrojów<sup>165</sup>. Kompleksowa ocena wag kalibracyjnych wykazała, że mogą one stanowić postawę uogólniania wyników z wykorzystaniem danych pochodzących z badania reprezentacyjnego.

W kontekście NSP 2021 Szymkowiak (2018) zwraca szczególną uwagę na problemy uogólniania wyników w starszych grupach wieku (85+) czy w starszych rocznikach wieku (85, 86, ...) oraz w grupach wieku (15–17) czy poszczególnych rocznikach z tej grupy. Zauważa on, że uogólniając wyniki na dużym poziomie agregacji (województwo), wpływ ekstremalnych wag kalibracyjnych będzie pomijalny, ale dla bardzo szczegółowo zdefiniowanych przekrojów terytorialnych czy innych domen ich znaczenie może wzrastać. Ponadto, uwzględniając potrzeby wynikające z szerszego wykorzystania wielu źródeł, Szymkowiak (2018) proponuje zastosowanie kalibracji zintegrowanej, umożliwiającej jednoczesne szacunki wag dla gospodarstw domowych i ich członków. Metoda ta umożliwia zapewnienie zgodności szacunków w obu przekrojach dla gospodarstw domowych oraz indywidualnych osób<sup>166</sup>. W tym zakresie należy postulować wykorzystanie dorobku pracowników Uniwersytetu Ekonomicznego w Poznaniu i Ośrodka Statystyki Małych Obszarów Urzędu Statystycznego w Poznaniu (GUS, 2016, 2017a; Szymkowiak, 2014; Szymkowiak i Beręsewicz, 2011; Szymkowiak i Józefowski, 2013).

Warto podkreślić, że zastosowanie metod kalibracji wymaga, by kluczowe zmienne dodatkowe spoza próby, np. z Bazy Spisowej/Operatu Spisowego, wykorzystywane w estymatorach kalibracyjnych były zweryfikowanej jakości, wyczyszczone z błędów i kompletne. Jakość tych zmiennych stanowi o rezultatach kalibracji. Istotne znaczenie jako źródło danych na potrzeby kalibracji oraz w ocenie jakości różnych badań może mieć tworzona w GUS statystyczna baza spisowa.

---

CALMAR (Calibration on Margins), napisane w języku 4GL w środowisku SAS, służące do wyznaczenia wag kalibracyjnych, które powstało na potrzeby INSEE (francuskiego urzędu statystycznego).

<sup>165</sup> Część tablic spisowych, szczególnie dla cech demograficznych, była konstruowana z wykorzystaniem danych rejestrów. Pozostałe tablice konstruowano na podstawie danych badania. Wagi wynikające ze schematu losowania próby nie zawsze idealnie odtwarzały dane rejestrów administracyjnych, zwłaszcza rejestru PESEL. Wynikało to z uwzględnienia wielu przekrojów terytorialnych, demograficznych i społecznych, tj. podziału na ponad 70,5 tys. warstw.

<sup>166</sup> Wszyscy członkowie gospodarstw domowych otrzymają te same wagi co gospodarstwa, do których należą.

## – Statystyka małych obszarów

Według Paradysza (2010, s. 47) ani tzw. badania wyczerpujące, czyli spisy, ani reprezentacyjne (oparte na próbie losowej) nie dostarczają pełnego obrazu rzeczywistości. Autor uważa, że znaczne korzyści w obu przypadkach można odnieść, sięgając do metod estymacji pośredniej, określanej również jako statystyka małych obszarów (SMO)<sup>167</sup>. Metodologia SMO znajduje zastosowanie szczególnie w sytuacji nielicznej reprezentacji próby dla danej jednostki terytorialnej, gdy szacunków nie można przeprowadzić za pomocą tradycyjnych technik metody reprezentacyjnej. Statystyka małych obszarów jest metodą estymacji dla domen na podstawie informacji dotyczących wybranych subpopulacji, ale zebranych w badaniu zaprojektowanym dla całej populacji, z wykorzystaniem wszelkich dostępnych informacji o badanej cesze, także pochodzących z innych źródeł. Ideę estymacji pośredniej Paradysz tłumaczy następująco (2010, s. 47): „Zgodnie z określeniem Rao (2003, s. 2) estymatory pośrednie są tymi, które «pożyczają siłę» (*borrow strength*) dla oszacowania danej wartości  $Y$  dzięki jej związkom w przestrzeni i (lub) czasie zwiększając tym samym efektywną wielkość próby”.

Kompleksowy przegląd metod estymacji pośredniej zawiera praca Rao (2003). Wesołowski (2003) wyszczególnia techniki estymacji w ramach podejścia randomizacyjnego (bezpośrednie, syntetyczne i złożone) oraz w ramach podejścia modelowego (empiryczny najlepszy liniowy predyktor nieobciążony – EBLUP, empiryczny estymator bayesowski – EB i hierarchiczny estymator bayesowski – HB). W Polsce własności wielu technik estymacji dla małych domen weryfikowano m.in. w projekcie EURAREA<sup>168</sup>. Metody statystyki małych obszarów są rozwijane i wprowadzane do praktyki statystycznej także w szacunkach spisowych (Klimanek, 2010; Münnich, Burgard, Gabler, Ganninger i Kolb, 2016). Obecnie w Ośrodku Statystyki Małych Obszarów Urzędu Statystycznego w Poznaniu są prowadzone prace, w których są testowane estymatory SPREE na potrzeby konstrukcji tabel spi-

<sup>167</sup> Określenia estymacja pośrednia i statystyka małych obszarów często są używane zamiennie. Jednak w rzeczywistości estymacja pośrednia jest pojęciem szerszym, powszechniej stosowanym, podczas gdy statystykę małych obszarów należy postrzegać jako jeden z działów metody reprezentacyjnej. Metodologia statystyki małych obszarów znajduje zastosowanie szczególnie w sytuacji małej reprezentacji próby, gdy estymatory bezpośrednie dostarczają ocen o niskiej precyzji (Bracha, 1996; Wywiół, 2010; Żądło, 2015).

<sup>168</sup> Projekt Eurostartu: *Enhancing Small Area Estimation Techniques to meet European Needs* (IST-2000–26290) wchodził w skład V Ramowego Programu Badań, Rozwoju Technicznego i Prezentacji Unii Europejskiej. W projekcie uczestniczyło sześć państw: Wielka Brytania, Finlandia, Szwecja, Włochy, Hiszpania i Polska, a głównym koordynatorem był Urząd Statystyczny Wielkiej Brytanii (Eurarea Consortium, 2004).

sowych<sup>169</sup> w kontekście niepełnosprawności (Szymkowiak, 2018; Szymkowiak i Józefowski, 2013). Zastosowanie estymatorów SPREE (wykorzystujących technikę iteracyjnego proporcjonalnego dopasowania), GLM (estymator SPREE bazujący na uogólnionych modelach liniowych) i GLSMM (SPREE bazujący na uogólnionych modelach mieszanych z efektami losowymi) daje bardzo dobre wyniki zarówno w kontekście precyzji, jak i obciążenia (Szymkowiak, 2018). Dlatego w tym miejscu warto sformułować postulat o rozszerzenie technik estymacji pośredniej planowanych do wykorzystania w NSP 2021, szczególnie na potrzeby szacunków dla niskich poziomów agregacji przestrzennej.

Pożyczanie mocy w estymacji pośredniej oznacza korzystanie z wielu źródeł danych, które dostarczają zmiennych wspomagających (o zweryfikowanej jakości), oraz zastosowanie zaawansowanych podejść analitycznych. W stosunku do tradycyjnych spisów estymacja pośrednia wprowadza korektę ewidentnych błędów systematycznych, braków odpowiedzi i nieobecności w badaniu. W porównaniu z tradycyjnymi badaniami reprezentacyjnymi zastosowanie estymatorów pośrednich pozwala na znaczne podwyższenie precyzji estymatorów lub przy danej dokładności obniżenie kosztów badań. Niestety, estymatory pośrednie może charakteryzować znaczne obciążenie. Pomimo rozwoju i ciągłego doskonalenia metod statystyki małych obszarów, nie należy ich traktować jako remedium na wszystkie problemy braku informacji.

## 9. Badanie jakości w NSP 2021

Ogromnym sukcesem NSP 2011 jest prowadzona w środowisku statystyków i demografów dyskusja nt. jakości wyników spisu. Bardzo wiele zrobiono w celu poprawy jakości wykorzystanych źródeł oraz precyzji szacunków. Planowane są również kolejne przedsięwzięcia stanowiące zręby kompleksowego systemu kontroli jakości, w tym:

### – Spis kontrolny

Według koncepcji NSP 2021 planowane jest przeprowadzenie niezależnego badania kontrolnego. Choć brakuje szczegółowych informacji o planowanej metodzie, bardzo pozytywnie należy ocenić świadomość potrzeby oraz już podjęte prace. Ma to zasadnicze znaczenie dla określenia przydatności spisu kontrolnego. Dotychczasowe doświadczenia sugerują, że pod tym względem jest szczególnie dużo do zrobienia. Do tej pory w Polsce nigdy nie przedstawiono korekty szacunków spisu właściwego na podstawie wyników badania kontrolnego.

<sup>169</sup> Projekt Eurostatu poświęcony problemom poprawienia jakości unijnych spisów w 2021 r. i po 2021 r.

- Kompleksowa ocena kompletności  
Przeprowadzenie spisu kontrolnego jest najlepiej ocenianą metodą oceny kompletności. Niezależnie można zastosować analizę rezydencji umożliwiającą weryfikację rzetelności szacunku liczby ludności oraz różne metody analizy demograficznej.
- Zasada redundancji – metryka zmiennych  
Należy zadbać o właściwe opisanie mechanizmu wzajemnej kontroli danych „wygenerowanych” na podstawie wielu źródeł. Zasada redundancji była stosowana w NSP 2011 i można zakładać jej wykorzystanie w NSP 2021, skoro planowane jest korzystanie z wielu źródeł. Ważnym elementem będzie uwzględnienie metryki dla każdej zmiennej spisowej (aneks IV.2). Propozycja metryki zawiera, obok definicji, między innymi źródło, algorytm i statystyki poziomu spójności. W koncepcji spisu sporo uwagi poświęcono kwestiom jakości. Przykładowo przedstawiona propozycja w odniesieniu do mikro danych operacyjnych (aneks IV.3) ma charakter kompleksowego ujęcia uwzględniającego wskaźniki w sześciu wymiarach jakości na etapie produkcji.
- Zapewnienie oceny precyzji estymacji z uwzględnieniem odmów oraz udostępnieniem opisu zastosowanej metody.
- Dbłość o zachowanie porównywalności podstawowych charakterystyk dla potrzeb:
  - badań i analiz demograficznych pomiędzy spisami,
  - zarządzania w skali kraju i regionu, a także samorządu lokalnego,
  - zgodności ze standardami międzynarodowymi.
- Badanie stopnia przydatności danych spisu, stopnia spełnienia zobowiązań i potrzeb użytkowników.
- Opracowanie schematu publikacji wyników spisu uwzględniającego:
  - zapewnienie powszechnego dostępu do danych spisu, także do danych jednostkowych,
  - pełną transparentność zastosowanych metod badawczych,
  - publikację metadanych, szczególnie ważną w spisie korzystającym z wielu źródeł,
  - wyczerpującą informację o metodach estymacji i oceny precyzji szacunku,
  - dostępność danych w różnych klasyfikacjach oraz przekrojach terytorialnych,
  - interpretację wyników.
- W kontekście oceny jakości wyników spisu, obok zastosowania i rozwoju metod integracji statystycznej, podkreślenia wymaga także potrzeba zachowania podstawowych charakterystyk populacji „wyprowadzonych” jedynie na podstawie zasobów rejestrów administracyjnych, oczywiście po weryfikacji ich jakości. Chodzi o „zachowanie” wartości referencyjnej zbiorów danych



administracyjnych przekształconych do postaci statystycznych baz spisowych. Dla każdej charakterystyki spisowej należy przedstawić pełną metrykę obejmującą informację o źródle i jego jakości, sposobie „wyprowadzania”, metodzie szacunku, precyzji szacunku, wielkości odmów itp.

#### 10. Upowszechnianie wyników spisu

- Powszechność dostępu do danych spisu

Spis jest badaniem finansowanym ze środków publicznych, dlatego dostęp do danych winien być powszechny i bezpłatny. W szczególności dla potrzeb badań naukowych winny być udostępniane dane jednostkowe umożliwiające konstrukcję różnych modeli i badanie zależności w dowolnie określonych przekrojach przestrzennych oraz dla różnych klasyfikacji. Zadaniem statystyki publicznej jest zapewnienie poufności danych, np. według cytowanego już austriackiego narodowego systemu identyfikacji dla potrzeb statystyki publicznej.

W tym kontekście proponuję uproszczenie procedur pozyskiwania danych. Utworzone przez GUS stanowisko „Naukowiec” w praktyce się nie sprawdziło. W chwili obecnej trzeba prowadzić długotrwałe, czasami wręcz kilkuletnie „zabiegi” o pozyskanie danych do badań naukowych, obfitą korespondencją, wielokrotne rozmowy, uzgodnienia, które kończą się fiaskiem. Absolutnie kuriozalnym zjawiskiem jest utrudniony dostęp do danych dla pracowników naukowych uczelni będących jednocześnie pracownikami urzędu statystycznego, którzy mają realizować badania często na zlecenie GUS i na potrzeby konkretnych projektów czy badań (np. spisu ludności).

- Statystyka sąsiedztwa

Wprowadzenie statystyki sąsiedztwa na wzór angielskiego systemu, który umożliwia pozyskanie podstawowych informacji społecznych, np. w zakresie służby zdrowia, systemu oświaty, pomocy społecznej, dla każdej lokalizacji (na podstawie kodów adresowych) (Roszka, 2013, s. 61 i nast.). System ten powinien umożliwiać także przeprowadzanie analiz porównawczych ułatwiających dokonywanie wyborów według określonych preferencji.

- Terminowość i punktualność upowszechniania informacji.
- Wykorzystanie nowoczesnych technik telekomunikacyjnych i informatycznych w celu upowszechnienia danych w formie spełniającej oczekiwania i potrzeby odbiorców (różne kanały dystrybucji, metadane, formy publikacji).

Przeprowadzenie spisu pełnego z wykorzystaniem danych rejestrów administracyjnych i innych pozastatystycznych źródeł ma szansę zyskać, podobnie jak w Austrii czy w państwach skandynawskich, pełną aprobatę społeczną. Oczywiście pod warunkiem zachowania wszystkich zasad Europejskiego Kodeksu Praktyk Statystycznych (Eurostat, 2011). Spis korzystający z istniejących zasobów systemów

informatycznych oznacza redukcję kosztów oraz obciążeń respondentów. Jeśli tak przeprowadzone badanie będzie w stanie zaspokoić potrzeby informacyjne społeczeństwa, przedsiębiorców, organów administracji państwowej, naukowców i jeszcze towarzyszyć mu będzie nieustanna poprawa jakości szacunków oraz wzrost zakresu informacji, można to tylko interpretować w kategoriach wzrostu efektywności wykorzystania środków publicznych.

---

## ZAKOŃCZENIE

---

Pomimo długiej tradycji i dobrze wypracowanej metodologii spisy ludności nie są badaniami dostarczającymi „idealnych” rezultatów (Baldrige i in., 1985). Przede wszystkim spisy ludności – jako badania wyczerpujące – mogą być obciążone błędami o nielosowym charakterze. Biorąc pod uwagę różnorodne metody przeprowadzania spisów, w tym metodę tradycyjną, wykorzystującą dane rejestrów administracyjnych, oraz metodę mieszaną, w rozważaniach uwzględniono różne źródła błędów. Poszukiwanie najwłaściwszej metody przeprowadzania spisów ludności wskazuje na brak rozwiązania idealnego z punktu widzenia celu badania, bez względu na uwarunkowania geograficzne, klimatyczne, tradycję, jak również koszty ekonomiczne i społeczne.

W opracowaniu przedstawiono obserwowane w ostatnim czasie **zmiany w zakresie metod przeprowadzania spisów ludności**. Opisano ich zasięg oraz zróżnicowanie. Jednocześnie ukazano konsekwencje wprowadzanych rozwiązań. Realizując określony cel główny opracowania udzielono odpowiedzi na zadania szczegółowe w kolejności umożliwiającej **kompleksową ocenę jakości spisów**, najpierw w odniesieniu do **kryteriów jakości według definicji Eurostatu** (rozdział 2), a następnie oceniając badanie **pod względem demograficznym, w szczególności w zakresie kompletności** (rozdział 3).

W pierwszej kolejności podjęto **dyskusję podstawowych pojęć i własności spisu ludności**. W rozważaniach zwrócono uwagę na właściwe rozumienie i interpretację właściwości spisów tak w kontekście metody ich przeprowadzania, jak również jakości. Wykazano bezzasadność twierdzeń o niespełnieniu, przez spisy oparte na rejestrach, wymogu powszechnego i indywidualnego spisywania. Wręcz przeciwnie, wykazano, że korzystanie ze źródeł administracyjnych dla każdej indywidualnej osoby ogranicza wielkość błędu respondenta grupowego.

Następnie przedstawiono **konsekwencje przyjęcia określonych definicji ludności spisowej**. Podniesiono tutaj zasadnicze dla jakości spisów kwestie związane z określeniem populacji celu. Omówiono różnice między przyjmowanymi w dotychczasowych spisach ludności w Polsce definicjami ludności faktycznej i stałej

oraz rezydującej, zgodnie z obowiązującymi regulacjami Eurostatu. Wskazano na sięgające 1,5 mln osób rozbieżności szacunków wynikające ze sposobu ujęcia migracji długookresowych oraz odwoływania się do obowiązujących regulacji prawnych. Podniesiono konsekwencje publikowanych szacunków, które dotyczą nie tylko obciążenia szacunku współczynników demograficznych, ale też wszystkich statystyk makroekonomicznych. Konsekwencje te są szczególnie poważne, gdyż odnoszą się do najbardziej mobilnych grup ludności, ludności w wieku produkcyjnym oraz o najwyższym natężeniu aktywności matrymonialnej i prokreacyjnej. Zasadna wydaje się przedstawiona propozycja uporządkowania przez GUS definicji ludności rezydującej, a co najmniej publikowania także szacunków z uwzględnieniem migracji.

Wyjaśnienie rozbieżności w zakresie podstawowych pojęć, definicji i własności warunkowało przejście do dyskusji **metod przeprowadzania spisów**. W tym zakresie odwołano się do zaleceń ONZ, ale przedstawiono także własną propozycję klasyfikacji metod z uwzględnieniem rodzajów wykorzystywanych źródeł danych, w tym administracyjnych i pozaadministracyjnych. Podsumowaniem tych rozważań jest dyskusja możliwości i ograniczeń, która stanowi jednocześnie wprowadzenie do oceny jakości spisu. **Dokonując oceny jakości spisu ludności, skorzystano z międzynarodowej definicji jakości badań statystycznych, oceniając je pod względem sześciu kryteriów: (i) przydatności, (ii) dokładności, (iii) terminowości i punktualności, (iv) dostępności i przejrzystości oraz (v) porównywalności i (vi) spójności.**

W ocenie **przydatności** wykazano, że na tle wcześniejszych spisów przeprowadzonych metodą tradycyjną, a więc z założenia bazujących na jednym źródle danych, przeprowadzony metodą mieszaną NSP 2011 należy ocenić jako najbardziej wyczerpujący. Większa liczba źródeł w sposób oczywisty oznacza możliwość poszerzenia zakresu tematycznego spisu.

Przestawiając kwestie **dokładności** spisu jako badania pełnego, ograniczono się do oceny błędów nielosowych. Zwrócono uwagę, że najbardziej uznane metody oceny błędów kompletności i treści w spisie tradycyjnym polegały na odwoływaniu się do wyników innych badań, przede wszystkim spisu kontrolnego. Badanie to jest określane jako niezależna, przeprowadzona metodą reprezentacyjną, replika spisu podstawowego. Odnośnie do tego referencyjnego badania zakłada się wysoki poziom merytoryczny w jego programowaniu, doborze próby oraz estymacji i ocenie precyzji, wykorzystanie najlepszej jakości operatu losowania, wysoko wykwalifikowanej kadry, najbardziej efektywnych metod organizacji badania w terenie, rzetelności przeprowadzonych wywiadów, dotarcia do respondentów, braku odmów itd. **Oznacza to, że ocena jakości tradycyjnego spisu ludności sprowadza się do porównania wyników niezależnych badań.** Procedura ta jest tożsama z ideą spisu opartego na rejestrach, która polega właśnie na ocenie dostępności danych w istniejących źródłach, porównaniach, ewaluacji i integracji wyników niezależnych badań.

W dyskusji problemów **terminowości i punktualności** skoncentrowano się na zagadnieniu **aktualności**. Powołując się na wyniki analiz przeprowadzonych przez Caldera i Teague (2012), wykazano korzyści z aktualności danych administracyjnych wobec spisu tradycyjnego. Rejestry administracyjne są w istocie badaniami ciągłymi, w których dane są na bieżąco aktualizowane. Przygotowując badanie oparte na rejestrach, można tak zaplanować i zautomatyzować szereg zadań, by skrócić czas potrzebny na udostępnienie danych i opublikować je stosunkowo bliżej daty referencyjnej. Uwzględnienie możliwości automatyzacji procedur procesu przekazywania danych oznacza ich większą aktualność oraz skrócenie czasu od momentu referencyjnego. Dodatkowo wykorzystanie danych ze źródeł administracyjnych umożliwia podejście zapewniające publikowanie danych o ludności z większą częstotliwością – np. roczną. Dzięki temu uniknie się dokonywanej wstecz po każdym spisie korekty stanu i struktury ludności. Wprowadzała ona nie tylko „aktualizację” na podstawie często z dwuletnim opóźnieniem publikowanych danych spisowych, ale także podkreślała ich nadrzędny charakter jednocześnie sankcjonując wszystkie błędy.

