

Table of contents

Strona tytułowa	2
Słownik pojęć i skrótów	3
Pojęcia	3
Skróty	14
1 Wprowadzenie	16
2 Cel i założenia dokumentu	17
3 Podstawy prawne tajemnicy statystycznej	20
4 Metodologiczne zasady ochrony tajemnicy statystycznej	35
4.1 Miejsce, rola i znaczenie ochrony tajemnicy statystycznej w przyjętym Modelu Procesu Produkcji Statystycznej	35
4.2 Zasady zachowania tajemnicy statystycznej poprzez zastosowanie odpowiednich metod kontroli ujawniania danych statystycznych stosowane w statystyce pub- licznej	38
4.2.1 Zasady ogólne	38
4.2.2 Zasady kontroli ujawniania zagregowanych danych statystycznych opra- cowywanych w tablicach	42
4.2.3 Zasady kontroli ujawniania danych statystycznych przy prezentacji danych w formie graficznej	48
4.2.4 Tryb pozyskiwania zgody na publikację określonych danych charak- teryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej	49
4.3 Zakres informacji dotyczących tajemnicy statystycznej wprowadzany w SMS . .	51

Strona tytułowa



Przewodnik dla jednostek służb statystyki publicznej
w zakresie kontroli ujawniania danych statystycznych

Opracowanie zbiorcze

Zespołu ds. metod kontroli ujawniania danych statystycznych

Warszawa, listopad 2023

Słownik pojęć i skrótów

Pojęcia

agregat (dane zagregowane) wynik przekształcenia danych na poziomie jednostkowym w ilościowe miary dla zbioru cech populacji (zob. “The Eurostat Concepts and Definitions Database” 2001);

anonimizacja trwałe i nieodwracalne usunięcie ze zbioru danych informacji umożliwiających powiązanie z konkretną osobą fizyczną lub podmiotem gospodarki narodowej albo inną osobą prawną bądź jednostką organizacyjną niemającą osobowości prawnej;

atrybuty niewrażliwe (*zmienne niewrażliwe, zmienne wynikowe niebędące poufne*) zmienne niemające takiego stopnia poufności jak atrybuty wrażliwe (np. miejscowość zamieszkania lub zawód); potencjalnie mogą się stać pseudoidentyfikatorami;

atrybuty wrażliwe (*zmienne wrażliwe, poufne zmienne wynikowe*) atrybuty, które są uznawane za należące wyłącznie do sfery prywatnej oraz poddane ochronie prawnej (wyznanie, stan zdrowia, poglądy polityczne, dochód osobisty itp.);

celowana wymiana rekordów (ang. *Targeted Record Swapping, TRS*) specyficzna metoda wymiany rekordów, polegająca na wykorzystaniu hierarchicznej struktury mikrodanych (zwłaszcza w kontekście geograficznym). Wymiana danych następuje pomiędzy jednostkami wyższego poziomu niż podstawowy. Bazuje ona na określaniu grup rekordów o największym ryzyku ujawnienia (wyznaczanym w oparciu o regułę -anonimowości) na każdym poziomie hierarchii i wymianie danych z zakresu zmiennych definiujących określone poziomy dokonywanej między grupami rekordów najbliższymi według odległości wyznaczonej na podstawie wartości określonych zmiennych (tzw. zmienne podobieństwa);

dane administracyjne informacje i dane zgromadzone w rejestrach urzędowych i systemach informacyjnych administracji publicznej [Art. 2 pkt 13a ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

dane jednostkowe dane jednostkowe identyfikowalne i dane jednostkowe nieidentyfikowalne [Art. 2 pkt 9 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)]; innymi słowy, są to dane dotyczące pojedynczego obiektu obserwacji statystycznej;

dane jednostkowe identyfikowalne dane statystyczne zawierające informacje dotyczące konkretnego podmiotu gospodarki narodowej albo osoby fizycznej, identyfikujące bezpośrednio ten podmiot albo osobę według nazwy, imienia i nazwiska, adresu lub publicznie dostępnego numeru identyfikacyjnego oraz pozwalające na pośrednią identyfikację tego podmiotu albo osoby z użyciem innych środków niż środki pozwalające na bezpośrednią identyfikację, z wyłączeniem środków wymagających nadmiernych kosztów, czasu lub działań [Art. 2 pkt 8a ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

dane jednostkowe nieidentyfikowalne dane statystyczne zawierające informacje dotyczące konkretnego podmiotu gospodarki narodowej albo osoby fizycznej, niepozwalające na bezpośrednią ani pośrednią identyfikację tego podmiotu albo osoby [Art. 2 pkt 8b ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

dane osobowe informacje o zidentyfikowanej lub możliwej do zidentyfikowania osobie fizycznej („osobie, której dane dotyczą”); możliwa do zidentyfikowania osoba fizyczna to osoba, którą można bezpośrednio lub pośrednio zidentyfikować, w szczególności na podstawie identyfikatora takiego jak imię i nazwisko, numer identyfikacyjny, dane o lokalizacji, identyfikator internetowy lub jeden bądź kilka szczególnych czynników określających fizyczną, fizjologiczną, genetyczną, psychiczną, ekonomiczną, kulturową lub społeczną tożsamość osoby fizycznej [Art. 4 pkt 1 Rozporządzenia Parlamentu Europejskiego i Rady (UE) 2016/679 w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych – RODO) (Dz. Urz. UE L 119 z 04.05.2016, str. 1, z późn. zm.)];

dane statystyczne dane dotyczące zjawisk, zdarzeń, obiektów i działalności podmiotów gospodarki narodowej oraz życia i sytuacji osób fizycznych, w tym dane osobowe, pozyskane bezpośrednio od respondentów albo z systemów informacyjnych administracji publicznej, rejestrów urzędowych lub niepublicznych systemów informacyjnych, od momentu ich zebrania na potrzeby wykonywania zadań statystyki publicznej [Art. 2 pkt 1a ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

dane statystyczne charakteryzujące wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą

dane pozyskane przez statystykę publiczną z różnych źródeł na potrzeby prowadzonych badań, które zarówno bezpośrednio jak i pośrednio mogą informować o wynikach ekonomiczno-finansowych osiągniętych przez podmiot; zmienne/zestawy danych, które w poszczególnych badaniach charakteryzują wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą w rozumieniu art. 38 ust. 2 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej, są wskazane w SMS;

dominacja (n, k) występuje, gdy udział n największych podmiotów gospodarczych jest większy niż $k\%$ całkowitej wartości cechy ilościowej w komórce; w Polsce przyjmuje się $n = 1$ i $k = 75$ [opracowanie własne na podstawie Art. 38 ust. 1 Ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

funkcje cyklometryczne funkcje odwrotne do funkcji trygonometrycznych; z powodu okresowości tych ostatnich określone tylko na pewnych przedziałach (na których funkcje trygonometryczne są różnowartościowe, co stanowi warunek konieczny istnienia funkcji odwrotnej); nazywa się je także funkcjami kołowymi;

identyfikator jednoznaczny, unikatowy wyróżnik osoby lub podmiotu w bazie danych (np. numer PESEL);

imputacja danych uzupełnienie braków danych poprzez zastosowanie odpowiednich mechanizmów lub modeli i zmiennych z kompletnymi informacjami (np. imputacja oparta na dawcach, imputacja oparta na modelach, imputacja losowa, itp.) celem zapewnienia odpowiedniej jakości udostępnianych danych jednostkowych lub oszacowań określonych wielkości agregatowych;

jednostka autorska komórka organizacyjna Głównego Urzędu Statystycznego lub urząd statystyczny (wiodący) odpowiedzialny za realizację badania statystycznego/pracy metodologicznej, wskazany w Planie Opracowań Statystycznych Głównego Urzędu Statystycznego dla danego badania statystycznego/ opracowania lub w Planie Prac Metodologicznych dla pracy metodologicznej na dany rok [*Rozdz. I pkt 15 Załącznika do zarządzenia wewnętrznego nr 32 z dnia 4 grudnia 2020 r. Prezesa GUS w sprawie ustanowienia Polityki postępowania z danymi statystycznymi (z późn. zm.)*];

kontrola udostępniania danych statystycznych proces polegający na weryfikacji różnej postaci wyników danych statystycznych przed ich opublikowaniem czy udostępnieniem w innej postaci;

kontrola ujawniania danych statystycznych działania prowadzone w odniesieniu do danych statystycznych w celu eliminacji (lub redukcji) ryzyka ujawnienia informacji o jednostkach statystycznych (respondentach); działania te zazwyczaj zmieniają lub ograniczają zakres publikowanych danych, minimalizując ryzyko ujawnienia i pozwalając jednocześnie na publikację możliwie szerokiego zakresu informacji (zob. Hundepool et al. 2012);

metoda deterministyczna metoda ukrywania komórek w kontekście ukrycia pierwotnego i wtórnego opierająca się na identyfikacji, opisie biznesowym oraz zalgorytmizowaniu przypadków stosowania ochrony przekrojów poprzez ich ukrycie;

metoda hipersześcianów (hiperkostki, HYPERCUBE) metoda kontroli ujawniania danych dla tablicy hierarchicznej lub łączonej polegająca na tym, że, podobnie jak dla podejścia modułowego, tablica zostaje podzielona na proste podtablice, a następnie poszukuje się rozwiązania dla każdej z podtablic w sposób iteracyjny. W odróżnieniu od podejścia modułowego, dla każdej z podtablic znalezione rozwiązanie nie musi być rozwiązaniem optimum, co z kolei z reguły prowadzi do bardziej restrykcyjnego ukrycia komórek, niż byłoby to konieczne;

metoda HITAS (podejście modułowe) heurystyczna metoda kontroli ujawniania danych dla tablicy hierarchicznej lub łączonej polegająca na tym, że tablicę tę dzieli się na tablice bez hierarchii i poszukuje optimum ukrycia dla każdej wydzielonej tablicy z osobna. Łącząc wyniki w odpowiedni sposób, uzyskuje się rozwiązanie dla całej tablicy, które nie musi

być optymalne. Celem tego podejścia jest szybsze znalezienie rozwiązania, które może jednak prowadzić do większej straty informacji;

metoda kluczy komórkowych (ang. *cell-key method*, CKM) post-tablicowe podejście kontroli ujawniania danych, polegające na dodawaniu do każdej komórki tablicy szumu losowanego określonym mechanizmem z ustalonego rozkładu. Istotne są tutaj klucze przyporządkowywane rekordom w zbiorze mikrodanych stanowiących podstawę konstrukcji tablicy – liczby losowe z rozkładu jednostajnego $X \sim U(0,1)$, które dodaje się odpowiednio do naliczania tablicy, uzyskując w efekcie klucz komórki oraz predefiniowane prawdopodobieństwa przejścia (tzw. *p-tablica* – ang. *p-table*) służące do wyznaczania szumu jako funkcji kluczy komórek i wartości komórek;

metoda lokalnego ukrywania danych (ang. *local suppression*) niezakłóceńowa metoda kontroli ujawniania polegająca na usuwaniu pewnych wartości niektórych zmiennych dla konkretnych jednostek celem uniknięcia ich identyfikacji; efekt lokalnego ukrywania stanowi zwiększenie liczby rekordów, dla których kombinacja określonych wartości pewnych innych zmiennych, uznanych za kluczowe, jest taka sama; podejście to ma zastosowanie przede wszystkim do danych jakościowych;

metoda OPTIMAL (metoda optymalna) metoda ochrony poufności danych w tablicy hierarchicznej poprzez traktowanie jej jak pojedynczą, bez dzielenia na mniejsze podtablice, co zwiększa wydajność działań w zakresie kontroli ujawniania danych;

metoda przekodowania (przekodowanie, restrukturyzacja, ang. *recoding*) zmiana sposobu kodowania określonych wrażliwych zmiennych; jeśli zmienne mają charakter jakościowy, wówczas polega na połączeniu kilku kategorii w jedną – bardziej zgrubną i o większej liczbie należących do niej jednostek, co pozwala ukryć informację wrażliwą; dla zmiennej ilościowej przekodowanie polega zaś na zastąpieniu owej zmiennej przez jej odpowiednik w postaci jakościowej; metoda ma zastosowanie zarówno do mikrodanych, jak i do tablic;

metoda SIMPLEHEURISTIC odmiana procedury heurystycznej HITAS, która może być sprawnie zastosowana do bardzo dużych rozmiarowo problemów;

mikrodane synonim pojęcia [dane jednostkowe](#);

minimalna kombinacja unikatowa (ang. *Minimal Sample Unique, MSU*) niebezpieczna (z punktu widzenia określonej reguły poufności, np. k -anonimowości) kombinacja wartości kategoryalnych quasi identyfikatorów, której żaden podzbiór właściwy nie stanowi kombinacji niebezpiecznej;

podmiot gospodarki narodowej osoba prawna, jednostka organizacyjna niemająca osobowości prawnej oraz osoba fizyczna prowadząca działalność gospodarczą [Art. 2 pkt 11 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

post-randomizacja (ang. *The Post–Randomization Method, PRAM*) probabilistyczna metoda kontroli ujawniania danych generująca określone zakłócenia; wartości zmiennych kategoryalnych dla pewnych rekordów zostają tutaj zamienione na inne z wykorzystaniem specyficznego mechanizmu probabilistycznego opartego na odpowiednio zdefiniowanej macierzy przejść Markowa, składającej się z prawdopodobieństw zmiany danej kategorii na inną oraz pozostawienia jej w niezmienionym kształcie;