Jakość w zakresie **udostępniania i przejrzystości danych** w przypadku korzystania z wielu źródeł wynika z konieczności zastosowania zaawansowanych metod, które wymagają nie tylko przedstawienia definicji i stosowanych klasyfikacji, ale przede wszystkim opisu algorytmu generowania wartości zmiennych, wyników porównań, badania spójności mikrodanych w trakcie realizacji badania. Wykonane operacje często się przekładają na sposób ich przedstawienia w materiałach źródłowych. Wskazano tutaj zawartość metryki zmiennej (aneks IV.2), czy proponowaną analizę jakości mikrodanych operacyjnych (aneks IV.3) w koncepcji NSP 2021, podkreślając, że podobnych informacji nie publikowano we wcześniejszych spisach w Polsce.

W ocenie jakości według kryterium **porównywalności** zwrócono uwagę na znaczenie regulacji międzynarodowych, których przestrzeganie jest konieczne i niezależne od sposobu badania. Podniesiono jednak problem zgodności deklaratywnej oraz konieczności odwołań do uwag metodycznych zawierających szczegółowy opis sposobu pozyskania zmiennych. Jest to element warunkujący wnikliwą analizę zgodności oraz wskazanie rozbieżności, co stanowi podstawę analiz porównawczych.

Zapewnienie **spójności** informacji w sytuacji wielu źródeł wymaga bardziej kompleksowego podejścia, uwzględniającego zastosowanie zaawansowanych metod edycji danych, integracji źródeł, estymacji, kalibracji, oceny precyzji itd. Stanowią one w rzeczywistości istotę wieloźródłowych badań statystycznych. Rozwój metod statystycznych odpowiadających wyzwaniom badań korzystających z różnych dostępnych zasobów informacji oraz ich właściwe zastosowanie jest warunkiem koniecznym wykazania wiarygodności szacunku opartego na wielu źródłach. Uzyskanie spójnych szacunków w badaniu bazującym na jednym źródle jest łatwiejsze, ale to nie znaczy, że automatycznie osiągalne. Wręcz przeciwnie, w spisie ludności,

który bez względu na rodzaje wykorzystanych źródeł oraz zastosowaną metodę jest badaniem niezwykle złożonym, osiągnięcie spójności szacunków zawsze stanowiło jedno z poważniejszych wyzwań.

Ocena tradycyjnych spisów ludności oraz spisów nowej generacji w świetle kryteriów międzynarodowej definicji jakości stanowiła pierwszy etap badania jakości spisów. Było to jednocześnie wprowadzenie do oceny spisu w kategoriach badania demograficznego, którego zasadniczym celem jest przedstawienie szacunku stanu i struktury ludności. Drugi etap oceny jakości spisu polegał na wykonaniu pozostałych zadań szczegółowych.

Kolejnym, piątym celem szczegółowym było określenie **rodzajów błędów i metod ewaluacji spisów ludności** jako szczególnych badań demograficznych i statystycznych o długiej historii i rozwiniętej metodologii. Również w tym zakresie zaprezentowano autorskie zestawienie rodzajów błędów nielosowych z uwzględnieniem źródeł danych wykorzystywanych w badaniu. Realizując szósty cel szczegółowy, podjęto próbę ewaluacji spisu z perspektywy demograficznej, jako badania stanu i struktury ludności, i przedstawiono **ocenę kompletności NSP 2011**.

Zawarta w tej części opracowania ocena jakości szacunków wykorzystuje różne metody analizy demograficznej. Ewaluacja w tym zakresie obejmuje wykorzystującą dane ewidencji urodzeń ocenę szacunku liczby niemowląt; przeprowadzoną na podstawie wskaźników „przetrwania” analizę kompletności liczby dzieci i osób starszych oraz całej populacji. Wykazano tutaj niedoszacowanie liczby niemowląt przy jednoczesnym przeszacowaniu liczby dzieci w wieku 1–9 lat. Brak niemowląt jest typowym błędem, powtarzającym się w kolejnych spisach. W spisie 2011 r. wykorzystującym dane rejestrów odnotowano poprawę kompletności w odniesieniu do najmłodszych niemowląt urodzonych w roku spisu. Z kolei przeszacowanie liczby dzieci jest sytuacją wyjątkową. Wyjaśnienia wynoszącej ponad 70 tys. rozbieżności między liczbą dzieci wykazanych w NSP 2011, które nie zostały objęte ewidencją urodzeń, poszukiwano, stosując metody statystyki lustrzanej na podstawie danych spisu ludności w Anglii i Walii.

W ocenie kompletności szacunku dla pozostałych grup wieku wykazano wyższe od jedności wskaźniki przetrwania, co potwierdza błędy niedoszacowania w spisie 2002 r. W szczególności wykazano braki osób w wieku 10 lat oraz 30–35 lat. Otrzymane wyniki są zbieżne z oceną spisu przedstawioną przez Paradysza (2010). Przeprowadzona analiza stanu i porządku wymierania 100-latków w zasadzie nie wykazała obserwowanych wcześniej „zgonów duchów”, przez co pozwala na wyższą ocenę jakości szacunków NSP 2011.

Całość rozważań kończy **dyskusja perspektyw spisu ludności w świetle dostępnych źródeł danych i rozwoju metod badawczych wobec międzynarodowych standardów statystyki publicznej**. W rozdziale czwartym przedstawiono rekomendacje organizacji międzynarodowych ze szczególnym uwzględnieniem zaleceń dotyczących krajów europejskich. Odwołano się tutaj do doświadczeń oraz

wynikających z nich planów diskutowanych w niektórych państwach, w szczególności w Stanach Zjednoczonych i Wielkiej Brytanii. Szczególną uwagę poświęcono jednak koncepcji spisu planowanego na 2021 r. w Polsce. Rozważania te stanowią ustosunkowanie się autorki do przedstawionych w dokumentach GUS propozycji.

Czy więc rzeczywiście przeżywamy zmierzch ery spisów ludności? Jeszcze w 1800 r. obejmowały one zaledwie 1% ludności świata, na początku XX w. zasięg spisów sięgał około połowy populacji (Kędelksi i Paradysz, 2006, s. 30), a według ostatnich statystyk ONZ spis rundy 2010 objął 97% ludności świata. Można się pokusić o wręcz przeciwne stwierdzenie – zarządzanie państwami i społeczeństwami jest współcześnie niemożliwe bez informacji, jakich dostarcza spis ludności. Przedstawiona w pracy dyskusja wskazuje, że zmiany metodologiczne w spisach powodują, że spisy te są istotnie różne od poprzednich. Transformacja spisów została spowodowana rozwojem informatyki z uwzględnieniem nowoczesnych technologii komunikacji, geograficznych systemów informacji GIS, dostępności do coraz bardziej nieograniczonych zasobów danych. Jednak sama informatyzacja oraz pojawienie się big data skutkują zalewem danych i są przyczyną wypierania dobrej informacji przez złą (Oleński, 2003, 2006). Uwzględnienie jedynie tych czynników prowadzi do refleksji o zbędności spisów. Dopiero wnikliwa analiza potrzeb i możliwości ich zaspokojenia, uwzględniająca obiektywną ocenę jakości danych, pozwala zauważyć konieczność dogłębnych przemian. Obejmują one rozwój metod badań statystycznych począwszy od statystycznej integracji danych, probabilistycznego łączenia rekordów, poprzez metody harmonizacji, imputacji, edycji, badania spójności, weryfikacji reprezentatywności danych ze źródeł pozastatystycznych, benchmarking, metody ważenia i kalibracji po rozwój metod statystyki małych obszarów, estymacji pośredniej, oceny jej jakości, precyzji i obciążenia, konstrukcję baz mikrodanych, zapewniających poufność oraz wizualizację i upowszechnienie danych dostosowane do indywidualnych potrzeb dla dowolnie definiowanych domen i obszarów. Połączenie nowoczesnych technologii informatycznych z odpowiednimi metodami statystycznymi prowadzi do spisu nowej generacji. Tylko takie kompleksowe podejście dzięki istocie synergii daje efekt przeobrażeń, który może być określany transformacją spisów ludności prowadzącą do nowej, lepszej jakości.

Od ponad 200 lat spisy ludności gromadzą dane o badanej populacji. Istotne znaczenie spisów wynika stąd, że ukazują i wyjaśniają długoterminowe trendy, ale przede wszystkim odzwierciedlają populację, w której żyjemy dzisiaj. Nie byłoby zasadne uwzględnienie starej klasyfikacji zawodów obejmującej latarnika, ale pomijającej informatyka czy analityka mediów społecznościowych. Zrozumiałe jest, że gromadzimy dane dotyczące problemów nieistniejących dla poprzednich pokoleń<sup>170</sup>. Otaczająca nas rzeczywistość nieustannie się zmienia, inne są procesy i zjawiska,

---

<sup>170</sup> <https://www.ons.gov.uk/census/censustransformationprogramme/consultations/the2021censusinitialviewoncontentforenglandandwales/updateon2021censusforenglandandwales>.

które kształtują rozwój gospodarczy i społeczny, rozwój ludności zamieszkującej nasze państwo, kraj naszych sąsiadów, Europę, cały świat. Zmianom podlega to, co mierzymy i w jaki sposób to czynimy. Jednak od zarania ludzkości warunkiem rozwoju było posiadanie informacji o tym, ilu nas jest, kim jesteśmy, co robimy, gdzie mieszkamy, jak żyjemy i pracujemy.

Statystyka publiczna powinna nadążać za zmieniającą się rzeczywistością i sprostać współczesnym wymaganiom. W dobie globalizacji statystyka musi dostarczać zharmonizowanych danych, które są porównywalne w czasie i przestrzeni, pomiędzy państwami i kontynentami. Międzynarodowe strategie inteligentnego i zrównoważonego rozwoju, np. Europa 2020, generują coraz większe potrzeby informacyjne oraz wskazują na tendencję kontynuowania inicjatyw wspólnego programu Europejskiego Systemu Statystycznego, w tym programu spisów. Modernizacja statystyki, wykorzystanie rejestrów administracyjnych i źródeł pozastatystycznych już w tej chwili tworzy ramy szerszych systemów statystyki demograficznej i społecznej. Program spisu powszechnego Unii Europejskiej, rundy 2020 można postrzegać jako etap pośredni w kierunku fundamentalnej restrukturyzacji europejskiego systemu statystyki ludności, w tym spisu ludności.



## **I. Informacje uzupełniające rozdział 1: Spis ludności jako źródło danych o ludności**

### **I.1. Definicje populacji według Rekomendacji ONZ (UN, 2015a, s. 46-47)**

- 2.50. It is recommended that countries apply a threshold of 12 months when considering place of usual residence according to one of the following two criteria:
- a. The place at which the person has lived continuously for most of the last 12 months (that is, for at least six months and one day), not including temporary absences for holidays or work assignments, or intends to live for at least six months;
  - b. The place at which the person has lived continuously for at least the last 12 months, not including temporary absences for holidays or work assignments, or intends to live for at least 12 months.<sup>47</sup>
- 2.51. Persons who move frequently and do not have a place of usual residence should be numerated at the place where they are found at the time of the census.
- 2.52. Regardless of the criteria used to define the 12-month period, countries should ensure that each person should have one and only one place of usual residence.
- 2.53. There are various population groups for which some uncertainty may arise about their inclusion in the usual resident population. The following persons would generally be considered in the usually resident population:
- (a) Persons found at the moment of enumeration that cannot identify their place of usual residence, such as those that move often;
  - (b) National military, naval and diplomatic personnel and their families, located outside the country;

- (c) Foreign persons working for international organizations (not including foreign diplomats or military forces), provided that they meet the criteria for the usual residence in the country;
  - (d) Merchant seafarers and fishers usually resident in the country but at sea at the time of the census (including those who have no place of residence other than their quarters aboard ship);
  - (e) Persons who may be illegal, irregular or undocumented migrants, as well as asylum seekers and persons who have applied for or been granted refugee status or similar types of international protections, provided that they meet the criteria for the usual residence in the country;
  - (f) Persons who cross a frontier daily or weekly to work or study in another country, provided that they meet the criteria for the usual residence in the country;
  - (g) Children born in the 12 months before the census reference time and whose families are usually resident in the country at the census reference time;
  - (h) Persons of minor age studying abroad for one year or more to attain the primary secondary level of education, regardless of the frequency of return to the family home located within the country. If the person is also working abroad, the same rules for cross-border workers apply;
  - (i) Persons who regularly live in more than one country during a year, if the reporting country is the one where they live most of the time, irrespective of whether or not they are present in the reporting country at the census reference time.
- 2.53a. On the other hand, the following group of persons need to be considered for being excluded from the usual resident population:
- (a) Foreign military, naval and diplomatic personnel and their families, located in the country, regardless of their place of usual residence;
  - (b) Persons of minor age attending the primary or secondary level of education whose family home is located abroad, regardless of the duration of their stay. However, if these persons are also working in the country, then the identification of the place of usual residence follows the same rules as for cross-border workers;
  - (c) Third-level students who are absent from the country for one year or more;
  - (d) Persons who regularly live in more than one country during a year, if the reporting country is not the one where they live most of the time, irrespective of whether or not they are present in the reporting country at the census reference time.
- 2.54. The concept of usual residence may be referred to as though it is synonymous with the concept of *de jure* residence. In certain circumstances, however, the term “*de jure*” may carry with it a requirement that the person’s

residence at that place has a basis in the legal system applicable to that specific place.

In turn this implies that people without such a legal basis should not be enumerated in that area. It is not recommended that censuses of population and housing enumerate only those people with a legal right to be in a place but rather, as described in section 2 below, should include either all those present at the place on census night or all those whose usual residence on census night was at the place of enumeration.

## **1.2. Rekomendacje dotyczące sposobu kwalifikacji ludności według miejsca zamieszkania w przypadku wątpliwości**

**There are population groups for which some uncertainty may arise in defining their place of usual residence within the country. The recommended conventional treatment of these cases is as follows (UN, 2015a, s. 192-193):**

- a) Persons who work away from home during the week and who return to the family home at weekends should consider the family home as their place of usual residence.
- b) Persons of minor age in primary and secondary education who are away from home during the school term should consider their family home as their place of usual residence
- c) Students in tertiary education who are away from home while at college or university should consider their term-time address as their place of usual residence regardless of whether this is an institution (such as a boarding school) or a private residence.
- d) The institution should be taken as the place of usual residence of all inmates who at the time of the census have spent, or are likely to spend, six months or more in the relevant institution.

Examples of inmates of institutions include patients in hospitals or hospices, old persons in nursing homes or convalescent homes, prisoners and those in juvenile detention centres.

- e) Where a person regularly lives in more than one residence within the country during the year, the one where he or she spends the majority of the week or year before the census should be taken as his or her place of usual residence. These persons are not considered to be persons with no usual residence.
- f) For the (national) military, naval and diplomatic personnel and their families located outside the country the following classification rules should be applied:
  - I. If they are residing abroad for less than 12 months and they are intending to return to the place of departure, they should be allocated within the country in accordance with the rules for usual residence. In particular, they could be allocated to (by decreasing order of priority):
    - i. The family home address within the country, if any;

- ii. The duty station within the country to which they were attached before leaving.
- II. If they are residing abroad for at least 12 months or if they are not intending to return to the place of departure (although returning in the country within a 12-month period), they should be attributed to a “virtual region” (extra-region) of the country of departure.
- g) The place of enumeration should be taken as the place of usual residence of homeless or roofless persons, nomads, vagrants and persons with no concept of usual residence.
- h) A child who alternates between two households within the country (for instance after his or her parents have divorced) should consider the household where he or she spends the majority of the year before the census as his or her place of usual residence. Where an equal amount of time is spent with both parents, the place of usual residence should be as for the parent or household with whom the child is staying at the census reference time.

### **I.3. Wyjaśnienia dotyczące szacunków dla populacji trudnych do spisania**

#### **Difficult-to-enumerate groups (UN, 2015a, s. 194)**

The following difficult-to-enumerate groups are relevant to the production of any population count:

- a) **Nomads and persons living in areas to which access is difficult.** Making contact with these groups to enumerate them can be difficult, particularly as part of a point-in-time count. Enumeration may need to be done at a different time, over an extended period, or by using alternative methods to enable contact with these groups. For example, countries might consider asking those who provide services to these groups to assist with their enumeration. Seasonal movements may be identified in advance and this information can be used by collectors to enable contact. There needs to be planning and consultation, particularly with influential members of these groups, prior to the census to organize for their enumeration. Communications that publicize the benefits of the census and engaging appropriate leaders in support of the census may assist coverage. Awareness of cultural issues relevant to specific groups should also be considered in developing enumeration strategies.
- b) **Civilian residents temporarily absent from the country.** As these persons will be absent from the country at the time of the census, they will be excluded from a usual resident population count. To produce a usual resident count countries may collect information on these people from another family or household member present at the time of the census, but where a complete family or household

- is outside the country at the time of the census, it may not be possible for the census to collect information about these people. Estimates for usual residents temporarily absent from the country based on other sources may be required to produce reliable estimates of usual residents for planning and policy purposes.
- c) **Civilian foreigners who do not cross a frontier daily and are in the country temporarily.** These include undocumented persons, or transients on ships in harbour at the time of the census. These groups may be in the country at the time of the census and therefore form part of the population present count. It is important to include these groups in the population count if their demand for services is to be considered for planning and policy development purposes. However, these groups may prefer not be counted, either because they fear ramifications from being counted or because they do not identify themselves as part of the population of the country. Language and communication may present challenges. Countries need to develop strategies, appropriate for their context, to include these groups in their enumeration.
  - d) **Refugees, asylum seekers and internally displaced persons.** Refugee populations, asylum seekers and internally displaced persons (in and outside camps) should be enumerated and their numbers presented separately, allowing calculation of country population excluding such groups, when such a population count is required for non-demographic purposes.
  - e) **Military, naval and diplomatic personnel and their families located outside the country and foreign military, naval and diplomatic personnel and their families located in the country.** Apart from the difficulties mentioned in (b) and (c) that are common to groups who are absent from their own country, enumeration of these groups is subject to diplomatic protocols. Detailed counts and characteristics of these groups may be considered sensitive on security grounds in some countries. Counts of these groups may be available from administrative records.
  - f) **Civilian foreigners who cross a frontier daily to work in the country.** This group should be excluded from a usual resident population count. The practice of counting people where they spend census night removes much ambiguity and reduces possible duplication. The difficulty then is trying to include them in a service population if countries want to consider this group in policy development and in planning service delivery.
  - g) **Civilian residents who cross a frontier daily to work in another country.** These persons are usual residents of the country and should be included in the population count.
  - h) **Merchant seafarers and fishers resident in the country but at sea at the time of the census.** This group includes those who have no place of residence other than their quarters aboard ship. Identifying that the ship will be at sea at the time of the census may be problematic, so countries will need to develop strategies to ensure inclusion of this group in the population count. This may

include providing this group with census forms before their ship goes to sea or enumerating the ship before the time of the census.

- i) **Homeless or roofless persons, vagrants and persons with no concept of usual residence.** These should be included in the population count, and the census office should work with local government agencies, charities and other supporting bodies that provide support for this population group to identify the best method of collecting census information from these people.
- j) **Persons living in buildings with restricted access.** Persons living in gated communities, condominiums and apartment buildings could be difficult to enumerate, particularly as part of a point-in-time count. Enumeration may need to be done at a different time, over an extended period, or by using alternative methods to enable contact with these groups.
- k) **Stateless persons.** These are individuals who are not considered as nationals by any State under the operation of its laws. They are often undocumented and may not wish to be enumerated. However, every effort should be made to include such persons in the census. The census office should work with responsible government agencies, non-governmental organizations familiar with this population group and the United Nations High Commissioner for Refugees (UNHCR) to establish the best method for identifying stateless persons and collecting census information on them. Country of citizenship is generally essential for the identification of this group, though the census office should consult with relevant ministries and agencies, including the UNHCR, to determine whether additional information (such as residence history or identity documentation) may be required to establish the status of a stateless person.

#### **I.4. Wyjaśnienie GUS dotyczące liczby mieszkańców Polski**

**Publikacja na stronie internetowej GUS 28.10.2014, <http://stat.gov.pl/index.php>  
Wyjaśnienie dotyczące liczby mieszkańców Polski**

Główny Urząd Statystyczny informuje, że dane o liczbie ludności Polski są opracowywane metodą bilansową. Polega ona na przyjęciu za punkt wyjścia wyników ostatniego Narodowego Spisu Powszechnego, do których dodawane są liczby urodzeń, a odejmowane liczby zgonów zaistniałych w ciągu badanego okresu – na podstawie informacji pozyskiwanych z urzędów stanu cywilnego. W bilansie ludności Polski ogółem uwzględniane jest ponadto saldo migracji zagranicznych na pobyt stały określone na podstawie badań GUS.

Według stanu na dzień 1 stycznia 2014 r. liczba ludności Polski wynosiła 38 495 659 osób. Takie same dane publikuje Eurostat w zakładce *Demography* (t\_pop) pod adresem: <http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&language=en&pcode=tps00001&plugin=1>.

Wielkość ta obejmuje wszystkie osoby, które w dniu 1 stycznia 2014 r. były stałymi mieszkańcami Polski. W liczbie tej ujmowane są również te osoby, które przebywają czasowo za granicą ale zachowały stałe zameldowanie w Polsce.

Dane przywoływane przez niektóre media i podawane jako szacunki Eurostat, upowszechnione w obszarze tematycznym *Labour Market* pod adresem Eurostat: ([http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsi\\_act\\_q&lang=en](http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsi_act_q&lang=en)), są faktycznie wynikami badań przekazywanymi przez GUS. Warto podkreślić, że wyniki te pochodzą z reprezentacyjnego Badania Aktywności Ekonomicznej Ludności (europejskie Labour Force Survey – LFS), poświęconego sytuacji na rynku pracy i pokazują liczbę ludności objętą BAEL, uzupełnioną danymi o ludności poniżej 15 roku życia (gromadzonymi wyłącznie na potrzeby wskaźnika dzieci (0–17 lat) w gospodarstwach domowych bez osób pracujących). Zgodnie z metodologią badanie jest prowadzone wyłącznie w wylosowanych gospodarstwach domowych. Prezentowana liczba ludności nie obejmuje osób bezdomnych, osób zamieszkałych w gospodarstwach zbiorowych tj. hotelach pracowniczych, domach studenckich, internatach, bursach, domach opieki społecznej itp. oraz członków gospodarstw domowych przebywających za granicą co najmniej 12 miesięcy. W związku z tym liczba ludności, wynikająca z uogólnienia wyników badania reprezentacyjnego, jest mniejsza od oficjalnych statystyk demograficznych obejmujących całą populację mieszkańców Polski. Dane prezentowane dla innych krajów we wspomnianej bazie Eurostat w zakładce *Labour Market* również wykazują znacznie niższe wartości niż liczby ludności dla tych krajów upowszechnione w zakładce *Demography*.