poufność informacji statystycznych oznacza ochronę poufnych danych odnoszących się do każdej jednostki statystycznej, które zostały pozyskane bezpośrednio do celów statystycznych lub pośrednio ze źródeł administracyjnych lub innych, i rozumiana jest jako zakaz wykorzystywania uzyskanych danych do celów innych niż statystyczne oraz ich bezprawnego ujawnienia [Art. 2 pkt 1e rozporządzenia Parlamentu Europejskiego i Rady (WE) nr 223/2009 z dnia 11 marca 2009 r. w sprawie statystyki europejskiej];

przeciętne odchylenie bezwzględne wartość bezwzględna różnicy pomiędzy wartością oryginalną a zaburzoną podzielona przez liczbę komórek lub jednostek wchodzących w skład danego agregatu; stosowane do wyznaczania straty informacji na skutek zastosowania kontroli ujawniania danych, zwłaszcza w przypadku danych zagregowanych;

pseudoidentyfikator (*quasi-identyfikator, zmienna kluczowa, domniemany identyfikator*) zmienna w zbiorze danych, która – w połączeniu z innymi dostępnymi źródłami danych, do których ma dostęp użytkownik owego zbioru – może prowadzić do ujawnienia tożsamości respondenta;

pseudonimizacja przetworzenie danych osobowych w taki sposób, by nie można ich było już przypisać konkretnej osobie, której dane dotyczą, bez użycia dodatkowych informacji, pod warunkiem że takie dodatkowe informacje są przechowywane osobno i są objęte środkami technicznymi i organizacyjnymi uniemożliwiającymi ich przypisanie zidentyfikowanej lub możliwej do zidentyfikowania osobie fizycznej [Art. 4 pkt 5 rozporządzenia Parlamentu Europejskiego I Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych – RODO) (Dz. Urz. UE L 119 z 04.05.2016, str. 1, z późn. zm.)];

rejstry urzędowe rejestry publiczne w rozumieniu ustawy z dnia 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne (t.j. Dz. U. 2023 poz. 57) oraz inne rejestry i ewidencje prowadzone na podstawie ustaw lub wydanych na ich podstawie aktów wykonawczych, zawierające informacje o podmiotach gospodarki narodowej i ich działalności, informacje o osobach fizycznych, ich życiu i sytuacji oraz dotyczące zjawisk, zdarzeń i obiektów [Art. 2 pkt 12 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

relatywne odchylenie bezwzględne wartość bezwzględna różnicy pomiędzy wartością oryginalną a zaburzoną podzielona przez wartość oryginalną; stosowane do wyznaczania straty informacji na skutek zastosowania kontroli ujawniania danych, zwłaszcza w przypadku danych zagregowanych;

respondent osoba fizyczna, osoba prawna lub jednostka organizacyjna niemająca osobowości prawnej, przekazująca dane dla celów statystycznych [Art. 2 pkt 11a ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

ryzyko globalne sumaryczne ryzyko identyfikacji jednostek dla całego zbioru danych jednostkowych; wyznacza się je najczęściej jako sumę lub średnią arytmetyczną ryzyk indywidualnych (czyli ryzyk identyfikacji poszczególnych jednostek);

ryzyko hierarchiczne ryzyko ujawnienia związane z hierarchiczną strukturą danych: jeżeli osoba, której udało się naruszyć tajemnicę statystyczną dokonała poprawnej identyfikacji choćby jednej jednostki niższego poziomu należącej do jednostki poziomu wyższego, to skutkiem wiedzy o tej przynależności może być identyfikacja pozostałych jednostek poziomu niższego wchodzących w skład tejże jednostki poziomu wyższego; ryzyko hierarchiczne wyznacza się zarówno na poziomie indywidualnym, jak i globalnym;

ryzyko indywidualne ryzyko identyfikacji pojedynczej jednostki;

służby statystyki publicznej Prezes Głównego Urzędu Statystycznego, podlegli mu dyrektorzy urzędów statystycznych, jednostki przez niego nadzorowane utworzone na podstawie ustawy z dnia 27 sierpnia 2009 r. o finansach publicznych (Dz. U. z 2021 r. poz. 305, z późn. zm.) oraz jednostki, o których mowa w art. 27 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej. [*Art. 22 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)*];

System Informatyczny Badania (SIB) dedykowany system dostępny dla uprawnionych osób realizujących badanie, służący do gromadzenia, przetwarzania, analizy i kontroli danych pozyskiwanych w badaniu;

System Metadanych Statystycznych (SMS) system służący do zarządzania metadanymi powstającymi podczas projektowania, planowania i realizacji badań statystycznych oraz ich udostępniania [GUS (2015), [Zasady dotyczące organizacji prac i podziału zadań związanych z prowadzeniem Systemu Metadanych Statystycznych – SMS oraz wykorzystaniem metadanych gromadzonych w Repozytorium Metadanych w jednostkach statystyki publicznej](#)];

systemy informacyjne administracji publicznej systemy zbierania, gromadzenia i przetwarzania informacji przez organy administracji publicznej, Zakład Ubezpieczeń Społecznych, Narodowy Fundusz Zdrowia, Komisję Nadzoru Finansowego, organy rejestrowe, inne państwowe lub samorządowe osoby prawne oraz inne podmioty prowadzące rejestry urzędowe [*Art. 2 pkt 13 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)*];

tablica częstości tablica, w której każda komórka reprezentuje licznosc (częstość), czyli liczbę jednostek należących do danej kategorii, tzn. do wymiaru tabeli, który dana komórka obrazuje. Tablice te są typowe dla badań społecznych;

tablica kontrolna tablica robocza, nieprzeznaczona do publikacji, która powstaje na etapie przetwarzania danych;

tablica publikacyjna tablica wynikowa opracowana w układzie odpowiadającym sposobowi prezentacji informacji wynikowych w publikacji, bazie danych lub innym opracowaniu statystycznym, przeznaczona do publikacji lub udostępnienia użytkownikom zewnętrznym, przy zachowaniu tajemnicy statystycznej poprzez zastosowanie odpowiednich metod kontroli ujawniania danych;

tablica statystyczna liczbowy opis struktury badanej zbiorowości, jest formą systematycznego uporządkowania danych liczbowych i stanowi zwięzłą i przejrzystą formę prezentacji danych statystycznych w postaci krzyżujących się kolumn i wierszy; ze względu na etapy pracy w badaniach, wyróżniamy tablice kontrolne, tablice wynikowe (analityczne) i tablice publikacyjne [Zajac K. (1994), Zarys metod statystycznych, PWE, Warszawa];

tablica wielkości tablica statystyczna, w której każda komórka reprezentuje wartość sumaryczną dla cechy ilościowej dla jednostek, które dana komórka reprezentuje;

tablica wynikowa (analityczna) tablica statystyczna, będąca wynikiem ostatecznego opisu badanego zjawiska, która może stanowić podstawę analiz statystycznych;

tablice łączone dwie lub więcej tablic statystycznych, w których jako zmienną grupującą (bądź zmienne grupujące) przyjęto tę samą zmienną (lub te same zmienne), przez co tablice te mają wspólne wszystkie lub niektóre kategorie wymiarów, a więc w tablicach takich prezentowane są dane dla tej samej zmiennej, a wszystkie te tablice mają pewien zakres wspólnych komórek;

tajemnica pierwotna zob. [ukrywanie komórek pierwotne](#);

tajemnica publikacyjna ochrona przed ujawnieniem informacji, stosowana w przypadku badań, w których wprowadzono możliwość występowania do podmiotów gospodarki narodowej (respondentów) o zgodę na publikowanie danych zgodnie z art. 38 ust. 3 ustawy o statystyce publicznej, zachowanie tajemnicy statystycznej jest kontrolowane po uwzględnieniu zgód udzielonych przez respondentów;

tajemnica statystyczna dane jednostkowe identyfikowalne zebrane w badaniach statystycznych podlegają bezwzględnej ochronie; dane te mogą być wykorzystywane wyłącznie do opracowań, zestawień i analiz statystycznych oraz do tworzenia przez Prezesa Głównego Urzędu Statystycznego operatu do badań statystycznych; udostępnianie lub wykorzystywanie tych danych dla innych niż podane w ustawie celów jest zabronione [Art. 10 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

tajemnica wtórna zob. [ukrywanie komórek wtórne](#);

ujawnienie dotyczy faktu poznania przez jednostkę (osobę, podmiot gospodarczy, instytucję, itd.) informacji na temat innej jednostki, do których nie uprawnia jej obowiązujące prawo lub inne względy (zob. Hundepool et al. 2012);

ukrywanie komórek pierwotne (potocznie nazywane **ochroną tajemnicy pierwotnej**) ochrona przed bezpośrednim poznaniem wartości cechy respondenta przez osoby do tego nieuprawnione realizowana zgodnie z art. 38 ust. 2 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej: *Nie mogą być publikowane ani udostępniane uzyskane w badaniach statystycznych dane statystyczne możliwe do powiązania i zidentyfikowania z konkretną osobą fizyczną oraz informacje i dane statystyczne charakteryzujące wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą, jeżeli na daną agregację składa się mniej niż trzy podmioty lub udział jednego podmiotu w określonym zestawieniu jest większy niż trzy czwarte całości.* Zasada ta oznacza konieczność ukrycia pozycji spełniających wymienione warunki z tablic, zbiorów oraz baz przeznaczonych do udostępniania. Zmienne, według których badany jest udział podmiotu w zestawieniu, określa jednostka autorska;

ukrywanie komórek wtórne (potocznie nazywane **ochroną tajemnicy wtórnej**) ochrona przed pośrednim ujawnieniem informacji chronionej, objętej tajemnicą statystyczną, po przeprowadzeniu pierwotnego ukrywania komórek. Ujawnienie takie może nastąpić np. poprzez wykorzystanie związków i zależności pomiędzy poszczególnymi informacjami czy danymi w tablicy lub bazie danych;

użytkownik zewnętrzny osoba lub jednostka spoza jednostek statystyki publicznej zamawiająca lub korzystająca z danych statystycznych;

ważność zmiennej istotność danej zmiennej określana jako jej wkład do globalnego ryzyka ujawnienia informacji wrażliwych z danej bazy;

wynikowe informacje statystyczne wyniki obliczeń, opracowań i analiz dokonanych na podstawie zebranych w badaniach statystycznych statystyki publicznej danych statystycznych [Art. 14 ust. 1 ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz. U. z 2023 r. poz. 773, z późn. zm.)];

zmienne ciągłe zmienne wyrażone na skali różnicowej (przedziałowej, interwałowej) lub ilorazowej; zmienne te nazywa się często także *zmiennymi ilościowymi*;

zmienne hierarchiczne szczególny przypadek zmiennych jakościowych; występują wówczas, gdy odnoszą się do domen, które mają charakter hierarchiczny, zagnieżdżony i które mogą być liniowo uporządkowane według stopnia tego zagnieżdżenia.

zmienne kartograficzne zmienne służące do określania topograficznego położenia obiektów w terenie lub ich identyfikacji w inny sposób przy tworzeniu wykresów mapowych.

zmienne kategoryjne zmienne wyrażone na skali nominalnej lub porządkowej; zmienne te nazywa się często także *zmiennymi jakościowymi*;

zmienne regionalne zmienne przestrzennie lokalizujące respondenta. Mogą dotyczyć zarówno jednostek administracyjnych (makroregion, województwo, region, podregion, powiat, gmina), jak i obszarów funkcjonalnych (np. wyznaczonych na podstawie poziomu urbanizacji czy oddziaływań społeczno-ekonomicznych).

zmienne ważące (*wagi*) specjalne zmienne występujące w zbiorach danych pochodzących z badań reprezentacyjnych, wykorzystywane w procesie estymacji do uogólniania wyników z próby na populację generalną. Są opracowywane na podstawie schematu doboru próby, przy uwzględnieniu poprawek związanych z jednostkowymi brakami odpowiedzi i ewentualnych innych kryteriów.

Skróty

ABM – Analityczna Baza Mikrodanych

Dz. U. – Dziennik Ustaw

ESS – (ang. *European Statistical System*) Europejski System Statystyczny

Euratom – (ang. *European Atomic Energy Community – EAEC*) Europejska Wspólnota Energii Atomowej

Eurostat – (ang. *European Statistical Office*) Urząd Statystyczny Unii Europejskiej

EWG – Europejska Wspólnota Gospodarcza

GSBPM – (ang. *Generic Statistical Business Process Model*) generyczny model procesu produkcji statystycznej

GUS – Główny Urząd Statystyczny

JSP – jednostki statystyki publicznej

MPPS – Model procesu produkcji statystycznej

MSU – (ang. *Minimal Sample Unique*) minimalna kombinacja unikatowa

OBM – Operacyjna Baza Mikrodanych

OBS – operat do badań społecznych

PBSSP – Program Badań Statystycznych Statystyki Publicznej

RODO – rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/ WE (ogólne rozporządzenie o ochronie danych) (Dz. Urz. UE L 199 z 04.05.2016, str. 1, z późn. zm.)