Jednocześnie informujemy, że w związku z wejściem w życie w grudniu 2013 r. Rozporządzenia PE i Rady nr 1260/2013 z dnia 20 listopada 2013 r. w sprawie statystyk europejskich w dziedzinie demografii, wszystkie kraje Unii Europejskiej są zobowiązane do przekazania danych o liczbie ludności rezydującej tj. uwzględniającej okres przebywania/nieobecności co najmniej 12 miesięcy. Zgodnie z rozporządzeniem dane te wg stanu na dzień 31 grudnia 2013 r. zostaną przekazane do Eurostatu (i upowszechnione na stronie GUS) do 31 grudnia 2014 r. Szczegółowy zakres danych zawiera Rozporządzenie Wykonawcze Komisji (UE) nr 205/2014 z dnia 4 marca 2014 r. ustanawiające jednolite warunki wykonania rozporządzenia PE i Rady (UE) nr 1260/2013 w sprawie statystyk europejskich w dziedzinie demografii odnośnie do podziałów danych, terminów przekazywania danych oraz zmian danych.

Dodatkowo informujemy, że kraje członkowskie UE – wypełniając zobowiązania wynikające z art. 4 wspomnianego Rozporządzenia nr 1260/2013 – przekazały w sierpniu br. na potrzeby ustalania większości kwalifikowanej (QMV) w Radzie UE liczbę ludności uwzględniającej kryterium 12 miesięcy. Szacowana liczba ludności Polski według stanu na dzień 1 stycznia 2014 r. wyniosła 38 018 tys. i została upowszechniona na stronie Eurostat pod adresem: [http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo\\_urespop&lang=en](http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_urespop&lang=en).

## **I.5. Definicje ludności w NSP 2011**

**Raport z wyników. Narodowy Spis Powszechny Ludności i Mieszkań 2011, GUS, Warszawa 2012 (GUS, 2012a, s. 32)**

Na podstawie wyników NSP<sup>2011</sup> zostały wyodrębnione trzy podstawowe kategorie ludności:

- 1) Ludność stała (stale zamieszkała)
- 2) Ludność faktycznie zamieszkała
- 3) Rezydenci (ludność rezydująca)

We wszystkich tablicach publikacji informacje dotyczą kategorii ludności faktycznie zamieszkałej – dalej określanej jako „ludność”.

### **Ludność stale zamieszkała (ludność stała)**

Kategoria obejmuje osoby mieszkające stale (z reguły zameldowane na pobyt stały) bez względu na to czy w momencie krytycznym spisu były obecne czy nieobecne w miejscu zamieszkania.

### **Ludność faktycznie zamieszkała (ludność faktyczna)**

Kategoria ta obejmuje następujące grupy osób:

1. **Mieszkające stale**, które:
  - a) były obecne w momencie spisu;
  - b) były nieobecne w momencie spisu, ale ich nieobecność trwała krócej niż 3 miesiące;
  - c) były nieobecne w momencie spisu przez okres dłuższy niż 3 miesiące, ale ich nieobecność wynikała z powodu:
    - przebywania w zakładzie karnym lub śledczym;
    - pobytu za granicą.
2. **Przebywające czasowo przez okres powyżej 3 miesięcy.**

Dotyczy to osób, które mieszkają na stałe w innym (są tam zameldowane na pobyt stały), natomiast w miejscu spisania przebywają czasowo z następujących powodów: nauka, praca, warunki rodzinne lub mieszkaniowe, leczenie lub rehabilitacja, przebywanie w domu opieki.

Jako czas okresowej nieobecności lub przebywania przyjęty został czas zamierzony. Do kategorii ludności faktycznej nie są zaliczane osoby przybyłe z zagranicy na pobyt czasowy, tj. te, które nie posiadają stałego pobytu w Polsce (pozwolenia na osiedlenie się).

### **Ludność rezydująca (rezydenci)**

Do rezydentów zalicza się:

- a) stałych mieszkańców, z wyjątkiem osób przebywających poza miejscem zamieszkania przez okres co najmniej 12 miesięcy – bez względu na ich miejsce



- przebywania (w kraju czy za granicą – wyjątek stanowią polscy dyplomaci na placówkach oraz żołnierze stacjonujący na zagranicznych misjach wojskowych);
- b) osoby mieszkające czasowo przez okres co najmniej 12 miesięcy, przybyłe z innego miejsca w kraju lub z zagranicy (cudzoziemcy bez stałego pobytu w Polsce).

### 1.6. Obszary tematyczne spisu, tematy i ocena stopnia zaspokojenia potrzeb, konsultacje społeczne przed spisem 2021 r.

Obszar tematyczny	Tematy/zmienne	Zaspokojenie potrzeb informacyjnych (%)		
		ogółem	organizacje	użytkownicy indywidualni
Stan i charakterystyka demograficzna ludności	płeć, wiek, miejsce zamieszkania, stan cywilny	89,0	91,3	85,7
Osoby w obiektach zbiorowego zamieszkania (oZZ) i osoby bezdomne (bez dachu nad głową)	rodzaj obiektu zbiorowego zakwaterowania, charakterystyka demograficzna ludności przebywającej w OZZ, osoby bezdomne wg cech demograficznych, okres pozostawania w bezdomności, przyczyna bezdomności	90,8	91,4	89,7
Edukacja	poziom wykształcenia, kontynuacja nauki	88,7	90,3	86,2
Aktywność ekonomiczna ludności	pracujący, status zatrudnienia pracujących, zawód wykonywany, rodzaj działalności miejsca pracy, bezrobotni, w tym długotrwale, bierni zawodowo, w tym według głównych przyczyn bierności, lokalizacja miejsca pracy, użytkowanie gospodarstwa rolnego, w tym pracujący w gospodarstwie rolnym, źródło utrzymania	90,4	91,7	88,4
Dojazdy do pracy	fakt dojazdu do pracy	82,1	84,7	78,7
Niepełnosprawność	subiektywna ocena stanu zdrowia (niepełnosprawność biologiczna), prawne potwierdzenie ograniczenia sprawności (niepełnosprawność prawna)	88,0	88,7	86,4
Kraj urodzenia i obywatelstwo	kraj urodzenia/miejsce urodzenia; kraj posiadanego obywatelstwa	92,9	94,4	90,4
Migracje wewnętrzne i zagraniczne	długookresowa imigracja na pobyt czasowy, krótkookresowa imigracja na pobyt czasowy, kraj poprzedniego zamieszkania, długookresowa emigracja na pobyt czasowy; krótkookresowa emigracja na pobyt czasowy; kraj wyjazdu, fakt zamieszkania kiedykolwiek za granicą, migracje wewnętrzne, miejsce zamieszkania rok przed spisem	92,0	93,3	89,7
Narodowość, język i wyznanie	przynależność narodowo-etniczna, język kontaktów domowych, język ojczysty, przynależność do wyznania religijnego			

Obszar tematyczny	Tematy/zmienne	Zaspokojenie potrzeb informacyjnych (%)		
		ogółem	organizacje	użytkownicy indywidualni
Gospodarstwa domowe i rodziny	wielkość gospodarstwa domowego (liczba osób w gospodarstwie domowym), typ gospodarstwa domowego (jednorodzinne, dwurodzinne, trzy i więcej rodzinne, nierodzinne: jednoosobowe, wieloosobowe), pozycja osoby w gospodarstwie domowym (osoby w gospodarstwach domowych: w rodzinach, nietworzące rodzin: mieszkające samotnie, z innymi osobami; osoby nietworzące gospodarstw domowych: w obiektach zbiorowego zakwaterowania, bezdomne), wielkość rodziny biologicznej (liczba osób w rodzinie), typ rodziny biologicznej (małżeństwa bez dzieci/z dziećmi, związki nieformalne bez dzieci/z dziećmi, samotni ojcowie z dziećmi, samotne matki z dziećmi), pozycja osoby w rodzinie (osoby pozostające w związkach: małżeńskich, nieformalnych; samotni rodzice; dzieci: żyjące z obojgiem rodziców, z jednym rodzicem; osoby nietworzące rodzin), tytuł prawny gospodarstwa domowego do zajmowanego mieszkania (ustalany na podstawie gospodarstw domowych zamieszkujących mieszkanie: własności mieszkania lub domu, spółdzielczego prawa do mieszkania, najmu, podnajmu, pokrewieństwa, innego)	94,1	94,4	93,7
Stan i charakterystyka zasobów mieszkaniowych (mieszkania i budynki)	adres zamieszkania, stan zamieszkania mieszkania, liczba osób zamieszkująca mieszkanie, powierzchnia użytkowa mieszkania, liczba izb w mieszkaniu, tytuł prawny do zajmowanego mieszkania, podmiot będący właścicielem mieszkania, wyposażenie mieszkania w instalacje sanitarno-techniczne, typ budynku, okres wybudowania budynku, podmiot będący właścicielem budynku, wyposażenie budynku w urządzenia techniczne	92,3	93,1	91,8

Źródło: Opracowano na podstawie (GUS, 2017b, s. 16–17).

### 1.7. Cechy zespołowe (1–3) oraz demograficzne (4–25) uwzględnione w poszczególnych spisach ludności w Polsce

Lp.	Cechy badane	Spis z roku								
		1921	1931	1950	1960	1970	1978	1988	2002	2011
1	Stosunek do reprezentanta/ki gospodarstwa <sup>d</sup> (pokrewieństwo z reprezentantem/ką rodziny)	×	×	×	×	×	×	×	×	× <sup>a</sup>
2	Podział na rodziny biologiczne	–	–	–	–	×	–	×	×	× <sup>a</sup>
3	Podstawowe źródło dochodu gospodarstwa domowego	–	–	–	–	–	×	×	×	× <sup>a</sup>

I. Informacje uzupełniające rozdział 1: Spis ludności jako źródło danych o ludności 197

cd. tabeli I.7

Lp.	Cechy badane	Spis z roku								
		1921	1931	1950	1960	1970	1978	1988	2002	2011
4	Miejsce zamieszkania, czasowa nieobecność i obecność	×	×	×	×	×	×	×	×	×
5	Płeć	×	×	×	×	×	×	×	×	×
6	Data urodzenia, wiek lub rok urodzenia	×	×	×	×	×	×	×	×	×
7	Stan cywilny	×	×	–	×	×	×	×	×	×
8	Stan małżeński	×	×	×	×	×	×	×	×	×
9	Miejsce urodzenia/kraj urodzenia	×	×	–	–	× <sup>a</sup>	× <sup>a</sup>	–	×	×
10	Poprzednie miejsce zamieszkania	–	–	×	×	×	× <sup>a</sup>	×	×	×
11	Czas zamieszkania w obecnym miejscu	–	–	–	–	×	× <sup>a</sup>	×	×	×
12	Obywatelstwo	×	×	–	–	–	–	–	×	×
13	Narodowość	×	–	–	–	–	–	–	×	×
14	Język	×	×	–	–	–	–	–	×	×
15	Wyznanie	×	×	–	–	–	–	–	–	× <sup>a</sup>
16	Analfabetyzm	×	×	×	×	×	×	–	–	–
17	Poziom wykształcenia	×	×	–	×	×	×	×	×	× <sup>a</sup>
18	Uczęszczanie do szkoły	–	–	×	×	×	×	×	×	× <sup>a,b</sup>
19	Sieroctwo	×	–	–	–	–	–	–	–	–
20	Inwalidztwo biologiczne/niepełnosprawność biologiczna	×	–	–	–	–	–	×	×	× <sup>a</sup>
21	Inwalidztwo prawne/niepełnosprawność prawna	–	–	–	–	–	×	×	×	× <sup>a</sup>
22	Dzietność kobiet	–	–	–	–	–	× <sup>a</sup>	× <sup>a</sup>	× <sup>a</sup>	× <sup>a</sup>
23	Przyczyna migracji	–	–	–	–	–	× <sup>a</sup>	–	×	× <sup>a</sup>
24	Zamierzenia co do przyszłych migracji	–	–	–	–	–	× <sup>a</sup>	–	×	× <sup>a</sup>
25	Charakterystyka dojazdów do pracy	–	–	–	–	–	×	–	×	× <sup>a,c</sup>

<sup>a</sup> Badanie metodą reprezentatywną.

<sup>b</sup> Badano fakt uczęszczania do szkoły zapewniającej podwyższenie posiadanego poziomu wykształcenia.

<sup>c</sup> Uwzględniono dojazdy do pracy.

<sup>d</sup> We wcześniejszych spisach posługiwano się określeniem „głowa gospodarstwa/rodziny”.

Źródło: (Krywuł-Albańska, 2012, s. 100–101).

### I.8. Cechy społeczno-zawodowe uwzględniane w powszechnych spisach ludności w Polsce

Lp.	Cechy badane	Spis z roku								
		1921	1931	1950	1960	1970	1978	1988	2002	2011
1	Główne i dodatkowe źródło utrzymania	–	–	–	–	×	×	×	×	×
Dla pracujących informacje dotyczące pracy głównej										
2	Adres i nazwa zakładu pracy / rodzaj działalności	× <sup>a</sup>	× <sup>a</sup>	× <sup>a</sup>	× <sup>a</sup>	× <sup>a</sup>	× <sup>a</sup>	× <sup>a</sup>	×	×
3	Forma gospodarki	–	–	×	×	– <sup>b</sup>	– <sup>b</sup>	– <sup>b</sup>	×	×

Lp.	Cechy badane	Spis z roku								
		1921	1931	1950	1960	1970	1978	1988	2002	2011
4	Charakter/status zatrudnienia	×	×	×	×	×	×	×	×	×
5	Zawód wykonywany / stanowisko	×	× <sup>c</sup>	× <sup>c</sup>	× <sup>c</sup>	×	×	×	×	×
6	Liczba pracowników najemnych (w prywatnych zakładach)	×	×	×	×	–	–	–	–	–
Dla pracujących informacje dotyczące pracy dodatkowej										
7	Nazwa zakładu pracy / rodzaj działalności	–	–	×	×	×	×	×	×	×
8	Charakter / status zatrudnienia	×	×	×	×	×	×	×	×	×
9	Zawód wykonywany	×	× <sup>d</sup>	–	× <sup>d</sup>	–	–	–	×	×
10	Stały lub sezonowy charakter pracy	–	–	–	×	×	– <sup>e</sup>	–	–	–
Dla bezrobotnych										
11	Poszukiwanie pracy	–	–	–	–	–	–	–	×	×
12	Gotowość do podjęcia pracy	–	–	–	–	–	–	–	×	×
13	Okres poszukiwania pracy	–	–	–	–	–	–	–	×	×
Dla mających niezarobkowe źródło utrzymania										
14	Rodzaj niezarobkowego źródła utrzymania	×	×	×	×	×	×	×	×	×
15	Grupa społeczno-ekonomiczna	–	–	–	–	×	×	–	–	–
16	Utrzymywanie przez rodzinę	×	×	×	×	×	×	×	×	×
17	Przyczyna bierności zawodowej	× <sup>f</sup>	–	–	–	–	×	–	×	×
Dla ludności związanej z rolnictwem										
18	Powierzchnia ogólna gospodarstwa rolnego	×	×	×	×	×	×	×	×	×
19	Wkład pracy w gospodarstwo rolne (liczba miesięcy pracy)	–	–	–	–	–	×	×	×	×
20	Rok podjęcia pracy poza gospodarstwem rolnym	–	–	–	–	–	×	–	–	–
21	Podstawowe źródło dochodów gospodarstwa rolnego	–	–	–	–	–	×	×	×	×

<sup>a</sup> Informacja pomocnicza do ustalenia działu i gospodarki narodowej.

<sup>b</sup> Ustala się na podstawie nazwy zakładu pracy.

<sup>c</sup> Zebranych informacji nie opracowano.

<sup>d</sup> Zebrano w celach pomocniczych do ustalenia działu gospodarki narodowej.

<sup>e</sup> Dla pracujących w rolnictwie indywidualnym informację tę uzyskuje się z pytania o wkład pracy w gospodarstwie rolnym.

<sup>f</sup> Dla bezrobotnych w zawodzie głównym zebrano informacje o czasie i przyczynie bezrobocia.

Źródło: (Krywult-Albańska, 2012, s. 102).

## I.9. Najczęściej zgłaszane dodatkowe potrzeby informacyjne i powiązania tematów, konsultacje społeczne przed spisem 2021 r.

Obszar tematyczny	Dodatkowe potrzeby informacyjne i powiązania z innymi tematami
Stan i charakterystyka demograficzna ludności	Stan cywilny rozszerzony o związki nieformalne (kohabitację), dzietność. Powiązania z tematami: niepełnosprawność, poziom wykształcenia, aktywność ekonomiczna, migracje, narodowość i wyznanie, cechy gospodarstwa domowego oraz typu mieszkań
Ludność w obiektach zbiorowego zamieszkania i osoby bezdomne	Miejsce zamieszkania, płeć, wiek. Powiązania z tematami: poziom wykształcenia, niepełnosprawność, korzystanie z usług zdrowotnych, źródło utrzymania, status zatrudnienia oraz czy osoba jest cudzoziemcem, wyposażenie oraz w urządzenia techniczne
Edukacja	Kształcenie ustawiczne, znajomość języków obcych, kwalifikacje na rynku pracy, typ uczelni, w której kontynuowana jest nauka (publiczna, prywatna) oraz monitorowanie losów absolwentów szkół, wykształcenie rodziców, wyniki egzaminów sprawdzających, poziom wykształcenia urzędników, pobieranie stypendiów, edukacja nieformalna
Aktywność ekonomiczna ludności	Aktywność ekonomiczna cudzoziemców; rodzaj umowy o pracę, szczegółowe grupy zawodów, w tym zawody pożądane dla pracodawców, rodzaj działalności, wielkość zarobków/dochodów, przyczyny braku aktywności ekonomicznej wąskich grup społecznych, np. usamodzielnianych wychowanków pieczy zastępczej, odbiorców pomocy społecznej; poziom zatrudnienia w szarej strefie; wymiar czasu pracy, pracujący w rolnictwie. Powiązania z tematami: poziom wykształcenia, niepełnosprawność, dojazdy do pracy, bezdomność, kraj urodzenia i obywatelstwa, typ rodziny, migracje wewnętrzne, wyposażenie budynku w urządzenia techniczne
Dojazdy do pracy	Czas dojazdu do pracy, środek lokomocji, godziny podróży, częstotliwość podróży, utrudnienia związane z infrastrukturą w podejmowaniu pracy na danym terenie, izochrony dojazdu do pracy, dojazd do pracy osoby niepełnosprawnej, dojazd do pracy za granicę, koszty dojazdu do pracy, dojazdy do szkół i uczelni, fakt dowozu/ odbioru dzieci z placówek edukacyjno-wychowawczych
Niepełnosprawność	Jakość życia osób niepełnosprawnych (także samoocena pod kątem zaspokojenia różnych potrzeb), charakterystyka społeczno-ekonomiczna osób niepełnosprawnych, korzystanie z usług zdrowotnych, realizacja praw osoby niepełnosprawnej, występowanie zachowań dyskryminacyjnych wobec osoby niepełnosprawnej, opiekunowie osób niepełnosprawnych (w tym osób starszych), warunki mieszkaniowe osób niepełnosprawnych, liczba osób niepełnosprawnych kierujących pojazdami uprawnionych do korzystania z miejsc postojowych dla poszczególnych dzielnic, osiedli, korzystanie przez osoby niepełnosprawne ze świadczeń społecznych, korzystanie z pomocy organizacji pozarządowych, uczestnictwo w kulturze, sporcie, bariery w codziennym funkcjonowaniu (techniczne, w komunikowaniu się itp.)
Kraj urodzenia i obywatelstwa	Rok przybycia do Polski (dotyczy repatriantów, reemigrantów, przymusowo przesiedlonych po zakończeniu drugiej wojny światowej), bardziej szczegółowe dane, np. sposób uzyskania obywatelstwa (małżeństwo, praca, status uchodźcy itp.), przyczyna migracji do Polski: chęć ubiegania się o polskie obywatelstwo, chęć podjęcia pracy, chęć podjęcia studiów, stypendium w ramach programów mobilności studentów (np. Erasmus), migracja za partnerem życiowym, poszukiwanie partnera życiowego, zamierzony czas przebywania, związki z krajem urodzenia, posiadanie wielokrotnego obywatelstwa wraz z informacją, jakie są to obywatelstwa. Powiązania z tematami: warunki mieszkaniowe, tytułu prawnego do mieszkania itp., tożsamość narodowa, status takiej osoby na rynku pracy, edukacja w kontekście zachowania własnej tożsamości narodowej

Obszar tematyczny	Dodatkowe potrzeby informacyjne i powiązania z innymi tematami
Migracje wewnętrzne i zagraniczne	Przyczyny migracji, migracje powrotne, rodziny emigrantów, przyczyny wyjazdu (główny powód i dodatkowy), czy planuje stałe zamieszkanie za granicą lub w innym regionie Polski, w tym sam czy z rodziną; czy emigracja była powodem rozpadu rodziny/rozvodu bądź więzi rodzinnych, bezdomność emigrantów, migracje wahańlowe, migracje w ramach aglomeracji (miasto – okolica poza miastem, kierunki). Powiązania z innymi tematami: poziom wykształcenia, źródła utrzymania migrantów, warunki życia i pracy imigrantów, warunki mieszkaniowe imigrantów, wpływ warunków mieszkaniowych na migracje wewnętrzne, status zatrudnienia migrantów przed migracją, zgodność pracy podejmowanej za granicą/w Polsce z wykształceniem i kwalifikacjami
Narodowość, język, wyznanie	
Gospodarstwa domowe i rodziny	Połączenie informacji o typach rodzin z informacjami na temat dzietności, samotne wychowywanie dzieci, niepełnosprawność członków gospodarstwa domowego/rodziny pozostawanie w związkach kohabitacyjnych, rodziny zrekonstruowane, w ramach typu rodziny biologicznej powinny pojawić się kategorie: związek osób różnej płci oraz związek osób tej samej płci (obydwa w podziale z dziećmi/bez dzieci). opieka nad dziećmi, opieka nad osobami starszymi. Powiązania z tematami: źródło utrzymania gospodarstwa domowego, dochody/wydatki gospodarstw domowych, charakterystyka mieszkaniowa (wielkość mieszkania, wyposażenie, dostęp do internetu), ogrzewanie, segregowanie śmieci i korzystanie z usług zdrowotnych
Stan i charakterystyka zasobów mieszkaniowych	Sposób ogrzewania mieszkania i rodzaj stosowanego paliwa, gospodarowanie komunalnymi zasobami mieszkaniowymi, dostęp do internetu, dostosowanie budynków do potrzeb osób starszych i niepełnosprawnych, wyposażenie mieszkań w dobra ruchome, np. sprzęt AGD

Źródło: (GUS, 2017, s. 18–19).

## I.10. Definicja jakości w statystyce publicznej

Definicja jakości – według normy ISO 8402–1986 jakość jest zdefiniowana jako: „ogół cech i właściwości produktu lub usługi, który decyduje o zdolności zaspokojenia potrzeb zadeklarowanych lub domyślnych”. Jakość w statystyce publicznej jest oparta na definicji jakości Europejskiego Systemu Statystycznego i określona na podstawie sześciu komponentów jakości, które obejmują:

- I. Przydatność.
- II. Dokładność.
- III. Terminowość i punktualność.
- IV. Dostępność i przejrzystość.
- V. Porównywalność.
- VI. Spójność.

Przy ocenie spełnienia zaleceń dotyczących jakości należy również uwzględnić koszty i obciążenia, jakie są związane z tworzeniem statystyk, oraz zagadnienia poufności, transparentności i bezpieczeństwa danych.

- I. Przydatność jest to stopień, w jakim statystyka spełnia obecne i potencjalne wymagania użytkowników. Mówi ona, czy wszystkie potrzebne statystyki zostały sporządzone oraz w jakim stopniu zastosowane pojęcia (definicje, klasyfikacje itp.) odzwierciedlają potrzeby użytkownika (np. osób fizycznych, ministerstw i urzędów centralnych, przedsiębiorców, naukowców itp.).
- II. Dokładność w ogólnym sensie statystycznym określa bliskość obliczeń lub oszacowań do wartości dokładnych lub rzeczywistych.
- III. Terminowość informacji odzwierciedla czas pomiędzy jej dostępnością a wydarzeniem lub zjawiskiem, jakie opisuje. Punktualność jest związana z opóźnieniem czasowym pomiędzy rzeczywistą datą dostarczenia danych oraz planowaną datą dostarczenia danych – na przykład, w związku z datami ogłoszonymi w oficjalnym kalendarzu publikacji, ustanowionymi w zarządzeniach lub uzgodnionymi wcześniej z klientami.
- IV. Dostępność odnosi się do fizycznych warunków, w jakich użytkownicy mogą uzyskiwać dane i obejmuje: kanały dystrybucji, procedury składania zamówień, czas potrzebny na realizację dostawy, politykę cenową, dogodne warunki marketingowe (prawa autorskie itp.), dostępność mikro – i makrodanych, media (wydruk, pliki, CD-ROM, internet) itp. Przejrzystość odnosi się do otoczenia informacyjnego danych, czyli faktu łączenia ich z odpowiednimi metadanymi (informacje tekstowe, wyjaśnienia, dokumentacja itp.) i formami graficznymi (wykresy, diagramy i mapy), faktu dostępności informacji dotyczących jakości danych (łącznie z ograniczeniami wykorzystania) oraz zakresu pomocy dodatkowej oferowanej przez instytucję statystyczną.
- V. Kryterium porównywalności ma na celu pomiar wpływu różnic w stosowanych pojęciach i definicjach statystycznych przy porównywaniu statystyk z różnych obszarów geograficznych, dziedzin pozageograficznych lub okresów bazowych. Można powiedzieć, że jest to zakres, w jakim różnice pomiędzy statystykami mogą zostać przypisane różnicom pomiędzy rzeczywistymi wartościami ujętymi w danej statystyce. Istnieją trzy główne rodzaje porównywalności statystyk:
  - porównywalność czasowa,
  - porównywalność geograficzna,
  - porównywalność dziedzin.Porównywalność czasowa jest związana z porównaniem wyników uzyskanych w normalnym trybie z analogicznych operacji statystycznych przeprowadzonych w różnym czasie. Aspekt geograficzny porównywalności kładzie nacisk na porównanie statystyk pomiędzy krajami i (lub) regionami w celu zapewnienia jednolitości zbiorczych statystyk na danym poziomie (międzynarodowym, europejskim, regionalnym). Porównywalność pomiędzy dziedzinami odnosi się do dziedzin pozageograficznych, na przykład sektorów przemysłowych, różnych typów gospodarstw domowych itp.

- VI. Spójność statystyki jest to zdolność statystyki do wiarygodnej zmiany jej konfiguracji na różne sposoby i dla różnych zastosowań. Niemniej łatwiej jest wskazać przypadki niespójności, niż wykazać spójność. Statystyki powstałe na bazie pojedynczego źródła są zazwyczaj spójne w takim sensie, że podstawowe wyniki uzyskane z rozważanego badania mogą podlegać wielu wiarygodnym kombinacjom w celu uzyskania wyników bardziej kompleksowych. Statystyki powstałe na bazie różnych źródeł, a zwłaszcza na podstawie badań statystycznych o odmiennym charakterze i (lub) częstotliwości, mogą nie być całkowicie spójne z tego względu, że u ich podstaw mogą leżeć różne podejścia, klasyfikacje i normy metodologiczne. Powyższe statystyki mogą dostarczać niezupełnie spójnego przekazu; użytkownicy powinni zostać jednoznacznie poinformowani o możliwych skutkach.

(Definicja jakości w statystyce prezentowana powyżej została opracowana na podstawie dokumentu Eurostat 4.2. – dokumenty metodologiczne – definicja jakości w statystyce dla potrzeb grupy roboczej „Ocena jakości w statystyce”, spotkanie szóste, Luksemburg, 2–3 października 2003).

#### Jakość w statystyce

Podmiot publikujący: Główny Urząd Statystyczny

Autor informacji: Jolanta Szutkowska (Departament Metodologii, Standardów i Rejestrów GUS)

Publikujący informację: Alicja Koszela (Departament Informacji GUS)

Wytworzenie informacji: 16.08.2016

Biuletyn Informacji Publicznej GUS: <http://bip.stat.gov.pl/dzialalnosc-statystyki-publicznej/jakosc-w-statystyce/definicja-jakosci-w-statystyce-publicznej/#>.



## II. Informacje uzupełniające rozdział II: Transformacja metod przeprowadzania spisów

### II.1. Rejestry administracyjne i systemy informacyjne wykorzystane w NSP 2011

Lp.	Nazwa systemu	Gestor
1	Ewidencja gruntów i budynków	Główny Urząd Geodezji i Kartografii
2	Geodezyjna ewidencja sieci uzbrojenia terenu	
3	Państwowy rejestr granic i powierzchni jednostek podziału terytorialnego państwa	
4	Rejestr cen i wartości nieruchomości	
5	System informacyjny KRUS	Kasa Rolniczego Ubezpieczenia Społecznego
6	Baza danych o podatnikach podatku dochodowego od osób fizycznych (PIT)	Ministerstwo Finansów
7	Krajowa Ewidencja Podatników (KEP)	
8	POLTAX – Bazy urzędów skarbowych	
9	Krajowy System Monitoringu Rynku Pracy ( <i>dawniej SI PULS</i> )	Ministerstwo Pracy i Polityki Społecznej
10	Krajowy System Pomocy Społecznej ( <i>dawniej SI POMOST</i> )	
11	Krajowy System Świadczeń Rodzinnych	
12	Powszechny Elektroniczny System Ewidencji Ludności PESEL – CBD	Ministerstwo Spraw Wewnętrznych i Administracji
13	PESEL2	
14	Rejestr gmin, na których obszarze są używane nazwy w języku mniejszości	
15	Urzędowy rejestr gmin, w których jest używany język pomocniczy	
16	Rejestry nieruchomości, udziałów i akcji nabytych lub objętych przez cudzoziemców	
17	Centralny Wykaz Ubezpieczonych	Narodowy Fundusz Zdrowia
18	System informacyjny PFRON	Państwowy Fundusz Rehabilitacji Osób Niepełnosprawnych
19	System Pobyt	Urząd d/s Cudzoziemców
20	Centralny Rejestr Ubezpieczonych	Zakład Ubezpieczeń Społecznych
21	Centralny Rejestr Członków Rodziny Ubezpieczonych Uprawnionych do Ubezpieczenia Zdrowotnego	
22	Centralny Rejestr Płatników Składek	
23	System Informacji Oświatowej	Ministerstwo Edukacji Narodowej
24	Zbiory danych jednostkowych w zakresie lokali użytkowników nie mających charakteru handlowego, usługowego lub produkcyjnego, wykorzystujących energię elektryczną na potrzeby mieszkaniowe	Jednostki zakładów energetycznych zajmujące się sprzedażą oraz świadczeniem usług przesyłowych energii elektrycznej

Lp.	Nazwa systemu	Gestor
25	<b>Ewidencja podatkowa nieruchomości</b>	urzędy gmin
26	<b>Gminne zbiory meldunkowe</b>	
27	Zbiory danych dotyczących przyrostów i ubytków budynków mieszkalnych, niemieszkalnych oraz obiektów zbiorowego zakwaterowania	wojewódzkie i powiatowe inspektoraty nadzoru budowlanego
28	Zbiory danych dotyczących pozwoleń na budowę obiektów budowlanych	organy administracji architektoniczno-budowlanej

Źródło: (Dygaszewicz 2009).

Na stronie internetowej GUS<sup>171</sup> zamieszczona jest poniższa informacja dotycząca danych pozyskanych z systemów informacyjnych administracji publicznej i systemów pozaadministracyjnych:

Podstawę prawną do pozyskania danych z systemów informacyjnych administracji publicznej i systemów pozaadministracyjnych stanowi Rozporządzenie Rady Ministrów z dnia 27 listopada 2008 r. w sprawie programu badań statystycznych statystyki publicznej na rok 2009 (Dz.U. Nr 221, poz. 1436), a po wejściu w życie ustawa o powszechnym spisie rolnym 2010 r.

Do spisu próbnego zostały zebrane dane z wybranych systemów informacyjnych administracji publicznej i systemów pozaadministracyjnych w zakresie, formie i terminach określonych w badaniu 1.45.29. „Spis próbny 2009 oraz prace przygotowawcze do powszechnego spisu rolnego 2010” programu badań statystycznych statystyki publicznej na rok 2009, a po wejściu w życie, w ustawie o powszechnym spisie rolnym 2010 r.

Zbiory danych administracyjnych, o których mowa w pkt 1, zostały pozyskane z następujących systemów centralnych: Ministerstwa Finansów/urzędów skarbowych – systemu informacyjnego prowadzonego przez organy podatkowe; Ministerstwa Spraw Wewnętrznych i Administracji – zbioru PESEL i centralnej ewidencji pojazdów CEP; Agencji Restrukturyzacji i Modernizacji Rolnictwa; Kasy Rolniczego Ubezpieczenia Społecznego; Narodowego Funduszu Zdrowia – Centralnego Wykazu Ubezpieczonych; Inspekcji Jakości Handlowej Artykułów Rolno-Spożywczych; Głównego Urzędu Geodezji i Kartografii – Państwowego rejestru granic oraz powierzchni jednostek podziału terytorialnego kraju i ortofotomapy dla obszaru Polski; oraz rozproszonych: urzędów marszałkowskich – Bazy Danych Obiektów Topograficznych; starostw powiatowych – ewidencji gruntów i budynków i systemów informacyjnych PUP; urzędów gmin – ewidencji podatkowej nieruchomości EPN i gminnych zbiorów meldunkowych.

Zbiory danych pozaadministracyjnych, o których mowa w pkt 1, zostały pozyskane z następujących systemów: przedsiębiorców wykonujących działalność

<sup>171</sup> <http://stat.gov.pl/spisy-powszechne/nsp-2011/pozyskanie-danych-z-innych-systemow/>.

gospodarczą w zakresie sprzedaży energii elektrycznej – systemów informacyjnych dotyczących odbiorców energii elektrycznej wykorzystujących energię elektryczną na potrzeby mieszkaniowe; dostawców publicznie dostępnych usług telekomunikacyjnych – systemów informacyjnych dotyczących abonentów.

## **II.2. Zakres tematyczny formularza do badania reprezentacyjnego w NSP 2011**

### **LUDNOŚĆ – Charakterystyka demograficzno-społeczna**

1. Płeć
2. Data urodzenia
3. Charakter przebywania/nieobecności
4. Okres przebywania/nieobecności
5. Zamierzony czas przebywania/nieobecności
6. Przyczyna czasowego przebywania lub czasowej nieobecności
7. Stan cywilny prawny
8. Data zawarcia/rozwiązania związku małżeńskiego
9. Data powstania aktualnego związku nieformalnego

### **Edukacja (dotyczy osób w wieku 13 lat i więcej)**

1. Poziom wykształcenia
2. Dziedzina wykształcenia
3. Kierunek wykształcenia
4. Zawód wyuczony
5. Kontynuacja nauki
6. Rodzaj szkoły/studiów, w której nauka jest kontynuowana.

### **Aktywność ekonomiczna (dotyczy osób w wieku 15 lat i więcej)**

1. Samoocena statusu na rynku pracy
2. Charakter zatrudnienia
3. Zawód wykonywany
4. Główne i dodatkowe miejsce pracy
5. Rodzaj działalności
6. Wymiar czasu pracy
7. Wykonywanie pracy kiedykolwiek
8. Rok zakończenia pracy
9. Przyczyna zaprzestania pracy
10. Okres pozostawania bez pracy
11. Gotowość podjęcia pracy
12. Użytkowanie gospodarstwa rolnego lub bycie członkiem gospodarstwa domowego z użytkownikiem
13. Powierzchnia użytków rolnych
14. Liczbę miesięcy pracy w swoim gospodarstwie rolnym w ciągu ostatnich 12 miesięcy.

### **Dojazdy do pracy**

1. Środek transportu i czas dojazdu do pracy
2. Dojazdy do miejsca nauki
3. Odległość od miejsca pracy lub miejsca nauki
4. Częstotliwość dojazdów do pracy.

### **Źródła utrzymania**

1. Rodzaj głównego i dodatkowego źródła utrzymania.

### **Niepełnosprawność (*pytania dobrowolne*)**

1. Samoocena posiadania ograniczonej zdolność wykonywania zwykłych czynności
2. Stopień ograniczenia zdolności w wykonywaniu zwykłych czynności
3. Okres ograniczenia zdolności w wykonywaniu zwykłych czynności?
4. Grupy schorzeń powodujących odczuwanie ograniczenia zdolności do wykonywania
5. zwykłych czynności
6. Posiadanie aktualnego orzeczenia o stopniu niepełnosprawności, niezdolności do pracy.

### **Plany prokreacyjne (*Pytania dobrowolne zadawane kobietom wieku 16–49 lat*)**

1. Fakt urodzenia dziecka/ci
2. Liczba urodzonych dzieci
3. Liczba urodzonych dzieci przez matkę kobiety
4. Zamierzenia odnośnie urodzenia dziecka/ci.

### **Kraj urodzenia i obywatelstwo**

1. Kraj urodzenia
2. Kraj urodzenia ojca i matki
3. Obywatelstwo
4. Obywatelstwo polskie od urodzenia.

### **Migracje wewnętrzne i zagraniczne**

1. Zamieszkiwanie od urodzenia w miejscowości aktualnego zamieszkania
2. Okres przybycia lub powrotu do miejscowości aktualnego zamieszkania
3. Miejsce poprzedniego zamieszkania (co najmniej rok)?
4. Przyczyna przybycia (powrotu) do miejscowości aktualnego zamieszkania.
5. Fakt kiedykolwiek przebywania za granicą przez co najmniej rok
6. Okres przebywania za granicą
7. Przyczyny migracji zagranicznych
8. Fakt wykonywania pracy za granicą
9. Zgodność wykonywanej pracy z kwalifikacjami
10. Zawód wykonywany przed wyjazdem za granicę
11. Fakt przebywania w Polsce po raz pierwszy
12. Znajomość języka polskiego przez cudzoziemców
13. Plany migracyjne.

### **Narodowość i język, wyznanie**

1. Narodowość
2. Przynależność do innego narodu lub wspólnoty etnicznej
3. Język używany w domu
4. Język ojczysty
5. Przynależność do wyznania religijnego (kościół lub związku wyznaniowego).

### **GOSPODARSTWA DOMOWE**

1. Stopień pokrewieństwa, powinowactwa lub innego związku z osobą uznaną za Reprezentanta
2. Główne i dodatkowe źródło utrzymania gospodarstwa domowego
3. Tytułu prawny do zajmowanego mieszkania.

### **BUDYNKI I MIESZKANIA**

#### **Mieszkania**

1. Mieszkanie indywidualne czy w budynku wielomieszkaniowym
2. Mieszkanie zamieszkane stale lub czasowo
3. Własność mieszkania
4. Powierzchnia użytkowa mieszkania
5. Powierzchnia użytkowa mieszkania wykorzystywana wyłącznie do prowadzenia działalności gospodarczej
6. Liczba pokoi w mieszkaniu
7. Rodzaj kuchni w mieszkaniu
8. Wyposażenie mieszkania w instalacje:
  - kran z wodą bieżącą
  - łazienka z wanną lub kabiną prysznicową
  - ustęp spłukiwany wodą bieżącą
  - gaz
9. Sposób ogrzewania mieszkania
10. Rodzaj stosowanego paliwa.

### II.3. Sposób wykorzystania centralnego rejestru ludności w statystyce ludności, wybrane państwa europejskie

Country	Statistical Office contributes to population register improvement	Frequency of data transfer	Statistical Office receives individual data	Type of data transferred	First year individual data concerning population stock were transferred	First year individual migration data were transferred
Austria	Yes	Quarterly	Yes	Identifiable	2002	2002
Belgium	No	Daily	Yes	Identifiable	1989	1989
Bulgaria	Yes	Monthly	Yes	Identifiable	–	1980
Czech Republic	No	–		Aggregated	–	–
Denmark	Yes	Weekly	Yes	Identifiable	1971	1973
Estonia	Yes	Monthly	Yes	Identifiable	2002	2001
Finland	Yes	Weekly	Yes	Identifiable	1975	1975
Hungary	No	Monthly	Yes	Anonymized	–	1984
Iceland	Yes	Other	Yes	Identifiable	1953	1953
Latvia	Yes	Monthly	Yes	Identifiable	2000	2002
Lithuania	Yes	Daily	Yes	Identifiable	2001	2001
Luxembourg	Yes	Monthly	Yes	Anonymized	1979	1987
Netherlands		Other	Yes	Aggregated	1995	1995
Norway	Yes	Daily	Yes	Identifiable	1996	1996
Poland	No	Quarterly	Yes	Identifiable	–	1990
Romania	No	Monthly	Yes	Identifiable	–	1996
Slovakia	No	–		–	–	–
Slovenia	Yes	Quarterly	Yes	Identifiable	1985	1985
Spain	Yes	Monthly	Yes	Identifiable	1996	1996
Sweden	Yes	Daily	Yes	Identifiable	1968	1968

Note: Statistical data collection from central population registers changes very quickly. The information contained in our survey may not reflect changes introduced since 2010.

Źródło: (Poulain i Herm, 2013, s. 198).