RP – Rzeczpospolita Polska

SDC – (ang. *Statistical Disclosure Control*) kontrola ujawniania danych statystycznych, kontrola ryzyka ujawnienia informacji, które powinny być chronione tajemnicą statystyczną oraz poziomu straty informacji związanych z ukrywaniem lub zniekształcaniem danych wrażliwych

SIB – System Informatyczny Badania

SISP – System Informacyjny Statystyki Publicznej

SMS – System Metadanych Statystycznych

SPDS – System Pozyskiwania, Przetwarzania i Integracji Danych Statystycznych; podstawową rolę SPDS jest zapewnienie jednolitego, spójnego i kompletnego sposobu zasilania hurtowni danych systemu SISP

SUDA – (ang. *Special Uniques Detection Algorithm*) algorytm wykrywania kombinacji unikalnych

TK – Trybunał Konstytucyjny

UE – Unia Europejska

UNECE – (ang. United Nations Economic Commission for Europe) Europejska Komisja Gospodarcza Organizacji Narodów Zjednoczonych

US – urząd statystyczny

WE – Wspólnota Europejska

WOBR – Wojewódzki Ośrodek Badań Regionalnych

1 Wprowadzenie

W *Przewodniku* zaprezentowano wybrane wyniki prac Zespołu ds. metod kontroli ujawniania danych statystycznych (zwanego dalej Zespołem), który został powołany zarządzeniem wewnętrznym nr 21 Prezesa Głównego Urzędu Statystycznego z dnia 19 lipca 2019 r. W poniższej publikacji zebrano najważniejsze informacje teoretyczne i praktyczne, które mogą być przydatne pracownikom polskiej statystyki publicznej. W trakcie prac Zespołu zostały opracowane również dokumenty ilustrujące wyniki prac, które udostępniono w Intranecie link.

W kolejnych rozdziałach *Przewodnika* zawarto zagadnienia związane z kontrolą udostępniania danych statystycznych. Omówiono w nich następujące zagadnienia: cel i założenia dokumentu, podstawy prawne tajemnicy statystycznej, metodologiczne zasady ochrony tajemnicy statystycznej, wytyczne w zakresie ukrywania przekrojów objętych tajemnicą statystyczną w SIB, rekomendacje w zakresie metod i narzędzi ukrywania przekrojów objętych tajemnicą statystyczną, metody oceny ryzyka ujawnienia danych objętych tajemnicą statystyczną oraz metody pomiaru utraty informacji w wyniku ukrycia przekrojów objętych tajemnicą statystyczną.

Ponadto dokument zawiera cztery załączniki zamieszczone na końcu: tabelę z symbolami oznaczania tajemnicy statystycznej stosowanymi przez Eurostat w polskiej i angielskiej wersji językowej oraz trzy wzory dokumentów związanych z pozyskiwaniem zgody na publikację określonych danych charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej.

2 Cel i założenia dokumentu

Przestrzeganie zasad tajemnicy statystycznej jest jednym z fundamentów, na których opiera się funkcjonowanie statystyki publicznej. Jest to kwestia zapewnienia ochrony danych, które krajowym instytucjom statystycznym są przekazywane przez respondentów (obywatele, podmioty gospodarcze, instytucje, jednostki samorządu terytorialnego i inne). Bez zapewnienia ochrony danych trudno mówić o zaufaniu respondentów do urzędu statystycznego, a w konsekwencji do *gwarancji rzetelnego, obiektywnego i systematycznego informowania społeczeństwa, organów państwa i administracji publicznej oraz podmiotów gospodarki narodowej o sytuacji ekonomicznej, demograficznej, społecznej oraz środowiska naturalnego* (art. 3 Ustawy z dnia 29 czerwca 1995 r. o statystyce publicznej, zwanej dalej Ustawą).

Głównym celem dokumentu, z punktu widzenia instytucji, jest zebranie i zaprezentowanie w jednym opracowaniu aktualnych – obecnie rozproszonych po różnych komórkach i jednostkach statystyki publicznej (departamenty Głównego Urzędu Statystycznego, urzędy statystyczne) – wiedzy i doświadczeń, jakimi dysponuje statystyka publiczna w Polsce w odniesieniu do zagadnień związanych z tajemnicą statystyczną oraz kontrolą ujawniania danych statystycznych. To rozproszenie wiedzy jest m.in. konsekwencją specjalizacji badawczej poszczególnych komórek organizacyjnych, która spowodowała wytworzenie odrębnych podejść do zagadnień ochrony tajemnicy statystycznej przez jednostki autorskie poszczególnych badań zawartych w programie badań statystycznych statystyki publicznej, wynikających ze specyfiki tych badań i obejmowanych nimi danych. W *Przewodniku* uwzględniono doświadczenia wynikające ze współpracy międzynarodowej w zakresie m.in. zagadnień związanych z ochroną danych objętych tajemnicą statystyczną i kontrolą jakości danych statystycznych oraz będące wynikiem śledzenia rozwoju metod i technik analitycznych związanych z tą problematyką. Opracowanie oraz wdrożenie jednolitego podejścia do ochrony tajemnicy statystycznej i stosowania metod kontroli ujawniania danych statystycznych, przy uwzględnieniu różnych typów danych i form ich udostępniania, przyczyni się do podniesienia jakości funkcjonowania statystyki publicznej w Polsce, a przede wszystkim zwiększenia bezpieczeństwa danych będących w jej posiadaniu.

Przewodnik stanowi zatem kompendium aktualnej wiedzy dla pracowników służb statystyki publicznej w zakresie poprawnego rozumienia i stosowania zasad ochrony tajemnicy statystycznej. W założeniu Zespołu będzie on w przyszłości aktualizowany i uzupełniany zgodnie z rozwojem wiedzy w tym zakresie. Stanowi on także element podnoszenia kompetencji oraz jakości kadr statystyki publicznej dzięki prezentacji nowoczesnych metod i narzędzi służących zapewnieniu ochrony danych statystycznych objętych tajemnicą statystyczną.

Aspekty prawne związane z tajemnicą statystyczną są szczegółowo przedstawione w trzecim rozdziale *Przewodnika* [Podstawy prawne tajemnicy statystycznej](#), natomiast warto zauważyć w tym miejscu, że dwie spośród wielu metod kontroli ujawniania danych zawarte są wprost w Ustawie. Jedną z nich jest **zasada minimalnej częstości** (*minimum frequency rule*), która stanowi, że komórka tablicy jest uważana jako niebezpieczna, jeśli liczba jednostek w tej komórce jest mniejsza od zdefiniowanej odgórnie częstości granicznej (w przypadku statystyki publicznej w Polsce wynosi ona 3. Z kolei druga z zasad – reguła (n, k) dominacji – stwierdza, że komórka obciążona jest ryzykiem pierwotnym, jeśli n największych podmiotów gospodarczych w komórce tablicy reprezentuje więcej niż $k\%$ całkowitej wartości cechy ilościowej w komórce. Zasady te znajdują odzwierciedlenie w zapisach art. 38 ust. 2 Ustawy (z $k = 3$ w pierwszym oraz $n = 1$ i $k = 75$ w drugim przypadku). Niejednokrotnie zapisy te utożsamiane są z definicją [tajemnicy statystycznej](#), jednak taka interpretacja jest błędna – te zapisy to zasady, których stosowanie eliminuje bądź ogranicza ryzyko ujawnienia danych, które powinny być chronione – są to zatem metody [kontroli ujawniania danych statystycznych](#).

Z ochroną tajemnicy statystycznej związane jest także pojęcie „ujawnienia”. Ujawnienie może nastąpić poprzez dostęp jednostek do danych statystycznych, np. znajdujących się w opublikowanych wynikach badań. Rozróżnia się trzy typy ujawnienia danych:

- identyfikacja, kiedy jednostka zostaje zidentyfikowana w zakresie danych chronionych, objętych tajemnicą statystyczną,
- ujawnienie atrybutu, kiedy na podstawie udostępnionych danych możliwe jest poznanie dodatkowej charakterystyki jednostki,
- ujawnienie poprzez wnioskowanie, jeżeli na podstawie publikowanych danych jest możliwe ustalenie szacunku dotyczącego charakterystyki podmiotu z większą dokładnością niż zakłada to publikacja.

Z pojęciem “ujawnienia” wiąże się nieodłącznie problem ryzyka ujawnienia, czyli prawdopodobieństwa tego, że dana jednostka pozna charakterystykę innej jednostki na podstawie udostępnionych danych, naruszając tym samym jej autonomię w zakresie prawa do ochrony danych. Obecnie już coraz bardziej powszechna jest wiedza, że najprostsza czynność polegająca na usunięciu bezpośrednich identyfikatorów jednostek, takich jak np. imię i nazwisko, numer PESEL, Numer Identyfikacji Podatkowej (NIP), etc. nie redukuje całkowicie prawdopodobieństwa identyfikacji. Współcześnie, w dobie wielu źródeł danych o znacznym rozmiarze i superkomputerów mogących analizować miliony kombinacji zmiennych, to o wiele za mało, a utrwalone w wieloletniej praktyce proste reguły ukrywania okazują się dalece niewystarczające. Dzieje się tak z uwagi na – zazwyczaj znaczną – liczbę zmiennych opisujących dane jednostki, co pociąga za sobą bardzo dużą liczbę możliwych kombinacji wariantów zmiennych wyrażonych na skali nominalnej lub porządkowej. Istnieje zatem spore ryzyko, że wystąpią wśród nich kombinacje unikatowe, w skrajnych sytuacjach odnoszące się tylko do jednej jednostki. Dlatego należy mieć świadomość, że każda publikacja danych statystycznych niesie za sobą pewne prawdopodobieństwo ujawnienia. Stąd ze strony krajowych urzędów statystycznych pojawiło się zapotrzebowanie na narzędzia i metody, które automatyzują proces [SDC](#), pozwalają go parametryzować oraz raportować prawdopodobieństwo ujawnienia

i stratę informacji. Takim remedium są metody kontroli ujawniania danych statystycznych, które zgodnie z definicją Komisji Europejskiej są określane jako narzędzia (metody, techniki) zastosowane na danych, które eliminują bądź redukują ryzyko ujawnienia informacji o analizowanych jednostkach. Należy je rozumieć jako kontrolę ryzyka ujawnienia informacji, które powinny być chronione tajemnicą statystyczną oraz poziomu straty informacji związanych z ukrywaniem lub zniekształcaniem danych wrażliwych.

Warto w tym miejscu zwrócić jeszcze uwagę na trzy aspekty związane z nomenklaturą stosowaną w zakresie tajemnicy statystycznej. Chodzi o sformułowania „**tajemnica pierwotna**”, „**tajemnica wtórna**” oraz „nakładanie tajemnicy statystycznej”. Są to sformułowania powszechnie funkcjonujące w praktyce resortu statystyki i wynikające raczej z kilkudziesięcioletnich doświadczeń operacjonalizowania zapisów użytych w Ustawie, niż wprost z samych jej zapisów. Definicje pojęć *primary/secondary confidentiality*, które można byłoby przyjąć jako odpowiedniki pierwotnej i wtórnej tajemnicy statystycznej, można znaleźć w opracowaniach EUROSTATU (por. (“The Eurostat Concepts and Definitions Database” 2001) oraz (Hundepool et al. 2010)). Dotyczą one danych w komórkach tablic statystycznych, których opublikowanie pozwoliłoby na ujawnienie atrybutów jednostek. Jako dwie główne przyczyny ryzyka ujawnienia atrybutów wskazuje się zbyt niską liczebność jednostek w komórce tablicy, bądź dominację jednej lub dwóch jednostek w komórce. Z pojęciami *primary/secondary confidentiality* wiążą się także pojęcia *primary/secondary suppression* (pierwotnego i wtórnego ukrywania komórek) - oznaczające techniki SDC, które mają zastosowanie w przypadku wystąpienia ryzyka ujawnienia tak zdefiniowanej pierwotnej i wtórnej tajemnicy statystycznej. Należy bowiem zauważyć, że „tajemnica pierwotna” oraz „tajemnica wtórna” są to zmodyfikowane polskie tłumaczenia pierwotnego i wtórnego ukrywania komórek w tablicach kontyngencji (ang. *primary suppression, secondary suppression*). Natomiast sformułowanie „nakładanie tajemnicy statystycznej” stosowane jest w celu opisanie technicznej czynności zaznaczenia tych przekrojów (komórek tablicy statystycznej, agregacji), które powinny być objęte ochroną ze względu na tajemnicę statystyczną. Należy zaznaczyć, że czynność ta wykonywana jest na tych etapach procesu produkcji statystycznej, które poprzedzają etap udostępniania. W tym przypadku Zespół sugeruje stosować sformułowanie „ukrywanie (pierwotne lub wtórne) przekrojów objętych tajemnicą statystyczną”. Uznając zatem pewien zwyczaj istniejący w resorcie statystyki, jeszcze jednym celem niniejszego *Przewodnika* będzie zatem promowanie właściwych określeń związanych z tajemnicą statystyczną i metodami kontroli ujawniania danych statystycznych.

3 Podstawy prawne tajemnicy statystycznej

Z punktu widzenia prawnego, **tajemnica** to określona przez przepisy wiadomość, informacja, dane, których poznanie lub ujawnienie jest zakazane przez prawo. Punktem wyjścia do stosowania tajemnicy statystycznej jest właściwe i efektywne stosowanie uregulowań formalno-prawnych, zarówno krajowych, unijnych, jak i międzynarodowych (np. OECD, ONZ), bowiem to na nich powinien opierać się każdy zorganizowany system statystyczny.

Ustawa z dnia 29 czerwca 1995 r. o statystyce publicznej

Tajemnica statystyczna jest pojęciem wprowadzonym do prawa polskiego art. 10 Ustawy, który stanowi, że *Dane jednostkowe identyfikowalne zebrane w badaniach statystycznych podlegają bezwzględnej ochronie. Dane te mogą być wykorzystywane wyłącznie do opracowań, zestawień i analiz statystycznych oraz do tworzenia przez Prezesa Głównego Urzędu Statystycznego operatu do badań statystycznych; udostępnianie lub wykorzystywanie tych danych dla innych niż podane w ustawie celów jest zabronione (tajemnica statystyczna).*

Konstrukcja tajemnicy statystycznej opiera się zasadniczo na trzech pojęciach zdefiniowanych w art. 2 Ustawy:

- **dane jednostkowe identyfikowalne** (art. 2 pkt 8a Ustawy), które oznaczają *dane statystyczne zawierające informacje dotyczące konkretnego podmiotu gospodarki narodowej albo osoby fizycznej, identyfikujące bezpośrednio ten podmiot albo osobę według nazwy, imienia i nazwiska, adresu lub publicznie dostępnego numeru identyfikacyjnego oraz pozwalające na pośrednią identyfikację tego podmiotu albo osoby z użyciem innych środków niż środki pozwalające na bezpośrednią identyfikację, z wyłączeniem środków wymagających nadmiernych kosztów, czasu lub działań;*
- **badania statystyczne** (art. 2 pkt 2 Ustawy) zdefiniowane jako *zbieranie, gromadzenie i opracowywanie danych statystycznych oraz ogłaszanie i udostępnianie wyników dokonanych obliczeń, opracowań i analiz, w tym podstawowych wielkości i wskaźników;*
- **dane statystyczne** (art. 2 pkt 1a Ustawy), które oznaczają *dane dotyczące zjawisk, zdarzeń, obiektów i działalności podmiotów gospodarki narodowej oraz życia i sytuacji osób fizycznych, w tym dane osobowe, pozyskane bezpośrednio od respondentów albo z systemów informacyjnych administracji publicznej, rejestrów urzędowych lub niepublicznych systemów informacyjnych, od momentu ich zebrania na potrzeby wykonywania zadań statystyki publicznej.*

Przepisy Ustawy opierają system statystyki publicznej na zaufaniu między służbami statystycznymi odpowiedzialnymi za zbieranie i opracowywanie informacji a respondentami. Podejście do tajemnicy statystycznej i samej statystyki publicznej ewoluowało wraz z demokratyzacją życia społecznego oraz wdrażaniem szeregu uregulowań wynikających z umów międzynarodowych. W stosunku do uprzednio obowiązującej ustawy z dnia 26 lutego 1982 r. o statystyce państwowej¹, Ustawa z 1995 r. wprowadziła rozwiązania właściwe dla społeczeństw demokratycznych, łagodząc przymus administracyjny w wypełnianiu obowiązków statystycznych na rzecz społecznego wymiaru statystyki publicznej i postrzegania statystyki jako dobra publicznego. Nawet w ustawie z dnia 26 lutego 1982 r. o statystyce państwowej, w brzmieniu pierwotnym art. 25 dane statystyczne podlegały ochronie: *Publikowanie wyników badań statystycznych oraz dostęp do informacji statystycznych podlegają ograniczeniom wynikającym z ochrony tajemnicy państwowej i gospodarczej oraz sfery praw osobistych obywateli.*

Statystyka publiczna (w tym tajemnica statystyczna) funkcjonuje w krajowym porządku w wielu aktach prawnych, jednak podstawy jej funkcjonowania określone są Ustawą.

Rola i cele statystyki publicznej określa art. 3 Ustawy: *Statystyka publiczna zapewnia rzetelne, obiektywne i systematyczne informowanie społeczeństwa, organów państwa i administracji publicznej oraz podmiotów gospodarki narodowej o sytuacji ekonomicznej, demograficznej, społecznej oraz środowiska naturalnego.* Badania statystyki publicznej mogą dotyczyć każdej dziedziny życia społecznego i gospodarczego oraz występujących w nim zjawisk dających się obserwować i analizować z wykorzystaniem metod statystycznych przy uwzględnieniu celowości, niezbędności oraz społecznej użyteczności danych (art. 4 Ustawy). Służby statystyki publicznej zbierają, gromadzą i opracowują dane od podmiotów gospodarki narodowej i o tych podmiotach oraz ich działalności, a także dane od osób fizycznych i o tych osobach, ich życiu i sytuacji, oraz dane dotyczące zjawisk, zdarzeń i obiektów, przechowują, łączą pozyskane dane i wtórnie je wykorzystują w celu realizacji zadań określonych w Ustawie (art. 5 ust. 1 Ustawy).

W art. 6 ust. 2 Ustawy wprowadzono przepis mający zapewnić możliwość prawidłowego wykonywania zadań statystyki publicznej, stanowiąc, że: *Badania statystyczne mogą być prowadzone na zasadzie obowiązku albo udziału dobrowolnego.* W przypadku, gdy badania statystyczne prowadzone są z udziałem osób fizycznych, z wyłączeniem osób fizycznych prowadzących działalność gospodarczą i spisów powszechnych, są prowadzone na zasadzie udziału dobrowolnego (art. 6 ust. 3 i 4 Ustawy). Udział w badaniach statystycznych, dotyczących działalności gospodarczej prowadzonej przez osoby fizyczne, jest obowiązkowy (art. 6 ust. 4 Ustawy).

W obowiązującym porządku prawnym szczególne miejsce zajmuje obejmowanie tajemnicą statystyczną danych osób fizycznych biorących udział w badaniach statystycznych. Zasady ustanowione w Ustawie głównie osadzone są na normach określonych w rozporządzeniu Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego

¹Dz.U. z 1989 r. Nr 40, poz. 221 z późn. zm.

przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych)², zwanego dalej Rozporządzeniem RODO lub RODO, w powiązaniu z normami wywiedzionymi z Rozporządzenia Parlamentu Europejskiego i Rady (WE) Nr 223/2009 z dnia 11 marca 2009 r. w sprawie statystyki europejskiej oraz uchylające rozporządzenie Parlamentu Europejskiego i Rady (WE, Euratom) nr 1101/2008 w sprawie przekazywania do Urzędu Statystycznego Wspólnot Europejskich danych statystycznych objętych zasadą poufności, rozporządzenie Rady (WE) – nr 322/97 w sprawie statystyk Wspólnoty oraz decyzję Rady 89/382/EWG, Euratom w sprawie ustanowienia Komitetu ds. Programów Statystycznych Wspólnot Europejskich – dalej Rozporządzenie Nr 223/2009.

Zagadnieniu przetwarzania danych osobowych dla celów statystycznych poświęcony został rozdział 4a Ustawy (obejmuje art. 35a-h), poprzedzony wyjątkiem opisanym w art. 8 Ustawy, który zawiera bezwzględny zakaz zbierania na zasadzie obowiązku w badaniach statystycznych prowadzonych z udziałem osób fizycznych danych osobowych, o których mowa w art. 9 ust. 1 i art. 10 RODO tj. *danych osobowych ujawniających pochodzenie rasowe lub etniczne, poglądy polityczne, przekonania religijne lub światopoglądowe, przynależność do związków zawodowych oraz przetwarzania danych genetycznych, danych biometrycznych w celu jednoznacznego zidentyfikowania osoby fizycznej lub danych dotyczących zdrowia, seksualności lub orientacji seksualnej tej osoby oraz danych osobowych dotyczących wyroków skazujących oraz czynów zabronionych lub powiązanych środków bezpieczeństwa.*

Należy zwrócić przy tym uwagę, że zakaz przetwarzania powyższych danych dotyczy tylko zbierania danych osobowych na zasadzie obowiązku w badaniach statystycznych prowadzonych z udziałem osób fizycznych, natomiast nie dotyczy przetwarzania danych osobowych zbieranych przez służby statystyki publicznej od respondentów na zasadzie dobrowolności oraz danych pozyskiwanych z systemów informacyjnych administracji publicznej, rejestrów urzędowych i niepublicznych systemów informacyjnych, przy jednoczesnym zapewnieniu gwarancji ich pełnej ochrony na zasadach określonych w Ustawie (art. 35c ust. 3 Ustawy).

W zakresie nakładania tajemnicy statystycznej rozdział ten wyraźnie wzmacnia i doprecyzowuje uprawnienia i obowiązki służb statystyki publicznej względem danych osób fizycznych zebranych w badaniach statystycznych i ich bezwzględnej ochrony poprzez bezpośrednie wskazanie objęcia danych osobowych tajemnicą statystyczną od momentu ich zebrania (art. 35aa Ustawy: *Dane osobowe od momentu ich zebrania bezpośrednio od respondentów albo z systemów informacyjnych administracji publicznej i rejestrów urzędowych lub niepublicznych systemów informacyjnych na potrzeby wykonywania zadań określonych w ustawie stają się danymi statystycznymi i objęte są tajemnicą statystyczną z wyłączeniem informacji zawartych w krajowym rejestrze urzędowym podmiotów gospodarki narodowej.*).

Dodatkowym wzmocnieniem świadomości co do tajemnicy statystycznej i praw wśród osób fizycznych, które udzielają bezpośrednio odpowiedzi w badaniach statystycznych jest rozwiązanie przyjęte w art. 35h ust. 3 w związku z art. 6 ust. 5 Ustawy. W sytuacji zbierania danych osobowych od osoby, której dane dotyczą, przed rozpoczęciem badania niezbędne jest

²Dz. Urz. UE L 119 z 04.05.2016, str. 1, z późn. zm.

przekazanie informacji istotnych z punktu widzenia procedury zbierania danych i ich dalszego przetwarzania (art. 13 ust. 1 i 2 RODO) razem z informacjami o podstawie prawnej badania, celu badania i gwarancjach zachowania tajemnicy, o której mowa w art. 10, oraz czy udział w badaniu jest obowiązkowy czy dobrowolny.

Jednocześnie w rozdziale 4a Ustawy zawarto cele przetwarzania danych osobowych (art. 35a ust. 1 Ustawy), zakres przetwarzanych danych osobowych (art. 35b ust. 1 Ustawy) oraz źródła przetwarzanych danych (art. 35c ust. 1 Ustawy).

W art. 35h ust. 1 Ustawy ograniczono prawa osób fizycznych poprzez wprowadzenie zastrzeżenia, że do przetwarzania danych osobowych w celu wykonywania zadań przez służby statystyki publicznej nie stosuje się następujących przepisów Rozporządzenia RODO dotyczących:

- art. 15 - prawa do informacji, czy przetwarzane są dane osobowe tej osoby i uzyskania do nich dostępu,
- art. 16 – prawa do sprostowania danych,
- art. 18 – prawa do ograniczenia przetwarzania,
- art. 21 – prawa do wniesienia sprzeciwu wobec przetwarzania dotyczących jej danych osobowych.

W art. 35d Ustawy wprowadzono postanowienia określające narzędzie wzmacniające ochronę danych osobowych i wspomagające tajemnicę statystyczną. Pseudonimizacja danych to pojęcie wprowadzone do prawa za pomocą art. 4 pkt 5 Rozporządzenia RODO. Pseudonimizacja jest odwracalnym procesem, który jest przeprowadzany w celu ograniczenia możliwości identyfikacji danej osoby. Dane poddane procesowi pseudonimizacji dalej podlegają ochronie prawnej. W zakresie statystyki publicznej *dane osobowe po ich zebraniu przez służby statystyki publicznej, jeżeli pozwala na to cel ich przetwarzania, są **pseudonimizowane***. Sposób i tryb pseudonimizacji danych osobowych określa, w drodze zarządzenia, Prezes Głównego Urzędu Statystycznego.

Z tajemnicą statystyczną związane jest udostępnianie i przechowywanie informacji statystycznych, o czym stanowi rozdział 5 Ustawy (art. 36 – art. 39).

Prowadzący badania statystyczne statystyki publicznej obowiązani są ogłaszać, udostępniać i rozpowszechniać wynikowe informacje statystyczne z zachowaniem przepisów o tajemnicy statystycznej (art. 14 Ustawy)³. Wynikowe informacje statystyczne są powszechnie dostępne, a ich rodzaje, formy i terminy ich udostępnienia są ustalane w Rozporządzeniu Rady Ministrów dotyczącego programu badań statystycznych statystyki publicznej.

Konsekwentnie do brzmienia definicji tajemnicy statystycznej, zawartej w art. 10 Ustawy, wprowadzono w zakresie publikowania i udostępniania, zgodnie z art. 38 Ustawy, zakaz publikowania i udostępniania danych jednostkowych identyfikowalnych uzyskanych w badaniach statystycznych. Dodatkowo, w celu zapewnienia zachowania poufności w rozumieniu przepisów wspólnotowych oraz tajemnicy statystycznej w rozumieniu prawa krajowego, przy publikacji

³Zob. również art. 21 Ustawy.

i udostępnianiu uzyskanych w badaniu statystycznym danych, których ochrona powinna być wzmocniona, doprecyzowano zakres objęcia tajemnicą statystyczną, wprowadzając w art. 38 ust. 2 Ustawy poniższy zakaz.

Zakaz dotyczy publikowania i udostępniania danych umożliwiających powiązanie i zidentyfikowanie konkretnej osoby fizycznej oraz informacji i danych statystycznych charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą, jeżeli na daną agregację składa się mniej niż trzy podmioty lub udział jednego podmiotu w określonym zestawieniu jest większy niż trzy czwarte całości.

Inne zastrzeżenie w kwestii udostępniania danych dotyczy operatu do badań statystycznych (OBS). Operat to wykaz objętych badaniami statystycznymi podmiotów gospodarki narodowej, osób fizycznych oraz innych podmiotów podlegających obserwacji statystycznej uporządkowany według określonych cech, zawierający dane jednostkowe wraz z ich identyfikacją teleadresową (art. 2 pkt 8 Ustawy).

W art. 39c ust. 2 Ustawy, wskazano ograniczenia w udostępnianiu (objęcie tajemnicą statystyczną) danych gromadzonych w OBS. W przypadku badania statystycznego prowadzonego przez inny wskazany naczelny lub centralny organ administracji państwowej albo Narodowy Bank Polski, o ile badanie to wymaga wiedzy specjalistycznej, którą dysponuje ten organ, albo jeżeli jest on w posiadaniu danych zgromadzonych w innym celu, Prezes Głównego Urzędu Statystycznego, na wniosek prowadzącego badanie udostępnia mu z OBS dane adresowe podmiotów wylosowanych lub wybranych na podstawie wskazanych przez niego cech, z wyłączeniem danych adresowych osób fizycznych nieprowadzących działalności gospodarczej.