### III. Informacje uzupełniające rozdział III: Jakość spisów

#### III.1. Ocena kompletności NSP 2011: Rozbieżności szacunków bilansu ludności generacji 1916 i wcześniejszych na podstawie danych ewidencji ruchu naturalnego i stanów ludności według roczników urodzenia

Rok kalendarzowy	Wyszczególnienie	Ogółem	Mężczyźni	Kobiety
2001	Liczba osób w wieku 85 i więcej lat (urodzonych w 1916 r. i wcześniej) na 31.12.2001	339 778	95 265	244 513
2002	Liczba zgonów w 2002 r. wśród osób urodzonych w 1916 r. i wcześniej wg roczników urodzenia	59 448	17 156	42 292
	Szacunek bilansu liczby osób urodzonych w 1916 r. i wcześniej na 31.12.2002	280 330	78 109	202 221
	Liczba osób urodzonych w 1916 r. i wcześniej na 31.12.2002	281 489	71 970	209 519
	<b>Różnica</b>	<b>1 159</b>	<b>-6 139</b>	<b>7 298</b>
	Liczba osób urodzonych w 1903 r. i wcześniej na 31.12.2002	2798	602	2 196
2003	Liczba zgonów w 2003 r. wśród osób urodzonych w 1903 r. i wcześniej wg roczników urodzenia	898	185	713
	Szacunek bilansu liczby osób urodzonych w 1903 r. i wcześniej na 31.12.2003	1 900	417	1 483
	Liczba osób urodzonych w 1903 r. i wcześniej na 31.12.2003	1 890	413	1 477
	Różnica	10	4	6
2004	Liczba zgonów w 2004 r. wśród osób urodzonych w 1904 r. i wcześniej wg roczników urodzenia	980	191	789
	Szacunek bilansu liczby osób urodzonych w 1904 r. i wcześniej na 31.12.2004	2 208	491	17 17
	Liczba osób urodzonych w 1904 r. i wcześniej na 31.12.2004	2 207	490	1 717
	Różnica	1	1	0
2005	Liczba zgonów w 2005 r. wśród osób urodzonych w 1905 r. i wcześniej wg roczników urodzenia	1 074	199	875
	Szacunek bilansu liczby osób urodzonych w 1905 r. i wcześniej na 31.12.2005	2457	544	1 913
	Liczba osób urodzonych w 1905 r. i wcześniej na 31.12.2005	2454	544	1 910
	Różnica	3	0	3
2006	Liczba zgonów w 2006 r. wśród osób urodzonych w 1906 r. i wcześniej wg roczników urodzenia	1 048	207	841
	Szacunek bilansu liczby osób urodzonych w 1906 r. i wcześniej na 31.12.2006	2 925	681	2 244
	Liczba osób urodzonych w 1906 r. i wcześniej na 31.12.2006	2 919	680	2 239
	Różnica	6	1	5
2007	Liczba zgonów w 2007 r. wśród osób urodzonych w 1907 r. i wcześniej wg roczników urodzenia	1 257	292	965

Rok kalen-darzowy	Wyszczególnienie	Ogółem	Mężczyźni	Kobiety
	Szacunek bilansu liczby osób urodzonych w 1907 r. i wcześniej na 31.12.2007	3 335	790	2 545
	Liczba osób urodzonych w 1907 r. i wcześniej na 31.12.2007	3 333	790	2 543
	Różnica	2	0	2
2008	Liczba zgonów w 2008 r. wśród osób urodzonych w 1908 r. i wcześniej wg roczników urodzenia	1 281	275	1 006
	Szacunek bilansu liczby osób urodzonych w 1908 r. i wcześniej na 31.12.2008	3 700	885	2 815
	Liczba osób urodzonych w 1908 r. i wcześniej na 31.12.2008	3 694	883	2 811
	Różnica	6	2	4
2009	Liczba zgonów w 2009 r. wśród osób urodzonych w 1909 r. i wcześniej wg roczników urodzenia	1 488	305	1 183
	Szacunek bilansu liczby osób urodzonych w 1909 r. i wcześniej na 31.12.2009	3 766	923	2 843
	Liczba osób urodzonych w 1909 r. i wcześniej na 31.12.2009	3 765	923	2 842
	Różnica	1	0	1
2010	Liczba zgonów w 2010 r. wśród osób urodzonych w 1910 r. i wcześniej wg roczników urodzenia	1 606	326	1 280
	Szacunek bilansu liczby osób urodzonych w 1910 r. i wcześniej na 31.12.2010	4 419	1 115	3 304
	Liczba osób urodzonych w 1910 r. i wcześniej na 31.12.2010	3 088	678	2 410
	<b>Różnica</b>	<b>1331</b>	<b>437</b>	<b>894</b>
2011	Liczba zgonów w 2011 r. wśród osób urodzonych w 1911 r. i wcześniej wg roczników urodzenia	1 553	328	1 225
	Szacunek bilansu liczby osób urodzonych w 1911 r. i wcześniej na 31.12.2011	3 440	715	2 725
	Liczba osób urodzonych w 1911 r. i wcześniej na 31.12.2011	3 437	715	2 722
	Różnica	3	0	3
2012	Liczba zgonów w 2012 r. wśród osób urodzonych w 1912 r. i wcześniej wg roczników urodzenia	1 760	370	1 390
	Szacunek bilansu liczby osób urodzonych w 1912 r. i wcześniej na 31.12.2012	3 790	721	3 069
	Liczba osób urodzonych w 1912 r. i wcześniej na 31.12.2012	3 786	720	3 066
	Różnica	4	1	3
2013	Liczba zgonów w 2013 r. wśród osób urodzonych w 1913 r. i wcześniej wg roczników urodzenia	1 815	351	1 464
	Szacunek bilansu liczby osób urodzonych w 1913 r. i wcześniej na 31.12.2013	4 244	845	3 399
	Liczba osób urodzonych w 1913 r. i wcześniej na 31.12.2013	4 242	844	3 398
	Różnica	2	1	1
2014	Liczba zgonów w 2014 r. wśród osób urodzonych w 1914 r. i wcześniej wg roczników urodzenia	1 850	352	1 498



III. Informacje uzupełniające rozdział III: Jakość spisów 211

cd. tabeli III.1

Rok kalen- darzowy	Wyszczególnienie	Ogółem	Mężczyźni	Kobiety
	Szacunek bilansu liczby osób urodzonych w 1914 r. i wcześniej na 31.12.2014	4 836	964	3 872
	Liczba osób urodzonych w 1914 r. i wcześniej na 31.12.2014	4 828	963	3 865
	Różnica	8	1	7
2015	Liczba zgonów w 2015 r. wśród osób urodzonych w 1915 r. i wcześniej wg roczników urodzenia	1 966	353	1 613
	Szacunek bilansu liczby osób urodzonych w 1915 r. i wcześniej na 31.12.2015	4 763	945	3 818
	Liczba osób urodzonych w 1915 r. i wcześniej na 31.12.2015	4 755	944	3 811
	Różnica	8	1	7
2016	Liczba zgonów w 2016 r. wśród osób urodzonych w 1916 r. i wcześniej wg roczników urodzenia	1 724	327	1 397
	Szacunek bilansu liczby osób urodzonych w 1916 r. i wcześniej na 31.12.2016	4 933	962	3 971
	Liczba osób urodzonych w 1916 r. i wcześniej na 31.12.2016	4 933	961	3 972
	Różnica	0	1	-1

Źródło: Szacunek własny na podstawie baz danych GUS: Demografia, Bank Danych Lokalnych i zasobów danych obszarów tematycznych: spisy powszechne.

### III.2. Ocena kompletności NSP 2011: Liczba osób starszych według Bazy Danych Demografia GUS oraz szacunku bilansu na podstawie danych ewidencji zgonów według roku urodzenia w latach 2006–2016

Rok urodz. <sup>a</sup>	Liczba osób wg roczników urodzenia, 31.12.2006 r.			Liczba zgonów wg roczników urodzenia w 2007 r.			Szacunek stanu wg roczników urodzenia, 31.12.2007 r.		
1911	7927	1776	6151	2157	488	1669	5770	1288	4482
1910	6023	1327	4696	1601	349	1252	4422	978	3444
1909	3539	763	2776	1158	237	921	2381	526	1855
1908	2381	538	1843	732	168	564	1649	370	1279
1907	1673	402	1271	511	125	386	1162	277	885
1906 –	2919	680	2239	746	167	579	2173	513	1660
Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2007 r.			Liczba zgonów wg roczników urodzenia w 2008 r.			Szacunek stanu wg roczników urodzenia, 31.12.2008 r.		
1912	8907	1921	6986	2345	514	1831	6562	1407	5155
1911	5768	1288	4480	1638	392	1246	4130	896	3234
1910	4420	978	3442	1246	269	977	3174	709	2465
1909	2381	527	1854	818	182	636	1563	345	1218
1908	1648	370	1278	455	101	354	1193	269	924
1907 –	3333	790	2543	826	174	652	2507	616	1891
Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2008 r.			Liczba zgonów wg roczników urodzenia w 2009 r.			Szacunek stanu ludności wg roczników urodzenia 31.12.2009 r.		
1913	8520	1956	6564	2347	567	1780	6173	1389	4784
1912	6559	1407	5152	1812	400	1412	4747	1007	3740
1911	4126	895	3231	1241	275	966	2885	620	2265
1910	3173	708	2465	913	190	723	2260	518	1742
1909	1560	345	1215	561	116	445	999	229	770
1908 –	3694	883	2811	927	189	738	2767	694	2073
Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2009 r.			Liczba zgonów wg roczników urodzenia w 2010 r.			Szacunek stanu ludności wg roczników urodzenia 31.12.2010 r.		
1914	8936	1979	6957	2434	555	1879	6502	1424	5078
1913	6174	1389	4785	1823	386	1437	4351	1003	3348
1912	4747	1007	3740	1345	265	1080	3402	742	2660
1911	2885	620	2265	882	175	707	2003	445	1558
1910	2260	518	1742	609	126	483	1651	392	1259
1909 –	3765	923	2842	997	200	797	2768	723	2045

III. Informacje uzupełniające rozdział III: Jakość spisów 213

cd. tabeli III.2

Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2010 r.			Liczba zgonów wg roczników urodzenia w 2011 r.			Szacunek stanu ludności wg roczników urodzenia 31.12.2011 r.		
1915	7192	1460	5732	1873	409	1464	5319	1051	4268
1914	6530	1344	5186	1762	368	1394	4768	976	3792
1913	4491	979	3512	1275	303	972	3216	676	2540
1912	3052	575	2477	937	199	738	2115	376	1739
1911	1905	365	1540	555	121	434	1350	244	1106
1910 –	3088	678	2410	998	207	791	2090	471	1619
Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2011 r.			Liczba zgonów wg roczników urodzenia w 2012 r.			Szacunek stanu ludności wg roczników urodzenia 31.12.2012 r.		
1916	6905	1450	5455	1782	391	1391	5123	1059	4064
1915	5320	1051	4269	1488	334	1154	3832	717	3115
1914	4767	975	3792	1382	317	1065	3385	658	2727
1913	3215	676	2539	941	199	742	2274	477	1797
1912	2113	376	1737	624	134	490	1489	242	1247
1911 –	3437	715	2722	1136	236	900	2301	479	1822
Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2012 r.			Liczba zgonów wg roczników urodzenia w 2013 r.			Szacunek stanu ludności wg roczników urodzenia 31.12.2013 r.		
1917	6533	1396	5137	1712	409	1303	4821	987	3834
1916	5120	1059	4061	1430	314	1116	3690	745	2945
1915	3830	717	3113	1125	208	917	2705	509	2196
1914	3385	658	2727	941	186	755	2444	472	1972
1913	2273	476	1797	670	149	521	1603	327	1276
1912 –	3786	720	3066	1145	202	943	2641	518	2123
Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2013 r.			Liczba zgonów wg roczników urodzenia w 2014 r.			Szacunek stanu ludności wg roczników urodzenia 31.12.2014 r.		
1918	6599	1395	5204	1624	377	1247	4975	1018	3957
1917	4820	987	3833	1312	288	1024	3508	699	2809
1916	3690	745	2945	991	237	754	2699	508	2191
1915	2705	509	2196	800	173	627	1905	336	1569
1914	2444	472	1972	682	132	550	1762	340	1422
1913 –	4242	844	3398	1168	220	948	3074	624	2450

Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2014 r.			Liczba zgonów wg roczników urodzenia w 2015 r.			Szacunek stanu ludności wg roczników urodzenia 31.12.2015 r.		
1919	9775	2164	7611	2630	622	2008	7145	1542	5603
1918	4975	1018	3957	1386	330	1056	3589	688	2901
1917	3505	698	2807	1002	203	799	2503	495	2008
1916	2699	508	2191	793	162	631	1906	346	1560
1915	1901	335	1566	576	93	483	1325	242	1083
1914 –	4828	963	3865	1390	260	1130	3438	703	2735
Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2015 r.			Liczba zgonów wg roczników urodzenia w 2016 r.			Szacunek stanu ludności wg roczników urodzenia 31.12.2016 r.		
1920	11209	2564	8645	2771	661	2110	8438	1903	6535
1919	7145	1542	5603	1938	463	1475	5207	1079	4128
1918	3586	687	2899	991	190	801	2595	497	2098
1917	2503	495	2008	706	127	579	1797	368	1429
1916	1902	345	1557	476	104	372	1426	241	1185
1915 –	4755	944	3811	1248	223	1025	3507	721	2786
Rok urodz.	Liczba osób wg roczników urodzenia, 31.12.2016 r.								
1921	12708	2765	9943						
1920	8438	1903	6535						
1919	5207	1079	4128						
1918	2595	497	2098						
1917	1796	368	1428						
1916 –	4933	961	3972						

Rok urodzenia określony jako „1906 –” oznacza 1906 i wcześniej.

Dwoma odcieniami szarości (po przekątnej tabeli dla dwóch kolejnych roczników) zaznaczono dane GUS oraz oszacowany bilans. Porównanie winno dotyczyć przykładowo szacunku liczby osób urodzonych w latach 1908–1911 wg stanu na 31.12.2007 r. z danymi GUS o liczbie osób urodzonych w latach 1908–1911.

Źródło: Szacunek własny na podstawie baz danych GUS: Demografia, Bank Danych Lokalnych i zasobów danych obszarów tematycznych: Spisy Powszechne

## IV. Informacje uzupełniające rozdział IV: Perspektywy spisów ludności

### IV.1. Wykaz pozastatystycznych systemów informacyjnych, rejestrów i baz danych stanowiących potencjalne źródło danych dla NSP 2021

Lp.	Nazwa systemu informacyjnego/rejestru/bazy
1	Rejestr PESEL
2	Centralna Ewidencja Kierowców – CEK
3	Krajowy zbiór rejestrów, ewidencji i wykazu w sprawach cudzoziemców – centralny rejestr danych o nabyciu i utracie obywatelstwa polskiego
4	Centralny Rejestr Podmiotów Krajowa Ewidencja Podatników – CRP KEP
5	Baza danych o płatnikach podatku dochodowego od osób fizycznych – PIT
6	Centralny Rejestr Płatników Składek – CRPS
7	Centralny Rejestr Ubezpieczonych – CRU
8	Centralny Rejestr Członków Rodziny Ubezpieczonych Uprawnionych do Ubezpieczenia Zdrowotnego
9	System emerytalno-rentowy – EiR
10	System kompleksowej obsługi ubezpieczenia społecznego rolników – KRUSNAL
11	System kompleksowej obsługi świadczeń emerytalno-rentowych – FARMER
12	Centralny Wykaz Ubezpieczonych – CWU
13	System informacyjny PFRON
14	Krajowy zbiór rejestrów, ewidencji i wykazu w sprawach cudzoziemców – POBYT
15	Centralna Baza Danych Ksiąg Wieczystych
16	System informacyjny o osobach pozbawionych wolności oraz przebywających w zakładach poprawczych
17	Zintegrowany System Informacji o Nauce i Szkolnictwie Wyższym – POL-on, wykaz studentów (z wyłączeniem studentów uczelni wojskowych będących żołnierzami), informacje z rocznych sprawozdań z działalności uczelni
18	Zintegrowany System Informacji o Nauce i Szkolnictwie Wyższym – POL-on, wykaz doktorantów (z wyłączeniem doktorantów będących żołnierzami)
19	Zintegrowany System Informacji o Nauce i Szkolnictwie Wyższym – POL-on, ogólnopolski wykaz nauczycieli akademickich i pracowników naukowych
20	Zintegrowany System Informacji o Nauce i Szkolnictwie Wyższym – POL-on, ogólnopolski wykaz osób, którym nadano stopień doktora lub doktora habilitowanego
21	Zintegrowany System Informacji o Nauce i Szkolnictwie Wyższym – POL-on, ogólnopolski wykaz osób, którym nadano tytuł profesora
22	System Informacji Oświatowej – SIO – dane o uczniach
23	System Informacji Oświatowej – SIO – dane o nauczycielach
24	Krajowy System Monitoringu Świadczeń Rodzinnych – KSMSR
25	Krajowy System Monitoringu Pomocy Społecznej – KSMPs
26	Elektroniczny Krajowy System Monitoringu Orzekania o Niepełnosprawności – EKSMOON

Lp.	Nazwa systemu informacyjnego/rejestru/bazy
27	Państwowy rejestr granic i powierzchni jednostek podziału terytorialnego kraju
28	System informacyjny powiatowych urzędów pracy PUP o osobach bezrobotnych i poszukujących pracy
29	Baza Danych Obiektów Topograficznych – BDOT
30	Ewidencja podatkowa nieruchomości – EPN
31	Rejestr mieszkańców (gminne zbiory)
32	Rejestr zamieszkania cudzoziemców (gminne zbiory)
33	System pomocy społecznej o świadczeniobiorcach i osobach ubiegających się o przyznanie świadczeń
34	System świadczeń rodzinnych
35	System świadczeń z funduszu alimentacyjnego

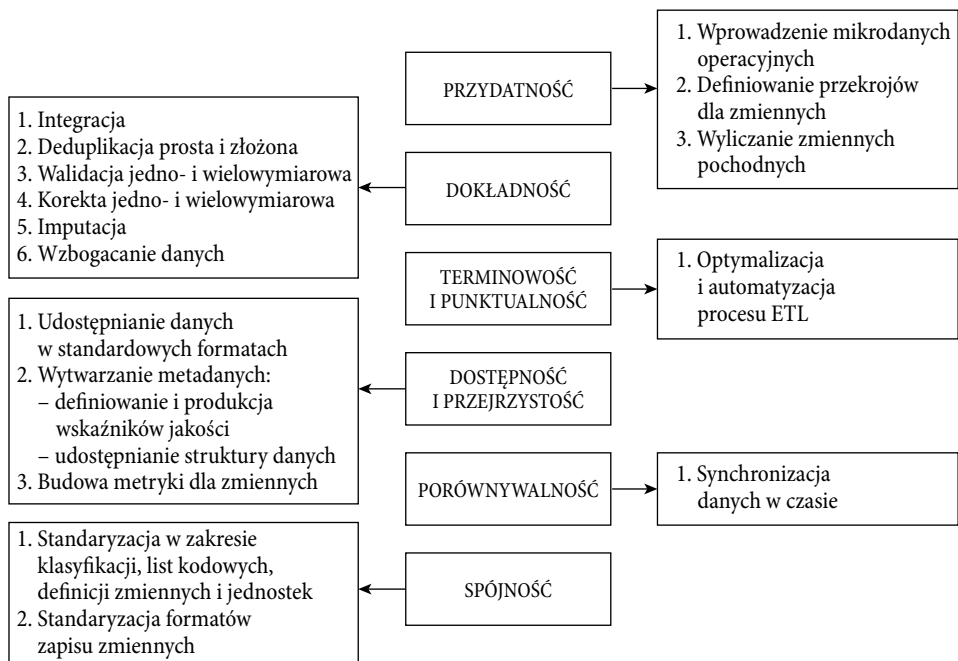
Źródło: (GUS, 2017d, s. 6–7).

## IV.2. Metryka zmiennej statystycznej według koncepcji spisu 2021 r.

Metryka zmiennej statystycznej	
I. Informacje o zmiennej	
Identyfikator zmiennej (ID)	Nadawany automatycznie numer porządkowy zmiennej
Nazwa zmiennej	należy podać nazwę zmiennej w strukturze Statystycznej Bazy Spisowej (SBS)
Definicja zmiennej	należy podać definicję zmiennej zgodną ze stosowaną w statystyce publicznej <u>źródło definicji</u>
Charakter zmiennej	należy podać charakter zmiennej: adresowa, identyfikacyjna, określająca czas, jakościowa, ilościowa
Format zmiennej	należy podać format zmiennej: tekstowa, numeryczna
Jednostki miary	należy podać jednostki miary stosowane w statystyce publicznej
Wymiary zmiennej	należy podać nazwy list kodowych, klasyfikacji, przekrojów terytorialnych lub też wybrane ich pozycje
Dziedzina	należy wskazać obszar tematyczny, którego dotyczy zmienna
Typ zmiennej	należy podać typ zmiennej: źródłowa, wyliczana <u>algorytm</u>
Źródło pochodzenia zmiennej	należy podać nazwy systemów informacyjnych, rejestrów, baz danych itd., z których pochodzi zmienna <u>opis źródeł</u>
Stan na	należy podać datę ostatniej aktualizacji zmiennej
Powiązane zmienne	należy podać zmienne powiązane: nadrzędne, podrzędne, składowe zmiennej <u>nazwa zmiennej</u>
II. Statystyki	
Stopień wypełnienia	źródło: Statystyki jakości danych źródłowych – powiązanie z Systemem Ewidencji Zmiennych (SEZ) Operat Spisowy: Statystyki jakości danych w Operacie Spisowym
Poziom spójności	źródło: Statystyki jakości danych źródłowych – powiązanie z SEZ Operat Spisowy: Statystyki jakości danych w Operacie Spisowym

Źródło: (GUS, 2017d, s. 19-20).