Istotnym dla zachowania tajemnicy statystycznej jest przypisanie Prezesowi Głównego Urzędu Statystycznego obowiązku zapewnienia przechowywania zgromadzonych danych zgodnie z zasadami określonymi w Ustawie (art. 39 Ustawy).

Ochronie tajemnicy statystycznej służy również art. 12 Ustawy, według którego: *Pracownicy służb statystyki publicznej, rachmistrze spisowi, ankieterzy statystyczni oraz inne osoby wykonujące czynności w imieniu i na rzecz statystyki publicznej, mający bezpośredni dostęp do danych indywidualnych i danych osobowych, są obowiązani do bezwzględnego przestrzegania tajemnicy statystycznej i mogą być dopuszczeni do wykonywania tych czynności po złożeniu w urzędzie statystycznym albo innej jednostce organizacyjnej służb statystyki publicznej pisemnego przyrzeczenia.*

Obowiązek respektowania zasady ochrony poufności statystycznych danych jednostkowych (tj. umożliwiających identyfikację respondenta) został nadto obwarowany sankcją karną, co wskazuje na znaczenie prawne tajemnicy statystycznej.

W art. 54 Ustawy naruszenie tajemnicy statystycznej spenalizowane zostało karą do trzech lat pozbawienia wolności.

W art. 55 Ustawy ustanowiono zaś przestępstwo polegające na wykorzystaniu w celu osiągnięcia korzyści majątkowej lub osobistej danych statystycznych, z którymi sprawca zapoznał

się w związku z wykonywaniem pracy lub odpowiedniego zlecenia, zagrożone sankcją do 5 lat pozbawienia wolności.

Zastrzec należy, że tajemnica statystyczna jako obowiązek ustawowy, nie może być postrzegana jedynie przez pryzmat samej definicji zawartej w art. 10 Ustawy, bowiem na jej charakter prawny wpływają nie tylko przepisy prawa krajowego na czele z ustawą zasadniczą (Konstytucją Rzeczypospolitej Polskiej), ale też normy europejskie i międzynarodowe, które w związku z art. 91 ust. 3 Konstytucji Rzeczypospolitej Polskiej stają elementem prawa krajowego⁴.

Konstytucja Rzeczypospolitej Polskiej z dnia 2 kwietnia 1997 r.⁵

- **art. 2 Konstytucji RP, który stanowi, że Rzeczpospolita jest demokratycznym państwem prawnym.**

Klauzula państwa prawnego wynikająca z art. 2 Konstytucji RP, dotyczy między innymi zasady zaufania obywatela do państwa. Zasada zaufania uznana została za fundament państwa prawnego. Według Trybunału Konstytucyjnego zasada zaufania obywateli do państwa i stanowionego przez nie prawa, nazywana także zasadą lojalności państwa względem obywateli i podmiotów, jest jedną z najważniejszych zasad pochodnych, wynikających z zasady demokratycznego państwa prawnego⁶.

W kontekście tajemnicy statystycznej konstytucyjna zasada zaufania do państwa i stanowionego przez nie prawa polega na zagwarantowaniu każdemu respondentowi pewności, że jego dane nie będą zbierane, gromadzone, przetwarzane i udostępniane na warunkach innych aniżeli wynika to z Ustawy tj. z zachowaniem tajemnicy statystycznej, skoro warunki zapewnienia tajemnicy zostały określone w Ustawie.

- **art. 7 Konstytucji RP**, w myśl którego organy władzy publicznej działają na podstawie i w granicach prawa. W najbardziej ogólnym ujęciu w art. 7 Konstytucji RP ustanowiono konstytucyjny nakaz działania organów władzy publicznej na podstawie i w granicach prawa. Kompetencji organów nie można domniemywać, a co nie jest im dozwolone, jest zabronione. To, co jest nakazane, powinno być zrealizowane. W kontekście art. 7 Konstytucji RP na szczególną uwagę zasługują przepisy dotyczące przestrzegania tajemnicy statystycznej (art. 10, art. 39, art. 39a, art. 39c Ustawy).
- **art. 31 ust. 2 i 3 Konstytucji RP**, który stanowi, że *ograniczenia w zakresie korzystania z konstytucyjnych wolności i praw mogą być ustanawiane tylko w ustawie i tylko wtedy, gdy są konieczne w demokratycznym państwie dla jego bezpieczeństwa lub porządku publicznego, bądź dla ochrony środowiska, zdrowia i moralności publicznej, albo wolności i praw innych osób. Ograniczenia te nie mogą naruszać istoty wolności i praw.*

⁴Konstytucja RP art. 91 ust. 3 Jeżeli wynika to z ratyfikowanej przez Rzeczpospolitą Polską umowy konstytuującej organizację międzynarodową, prawo przez nią stanowione jest stosowane bezpośrednio, mając pierwszeństwo w przypadku kolizji z ustawami.

⁵Dz.U. z 1997 r. Nr 78, poz. 483.

⁶zob. np. wyr. TK z 20.12.1999 r., K 4/99, OTK 1999, Nr 7, poz. 165.

Przykładem ograniczeń może być obowiązek uczestniczenia obywateli w spisie powszechnym na podstawie ustawy spisowej (art. 9 Ustawy), co nie zmienia faktu, że obywatel ma prawo oczekiwać od państwa proporcjonalności działań oraz stosownej ochrony.

- **art. 47 Konstytucji RP**, w myśl którego każdy ma prawo do ochrony prawnej życia prywatnego i rodzinnego.

Nie ulega wątpliwości, że z istoty gwarancji określonej w art. 47 Konstytucji RP, z jej ścisłego związku z godnością człowieka oraz z użytego pojęcia “każdy” wynika, że dotyczy ona wszystkich osób fizycznych⁷.

Życie prywatne to przymioty, wewnętrzne przeżycia osobiste (jednostkowe) człowieka i ich oceny, refleksje dotyczące wydarzeń zewnętrznych i jego wrażenia zmysłowe, a także stan zdrowia⁸ oraz sytuacja majątkowa⁹. Nie są one w założeniu przeznaczone do upowszechniania i sam zainteresowany decyduje o kręgu osób, z którymi zechce się nimi podzielić. Samo życie prywatne jako element prawa do prywatności zaczęło dopiero niedawno odgrywać poważniejszą rolę w regulacjach konstytucyjnych i orzecznictwie sądowym. Od dłuższego czasu wywierało jednak coraz silniejszy wpływ na pozycję jednostki we współczesnych państwach demokratycznych. Tak rozumiane życie prywatne odnosi się generalnie do życia: osobistego, towarzyskiego, nienaruszalności mieszkania, tajemnicy korespondencji i ochrony informacji dotyczących danej osoby.

Życie rodzinne, aczkolwiek związane z prywatnym, dotyczy już szerszej sfery przeżyć i zdarzeń związanych z rodziną pojętą nie tylko jako zespół osób prowadzących z danym człowiekiem wspólne gospodarstwo domowe.

Ustawa tę ochronę ogranicza, co jest dopuszczalne, jednakże pod warunkiem zapewnienia anonimowości ingerencji dzięki tajemnicy statystycznej.

W tym zakresie Ustawa gwarantuje respondentom prawa oparte na dobrowolności uczestniczenia w badaniach statystycznych (art. 6 ust. 3 i 5 w związku z art. 10 oraz art. 8 Ustawy) z wyłączeniem spisów powszechnych (art. 9 Ustawy) oraz pewność, że dane podane w badaniach będą objęte bezwzględną tajemnicą statystyczną.

Jako pewne uproszczenie należy jednak uznać pogląd, jakoby tylko osoby fizyczne były zdolne do posiadania życia prywatnego, czci i dobrego imienia w rozumieniu wskazanego przepisu.

Gwarancje art. 47 Konstytucji RP są bowiem rozciągane również na jednostki organizacyjne. Już w pierwszym wyroku poświęconym prawu do prywatności TK, oceniając zgodność z Konstytucją RP ingerencji w tajemnicę bankową, przyznawał *każdemu podatnikowi prowadzącemu działalność gospodarczą* prawo do pewnej prywatności tej działalności, o ile nie ma ona na

⁷tak np. Banaszak B., , Konstytucja Rzeczypospolitej Polskiej. Komentarz, Warszawa, 2012, komentarz do art. 47, nb. 1.

⁸wyr. TK z 19.05.1998 r., U 5/97, OTK 1998, Nr 4, poz. 46.

⁹orz. TK z 24.06.1997 r., K 21/96, OTK 1997, Nr 2, poz. 23.

celu ukrycia działalności przestępczej lub uchylenia się od określonych zobowiązań publiczno-prawnych¹⁰. Pojęcie podatnika obejmuje zarówno osoby fizyczne jak i podmioty niebędące osobami fizycznymi.

- **art. 51 ust. 2 Konstytucji RP**, w myśl którego władze publiczne nie mogą pozyskiwać, gromadzić i udostępniać innych informacji o obywatelach niż niezbędne w demokratycznym państwie prawnym. Zakres podmiotowy gwarantowanego w art. 51 ust. 2 Konstytucji RP prawa dotyczy obywateli – z czego można wywieść wniosek, iż władze publiczne mogą przetwarzać informacje dotyczące innych osób i dane poddane przetworzeniu nie muszą spełniać jednocześnie warunku niezbędności w demokratycznym państwie prawnym (np. uchodźcy, cudzoziemcy).

Sformułowanie *informacje o obywatelach (...) niezbędne w demokratycznym państwie prawnym* można określić jako te dane, które umożliwiają normalne funkcjonowanie jednostki w zorganizowanym w państwo społeczeństwie. Zaś z punktu widzenia organów władz publicznych za **niezbędną** uznać należy każdą informację, bez posiadania której nie będą one zdolne do podjęcia (czy zakończenia) działań w ramach przyznanych im kompetencji (np. art. 4 ust. 2 oraz rozdziały 2 i 4a Ustawy).

Przetwarzanie i gromadzenie informacji o osobach spełni warunek konstytucyjności, jeżeli: nie powstaje ryzyko wykorzystywania tych danych także do celów pozaustawowych lub celów pozaumownych, gdy chodzi o umowę międzynarodową, której stroną jest RP. Informacje powinny być chronione przed niepowołanym dostępem lub przed wykorzystywaniem w celach niezgodnych z prawem i pierwotnym celem ich pozyskania, a standardy ich przetwarzania powinny być zgodne z przepisami prawa.

W ocenie TK nie jest zakazane zbieranie informacji, nawet bez wiedzy osoby, której te informacje dotyczą, aczkolwiek sposób i zakres pozyskiwanych informacji musi sprostać testowi *konieczności w demokratycznym państwie prawnym*¹¹. Usprawiedliwione jest stosowanie środków niezbędnych w tym sensie, że chronią wartości istotne w państwie demokratycznym, i to w sposób (bądź w stopniu), który nie mógłby zostać osiągnięty przy stosowaniu innych środków (np. art. 5a Ustawy).

Art. 51 ust. 2 Konstytucji RP wprowadza powołany wymóg niezbędności. Pozyskanie (zbieranie) informacji o osobie jest działaniem państwa, które jest wyłomem w sferze wolności jednostki, a takie działanie musi podlegać ścisłej kontroli pod względem proporcjonalności ograniczeń, zakazu naruszania istoty prawa, a także określoności normy kompetencyjnej zezwalającej na pozyskiwanie informacji o obywatelach (art. 2, art. 31 ust. 2, art. 51 Konstytucji RP). Niezbędność musi być jednak rozumiana ściśle, a więc dane osobowe muszą być warunkiem sine qua non wykonania obowiązku ciążącego na władzy publicznej.

Prawo Unii Europejskiej

¹⁰orz. TK z 24.06.1997 r., K 21/96, OTK 1997, Nr 2, poz. 23.

¹¹orz. z 23.06.2009 r., K 54/07, OTK 2009, Nr 6A, poz. 86.

Akty prawne wydawane na poziomie UE wpływają na krajowy porządek prawny państw członkowskich zależnie od ich charakteru – albo wchodząc do niego wprost albo też wymuszając podjęcie działań prawnych, w tym krajowych, w nich określonych. Zasady ogólnie nakreślające charakter poszczególnych narzędzi prawnych UE zawiera art. 288 Traktatu o funkcjonowaniu Unii Europejskiej¹².

Instrumentem najgłębszej ingerencji prawodawcy wspólnotowego w porządku prawne państw członkowskich są rozporządzenia. Mają one zasięg ogólny, generalny. Obowiązują w całości i są stosowane bezpośrednio we wszystkich państwach członkowskich.

Dla Trybunału Sprawiedliwości Unii Europejskiej „zasięg ogólny” jest równoznaczny z normatywnością aktu, co potwierdza jego status, jako źródła prawa wspólnotowego. Oznacza to, że państwa członkowskie nie mogą podejmować żadnych środków krajowych ustawodawczych, które byłyby sprzeczne z rozporządzeniem. Związanie rozporządzeniem staje się aktualne z chwilą jego wejścia w życie.

Cechą swoistą rozporządzeń jest to, że są one bezpośrednio stosowalne. Oznacza to, że akty te ze swej natury stają się częścią krajowych systemów prawnych bez potrzeby dokonywania jakichkolwiek czynności transpozycyjnych i wywierają skutki bezpośrednie w stosunku do jednostek (tak jak omawiane niżej Rozporządzenie Nr 223/2009 i Rozporządzenie RODO). Jeśli unijny akt prawny upoważnia państwa członkowskie do podejmowania środków wykonawczych, to sposób wykonywania ma określać prawo publiczne państwa członkowskiego.

Mimo, że kluczowym dla statystyki europejskiej jest rozporządzenie 223/2009 kwestię z nią związane regulują też inne unijne akty prawne, odnoszące się również do zasady poufności (np. wspomniane wcześniej RODO).

Rozporządzenie Parlamentu Europejskiego i Rady (WE) Nr 223/2009 z dnia 11 marca 2009 r. w sprawie statystyki europejskiej oraz uchylające rozporządzenie Parlamentu Europejskiego i Rady (WE, Euratom) nr 1101/2008 w sprawie przekazywania do Urzędu Statystycznego Wspólnot Europejskich danych statystycznych objętych zasadą poufności, rozporządzenie Rady (WE) – nr 322/97 w sprawie statystyk Wspólnoty oraz decyzję Rady 89/382/EWG, Euratom w sprawie ustanowienia Komitetu ds. Programów Statystycznych Wspólnot Europejskich¹³– Rozporządzenie Nr 223/2009.

Zgodnie z Traktatem ustanawiającym Wspólnotę Europejską, a w szczególności z jego art. 285 ust. 2, Polska jako członek Unii Europejskiej przyjęła na siebie zobowiązanie stosowania

¹²Art. 288 TFUE W celu wykonania kompetencji Unii instytucje przyjmują rozporządzenia, dyrektywy, decyzje, zalecenia i opinie. Rozporządzenie ma zasięg ogólny. Wiąże w całości i jest bezpośrednio stosowane we wszystkich Państwach Członkowskich. Dyrektywa wiąże każde Państwo Członkowskie, do którego jest kierowana, w odniesieniu do rezultatu, który ma być osiągnięty, pozostawia jednak organom krajowym swobodę wyboru formy i środków. Decyzja wiąże w całości. Decyzja, która wskazuje adresatów, wiąże tylko tych adresatów. Zalecenia i opinie nie mają mocy wiążącej.

¹³Dz.Urz. UE L 2009 Nr 87, str. 164.

praw obowiązujących we Wspólnocie, w tym bezwzględnego przestrzegania tajemnicy statystycznej.

W art. 3 Rozporządzenia Nr 223/2009 zawarto kluczowe w zakresie poufności danych statystycznych definicje:

- **dane statystyczne:** ilościowe i jakościowe, zagregowane i reprezentatywne informacje opisujące złożone zjawisko w rozpatrywanej populacji,
- **opracowywanie danych:** działania zmierzające do utworzenia, usprawnienia i udoskonalenia metod, standardów i procedur statystycznych stosowanych przy tworzeniu i rozpowszechnianiu danych statystycznych, jak również do projektowania nowych danych i wskaźników,
- **tworzenie informacji statystycznych:** wszelka działalność związana ze zbieraniem, przechowywaniem, przetwarzaniem, zestawianiem i analizą, niezbędnymi do zestawienia danych statystycznych,
- **rozpowszechnianie danych:** działanie, którego celem jest udostępnianie danych statystycznych i analiz statystycznych użytkownikom,
- **zbieranie danych:** badania oraz wszelkie inne formy pozyskiwania informacji z różnych źródeł, w tym ze źródeł administracyjnych,
- **jednostka statystyczna:** podstawowa jednostka objęta obserwacją – mianowicie osoba fizyczna, gospodarstwo domowe, podmiot gospodarczy czy inny podmiot – do której odnoszą się dane,
- **dane poufne:** dane umożliwiające bezpośrednią lub pośrednią identyfikację jednostek statystycznych, co skutkuje ujawnieniem informacji indywidualnych. Aby określić, czy możliwa jest identyfikacja danej jednostki statystycznej, należy wziąć pod uwagę wszystkie przewidywalne środki, które mogą być użyte przez osobę trzecią do zidentyfikowania jednostki statystycznej,
- **wykorzystanie do celów statystycznych:** wykorzystanie wyłącznie do celów opracowywania i tworzenia wyników oraz analiz statystycznych,
- **identyfikacja bezpośrednia:** identyfikacja jednostki statystycznej według jej nazwy lub adresu lub według publicznie dostępnego numeru identyfikacyjnego,
- **identyfikacja pośrednia:** identyfikacja jednostki statystycznej z użyciem wszelkich innych środków niż środki identyfikacji bezpośredniej.

Kwestia poufności danych statystycznych została doprecyzowana w rozdziale piątym Rozporządzenia Nr 223/2009. Stanowi on, że krajowe urzędy statystyczne i inne organy odpowiedzialne za opracowywanie, tworzenie i rozpowszechnianie statystyki europejskiej oraz Eurostat dla zapewnienia, że poufne dane są wykorzystywane wyłącznie do celów statystycznych, a także w celu zapobieżenia ich bezprawnemu ujawnieniu, stosują następujące zasady i środki:

- wykorzystują poufne dane uzyskane wyłącznie do celów tworzenia statystyki europejskiej wyłącznie do celów statystycznych, chyba że jednostka statystyczna wyrazi jednoznaczną

zgoda na wykorzystanie tych danych do innych celów - art. 20 ust. 1 Rozporządzenia Nr 223/2009,

- wyniki statystyczne, dzięki którym możliwa jest identyfikacja jednostki statystycznej, mogą być rozpowszechniane w następujących wyjątkowych przypadkach (art. 20 ust. 3 Rozporządzenia Nr 223/2009):
 - jeżeli szczególne warunki i tryb zostały określone w akcie prawnym Parlamentu Europejskiego i Rady działających zgodnie z art. 251 Traktatu, a wyniki statystyczne wprowadzane są w taki sposób, że ich rozpowszechnianie nie stanowi uszczerbku dla poufności informacji statystycznych,
 - lub jeżeli jednostka statystyczna w jednoznaczny sposób wyraziła zgodę na ujawnienie tych danych,
- w zakresie odpowiednich kompetencji podejmują wszelkie niezbędne środki regulacyjne, administracyjne, techniczne i organizacyjne działania w celu zapewnienia ochrony fizycznej i logicznej poufnych danych (kontrola ujawniania danych statystycznych)¹⁴ – art. 20 ust. 4 Rozporządzenia Nr 223/2009,
- stosują wszelkie środki konieczne do zapewnienia harmonizacji zasad i wytycznych dotyczących fizycznej i logicznej ochrony poufnych danych - art. 20 ust. 4 Rozporządzenia Nr 223/2009,
- urzędnicy i inny personel, mający dostęp do poufnych danych, obowiązani są przestrzegać wymogu poufności nawet po zakończeniu pełnienia swoich funkcji - art. 20 ust. 5 Rozporządzenia Nr 223/2009.

Rozporządzenie przewiduje możliwość przyznawania prawa dostępu do poufnych danych pracownikom naukowym prowadzącym analizy statystyczne do celów prowadzonych badań - art. 23 Rozporządzenia Nr 223/2009. Warunkiem przyznania dostępu do badań jest tylko pośrednia identyfikacja jednostek statystycznych, uzyskiwana w praktyce poprzez pozbawienie jednostkowych cech umożliwiających bezpośrednie zidentyfikowanie osób, których dane dotyczą np. imię, nazwisko, numer identyfikacyjny ewidencji ludności, adres zamieszkania. Również w tej sytuacji winny być podjęte działania uniemożliwiające pośrednią identyfikację jednostki statystycznej. Należy również zwrócić uwagę, że w przypadku, gdy dane zostały przekazane Komisji (Eurostatowi), wymagana jest zgoda krajowych urzędów statystycznych (KUS) lub innych organów krajowych, które dostarczyły te dane. Uwzględniając Rozporządzenia Nr 223/2009 w szczególności jego art. 23 wydane zostało Rozporządzenie Komisji (UE) Nr 557/2013 z dnia 17 czerwca 2013 r. w sprawie wykonania rozporządzenia (WE) nr 223/2009 Parlamentu Europejskiego i Rady w sprawie europejskiej statystyki w zakresie dostępu do

¹⁴Pojęcie kontrola ujawniania danych statystycznych zostało formalnie zdefiniowane dopiero w art. 2 Rozporządzenia Komisji (UE) Nr 1151/2010 z dnia 8 grudnia 2010 r. w sprawie wykonania rozporządzenia Parlamentu Europejskiego i Rady (WE) nr 763/2008 w sprawie spisów powszechnych ludności i mieszkań w odniesieniu do ustaleń dotyczących raportów jakości i ich struktury oraz formatu technicznego przekazywania danych (Dz.Urz. UE L Nr 324, str. 1). Według tego rozporządzenia kontrola ujawniania danych statystycznych oznacza metody i procesy zastosowane w celu zminimalizowania ryzyka ujawnienia informacji na temat poszczególnych jednostek statystycznych, przy jednoczesnym udostępnianiu jak największej ilości informacji statystycznych.

poufnych danych do celów naukowych i uchylające rozporządzenie Komisji (WE) nr 831/2002¹⁵. Dokument ten określa warunki przyznawania dostępu do poufnych danych przekazywanych w celach naukowych, a także zasady współpracy pomiędzy Unią a władzami krajowymi w tym zakresie.

Ponadto, Rozporządzenie Nr 223/2009 w art. 25 statuuje, że *dane pozyskane ze źródeł, które zgodnie z prawem są ogólnodostępne i które pozostają ogólnodostępne zgodnie z przepisami krajowymi, nie są uznawane za poufne dla celów rozpowszechniania danych statystycznych uzyskanych na podstawie tych danych.*

Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych)¹⁶, zwane dalej Rozporządzeniem RODO lub RODO.

Rozporządzenie RODO nie wprowadza pełnej harmonizacji w zakresie regulacji zasad ochrony danych osobowych rozumianej jako zupełne ukształtowanie regulacji przedmiotowej bez dopuszczalności lub konieczności stosowania regulacji krajowych. Na gruncie RODO mamy do czynienia z harmonizacją częściową, w przypadku której państwa członkowskie są uprawnione do swobodnej, jednostronnej regulacji w zakresie nieunormowanym przepisami RODO, przy jednoczesnym zakazie stosowania rozwiązań sprzecznych z RODO.

Przepisy RODO mają w większości charakter prawnomaterialny, których przyjęcie jest wymagane przez państwa członkowskie, dotyczą kwestii nieuregulowanych w RODO, a więc przepisów dotyczących organów nadzorczych oraz kwestii proceduralnych.

Analizując przepisy prawa krajowego przyjęte na podstawie Rozporządzenia RODO, w pierwszej kolejności należy wskazać na regulacje obligatoryjne. Oznacza to, że każde z państw członkowskich, w tym Rzeczpospolita Polska, zostało zobowiązane do przyjęcia krajowych przepisów regulujących ustrój organów nadzorczych oraz zasady postępowania w sprawach uregulowanych przepisami RODO, które to przepisy umożliwiają stosowanie RODO w systemie prawnym danego państwa członkowskiego. Przyjmowane przepisy krajowe muszą przestrzegać zasad niezależności i pozostawać w zgodzie z regulacjami Rozporządzenia RODO.

Drugą grupę przepisów krajowych, których przyjęcie zostało przewidziane w RODO, stanowią przepisy, których przyjęcie przez państwa członkowskie jest fakultatywne. W tym zakresie wyróżnia się trzy kategorie tego rodzaju przepisów:

- odesłania, w których RODO wprost przewiduje możliwość regulowania przez państwo członkowskie określonych kwestii w stosunku do tych zawartych w RODO (upoważnienia przewidujące możliwość wydania szczególnych przepisów krajowych);

¹⁵Dz.Urz. UE L 2013 Nr 164, str. 16, European Communities (Wspólnoty Europejskie), <http://eur-lex.europa.eu/>

¹⁶Dz. Urz. UE L 119 z 04.05.2016, str. 1, z późn. zm.

- odesłania, w których od regulacji prawa krajowego RODO uzależnia obowiązywanie przepisu RODO lub zawężenie stosowania przepisu RODO;
- odesłania określające dopuszczalny zakres regulacji krajowej co do podstaw przetwarzania danych osobowych, które doprecyzowują RODO.

Przypadków, w których państwa członkowskie mogą przyjąć przepisy krajowe dotyczące ochrony danych osobowych, należy poszukiwać w szczególności w Rozdziale IX Rozporządzenia RODO. Są to sytuacje dotyczące:

- przetwarzania danych osobowych w kontekście wolności wypowiedzi i informacji;
- przetwarzania danych zawartych w dokumentach urzędowych udostępnianych publicznie;
- przetwarzania krajowego numeru identyfikacyjnego;
- przetwarzania danych osobowych w kontekście zatrudnienia;
- przetwarzania danych osobowych do celów archiwalnych w interesie publicznym, do celów badań naukowych lub historycznych lub do celów statystycznych;
- przetwarzania danych osobowych objętych tajemnicą zawodową¹⁷;
- przetwarzania danych osobowych przez kościoły i związki wyznaniowe.

Poza przypadkami, w których przyjęcie krajowych przepisów o ochronie danych osobowych wynika z konieczności zapewnienia stosowania RODO, w ustawie z 10 maja 2018 r. o ochronie danych osobowych można znaleźć wiele przepisów, co do których decyzja o ich przyjęciu nie jest związana bezpośrednio z rozpoczęciem stosowania przepisów RODO.

Rozpoczęcie stosowania RODO wymusiło także zmiany wielu aktów prawnych dotyczących przetwarzania danych osobowych. Wynikają one czy to z prostej zmiany nazw poszczególnych instytucji prawa ochrony danych osobowych, ze zmian ustrojowych i proceduralnych, czy też z konieczności uwzględnienia konkretnych regulacji RODO, które – jako przepisy prawa wspólnotowego – mają pierwszeństwo przed przepisami prawa krajowego.