### IV.3. Planowana analiza jakości mikrodanych operacyjnych według koncepcji spisu 2021 r.



Źródło: (GUS, 2017d, s. 19–20)

---

## BIBLIOGRAFIA

---

- Ahas, R., Armoogum, J., Esko, S., Ilves, M., Karus, E., Madre, J.-L., ... Tiru, M. (2014). *Feasibility study on the use of mobile positioning data for tourism statistics – Consolidated Report*. Luxembourg: Publications Office of the European Union. <https://doi.org/10.2785/55051>
- Acelikli, K., Basarbowicz, K., Filas-Przybył, S., Franków, K., Kaźmierczak, M., Kruszka, K. i Stachowiak, D. (2010). *Dojazdy do pracy w Polsce. Terytorialna identyfikacja przepływów ludności związanych z zatrudnieniem*. (K. Kruszka, red.). Poznań: US Poznań. Pobrane z [http://stat.gov.pl/cps/rde/xbcr/poznan/ASSETS\\_dojazdy\\_do\\_pracy\\_w\\_polsce\\_2010.pdf](http://stat.gov.pl/cps/rde/xbcr/poznan/ASSETS_dojazdy_do_pracy_w_polsce_2010.pdf)
- Alho, J. i Spencer, B. (2005). *Statistical demography and forecasting*. Berlin – Heidelberg: Springer.
- Asamer, E.-M., Astleithner, F., Cetkovic, P., Humer, S., Lenk, M., Moser, M. i Rechta, H. (2016). Quality assessment for register-based statistics – Results for the Austrian census 2011. *Austrian Journal of Statistics*, 45(2), 3. <https://doi.org/10.17713/ajs.v45i2.97>
- Asamer, E.-M., Bachholz, N., Bauer, A., Bettel, I., Blauensteiner, S., Einfalt, J., ..., Wanek-Zajic, B. (2014). *Census 2011 Austria. Results of the register-based census*. (S. Blauensteiner, Ed.). Vienna: Statistics Austria. Pobrane z [http://ibestat.caib.es/ibestat/estadistiques/9cc677ad-ced2-475a-9e57-e0193466a532/2bfccc02-1a64-4ea1-bbcd-2ffe5bf0727f/es/I101004\\_1107.px](http://ibestat.caib.es/ibestat/estadistiques/9cc677ad-ced2-475a-9e57-e0193466a532/2bfccc02-1a64-4ea1-bbcd-2ffe5bf0727f/es/I101004_1107.px)
- Baffour, B., King, T., Valente, P. (2013). The modern census: evolution, examples and evaluation. *International Statistical Review*, 81(3), 407–425. <https://doi.org/10.1111/insr.12036>
- Baldrige, M., Brown, C. J., Jones, S. i Keane, J. G. (1985). *Evaluating censuses of population and housing*. New York: US Bureau of the Census.
- Barwiński, M. (2014). Struktura narodowościowa Polski w świetle wyników spisu powszechnego z 2011 roku. *Przegląd Geograficzny*, 86(2), 217–241.
- Basarbowicz, K., Dąbrowski, A., Filas-Przybył, S., Kaźmierczak, M., Klimanek, T., Kowalewski, J., ..., Stawikowska, M. (2015). *Identyfikacja obszarów specjalnych wewnątrz miast wojewódzkich oraz na ich obszarach funkcjonalnych uwzględniająca sytuację demograficzną i ekonomiczną ich mieszkańców na podstawie analiz przestrzennych z wykorzystaniem Geographic Information System (GUS)*. Warszawa, Poznań. Pobrane z [https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultstronaopisowa/5850/1/1/raport\\_obszary\\_specjalne\\_gis.pdf](https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultstronaopisowa/5850/1/1/raport_obszary_specjalne_gis.pdf)
- Bell, M., Charles-Edwards, E., Ueffing, P., Stillwell, J., Kupiszewski, M. i Kupiszewska, D. (2015). Internal migration and development: comparing migration intensities around



- the world. *Population and Development Review*, 41(1), 33–58. <https://doi.org/10.1111/j.1728-4457.2015.00025.x>
- Beręsewicz, M. (2017). A two-step procedure to measure representativeness of internet data sources. *International Statistical Review*, (5), 1–21.
- Beręsewicz, M., Pawlikowski, D., Filas-Przybył, S., Józefowski, T., Klimanek, T., Szymkowiak, M. i Kaźmierczak, M. (2018). *Poprawa wykorzystania administracyjnych źródeł danych*. Poznań.
- Beręsewicz, M. i Szymkowiak, M. (2015). Big Data w statystyce publicznej – nadzieje, osiągnięcia, wyzwania i zagrożenia. *Ekonometria*, 2(48), 9–22. <https://doi.org/10.15611>
- Berger, J. (2002). Spisy ludności na Ziemiach Polskich do 1918 r. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 87–95). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Berka, C., Humer, S., Moser, M., Lenk, M., Rechta, H. i Schwerer, E. (2012). Combination of evidence from multiple administrative data sources: Quality assessment of the Austrian register-based census 2011. *Statistica Neerlandica*, 66(1), 18–33. <https://doi.org/10.1111/j.1467-9574.2011.00506.x>
- Biemer, P. P., Trewin, D., Bergdahl, H. i Japac, L. (2014). A system for managing the quality of official statistics. *Journal of Official Statistics*, 30(3), 381–415. <https://doi.org/10.2478/JOS-2014-0022>
- Bijak, J. (2008a). *Metodologia procedur przejścia z różnych stanów poszczególnych systemów przewidzianych do wykorzystania w spisach na dany moment spisu*. Warszawa, Poznań.
- Bijak, J. (2008b). *Parametryzacja tablic migracji i ich dostosowanie do procedury przejścia z różnych stanów poszczególnych systemów, przewidzianych do wykorzystania w spisach, na dany moment spisu*. Warszawa, Poznań.
- Bijak, J., Kicinger, A. i Kupiszewski, M. (2007). *Studium metodologiczne oszacowania rzeczywistej liczby ludności Warszawy* (CEFMR Working Paper No. 2/2007). Warszawa. Pobrane z [http://www.cefmr.pan.pl/docs/cefmr\\_wp\\_2007-02.pdf](http://www.cefmr.pan.pl/docs/cefmr_wp_2007-02.pdf)
- Biraben, J.-N. (1963). Inventaire des listes nominatives de recensement en France. *Population*, (2), 305–328.
- Borys, T. (2012). Interdyscyplinarność nauk o jakości. *Zarządzanie i Finanse*, 10(3(1)), 7–23. Pobrane z [http://zif.wzr.pl/pim/?htm=2012\\_3\\_1.htm](http://zif.wzr.pl/pim/?htm=2012_3_1.htm)
- Bracha, C. (1996). *Teoretyczne podstawy metody reprezentacyjnej*. Warszawa: WN PWN.
- Brackstone, G. (1999). *Managing data quality in a statistical agency*. *Survey Methodology*, 25(2), 139–149.
- Brackstone, G. (2003). *Managing data quality in a statistical agency*. Santiago, Chile.
- Brown, J., Abbott, O. i Diamond, I. (2006). Dependence in the 2001 one-number census project. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 169(4), 883–902. <https://doi.org/10.1111/j.1467-985X.2006.00431.x>
- Brown, L. D., Cohen, M. L., Cork, D. L., Citro, C. F. i Lawrence D. (2010). *Envisioning the 2020 Census. Sciences – New York*. National Academies Press. <https://doi.org/10.17226/12865>
- Bul, R. (2016). Wpływ infrastruktury transportowej na zmiany dostępności czasowej Poznania z obszaru województwa wielkopolskiego w latach 2010–2016. *Prace Komisji Geografii Komunikacji PTG*, 19(2), 16–30. <https://doi.org/10.4467/2543859XPKG.16.008.6306>

- Buławski, R. (2002a). Kwestia narodowościowa w programie drugiego powszechnego spisu ludności Rzeczypospolitej Polskiej. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 205–224). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Buławski, R. (2002b). Metody statystyczne badania stosunków narodowościowych w spisach ludności. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 185–205). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Bussolo, M., Koettl, J. i Sinnott, E. (2015). *Golden ageing*. Pobrane z file:///C:/Users/1/Downloads/9781464803536 (1).pdf
- Calder, A. i Teague, A. (2012). Beyond 2011 – the future of population statistics? In *UNECE Seminar on New Frontiers for Statistical Data Collection* (s. 1–9). Geneva.
- Cellary, W. (2015, March 30). Potrzebujemy mniejszej chmury. Bezpieczniejszej. *Gazeta Prawna*. Poznań. Pobrane z [http://wyborcza.pl/1,155290,17678098,Potrzebujemy\\_mniejszej\\_chmury\\_\\_Bezpieczniejszej.html?disableRedirects=true](http://wyborcza.pl/1,155290,17678098,Potrzebujemy_mniejszej_chmury__Bezpieczniejszej.html?disableRedirects=true)
- Cellary, W. (2016, May 12). Inwigilacja w warunkach terroryzmu i gigadanych. *Gazeta Wyborcza*. Poznań. Pobrane z <http://wyborcza.pl/1,75400,20059513,inwigilacja-w-epoce-terroryzmu.html>
- Cetkovic, P., Humer, S., Lenk, M., Moser, M., Schnetzer, M. i Schwerer, E. (2012). A quality monitoring system for statistics based on administrative data. W: *Conference of European Statisticians, Seminar on New Frontiers for Statistical Data Collection* (s. 1–8). Pobrane z <https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2012/mtg2/WP11.pdf>
- Chipperfield, J., Brown, J. i Bell, P. (2017). Estimating the count error in the Australian census. *Journal of Official Statistics*, 33(1), 43–59. <https://doi.org/10.1515/jos-2017-0003>
- Cieślak, M. (1992). *Demografia : metody analizy i prognozowania*. (M. Cieślak, red.). Warszawa: Wydawnictwo Naukowe PWN.
- Coleman, D. (2013). The twilight of the census. *Population and Development Review*, 38(1), 334–351. <https://doi.org/10.1111/j.1728-4457.2013.00568.x>
- Courgeau, D. (2012). *Probability and social science: Methodological relationships between the two approaches*. Dordrecht, Heidelberg, London, New York: Springer.
- CSR. (2008). *Wielowymiarowa analiza błędów systematycznych w NSP 2002 oraz statystyczna analiza zmiennych wspomagających z NSP 2002 pod wykorzystanie dla małych obszarów. Sformułowanie wymagań funkcjonalnych bazy mikrodanych spisowych pod względem potrzeb estymacji dl*. Poznań, Warszawa.
- CSR. (2009a). *Opracowanie dla wybranych metod integracji danych reguł, procedur integracji danych z różnych źródeł: dla małych obszarów łączenie danych z różnych systemów informacyjnych z danymi z badań statystycznych*. Poznań, Warszawa.
- CSR. (2009b). *Opracowanie modelu estymacji danych dla małych obszarów na potrzeby PSR 2010 i NSP 2011 z wykorzystaniem danych systemów administracyjnych i badań statystycznych*. Poznań, Warszawa.
- D’Orazio, M., Di Zio, M. i Scanu, M. (2006). *Statistical matching: Theory and practice*. John Wiley i Sons.

- Daas, P. J. H., Puts, M. J., Buelens, B. i van den Hurk, P. A. M. (2014). Big data as a source for official statistics. *Journal of Official Statistics*, 31(2), 249–262. <https://doi.org/10.1515/jos-2015-0016>
- Dehnel, G. i Gołata, E. (2006). Attempts to estimate basic information for small business in Poland. *Statistics in Transition – New Series*, 7(4), 785–803.
- Dehnel, G. i Gołata, E. (2012). Wykorzystanie rejestrów administracyjnych w statystyce przedsiębiorstw. *Zeszyty Naukowe / Uniwersytet Ekonomiczny W Poznaniu*, (227), 63–83.
- Ditchburn, J. (2010). Former StatsCan head slams census decision by Tories. *Winnipeg Free Press*. Winnipeg.
- Domański, C. (2014). Wyzwania wobec statystyki jako nauki. *Wiadomości Statystyczne*, (5), 1–9.
- Dygaszewicz, J. (2009). Spisy powszechnie XXI wieku. *Wiadomości Statystyczne*, (6), 13–26.
- Esping-Andersen, G. i Billari, F. C. (2015). Re-theorizing Family Demographics. *Population and Development Review*, 41(1), 1–31. Pobrane z <http://web.a.ebscohost.com.0000892t1ba2.han3.ue.poznan.pl/ehost/pdfviewer/pdfviewer?vid=1&sid=12da3b00-150f-4c88-8132-cd864e85e738%40sessionmgr4007>
- EU. (1998). *Recommendation on statistics of international migration*. New York. Pobrane z [https://unstats.un.org/unsd/publication/SeriesM/SeriesM\\_58\\_rev1e.pdf](https://unstats.un.org/unsd/publication/SeriesM/SeriesM_58_rev1e.pdf)
- EU. (2001). Deklaracja Jakości Europejskiego Systemu Statystycznego. Luksemburg; Aneks 1 do Raportu końcowego Grupy wiodących ekspertów ds. jakości (LEG) przyjętego na 42 spotkaniu Komitetu ds. Programu Statystycznego.
- EU. (2011). *EU legislation on the 2011 population and housing censuses* (Methodologies i Working Papers). Luxembourg. <https://doi.org/10.2785/12778>
- EU. (2014). *Population and housing censuses*. Pobrane z [http://www.knbs.or.ke/index.php?option=com\\_phocadownload&view=category&id=100:population-and-housing-census&Itemid=1176](http://www.knbs.or.ke/index.php?option=com_phocadownload&view=category&id=100:population-and-housing-census&Itemid=1176)
- EU. (2015). *Ageing report*. Brussels. Pobrane z <http://www.aal-europe.eu/wp-content/uploads/2015/08/Ageing-Report-2015.pdf>
- EU. (2016). *EU census post-2021 programme*. Luxembourg.
- Eurarea Consortium. (2004). *EURAREA. Project Reference Volume. Deliverable. D7.1.4. Version. Final*. Pobrane z <https://www.ons.gov.uk/.../eurarea/.../project-reference-volume--volume-one--main-re>
- Eurostat. (2000). *Assessment of the quality in statistics*, 8.
- Eurostat. (2003). *Assessment of the quality in statistics*.
- Eurostat. (2011). *Europejski kodeks praktyk statystycznych*.
- Federal Law Gazette. (2015). Federal Act concerning the protection of personal data (DSG 2000). *Federal Law Gazette*, 132, 67. Pobrane z [https://www.ris.bka.gv.at/Dokument.wxe?Abfrage=Erv&Dokumentnummer=ERV\\_1999\\_1\\_165](https://www.ris.bka.gv.at/Dokument.wxe?Abfrage=Erv&Dokumentnummer=ERV_1999_1_165)
- Feeney, G. (1975). *The own-children method of estimating age-specific fertility rates*. Kula Lumpur.
- Fihel, A. (2015). Wpływ czasowych migracji zagranicznych na perspektywy demograficzne Polski. *Wiadomości Statystyczne*, (7), 74–89.
- Filas-Przybył, S. (2017). Wpływ migracji na lokalizację klastrów o wysokim poziomie starości demograficznej w przestrzeni miasta Poznania. Prezentacja podczas Ogólnopolskiej

- Konferencji Naukowej: Procesy integracji różnych źródeł danych w badaniach statystycznych. Źródła administracyjne, pozaadministracyjne i Big Data w statystyce publicznej. Uniwersytet Szczeciński. Szczecin 31.05–1.06.2017 r.
- Frątczak, E., Kotowska, I. E., Kurkiewicz, J., Kuroпка, I. i Paradysz, J. (2010). *Procesy demograficzne i metody ich analizy*. (J. Kurkiewicz, red.). Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Gerards, E. (2012). *Dutch virtual census*. Voorburg. Pobrane z <https://unstats.un.org/unsd/.../meetings/.../s04-5-1-Netherlands.ppt>
- Gołata, E. (2009). Aktywność ekonomiczna ludności w NSP' 2011 a zasoby rejestrów administracyjnych. W: E. Gołata (red.), *Metody i źródła pozyskiwania informacji w statystyce publicznej*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Gołata, E. (2012). Spis ludności i prawda. *Studia Demograficzne*, 1(161).
- Gołata, E. (2013a). Ludność rezydująca a faktyczna. W: P. Ulman (red.), *Zjawiska społeczne w badaniach statystycznych. Księga jubileuszowa dedykowana Profesor Barbarze Podolec*. Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Gołata, E. (2013b). Metodologia Powszechnego Spisu Ludności i Mieszkań 2011 w zakresie pytań etnicznych. *Studia Demograficzne*, 1(163), 97–133.
- Gołata, E. (2014). Jakość spisów nowej generacji. *Wiadomości Statystyczne*, (5), 26–38.
- Gołata, E. (2015). Metodologia Powszechnego Spisu Ludności i Mieszkań 2011 roku w zakresie pytań etnicznych. W: S. Łodziński, K. Warmińska i G. Gudaszewski (red.), *Mniejszości narodowe i etniczne w Polsce w świetle Narodowego Spisu Powszechnego Ludności z 2011 roku* (s. 126–159). Warszawa: Wydawnictwo Naukowe Scholar – Instytut Socjologii UW.
- Gołata, E. (2016a). Estimation of fertility in Poland and of Polish born women in the United Kingdom. *Studia Demograficzne*, 169(1), 13–38. Pobrane z <http://journals.pan.pl/dlibra/publication/110814/edition/96114/content>
- Gołata, E. (2016b). Shift in methodology and population census quality. *Statistics in Transition*, 17(4). Pobrane z <http://stat.gov.pl/en/sit-en/issues-and-articles-sit/current-issuse/>
- Gołata, E. i Kuroпка, I. (2016). Large cities in Poland in face of demographic changes. *Bulletin of Geography*, 34(34). <https://doi.org/10.1515/bog-2016-0032>
- Grand, T. (2010). An independent mind has always characterized Sheikh. *Globe and Mail*. Toronto. Pobrane z <https://beta.theglobeandmail.com/news/politics/an-independent-mind-has-always-characterized-sheikh/article1389402/?ref=http://www.theglobeandmail.com&>
- Grotowska-Leder, J. (2002). *Fenomen wielkomiejskiej biedy : od epizodu do "underclass"*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Grotowska-Leder, J. (2014a). *Nowe niepokojące zjawiska społeczne*. (J. Grotowska-Leder, red.), *Serie: Acta Universitatis Lodziensis. Folia Sociologica*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Grotowska-Leder, J. (2014b). Rzecz o kształceniu dorosłych. Lifelong learning w Polsce, w perspektywie Unii Europejskiej. *Acta Universitatis Lodziensis. Folia Sociologica*, 50, 117–135. Pobrane z <http://hdl.handle.net/11089/6996>
- Grotowska-Leder, J. (2014c). Zbiorowość NEET – niewykorzystane zasoby pracy młodzieży. Skala, dynamika i wybrane cechy zjawiska – w perspektywie Unii Europejskiej. W: L. Machol-Zajda i C. Sadowska-Snarska (red.), *Rynek pracy i polityka społeczna*

- w XXI wieku. *Aktualne problemy* (s. 13–28). Warszawa – Białystok: Instytut Pracy i Spraw Socjalnych, Uniwersytet w Białymstoku.
- Grotowska-Leder, J. i Rokicka, E. (2015). *Przemiany społeczne w Polsce i ich konsekwencje : perspektywa socjologiczna*. (J. Grotowska-Leder i E. Rokicka, red.). Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- GUS. (2003). *Stan i struktura ludności Polski. Narodowy Spis Ludności i Mieszkań 2002*. Warszawa. Pobrane z <http://stat.gov.pl/spisy-powszechno/narodowe-spisy-powszechno/narodowy-spis-powszechny-2002/ludnosc-stan-i-struktura-demograficzno-spoeczna-nsp-2002,4,1.html>
- GUS. (2011a). *Metodologia Narodowego Spisu Powszechnego Ludności i Mieszkań 2011 r*. Warszawa.
- GUS. (2011b). *Wyniki wstępne Narodowego Spisu Powszechnego Ludności i Mieszkań 2011*. Warszawa.
- GUS. (2012a). *Raport z wyników, Narodowy Spis Powszechny Ludności i Mieszkań 2011*. Warszawa.
- GUS. (2012b). *Wyniki Narodowego Spisu Powszechnego Ludności i Mieszkań 2011. Podstawowe informacje o sytuacji demograficzno-społecznej ludności Polski oraz zasobach mieszkaniowych*. Warszawa.
- GUS. (2012c). *Zasady opracowywania wyników Narodowego Spisu Powszechnego Ludności i Mieszkań 2011 w zakresie mniejszości narodowych i etnicznych oraz języka regionalnego*. Warszawa – Białystok.
- GUS. (2013a). *Ludność, stan i struktura demograficzno-społeczna. Narodowy Spis Powszechny Ludności i Mieszkań 2011*. Warszawa.
- GUS. (2013b). *Migracje zagraniczne ludności. Narodowy Spis powszechny ludności i mieszkań 2011*. Warszawa. Pobrane z <http://stat.gov.pl/spisy-powszechno/nsp-2011/nsp-2011-wyniki/migracje-zagraniczne-ludnosc-nsp-2011,13,1.html>
- GUS. (2014a). *Prognoza ludności na lata 2014–2050*. Warszawa: GUS.
- GUS. (2014b). *Spisy powszechno. Monografia poświęcona Powszechnemu Spisowi Rolnemu 2010 r. oraz Narodowemu Spisowi Powszechnemu Ludności i Mieszkań 2011 r.* (J. Dygaszewicz, red.).
- GUS. (2015). *Prognoza ludności rezydującej dla Polski 2015–2050*. Warszawa: GUS. Pobrane z <http://stat.gov.pl/obszary-tematyczne/ludnosc/prognoza-ludnosc/prognoza-ludnosc-rezydujacej-dla-polski-na-lata-2015-2050,8,1.html>
- GUS. (2016). *Metoda powtórnego ważenia w spisach powszechnych*. Warszawa – Poznań.
- GUS. (2017a). *Kalibracja zintegrowana w spisach realizowanych metodą mieszaną*. Warszawa – Poznań.
- GUS. (2017b). *Raport z konsultacji społecznych do powszechnego spisu ludności i mieszkań 2021*. Warszawa. Pobrane z <http://stat.gov.pl/spisy-powszechno/nsp-2021/>
- GUS. (2017c). *Raporty tematyczne z konsultacji społecznych do spisu ludności i mieszkań 2021*. Warszawa. Pobrane z <http://stat.gov.pl/spisy-powszechno/nsp-2021/>
- GUS. (2017d). *Zakres informacyjny a metoda realizacji spisu ludności i mieszkań 2021. Materiał przygotowany na posiedzenie Komisji Metodologicznej Warszawa, 30.06.2017 r.* Warszawa.
- Hogan, H., Cantwell, P. J., Devine, J., Mule, V. T. i Velkoff, V. (2013). Quality and the 2010 census. *Population Research and Policy Review*, 32(5), 637–662. <https://doi.org/10.1007/s11113-013-9278-5>

- Holt, T. D. (2007). The official statistics olympic challenge: Wider, Deeper, Quicker, Better, Cheaper. Discussion. *The American Statistician*, 61(1)(1), 1–8. <https://doi.org/10.1198/000313007X168173>
- Holzer, J. Z. (2003). *Demografia* (VI zmien.). Warszawa: Polskie Wydawnictwo Ekonomiczne.
- INEE. (2005). *Outline of the census type planned for 2011 in Spain*. Pobrane z <https://unstats.un.org/unsd/demographic/sources/census/Spainpdf.pdf>
- INEE. (2011). *Demographic Census Project 2011*. Madrid. Pobrane z [http://www.ine.es/en/censos2011/censos2011\\_proyecto\\_en.pdf](http://www.ine.es/en/censos2011/censos2011_proyecto_en.pdf)
- INEI. (2005). *Peru's 2005 population and housing censuses*. Lima. Pobrane z <https://unstats.un.org/unsd/demographic/sources/census/perupdf.pdf>
- Jo, Z., Dormon, O. i Jefferies, J., (2012). *Childbearing among UK born and non-UK born women living in the UK*. Office for National Statistics. <https://doi.org/10.1080/00343404.2011.581276>
- Jończy, R. (2010). *Migracje zagraniczne z obszarów wiejskich województwa opolskiego po akcesji Polski do Unii Europejskiej. Wybrane aspekty ekonomiczne i demograficzne*. Opole-Wrocław: Wydawnictwo Instytut Śląski.
- Jończy, R. (2017). *Sytuacja demograficzna i zatrudnienie na obszarach wiejskich województwa opolskiego*. Opole-Wrocław: Wydawnictwo Instytut Śląski.
- Józefowski, T. i Rynarzewska-Pietrzak, B. (2010). Ocena możliwości wykorzystania rejestru PESEL w spisie ludności. w: E. Gołata (red.), *Pomiar i informacja w gospodarce*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Jóźwiak, J. (2002). Standardy międzynarodowe w metodologii NSP 2002 w Polsce. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 674–681). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Keilman, N. i Kučera, T. (1991). The impact of methodology on accuracy of population forecasts. Evidence from the Netherlands and Czechoslovakia. *Journal of Forecasting*, 10, 371–398. Pobrane z <https://folk.uio.no/keilman/KeilmanKucera1991.pdf>
- Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., ..., Velez, E. D. (2015). The condition of education (2015). *U. S. Department of Education, National Center for Education Statistics*, 1–279. <https://doi.org/10.4135/9781412963916.n394>
- Kenett, R. S. i Shmueli, G. (2016). From quality to information quality in official statistics. *Journal of Official Statistics*, 32(4), 867–885. <https://doi.org/10.1515/jos-2016-0045>
- Kędelski, M. (1990). Fikcja demograficzna w Polsce i RFN (Ze studiów nad migracjami zagranicznymi). *Studia Demograficzne*, 99(1), 21–55.
- Kędelski, M. i Paradysz, J. (1984). Uwagi o kolejnych edycjach Rocznika Demograficznego. *Wiadomości Statystyczne*, (5), 41–43.
- Kędelski, M. i Paradysz, J. (2006). *Demografia*. Poznań: Wydawnictwo Akademii Ekonomicznej w Poznaniu.
- Klimanek, T. (2010). Zastosowanie estymatora EBLUP uwzględniającego korelację przestrzenną na danych ze Spisu Rolnego w 2002 roku. *Zeszyty Naukowe. Uniwersytet Ekonomiczny W Poznaniu*, (149), 105–121.
- Kobus, P., Smolka, M. i Nowakowska, G. (2009). *Raport z prac na zbiorach kompleksowego systemu informatycznego ZUS*. Warszawa.

- Kordos, J. (1987). *Dokładność danych w badaniach społecznych*. Warszawa: GUS.
- Kordos, J. (1988). *Jakość danych statystycznych*. Warszawa: Państwowe Wydawnictwo Ekonomiczne.
- Kordos, J. (2003). Program poprawy jakości statystyki. *Wiadomości Statystyczne*, (7–8), 64–77.
- Kordos, J. (2007). Some aspects of post-enumeration surveys in Poland. *Statistics in Transition – New Series*, 8(3), 563–576.
- Kordos, J. (2016). Zarys teoretycznego modelu TQM w statystyce oficjalnej. *Wiadomości Statystyczne*, (5), 1–16.
- Kordos, J. (2017). The challenges of the population census round of 2020. Outline of the methods of quality assessment of population census data. *Statistics in Transition – New Series*, 18(1), 115–138. Pobrane z <http://stat.gov.pl/en/sit-en/issues-and-articles-sit/previous-issues/volume-18-number-1-march-2017/>
- Kotowska, I. E. (2009a). Spisy powszechne ludności jako źródło informacji o przemianach demograficznych i społecznych. *Wiadomości Statystyczne*, (9), 1–15.
- Kotowska, I. E. (2009b). Zmiany modelu rodziny a zmiany aktywności zawodowej kobiet w Europie. W: I. E. Kotowska (red.), *Strukturalne i kulturowe uwarunkowania aktywności zawodowej kobiet w Polsce*. Warszawa: Scholar.
- Kotowska, I. E. (2010). Przemiany demograficzne w Polsce po 1989: Polska – Europa. W: H. Jastrzębska-Smoląga (red.), *Gospodarka rynkowa w Polsce w latach 1989–2009*. Kielce: Wydawnictwa Politechniki Świętokrzyskiej w Kielcach.
- Krapf, S. i Kreyenfeld, M. (2015). Fertility assessment with the own-children method: A validation with data from the German Mikrozensus. *MPIDR Technical Report TR-2015-003*, 49(0), 0–13.
- Krywułt-Albańska, M. (2012). Spis powszechny jako źródło informacji o ludności. *Studia Socjologiczne*, 4(207), 87–107.
- Krzyk, J. (2012). Naukowcy skrytykowali GUS za spis. *Gazeta Wyborcza*. Katowice. Pobrane z [http://katowice.gazeta.pl/katowice/1,35063,11617131,Naukowcy\\_skrytykowali\\_GUS\\_za\\_spis.html](http://katowice.gazeta.pl/katowice/1,35063,11617131,Naukowcy_skrytykowali_GUS_za_spis.html)
- Krzywicki, L. (2002). Wartość wyników spisu jednodniowego. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 131–143). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Kuc, M. (2014). Najstarszy człowiek świata żyje w Polsce w bazie PESEL. *Gazeta Prawna*, 4.06.2014.
- Kuklo, C. (2009). *Demografia Rzeczypospolitej Przedrozbiorowej*. Warszawa: Wydawnictwo DIG.
- Kukutai, T., Thompson, V. i McMillan, R. (2015). Whither the census? Continuity and change in census methodologies worldwide, 1985–2014. *Journal of Population Research*, 32(1), 3–22. <https://doi.org/10.1007/s12546-014-9139-z>
- Kupiszewska, D. i Kupiszewski, M. (2012). Long-term regional consequences of migration in Europe – assessment method and selected results. *Studia Migracyjne – Przegląd Polonijny*, 38(2), 105–134.
- Kupiszewski, M. (2004). Consequences of EU enlargement for freedom of movement between the Council of Europe member states. *Europa XXI*, 45–51. Pobrane z <http://rcin.org.pl/igipz/dlibra/docmetadata?id=3094&from=publication>

- Kupiszewski, M. (2005). *Migration in Poland in the period of transition – the adjustment to the labour market change*. IDEAS Working Paper Series from RePEc. St Louis. Pobrane z <http://ideas.repec.org/p/hit/piedp1/266.html>
- Kupiszewski, M. i Bijak, J. (2012). *Ocena założeń aktualnej prognozy demograficznej GUS na lata 2003–2030. Ocena wpływu migracji zagranicznej na bilans demograficzny Polski. Zmiany i problemy demograficzne a rozwój przestrzenny, IV*. Warszawa.
- Kupiszewski, M. i Kupiszewska, D. (2014). *Demographic trends and scenarios. Territorial scenarios and visions for Europe (ET2050)*. Luxembourg.
- Kurkiewicz, J. (2003). Przemiany struktury populacji w świetle wybranych teorii demograficznych. W: J. Pocięcha (red.) *Ekonomiczne konsekwencje osiągnięcia wieku emerytalnego przez generacje powojennego wyżu demograficznego* (s. 21–43). Kraków: Wydawnictwo Akademii Ekonomicznej w Krakowie.
- Kurkiewicz, J. (2007). Starzenie się ludności i jego ekonomiczno-społeczne konsekwencje. W: J. Kurkiewicz (red.) *Ludzie starsi w rodzinie i społeczeństwie* (s. 15–30). Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Kurkiewicz, J. (red.). (2010). *Procesy demograficzne i metody ich analizy*. Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Kurkiewicz, J. (red.). (2012). *Demograficzne uwarunkowania i wybrane społeczno-ekonomiczne konsekwencje starzenia się ludności w krajach europejskich*. Kraków, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Kuropka, I. (2011). Nierówność wobec śmierci w krajach Unii Europejskiej. *Nauki Społeczne*, 209, 68–84.
- Kuropka, I. i Pisz, Z. (2008). Równe traktowanie kobiet i mężczyzn w miejscu pracy szansą na wzrost aktywności zawodowej kobiet. W: A. Zagórska (red.), *Perspektywy demograficzne Śląska do 2030 r.* (s. 131–146). Opole: Oficyna Wydawnicza Politechniki Opolskiej.
- Lamla, K. (2010). Zensus 2011: Auswertung der Ergebnisse. *Statistisches Monatsheft Baden-Württemberg*, (11), 25–30.
- Lenk, M. (2008). *Methods of register-based census in Austria*. Wien. Pobrane z [https://www.researchgate.net/publication/265155437\\_Methods\\_of\\_Register-based\\_Census\\_in\\_Austria](https://www.researchgate.net/publication/265155437_Methods_of_Register-based_Census_in_Austria)
- Lenk, M. (2014). *Register-based census in Austria*. Wien. Pobrane z <https://www.stat.ee/dokumentid/79783>
- Lenk, M., Asamer, E.-M., Rechta, H., Schwerer, E., Astleithner, F., Humer, S. i Moser, M. (2014). *Quality assessment of administrative data*. Wien. Pobrane z [http://www.statistik.at/web\\_de/Redirect/index.htm?dDocName=076880](http://www.statistik.at/web_de/Redirect/index.htm?dDocName=076880)
- Lesthaeghe, R. i van de Kaa, D. J. (1986). Twee demografische transitities? W: D. J. van de Kaa i R. Lesthaeghe (Eds.), *Bevolking: groei en krimp* (s. 9–24). Deventer: Van Loghum Slaterus.
- Leventhal, B. (2013). Census-taking in the United Kingdom: 2011 and beyond. *Journal of Direct, Data and Digital Marketing Practice*, 14(3), 201–213. <https://doi.org/10.1057/ddmp.2013.4>
- Little, R. J. A. i Rubin, D. B. (2002). *Statistical analysis with missing data*. John Wiley i Sons.
- Lynn, S. (2001). *Developing quality standards for cross-national survey research: five approaches* (ISER Working Papers No. 2001–21). Colchester. Pobrane z <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.198.9357&rep=rep1&type=pdf>



- Lynn, S. (2003). Developing quality standards for cross-national survey research: five approaches. *International Journal of Social Research Methodology*, 6, 323–336. <https://doi.org/10.1080/13645570210132848>
- Łodziński, S. (2012). *Analiza metodologii Narodowego Spisu Powszechnego Ludności i Mieszkań z 2011 roku w zakresie pytań o narodowość, etniczność oraz język domowy i ojczysty*. Warszawa.
- Łodziński, S., Warmińska, K. i Gudaszewski, G. (red.). (2015). *Mniejszości narodowe i etniczne w Polsce w świetle Narodowego Spisu Powszechnego Ludności z 2011 roku*. Warszawa: Wydawnictwo Naukowe Scholar – Instytut Socjologii UW.
- Matheson, J. (2014). *The census and future provision of population statistics in England and Wales: Recommendation from the national statistician and chief executive of the UK statistics authority*.
- Matysiak, A. (red.). (2014). *Nowe wzorce formowania i rozwoju rodziny w Polsce. Przyczyny oraz wpływ na zadowolenie z życia*. Warszawa: Wydawnictwo Naukowe Scholar.
- Matysiak, A. i Mynarska, M. (2013). *Women's self-employment in Poland: A strategy for combining work and childcare? (Zeszyty Naukowe No. 28)*. Warszawa. Pobrane z <https://econpapers.repec.org/RePEc:isd:wpaper:68>
- Münnich, R. T., Burgard, J. P., Gabler, S., Ganninger, M. i Kolb, J.-P. (2016). Small area estimation in the German Census 2011. *Statistics in Transition – New Series*, (1), 25–40.
- Mynarska, M., Baranowska-Rataj, A. i Matysiak, A. (2014). Free to stay, free to leave: Insights from Poland into the meaning of cohabitation. *Demographic Research*, 31(1), 1107–1136. <https://doi.org/10.4054/DemRes.2014.31.36>
- Myrskylä, S. (Ed.). (2004). *Use of registers and administrative data sources for statistical purposes. Best practices of statistics Finland*. Helsinki: Statistics Finland. Pobrane z [http://www.doria.fi/bitstream/handle/10024/136324/xksk45\\_200400\\_2004\\_dig.pdf?sequence=1](http://www.doria.fi/bitstream/handle/10024/136324/xksk45_200400_2004_dig.pdf?sequence=1)
- Newell, R. i Smallwood, S. (2010). A cross country review of the validation and/or adjustment of census data. *Population Trends*, (141), 115–126. Pobrane z <http://search.proquest.com/docview/851893712?accountid=12037%5Cnhttp://linksource.ebsco.com/linking.aspx?sid=ProQ%3Aenvscijournals&fmt=journal&genre=article&issn=03074463&volume=&issue=141&date=2010-10-01&spage=115&title=Population+Trends&atitle=A+cross+c>
- Nordholt, E. S. (2005). *Brief description of the methodology plan for the 2011 Census of population and housing in the Netherlands*. Pobrane z <https://unstats.un.org/unsd/demographic/sources/census/netherlandspdf.pdf>
- Nordholt, E. S., Hartgers, M. i Gircour, R. (Eds.). (2004). *The Dutch virtual census of 2001 analysis and methodology*. Voorburg, Heerlen: Statistics Netherlands. Pobrane z <https://www.cbs.nl/NR/rdonlyres/D1716A60-0D13-4281-BED6-3607514888AD/0/b572001.pdf>
- Nowak, L. (1998). Jakość danych spisowych. W: *Tendencje zmian w strukturze ludności, gospodarstw domowych i rodzin w latach 1988–1995*. Warszawa: GUS.
- Obwieszczenie Marszałka Sejmu Obwieszczenie Marszałka Sejmu Rzeczypospolitej Polskiej z dnia 27 czerwca 2016 r. w sprawie ogłoszenia jednolitego tekstu ustawy o statystyce publicznej (Dz.U. z 2016 r., poz. 60).
- Okólski, M. (1990). *Teoria przejścia demograficznego*. (M. Okólski, red.). Warszawa: PWE.

- Okólski, M. (2004). *Demografia. Podstawowe pojęcia, procesy i teorie w encyklopedycznym zarysie*. Warszawa: Wydawnictwo Naukowe Scholar.
- Okólski, M. i Fihel, A. (2012). *Demografia*. Warszawa: Wydawnictwo Naukowe Scholar.
- Oleński, J. (2003). *Ekonomika informacji. Metody*. Warszawa: Polskie Wydawnictwo Ekonomiczne.
- Oleński, J. (2006). *Infrastruktura informacyjna państwa w globalnej gospodarce*. Warszawa: Wydział Nauk Ekonomicznych, Uniwersytet Warszawski ; Wydział Nauk.
- ONS. (1999). The 2001 census of population, White Paper (March 1999).
- ONS. (2008). *Helping to shape tomorrow. The 2011 Census of Population and Housing in England and Wales*. London: ONS.
- ONS. (2013). Detailed country of birth and nationality analysis from the 2011 Census of England and Wales. *Census*, (2011), 1–30. Pobrane z [http://www.ons.gov.uk/ons/dcp171776\\_310441.pdf](http://www.ons.gov.uk/ons/dcp171776_310441.pdf)
- Ostasiewicz, W. (2012). *Myślenie statystyczne*. Warszawa: Oficyna Wolters Kluwer.
- Paradysz, J. (1992). *Dzietność kobiet w Polsce*. Warszawa: GUS.
- Paradysz, J. (1999). Rekonstrukcja dzietności małżeńskiej kobiet w późniejszym wieku na podstawie ankiety retrospektywnej. *Studia Demograficzne*, (1), 13–34.
- Paradysz, J. (2002a). Badanie małżeńskości i dzietności kobiet w narodowych spisach powszechnych. *Wiadomości Statystyczne*, (1), 77–87.
- Paradysz, J. (2002b). O błędach nielosowych w badaniu dzietności kobiet w ramach Narodowego Spisu Powszechnego 1970. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 479–482). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Paradysz, J. (2007a). Recenzja opracowania „Studium metodologiczne oszacowania rzeczywistej liczby ludności Warszawy”, przygotowanego w Środkowoeuropejskim Forum Badań Migracyjnych przez zespół w składzie: Jakub Bijak, Anna Kicingier, Marek Kupiszewski. Poznań, Warszawa: Raport przygotowany na zlecenie Biura Naczelnego Architekta Miasta Urzędu m.st. Warszawy, umowa nr AM/001/1.1/B/275/06/2162.
- Paradysz, J. (2007b). Rejestry administracyjne jako źródło zasilania w statystyce regionalnej. W: J. Paradysz (red.), *Statystyka regionalna w jednoczącej się Europie* (s. 193–206). Poznań: Internetowa Oficyna Wydawnicza Centrum Statystyki Regionalnej.
- Paradysz, J. (2009). Ocena dobroci estymacji dla małych obszarów. W: E. Gołata (red.), *Metody i źródła pozyskiwania informacji w statystyce publicznej* (s. 19–28). Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Paradysz, J. (2010). Konieczność estymacji pośredniej na użytek spisów powszechnych. W: E. Gołata (red.), *Pomiar i informacja w gospodarce*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Perelli-Harris, B., Mynarska, M., Berrington, A., Berghammer, C., Evans, A., Isupova, O., ..., Vignoli, D. (2014). Towards a new understanding of cohabitation: Insights from focus group research across Europe and Australia. *Demographic Research*, 31(1), 1043–1078. <https://doi.org/10.4054/DemRes.2014.31.34>
- Platek, R. i Särndal, C. E. (2001). Can a statistician deliver? *Wiadomości Statystyczne*, (4).
- Plewis, I., Simpson, L. i Williamson, S. (2011). *CENSUS 2011: Independent review of coverage assessment, adjustment and quality assurance*. London. Pobrane z <http://hummedia>.

- manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2011/2011-01-census-2011.pdf
- Pociecha, J. (2003a). Charakterystyka powojennego wyżu demograficznego. W: J. Pociecha (red.), *Ekonomiczne konsekwencje osiągnięcia wieku emerytalnego przez generację powojennego wyżu demograficznego* (s. 11–20). Kraków: Wydawnictwo Akademii Ekonomicznej w Krakowie.
- Pociecha, J. (red.). (2003b). *Ekonomiczne konsekwencje osiągnięcia wieku emerytalnego przez generację powojennego wyżu demograficznego*. Kraków: Wydawnictwo Akademii Ekonomicznej w Krakowie.
- Pociecha, J. (2007). Starzenie się, stan zdrowia i umieralność ludności Polski w latach 1990–2004. W: J. Kurkiewicz (red.), *Ludzie starsi w rodzinie i społeczeństwie* (s. 31–53). Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Poulain, M. i Herm, A. (2013). Central population registers as a source of demographic statistics in Europe. *Population*, 68(2), 183–212. <https://doi.org/10.3917/popu.1302.0215>
- Pressat, R. (1966). *Analiza demograficzna. Metody. Wyniki. Zastosowania*. Warszawa: PWN.
- Prewitt, K. (2004). What if we give a census and no one comes? *Science*, 304(5676), 1452–1453.
- Rao, C. R. (1994). *Statystyka i prawda*. Warszawa: PWN.
- Rao, J. N. K. (2003). *Small area estimation*. New Jersey: Wiley i Sons, Ltd.
- Raport. Narodowy Spis Powszechny Ludności i Mieszkań 2011 w ocenie mniejszości ukraińskiej. (2011). Warszawa.
- Rässler, S. (2002). *Statistical matching. A frequentist theory, practical applications, and alternative bayesian approaches*. New York: Springer.
- Roszka, W. (2013). *Statystyczna integracja danych w badaniach społeczno-ekonomicznych*. Uniwersytet Ekonomiczny w Poznaniu. Pobrane z [http://www.wbc.poznan.pl/Content/265243/Roszka\\_Wojciech\\_doktorat.pdf](http://www.wbc.poznan.pl/Content/265243/Roszka_Wojciech_doktorat.pdf)
- Rozporządzenie (WE) Parlamentu Europejskiego i Rady nr 763/2008 w sprawie spisów powszechnych ludności i mieszkań. European Union (Dz.Urz. UE z 9.07.2008 r.).
- Ruggles, S. (2014). Big Microdata for Population Research. *Demography*, 51(1), 287–297. <https://doi.org/10.1007/s13524-013-0240-2>
- Sakson, B. (2002). *Wpływ „niewidzialnych” migracji zagranicznych lat osiemdziesiątych na struktury demograficzne Polski*. Warszawa: Szkoła Główna Handlowa.
- Skinner, C., Hollis, J. i Murphy, M. (2013). *Beyond 2011: Independent Review of Methodology*. London.
- Slany, K. (2014). Osoby niepełnosprawne w świetle Narodowego Spisu Powszechnego Ludności i Mieszkań z 2011 r. – wybrane aspekty. *Niepełnosprawność*, (11), 44–62. Pobrane z <http://www.pfron.org.pl/download/5/503/04-KrystynaSlany.pdf>
- Slany, K., Ślusarczyk, M. i Krzyżowski, Ł. (2014). *Wpływ współczesnych migracji Polaków na przemiany więzi społecznych, relacje w rodzinie i relacje międzygeneracyjne*. Warszawa.
- Snipp, M. C. (2003). Racial measurement in the American census: past practices and implications for the future. *Annual Review of Sociology*, 29, 563–588. <https://doi.org/10.1146/annurev.soc.29.010202.100006>
- Stachurski, T. i Żądło, T. (2017). On accuracy of calibration estimators supported by auxiliary variables from past periods based on simulation analyses. *Acta Universitatis Lodziensis. Folia Oeconomica*, 4(330), 39–53. <https://doi.org/10.18778/0208-6018.330.03>