Europejski Kodeks Praktyk Statystycznych

Istotne znaczenie dla podstaw prawnych zachowania poufności w statystyce publicznej ma tzw. miękkie prawo (ang. *soft law*). Miękkie prawo są to pokrewne prawu instrumenty, które nie posiadają żadnej mocy prawnie wiążącej albo moc wiążąca jest w jakiś sposób „słabsza” w porównaniu do mocy wiążącej tradycyjnego prawa określanego jako „twarde prawo” (ang. *hard law*).

Termin „miękkie prawo” jest także często używany w odniesieniu do różnych rodzajów quasi-prawnych instrumentów Unii Europejskiej takich jak „kodeksy postępowania” (ang. *codes of conduct*), „kodeksy praktyk” (ang. *codes of practice*), „wytycznych” (ang. *guidelines*), „komunikatów” (ang. *communications*) itp.

W domenie prawa Unii Europejskiej, instrumenty „miękkiego prawa” są często używane w celu zasygnalizowania, w jaki sposób Komisja Europejska będzie używała swoich kompetencji, w

¹⁷Zob. art. 89 Rozporządzenia RODO.

jaki sposób Komisja będzie wykonywała swoje zadania w ramach przysługującego jej prawa. Kluczym instrumentem „soft law” w zakresie statystyki publicznej jest Europejski Kodeks Praktyk Statystycznych, przyjęty przez Komitet do spraw Europejskiego Systemu Statystycznego w dniu 28 sierpnia 2011 r.¹⁸ - dalej zwany Kodeksem. Jego moc została wzmocniona przez przepisy rozporządzenia 223/2009, a wdrożenie zasad objętych Kodeksem w krajach unijnych podlega cyklicznemu monitoringowi (*peer reviews*).

Do Kodeksu na gruncie prawa krajowego odwołuje się art. 25a Ustawy, który statuuje zasady realizacji zadań Prezesa Głównego Urzędu Statystycznego. Wykonując zadania określone w Ustawie, Prezes GUS kieruje się zasadami zgodnymi z Kodeksem wskazanym w art. 11 Rozporządzenia Nr 223/2009.

Kodeks opiera się na dwóch zasadniczych filarach wywodzących się z rozporządzenia 223/2009: ochronie poufności danych (tj. nieujawnianie danych indywidualnych, zminimalizowanie ryzyka ich odtworzenia z mikrodanych zanonimizowanych lub zbiorczych) i wykorzystywania danych jednostkowych wyłącznie do celów statystycznych lub naukowych, które również stanowią podwaliny Ustawy.

Kodeks składa się z 15 zasad. W warstwie zasadniczej o poufności statystyk i ochronie danych traktuje Zasada 5. W tym obszarze zasada ta głosi, że należy bezwzględnie zagwarantować prywatność dostawcom danych, poufność informacji przez nich przekazywanych, wykorzystanie tych informacji wyłącznie do celów statystycznych, a także bezpieczeństwo danych. W tym celu przepisy prawa powinny gwarantować poufność i bezpieczeństwo danych zarówno w zakresie przepisów prawnych a także , jak i warunków technicznych i fizycznych ochrony. Stąd też analogicznie do art. 12 ust. 1 Ustawy, pracownicy statystyki podejmując zatrudnienie podpisują zobowiązanie do zachowania poufności statystyk i ponoszą odpowiedzialność karną za naruszenia poufności. Kodeks wskazuje również na konieczność zapewnienia pracownikom w zakresie poufności i ochrony danych dostępności do prawnych instrukcji i wytycznych dotyczących w procesie ich przetwarzania i udostępniania statystyk i danych. Jednocześnie przyjęto, że informacje o polityce zapewnienia poufności winny być podawane do wiadomości publicznej. Kodeks zakłada także, że obowiązują przepisy regulujące kwestie fizycznego i technicznego bezpieczeństwa i integralności statystycznych baz danych oraz protokoły, które regulują dostęp do danych statystycznych dla użytkowników zewnętrznych wykorzystujących je do celów naukowych.

Do tych postulatów nawiązują również postanowienia zawarte w zasadzie 14 Kodeksu. W zasadzie 14 (spójność i porównywalność) zalecono, że zastosowanie odpowiednich mechanizmów udostępniania danych pochodzących z różnych źródeł w przypadku ich łączenia i zbiorczego wykorzystania.

¹⁸Europejski Kodeks Praktyk Statystycznych dla krajowych i wspólnotowych organów statystycznych. Przyjęty przez Komitet ds. Europejskiego Systemu Statystycznego 28 sierpnia 2011 r., znowelizowany w 2017 r. <https://ec.europa.eu/eurostat/documents/4031688/9394304/KS-02-18-142-PL-N.pdf/7ad9346a-5019-44ee-bb16-c7b7dfdb8779?t=1542711876000>

Z kolei zasada 15 (dostępność i przejrzystość) stanowi między innymi, że można zezwolić na dostęp do danych dla celów badawczych (dostęp ten winien jednak być uregulowany ścisłymi zasadami lub protokołami – punkt 15.4).

4 Metodologiczne zasady ochrony tajemnicy statystycznej

4.1 Miejsce, rola i znaczenie ochrony tajemnicy statystycznej w przyjętym Modelu Procesu Produkcji Statystycznej

Model Procesu Produkcji Statystycznej (MPPS) to krajowa implementacja generycznego modelu procesu produkcji statystycznej – ang. *Generic Statistical Business Process Model* (GSBPM)¹ – wzbogacona o komponenty geoprzestrzenne (oznaczone na poniższym schemacie na żółto), podprocesy weryfikacji i oceny jakości (oznaczone na zielono) oraz podprocesy planowania (oznaczone na niebiesko). GSBPM jest uznawany za branżowy standard wypracowany przez UNECE i wdrażany na świecie w ponad 80 krajach przez instytucje statystyki publicznej. GSBPM opiera się na podejściu procesowym, charakteryzującym się elastycznością i innowacyjnością. Koncentruje się na całych procesach, a nie na wybranych aspektach funkcjonowania organizacji. Model Procesu Produkcji Statystycznej umieszczony w załączniku do zarządzenia wewnętrznego Prezesa GUS w sprawie ustanowienia Polityki postępowania z danymi statystycznymi jest następujący:

Źródło: opracowanie własne na podstawie Polityki postępowania z danymi statystycznymi (Załącznik do zarządzenia wewnętrznego nr 32 z dnia 4 grudnia 2020 r. Prezesa GUS w sprawie ustanowienia Polityki postępowania z danymi statystycznymi (z późn. zm.)).

W powyższym modelu czerwonymi obwódkami zaznaczono podprocesy produkcji statystycznej, w których, zdaniem Zespołu, konieczność uwzględnienia metod kontroli ujawniania danych statystycznych jest najbardziej oczywista. Ponadto metody te mogą znaleźć swoje miejsce również m.in. w podprocesach 1.7, 1.8, 1.9, 2.3, 2.6, 2.7, 2.8, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 5.1, 5.2, 5.8, 6.1, 6.2, 6.3, 7.1, 7.2a, 7.5, 7.6, 7.7, 8.1, 8.2, 8.3.

Podkreślić także należy, że GSBPM umieszcza obecnie uwzględnienie metod kontroli ujawniania danych w następujących jego fazach: projektowania, przetwarzania, analizy i udostępniania. Ma to także odzwierciedlenie w polskiej wersji modelu procesu produkcji statystycznej, który dla fazy 2. – **PROJEKTOWANIE** wskazuje m.in. na następujące działania:

- zaprojektowanie rozwiązań i narzędzi informatycznych pozwalających na wytworzenie informacji wynikowych (ust. 3),

¹Generic Statistical Business Process Model GSBPM (Version 5.1, January 2019) <https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1> (dostęp 16.12.2022).

Figure 4.1: Model Procesu Produkcji Statystycznej (MPPS) z oznaczonymi wybranymi podprocesami zawierającymi elementy metod kontroli ujawniania danych statystycznych.

1 Specyfikacja potrzeb	2 Projektowanie	3 Budowa	4 Zbieranie danych	5 Przetwarzanie	6 Analiza	7 Udostępnianie	8 Ocena
1.1 Identyfikacja potrzeb	2.1 Projektowanie produktów wyjściowych	3.1 Ponowne użycie lub budowa narzędzi zbierania danych	4.1 Utworzenie wykazu i wybór próby	5.1 Integracja danych	6.1 Opracowanie wyników wstępnych	7.1 Uaktualnienie systemów udostępniania	8.1 Zbieranie ocen
1.1a Sformułowanie programu badań statystycznych	2.2 Projektowanie opisów zmiennych	3.2 Ponowne użycie lub budowa komponentów przetwarzania i analizy	4.1a Geokodowanie wykazu i próby losowej	5.2 Klasyfikacja i kodowanie	6.2 Walidacja wyników	7.2 Utworzenie produktów udostępniania	8.2 Przeprowadzenie oceny
1.2 Konsultacje i potwierdzanie potrzeb	2.3 Projektowanie zbierania danych	3.3 Ponowne użycie lub budowa komponentów udostępniania	4.2 Konfigurowanie zbierania danych	5.3 Przegląd i walidacja	6.2a Opracowanie analiz przestrzennych i map	7.2a Zarządzanie analizami przestrzennymi i mapami przy pomocy GIS	8.3 Przygotowanie planu naprawy
1.3 Określanie zestawu danych wynikowych	2.4 Projektowanie wykazu i próby losowej	3.4 Konfiguracja procesów pracy	4.3 Zbieranie danych	5.4 Edycja i imputacja	6.3 Przegląd i objaśnienia danych wynikowych	7.3 Zarządzanie rozpowszechnianiem udostępnianych produktów	
1.4 Identyfikacja pojęć	2.5 Projektowanie przetwarzania i analizy	3.5 Testowanie systemu produkcyjnego	4.3a Geokodowanie zebranych danych	5.5 Wyprowadzanie nowych zmiennych i jednostek	6.4 Nałożenie tajemnicy statystycznej	7.4 Promocja udostępnianych produktów	
1.5 Sprawdzenie dostępności danych	2.5a Projektowanie geokodowania wykazu, próby oraz zbieranych danych	3.6 Testowanie procesu produkcji statystycznej	4.4 Zakończenie zbierania danych	5.6 Wylczanie wag	6.5 Zakończenie opracowania produktów wynikowych	7.5 Zarządzanie wsparciem użytkowników	
1.6 Przygotowanie i zgłoszenie uzasadnienia biznesowego	2.6 Projektowanie systemu produkcyjnego i procesów pracy	3.7 Zakończenie budowy systemu produkcyjnego	4.5 Weryfikacja planów	5.7 Wylczanie agregatów		7.6 Badanie satysfakcji	
1.7 Opracowanie planu ogólnego realizacji procesu produkcji	2.7 Projektowanie wskaźników jakościowych	3.8 Weryfikacja projektu i planów budowy		5.8 Utworzenie końcowych zbiorów danych		7.7 Przygotowanie planu naprawczego dla kolejnego cyklu badań	
1.8 Opracowanie szczegółowych planów dla kolejnych faz	2.8 Weryfikacja planu ogólnego i planów szczegółowych						
1.9 Weryfikacja koncepcji w celu poprawy satysfakcji							

- opracowywanie założeń, w celu budowy lub modyfikacji narzędzi informatycznych wspierających realizację badań, zawierających m.in. wyszczególnienie kategorii danych występujących w badaniu (inaczej status danych), np. chronione w trakcie opracowania dane jednostkowe (identyfikowalne, nieidentyfikowalne), dane osobowe, a także podanie wytycznych dotyczących ochrony poszczególnych kategorii danych, np. podstawę prawną ochrony danych, algorytmy służące do zapewnienia ochrony tajemnicy statystycznej (ust. 4),
- dokonanie rozpoznania i doboru właściwej metody bądź metod przez jednostkę autorską, zapewniających stosowanie zasad tajemnicy statystycznej w danych wynikowych planowanych do udostępnienia, w celu zapewnienia poufności danych statystycznych, w tym wskazanie zmiennych charakteryzujące wyniki ekonomiczno-finansowe, według których badany jest udział podmiotu w agregacie, a także poprzedzenie wyboru ww. metody bądź metod przez jednostkę autorską oceną ryzyka ujawnienia informacji poufnych z jednoczesną oceną ponoszonej straty informacji (ust. 5).