- Statistics Canada. (2017). *Statistics Canada's quality assurance framework*.
- Statistics Netherlands. (2011). *Dutch Census 2011*.
- Statistics Netherlands. (2014). Quality declaration of statistics Netherlands. Pobrane z <https://www.cbs.nl/en-gb/about-us/organisation/quality>
- Strzelecki, Z. i Toczyński, T. (red.). (2002). *Spisy ludności Rzeczypospolitej Polskiej 1921–2002*. Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Szablowski, P., Wieczorkowski, R. i Wesołowski, J. (1996). Indeks zgodności jako miara jakości danych. *Wiadomości Statystyczne*, (4), 41–44.
- Szreder, M. (2009). Statystyka w państwie demokratycznym. *Wiadomości Statystyczne*, (6), 6–13.
- Szreder, M. (2010). *Metody i techniki sondażowych badań opinii* (Wyd. 2 zm.). Warszawa: Polskie Wydawnictwo Ekonomiczne.
- Szreder, M. (2013, 27 sierpnia). Jak rodzi się błąd. *Rzeczpospolita*, s. A11. Warszawa.
- Szreder, M. (2015). Big data wyzwaniem dla człowieka i statystyki. *Wiadomości Statystyczne*, (8), 1–11. Pobrane z [pts.stat.gov.pl/download/gfx/pts/pl/defaultstronaopisowa/52/2/1/w08\\_15\\_internet.pdf](https://pts.stat.gov.pl/download/gfx/pts/pl/defaultstronaopisowa/52/2/1/w08_15_internet.pdf)
- Szreder, M. (2016). O niektórych nowych wyzwaniach i oczekiwaniach wobec statystyki. *Wiadomości Statystyczne*, (6), 1–9. Pobrane z <http://stat.gov.pl/czasopisma/wiadomosci-statystyczne/archiwum/wiadomosci-statystyczne-nr-62016-661,7,16.html>
- Szreder, M. (2017). Nowe źródła informacji i ich wykorzystywanie w podejmowaniu decyzji. *Wiadomości Statystyczne*, 7(674), 5–17.
- Szturm de Szterm, E. (2002). Prawdziwa statystyka. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 228–231). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Szymańska, D. (2013). *Geografia osadnictwa* (Wydanie II). Warszawa: PWN.
- Szymkowiak, M. (2009). Calibration estimators for quantiles in surveys with nonresponse. W: L. Wywił, Janusz i T. Żądło (Eds.). *Survey sampling in economic and social research* (s. 60–75). Katowice: Wydawnictwo Akademii Ekonomicznej im. Karola Adameckiego w Katowicach.
- Szymkowiak, M. (2011). *Assessing the feasibility of using information from administrative databases for calibration in short-term and annual business statistics*. W: *Use of Administrative Data for Business Statistics*. Warszawa.
- Szymkowiak, M. (2012). *Kalibracja wag w NSP 2011*. Poznań.
- Szymkowiak, M. (2014). Estymatory kalibracyjne stosowane w Narodowym Spisie Powszechnym w 2011 r. *Wiadomości Statystyczne*, (11), 1–16.
- Szymkowiak, M. (2018). *Propozycja praktycznego zastosowania wyników prac w NSP 2021 i w innych badaniach statystycznych*. Poznań, Warszawa.
- Szymkowiak, M. i Beręsewicz, M. (2011). Kilka uwag na temat metod ważenia w badaniach statystycznych z brakami odpowiedzi. *Studia Oeconomica Posnaniensia*, 204, 116–124.
- Szymkowiak, M. i Józefowski, T. (2013). Zastosowanie estymatora typu spree w szacowaniu liczby osób bezrobotnych w przekroju podregionów. *Studia Oeconomica Posnaniensia*, 1(10), 120–135.
- Śleszyński, S. (2004a). Regionalne różnice pomiędzy liczbą ludności według narodowego spisu powszechnego w 2002 r. i rejestrowaną na podstawie ewidencji bieżącej. *Studia Demograficzne*, 145(1), 93–103.

- Śleszyński, S. (2004b). Różnice liczby ludności wykazane w NSP 2002 – Suplement. *Studia Demograficzne*, 146(2), 104–109.
- Śleszyński, S. (2005). Różnice w spisie ludności ujawnione w Narodowym Spisie Powszechnym 2002. *Przegląd Geograficzny*, 77(20), 193–212.
- Śleszyński, S. (2014). W sprawie prognozy demograficznej i jej niektórych skutków. W: Z. Strzelecki i E. Kowalczyk (red.), *Przemiany ludnościowe w Polsce. Przyszłość demograficzna* (s. 152–156). Warszawa: Rządowa Rada Ludnościowa.
- Tanton, R., Vidyattama, Y., Nepal, B. i McNamara, J. (2011). Small area estimation using a reweighting algorithm. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 174(4), 931–951. <https://doi.org/10.1111/j.1467-985X.2011.00690.x>
- Taranowicz, I. i Grotowska, S. (2015). *Rodzina wobec wyzwań współczesności*. (I. Taranowicz i S. Grotowska, red.). Wrocław: Oficyna Wydawnicza Arboretum.
- Tønder, J.-K. (2008). The Register-based statistical system preconditions and processes. Shanghai: International association for official statistics conference Shanghai October 14–18, 2008. Pobrane z [www.iaos-isi.org/papers/CS\\_1\\_2\\_Tonder.doc](http://www.iaos-isi.org/papers/CS_1_2_Tonder.doc)
- UE. (2008). Rozporządzenie Parlamentu Europejskiego i Rady nr 763/2008 z dnia 9 lipca 2008 r. w sprawie spisów powszechnych ludności i mieszkań, 7. Pobrane z <http://eur-lex.europa.eu/legal-content/PL/TXT/?uri=CELEX%3A32008R0763>
- UE. (2013) Rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 1260/2013 z dnia 20 listopada 2013 r. w sprawie statystyk europejskich w dziedzinie demografii. (2013). Pobrane z <http://stat.gov.pl/obszary-tematyczne/ludnosc/regulacje-prawne-unii-europejskiej/rozporzadzenie-parlamentu-europejskiego-i-rady-ue-nr-12602013-z-dnia-20-listopada-2013-r-w-sprawie-statystyk-europejskich-w-dziedzinie-demografii-,8,1.html>
- UN. (1949). *Population census methods*. New York: United Nations.
- UN. (1958a). *General principles for a housing census*. New York.
- UN. (1958b). *Principles and recommendations for national population censuses*. New York: United Nations.
- UN. (1969). *Principles and recommendations for the 1970 population censuses*. New York: United Nations.
- UN. (1980). *Principles and recommendations for population and housing censuses* (Vol. 3). New York: United Nations.
- UN. (1990). *Supplementary principles and recommendations for population and housing censuses*. New York: United Nations.
- UN. (1997). *Principles and recommendations for population and housing censuses – revision 1* (Vol. 3). New York: United Nations.
- UN. (2008). *Principles and recommendations for population and housing censuses revision 2*. New York: United Nations.
- UN. (2009). *Manual on census evaluation. Post enumeration surveys*.
- UN. (2010). *Post enumeration surveys. Operational guidelines*.
- UN. (2013a). *Census quality and coverage: Key results of the UNECE survey on national census practices*.
- UN. (2013b). *Overview of national experiences for population and housing censuses of the 2010 round. Working Paper*. New York.
- UN. (2013c). *Report on the results of a survey on census methods used by countries in the 2010 census round*.

- UN. (2015a). *Principles and recommendations for population and housing censuses – Revision 3. Statistical Papers Series M.No.67/Rev.3*. New York: United Nations.
- UN. (2015b). *World population ageing*. New York. Pobrane z [http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015\\_Report.pdf](http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf)
- UNECE. (2006a). *Conference of European Statisticians recommendations for the 2010 censuses of population and housing*. New York, Geneva: United Nations. Pobrane z [http://www.csb.gov.lv/sites/default/files/item\\_file\\_5233\\_conference\\_recom.pdf](http://www.csb.gov.lv/sites/default/files/item_file_5233_conference_recom.pdf)
- UNECE. (2006b). *United Nations Economic Commission for Europe Conference of European Statisticians Recommendations for the 2010 censuses of Geneva*: United Nations. Pobrane z [http://www.unece.org/fileadmin/DAM/stats/publications/CES\\_2010\\_Census\\_Recommendations\\_English.pdf](http://www.unece.org/fileadmin/DAM/stats/publications/CES_2010_Census_Recommendations_English.pdf)
- UNECE. (2012). *Overview of the 2010 round of population and housing censuses in the United Nations Economic Commission for Europe region*.
- UNECE. (2015). *Conference of European Statisticians recommendations for the 2020 censuses of population and housing*. New York, Geneva: United Nations. Pobrane z [https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECECES41\\_EN.pdf](https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECECES41_EN.pdf)
- US Census Bureau. (2005). *Traditional enumeration with yearly updates of characteristics*. Pobrane z <https://unstats.un.org/unsd/demographic/sources/census/usapdf.pdf>
- Ustawa o mniejszościach narodowych i etnicznych oraz o języku regionalnym z dnia 6 stycznia 2005 r. (Dz.U. z 2005 r., Nr 17, poz. 141).
- Ustawa z dnia 4 marca 2010 r. o narodowym spisie powszechnym ludności i mieszkań w 2011 r. (Dz.U. z 2010 r., Nr 47, poz. 277).
- Valente, S. (2010). Census taking in Europe: how are populations counted in 2010? *Population and Societies Del'Institut National D'Études Démographiques INED, Paris*, (467).
- van de Kaa, D. J. (1987). Europe's second demographic transition. *Population Bulletin*, 42(1).
- Vega, J. i Argüeso, A. (2016). Spain 2021 . Why will this census have more quality than the previous one?, European Conference on Quality and Official Statistics, Madrid, 31 May–3 June.
- Vries, W. De. (2002). *Dimensions of statistical quality*. New York.
- Wallgren, A. i Wallgren, B. (2013). Quality assessment in systems with registers and sample surveys. Pobrane z <http://www.statistics.gov.hk/wsc/IPS078-P2-S.pdf>
- Wallgren, A. i Wallgren, B. (2014). *Register-based statistics. Statistical methods for administrative data* (Second Ed.). John Wiley i Sons, Ltd.
- Wallgren, A. i Wallgren, B. (2015). Quality assessment in production systems with registers and sample surveys. *Statistical Journal of the IAOS*, 31(2), 241–247. <https://doi.org/10.3233/sji-150895>
- Waszak, S. (2002). O spisach ludności. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 309–311). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Wesołowski, J. (2003). Problemy estymacji dla małych obszarów. *Wiadomości Statystyczne*, (9).
- Winkler, W. (2005). *Overview of record linkage and current research directions*. Washington: U. S. Bureau of the Census.
- Wiśniowski, A., Bijak, J., Christiansen, S., Forster, J. J., Keilman, N., Raymer, J. i Smith, P. W. F. (2013). Utilising expert opinion to improve the measurement of international mi-

- gration in Europe. *Journal of Official Statistics*, 29(4), 583–607. <https://doi.org/10.2478/jos-2013-0041>
- Witkowski, J. (2002a). Badanie aktywności ekonomicznej ludności w NSP 2002. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 708–715). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Witkowski, J. (2002b). Migracje jako element ruchliwości społecznej i pracowniczej. W: Z. Strzelecki i T. Toczyński (red.), *Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów* (s. 525–531). Warszawa: Polskie Towarzystwo Demograficzne, Główny Urząd Statystyczny.
- Witkowski, J. (2010). Rola statystyki publicznej we współczesnym świecie. *Wiadomości Statystyczne*, (2), 1–17.
- Witkowski, J. (2014). Statystyka oficjalna wobec wyzwań globalnych. *Wiadomości Statystyczne*, (4), 1–16.
- World Economic Forum. (2012). *Global population ageing*. Geneva.
- Wróblewska, W. (2006). Analiza umieralności w Polsce w latach 1970–2003. Dekompozycja zmian w oczekiwanim trwaniu życia noworodka. *Studia Demograficzne*, 1 (149).
- Wunsch, G. i Gourbin, C. (2018). Mortality, morbidity and health in developed societies: a review of data sources. *Genus*, 74(1), 2. <https://doi.org/10.1186/s41118-018-0027-9>
- Wywiał, J. (2010). *Wprowadzenie do metody reprezentacyjnej*. Katowice: Wydawnictwo Akademii Ekonomicznej w Katowicach.
- Yu, T. i Chen, S.-H. (2018). Big Data, Scarce Attention and Decision-Making Quality. *Computational Economics*, (1), 1–30. <https://doi.org/10.1007/s10614-018-9798-5>
- Zasępa, R. (1972). *Metoda reprezentacyjna*. Warszawa: Polskie Wydawnictwo Ekonomiczne.
- Zasępa, R. (1988). *Modele demograficzne na potrzeby analizy i sterowania rozwojem społeczno-gospodarczym*. Warszawa: Szkoła Główna Planowania i Statystyki.
- Zasępa, R. (1993). Use of sampling methods in population censuses in Poland. *Statistics in Transition*, 1(1), 69–78.
- Zhang, L.-C. (2011). A unit-error theory for register-based household statistics. *Journal of Official Statistics*, 27(3), 415–432.
- Zhang, L.-C. (2012). Topics of statistical theory for register-based statistics and data integration. *Statistica Neerlandica*, 66(1), 41–63.
- Zhang, L.-C. (2013). *Population size estimation based on multiple lists. Uncertainty analysis for categorical data fusion*. Wykład otwarty 24.10.2013. Uniwersytet Ekonomiczny w Poznaniu. Pobrane z [http://kdg.ue.poznan.pl/att/Wyklady\\_otwarte/LiChun.pdf](http://kdg.ue.poznan.pl/att/Wyklady_otwarte/LiChun.pdf).
- Zhang, L.-C. (2015). On modelling register coverage errors. *Journal of Official Statistics*, 31(3), 381–396. <https://doi.org/10.1515/jos-2015-0023>.
- Zych, M. i Czerny, A. (2015). *Urzędowy wykaz nazw państw i terytoriów niesamodzielnych (III)*. Warszawa: Główny Urząd Geodezji i Kartografii. Pobrane z [http://www.gugik.gov.pl/\\_data/assets/pdf\\_file/0005/24935/Urzedowy-wykaz-panstw.pdf](http://www.gugik.gov.pl/_data/assets/pdf_file/0005/24935/Urzedowy-wykaz-panstw.pdf)
- Żądło, T. (2015). *Statystyka małych obszarów w badaniach ekonomicznych. Podejście modelowe i mieszane*. Katowice: Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.

---

## SPIS TABEL I RYSUNKÓW

---

### Spis tabel

1.1. Ludność faktyczna i rezydująca w Narodowych Spisach Powszechnych 2002 i 2011 .....	36
1.2. Ludność faktyczna i rezydująca w Narodowym Spisie Powszechnym 2011 ....	37
1.3. Tematy podstawowe i dodatkowe według rekomendacji ONZ dla spisów ludności rundy 2020 .....	39
1.4. Obszary tematyczne wyróżnione w konsultacjach społecznych przed NSP 2021 .....	42
1.5. Cel wykorzystania danych spisu ludności deklarowany przez organizacje i osoby indywidualne według obszarów tematycznych, konsultacje społeczne przed spisem 2021 r. ....	44
2.1. Metody przeprowadzania spisu ludności <sup>a</sup> .....	55
2.2. Metody przeprowadzania spisu ludności a źródła danych i sposób ich gromadzenia .....	56
2.3. Opis metod przeprowadzania spisów ludności .....	58
2.4. Liczba państw według metody przeprowadzania spisu ludności w latach 1985–2014 .....	60
2.5. Rok utworzenia rejestru i rok pierwszego wykorzystania w spisie ludności według typu rejestru, państwa skandynawskie .....	65
2.6. Ilustracja zasady redundancji w spisie opartym na rejestrach, Austria 2011 ..	71
2.7. Ilustracja analizy rezydencji w spisie opartym na rejestrach, Austria 2011 .....	72
2.8. Nazwy centralnych rejestrów ludności i instytucji je prowadzących w trzydziestu państwach europejskich .....	85
2.9. Metody przeprowadzania spisów ludności w wybranych państwach w latach 2000–2010 .....	87
2.10. Metody gromadzenia danych w spisach rundy 2010 według regionów .....	89
3.1. Metody oceny błędów spisu w zależności od liczby źródeł danych wykorzystywanych w procesie ewaluacji .....	108
3.2. Tablica podstawowa – schemat podwójnego systemu estymacji .....	112
3.3. Rozbieżności szacunków liczby ludności według spisów oraz bilansów (w tys.) .....	121
3.4. Szacunek liczby ludności rezydującej ogółem, NSP 2011 .....	125



3.5. Stan ludności w świetle prognozy GUS 2013–2050 oraz jej modyfikacji .....	128
3.6. Ocena dokładności szacunku liczby niemowląt w NSP 2002 i 2011 bez uwzględnienia ich zgonów i migracji zagranicznych .....	133
3.7. Ocena dokładności szacunku liczby niemowląt w NSP 2002 i 2011 z uwzględnieniem zgonów .....	135
3.8. Ocena dokładności szacunku liczby dzieci w NSP 2011 z uwzględnieniem zgonów .....	136
3.9. Współczynniki „przetrwania dla osób danej generacji” pomiędzy NSP 2002 i 2011, z uwzględnieniem migracji i bez niej .....	141
3.10. Analiza zgonów osób generacji 1916 r. i wcześniejszych w latach 2012–2016	143
3.11. Liczba osób urodzonych w latach 1921 i wcześniej według bazy danych Demografia GUS oraz bilansu liczby osób urodzonych w latach 1921 i wcześniej na podstawie stanu ludności według wieku na 31.12.2001 r. i liczby zgonów według wieku i roku urodzenia w 2002 r. ....	144
3.12. Analiza zgonów w najstarszych grupach wieku .....	146
3.13. Liczba osób urodzonych w 1920 r. i wcześniej oraz zgony według roczników urodzenia .....	147
3.14. Ocena efektywności schematu losowania: rozkład powiatów według absolutnej różnicy względnej między bezpośrednią oceną estymatora a wartością w operacie losowania dla wybranych charakterystyk, badanie reprezentacyjne NSP 2011 .....	150
I.6. Obszary tematyczne spisu, tematy i ocena stopnia zaspokojenia potrzeb, konsultacje społeczne przed spisem 2021 r. ....	195
I.7. Cechy zespołowe (1–3) oraz demograficzne (4–25) uwzględnione w poszczególnych spisach ludności w Polsce .....	196
I.8. Cechy społeczno-zawodowe uwzględniane w powszechnych spisach ludności w Polsce .....	197
I.9. Najczęściej zgłaszane dodatkowe potrzeby informacyjne i powiązania tematów, konsultacje społeczne przed spisem 2021 r. ....	199
II.1. Rejestry administracyjne i systemy informacyjne wykorzystane w NSP 2011	203
II.3. Sposób wykorzystania centralnego rejestru ludności w statystyce ludności, wybrane państwa europejskie .....	208
III.1. Ocena kompletności NSP 2011: Rozbieżności szacunków bilansu ludności generacji 1916 i wcześniejszych na podstawie danych ewidencji ruchu naturalnego i stanów ludności według roczników urodzenia .....	209
III.2. Ocena kompletności NSP 2011: Liczba osób starszych według Bazy Danych Demografia GUS oraz szacunku bilansu na podstawie danych ewidencji zgonów według roku urodzenia w latach 2006–2016 .....	212
IV.1. Wykaz pozastatystycznych systemów informacyjnych, rejestrów i baz danych stanowiących potencjalne źródło danych dla NSP 2021 .....	215
IV.2. Metryka zmiennej statystycznej według koncepcji spisu 2021 r. ....	216

**Spis rysunków**

2.1.	Źródła danych w spisie opartym na rejestrach, Austria 2011 .....	68
2.2.	System zabezpieczenia poufności danych rejestrów w spisie opartym na rejestrach, Austria 2011 .....	70
2.3.	Schemat przetwarzania danych spisowych, NSP 2011 .....	81
2.4.	Informatyczny system spisowy z wyróżnieniem OBM oraz ABM wraz z podsystemem Metainformacji .....	82
2.5.	Metody przeprowadzania spisów ludności w krajach europejskich, 2010–2011 .....	89
2.6.	Potencjalne profile kosztów spisu tradycyjnego i wykorzystującego dane administracyjne .....	96
2.7.	Potencjalne profile korzyści tradycyjnego spisu i wykorzystującego dane administracyjne .....	96
3.1.	Klasyfikacja błędów nielosowych w spisach wykorzystujących różne źródła danych .....	107
3.2.	Zmiana liczby ludności Polski w latach 1970–2050 według rejestrowanych i szacowanych przez GUS stanów liczby ludności oraz trzech ostatnich prognoz GUS: P 2003, P 2007, P 2014 .....	120
3.3.	Liczba ludności Polski w latach 1989–2014: rozbieżności szacunków według spisów i bilansów .....	120
3.4.	Różnica między liczbą ludności rezydującej a faktycznej według płci i wieku, NSP 2011 .....	124
3.5.	Liczba ludności rezydującej i faktycznej według płci i wieku, NSP 2011 oraz szacunki własne .....	127
3.6.	Piramida wieku ludności polskiej narodowości oraz urodzonej w Polsce ludności rezydentów w Anglii i Walii, 2011 .....	137
3.7.	Współczynniki „przetrwania dla osób danej generacji” pomiędzy NSP 2002 i 2011, bez uwzględniania migracji .....	139
3.8.	Współczynniki „przetrwania dla osób danej generacji” pomiędzy NSP 2002 i 2011, z uwzględnieniem migracji .....	141
3.9.	Rozkład powiatów według względnych różnic między oceną bezpośrednią wartości globalnej oraz wartością w operacie losowania dla wybranych charakterystyk, badanie reprezentacyjne NSP 2011 .....	149
3.10.	Rozkład powiatów według absolutnych różnic względnych między oceną bezpośrednią wartości globalnej oraz wartością w operacie losowania dla wybranych charakterystyk, badanie reprezentacyjne NSP 2011 .....	150
4.1.	Schemat opracowywania wyników spisu ludności opartego na rejestrach administracyjnych .....	160
IV.3.	Planowana analiza jakości mikrodanych operacyjnych według koncepcji spisu 2021 r. ....	217