W obecnie obowiązującej *Polityce postępowania z danymi...* fazy 5 i 6 **PRZETWARZANIE** i **ANALIZA** są opisane łącznie i spośród szeregu działań podejmowanych w tych fazach, istotne znaczenie z punktu widzenia ochrony tajemnicy statystycznej mają m. in.:

- pseudonimizacja danych (ust. 2),
- ochrona danych osobowych (ust. 3),
- stosowanie reguł zapewniających zachowanie tajemnicy statystycznej (ust. 7),
- zapewnienie zachowania tajemnicy statystycznej poprzez stosowanie odpowiednich metod kontroli ujawniania danych przez jednostkę autorską, a także prawidłową ochronę tajemnicy statystycznej zgodnie z wymogami jednostki autorskiej przez jednostkę publikującą dane, która nie jest jednostką autorską badania (ust. 15),
- dostosowanie stosowanych metod zapewniających zachowanie tajemnicy statystycznej i zakresu ich wykorzystania do sposobu i celu udostępniania danych (ust. 16),
- ocena ryzyka ujawnienia informacji podlegających tajemnicy oraz ocena straty informacji danych wynikowych ponoszonej w wyniku zastosowania metod kontroli ujawniania danych (ust. 16),
- pokazanie obok oryginalnych wartości, odpowiedniego oznaczenia danych po zastosowaniu metod kontroli ujawniania danych, jeżeli w zbiorach danych przeznaczonych dla użytkowników wewnętrznych zostały one zastosowane (ust. 16),
- przekazanie przez jednostkę autorską badania jednostce naliczającej dane wszelkich dodatkowych wytycznych metodologicznych dotyczących realizowanego naliczenia, w tym w szczególności dotyczących zastosowanych reguł służących ochronie tajemnicy statystycznej w przypadku wystąpienia konieczności dodatkowego naliczenia danych na zbiorze krajowym (tj. w przekrojach innych, niż określone w założeniach do systemu informatycznego badania) (ust. 29),
- naliczenie zbiorów przez jednostkę realizującą prace w oparciu o art. 21 ust. 2 ustawy o statystyce publicznej dla działań dotyczących przygotowania zbiorów poniżej szczebla administracyjnego – województwo oraz statystycznego – NUTS 2 oraz przyjęcie przez

nią odpowiedzialności w zakresie zastosowanych zasad i metod ochrony tajemnicy statystycznej, a także zachowanie spójności z danymi na wyższych poziomach agregacji przestrzennej udostępnianymi przez jednostkę autorską (ust. 30).

Działania związane z ochroną tajemnicy statystycznej podejmowane w fazie 7 – **UDOSTĘPNIANIE** obejmują stosowanie zasad² rozpowszechniania wynikowych informacji statystycznych opierających się na:

- Rozporządzeniu Parlamentu Europejskiego i Rady (WE) Nr 223/2009 z dnia 11 marca 2009 r. w sprawie statystyki europejskiej,
- Europejskim Kodeksie Praktyk Statystycznych (EKPS)³,
- Ustawie.

W kontekście stosowania metod kontroli ujawniania danych należy także wskazać na ważne znaczenie jednego z działań opisanych w fazie 8 – **OCENA**. Chodzi tutaj o zastosowanie mierników jakości do pomiaru, oceny oraz monitorowania jakości badań statystycznych w służbach statystyki publicznej. Wśród tych mierników szczególne miejsce zajmują mierniki straty informacji ponoszonej na skutek zastosowania kontroli ujawniania danych (ust. 2 pkt 2).

W odniesieniu do przywołanego powyżej modelu procesu produkcji statystycznej należy zaznaczyć, że wymienione powyżej działania mogą obejmować nie tylko pierwotne i wtórne ukrywanie wrażliwych przekrojów tablic, ale także zastosowanie metod zakłóceńowych oraz kontrolę ujawniania wyników analiz prezentowanych w różnych formach (w tym analiz przestrzennych i map). Inne podejście będzie także konieczne w przypadku danych jednostkowych, inne w przypadku danych zagregowanych prezentowanych w tablicach statystycznych, a jeszcze inne w przypadku komentarzy analitycznych, infografik, czy kartogramów (przy zachowaniu spójności wyników wykonanego ukrycia danych prezentowanych w różnych formach – tablicach czy np. formie graficznej).

4.2 Zasady zachowania tajemnicy statystycznej poprzez zastosowanie odpowiednich metod kontroli ujawniania danych statystycznych stosowane w statystyce publicznej

4.2.1 Zasady ogólne

W praktyce za przygotowanie publikowanych i udostępnianych użytkownikom zewnętrznym wynikowych informacji statystycznych, zgodnie z zasadami zachowania tajemnicy statystycznej, o których mowa w Ustawie, odpowiadają jednostki autorskie (w zakresie pierwotnego

²Przewodnik rozdział 3. Podstawy prawne tajemnicy statystycznej.

³Europejski Kodeks Praktyk Statystycznych: https://bip.stat.gov.pl/files/gfx/bip/pl/defaultstronaopisowa/667/1/1/europejski_kodeks_praktyk_statystycznych_2.pdf (dostęp 27.12.2022).

i wtórnego ukrywania danych udostępnianych przez te jednostki). Oznaczanie tajemnicy statystycznej lub ukrywanie przekrojów objętych tajemnicą dla danych przekazywanych do Eurostatu jest wykonywane zgodnie z odpowiednimi rozporządzeniami i wytycznymi UE. Udostępnianie danych przez jednostki autorskie innym jednostkom i komórkom organizacyjnym GUS zostało opisane w rozdziale 4.2.2.

Jednostki służb statystyki publicznej, inne niż jednostka autorska, które przygotowują do publikowania i udostępnienia wynikowe informacje statystyczne, dla których wcześniej nie wykonano kontroli ujawniania danych statystycznych (inne niż otrzymane od jednostki autorskiej wynikowe informacje statystyczne z oznaczoną tajemnicą statystyczną do celów publikacyjnych, o których mowa w rozdziale 4.2.2.), wyznaczając przekroje objęte tajemnicą statystyczną, uwzględniają ogólne zasady stosowania metod kontroli ujawniania danych statystycznych zawarte w niniejszym dokumencie (pkt 4.2) oraz stosują się do zasad wprowadzanych przez jednostki autorskie w SMS. Ponadto biorą pod uwagę wyniki kontroli ujawniania danych statystycznych przeprowadzonej dla agregatów w podstawowych zestawach danych publikowanych przez jednostki autorskie (określonych w PBSSP), zarówno np. w publikacjach, jak i w systemach udostępniania danych. W razie braku informacji wystarczających do przeprowadzenia kontroli ujawniania danych statystycznych lub odpowiedniego punktu odniesienia w publikacjach jednostki autorskiej, jednostka służb statystyki publicznej, która planuje opublikowanie wynikowych informacji statystycznych może zwrócić się z prośbą o pomoc lub konsultacje w tym zakresie do jednostki autorskiej.

Kontrola ujawniania danych statystycznych dla danych imputowanych jest wykonywana na zasadach przyjętych dla danych, które nie są imputowane.

Kontrola ujawniania danych statystycznych (czy dane podlegają pierwotnemu lub wtórnemu ukrywaniu) jest przeprowadzana dla wszystkich agregatów, w których występuje ryzyko ujawnienia danych podlegających ochronie tajemnicy statystycznej, bez względu na poziom agregacji i stosowane jednostki miary.

W badaniach zjawisk występujących tylko w części populacji osób fizycznych i podmiotów gospodarki narodowej objętych danym badaniem (np. nakłady inwestycyjne w badaniu bieżących wyników finansowych i nakładów na środki trwałe) lub występujących okresowo, analiza kontroli ujawniania danych statystycznych dla danego agregatu powinna objąć zbiorowość wszystkich osób lub podmiotów objętych badaniem w tym agregacie, a nie tylko te osoby lub podmioty, w których zjawisko wystąpiło.

Dane dla zmiennej np. „Liczba podmiotów gospodarki narodowej” lub „Liczba osób fizycznych”, informującej o liczebności zbiorowości, której dotyczą poszczególne zmienne, mogą pozostać jawne, o ile informacja o liczbie jednostek w szczegółowo określonym przekroju nie prowadzi do ujawnienia danych podlegających ochronie tajemnicy statystycznej, w tym wyników ekonomiczno-finansowych podmiotu prowadzącego działalność gospodarczą lub pozwalających na zidentyfikowanie konkretnej osoby fizycznej.

Wskaźniki dynamiki oraz wskaźniki ilorazowe (np. dynamika produkcji sprzedanej przemysłu, wskaźnik rentowności) mogą być publikowane, jeśli oba elementy (licznik i mianownik) wyko-

rzystane do obliczeń danego wskaźnika nie podlegają ochronie przed ujawnieniem, tzn. nie można publikować wskaźnika, jeśli licznik lub mianownik, lub zarówno licznik, jak i mianownik, są objęte tajemnicą statystyczną (podlegają pierwotnemu lub wtórnemu ukrywaniu).

W przypadku wskaźników struktury, wskaźniki te mogą być publikowane, jeśli wartość ogółem oraz wszystkie elementy składowe nie są objęte tajemnicą statystyczną.

W przypadku danych dotyczących szeregów czasowych skumulowanych (tzw. danych narastających) należy uwzględniać wyniki kontroli ujawnienia danych statystycznych wykonanej w poprzednich okresach objętych kumulacją, w tym w zakresie tajemnicy podlegającej wtórnemu ukrywaniu.

Dane charakteryzujące wyniki ekonomiczno-finansowe działalności podmiotów gospodarczych prowadzących działalność gospodarczą udostępniane są zgodnie z art. 38 ust. 2. Oznacza to zakaz udostępniania jednostkowych danych identyfikowalnych i nieidentyfikowalnych charakteryzujących wyniki ekonomiczno-finansowe działalności podmiotów gospodarczych prowadzących działalność gospodarczą. Zasady te stosuje się również do udostępniania informacji o produkcji wyrobów w ujęciu wartościowym. Zmienne/zestawy danych, które w poszczególnych badaniach charakteryzują wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą w rozumieniu art. 38 ust. 2 Ustawy, są wskazane w SMS.

Dane dotyczące działalności podmiotów gospodarki narodowej prowadzących działalność gospodarczą odnoszące się do rzeczowych wyników działalności, jako dane niestanowiące charakterystyki wyników ekonomiczno-finansowych, nie są obejmowane powyższymi zasadami kontroli ujawniania danych statystycznych, ale – jak każde dane statystyczne – są objęte ochroną przed ujawnieniem. W związku z powyższym, powinny być chronione przed ujawnieniem, ale nie podlegają kryteriom zapisanym w art. 38 ust. 2 Ustawy. Dane te w postaci agregatów mogą więc być udostępniane bez względu na liczbę podmiotów wchodzących w skład agregatu lub występowanie dominacji na poziomie ogólnopolskim, ale bez prezentacji informacji identyfikujących respondenta (takich jak np. nazwa, adres, numer identyfikacyjny, itp.). Jednak aby zachować ochronę tych danych statystycznych przed ujawnieniem (bezpośrednim lub pośrednim), w ujęciu terytorialnym, gdzie ryzyko identyfikacji producenta jest bardziej znaczące, stosowane są kryteria zapisane w art. 38 ust. 2 Ustawy. Zmienne/zestawy danych, które w poszczególnych badaniach nie charakteryzują wyników ekonomiczno-finansowych podmiotów gospodarki narodowej prowadzących działalność gospodarczą w rozumieniu art. 38 ust. 2 Ustawy, są wskazane w SMS.

Mogą być rozpowszechniane dane z zakresu międzynarodowego handlu towarami umożliwiające identyfikację importera lub eksportera o ile podmiot nie zgłosi braku zgody na ich ujawnianie, zgodnie z art. 19 rozporządzenia Parlamentu Europejskiego i Rady (WE) 2019/2152 z dnia 27 listopada 2019 r. w sprawie europejskiej statystyki gospodarczej: *Wyłącznie na wniosek importera lub eksportera towarów krajowy organ statystyczny decyduje o tym, czy rozpowszechniać wyniki statystyczne odnoszące się do odpowiedniego przywozu lub wywozu bez jakiegokolwiek zmiany lub czy – na uzasadniony wniosek tego importera lub eksportera – zmienić*

wyniki statystyczne, tak aby uniemożliwić identyfikację tego importera lub eksportera w celu zapewnienia poszanowania zasady poufności informacji statystycznych, zgodnie z art. 20 ust. 3 lit. a) rozporządzenia (WE) nr 223/2009.

Dane z systemów informacyjnych administracji publicznej i rejestrów urzędowych lub niepublicznych systemów informacyjnych, wykorzystywane w danym badaniu statystycznym, podlegają zasadom kontroli ujawniania danych statystycznych przyjętym dla tego badania i danego rodzaju zmiennych. Jeśli wskazane powyżej zasady tego wymagają, dla wynikowych informacji statystycznych przygotowanych z wykorzystaniem danych z systemów informacyjnych administracji publicznej i rejestrów urzędowych lub niepublicznych systemów informacyjnych wykonywana jest kontrola ujawniania danych statystycznych (z uwzględnieniem pierwotnego i wtórnego ukrywania przekrojów). Zasady kontroli ujawniania danych statystycznych dla otrzymanych danych są ustalane przez jednostkę autorską badania, w którym są one wykorzystywane.

Jeśli jednostki statystyki publicznej prowadzą w ramach PBSSP badania wspólne wraz z jednostkami spoza statystyki publicznej, wówczas zasady kontroli ujawniania danych statystycznych powinny być w spójny sposób stosowane również przez jednostki spoza statystyki publicznej przy opracowaniu i publikacji wyników tych badań.

Jeśli jednostka zewnętrzna występuje o udostępnienie danych zagregowanych, opracowania lub analizy dotyczącej przedstawionej przez nią listy jednostek i dane te dotyczą zmiennych, które są obejmowane kontrolą ujawniania danych statystycznych, to takie informacje nie mogą być udostępnione dla pełnej listy jednostek przekazanej przez zamawiającego. Mogą być udostępniane agregaty opracowane tylko dla podzbioru przedstawionej listy, a nie dla 100% wskazanych na niej jednostek. W przypadku danych charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą, dane naliczone dla powyższego podzbioru nie powinny zawierać informacji dotyczących mniej niż trzech jednostek z przedstawionej listy ogółem lub agregatu, w którym udział jednego podmiotu wynosi więcej niż 75% całości na poziomie ogółem. Ponadto, zamawiający nie może być informowany, których jednostek z listy dane dotyczą (lub nie dotyczą). Przy udostępnianiu zamawiającemu powyższych agregatów trzeba wykonać kontrolę ujawniania danych statystycznych (z uwzględnieniem pierwotnego i wtórnego ukrywania przekrojów).

Użytkownik zewnętrzny otrzymuje dane, w tym przygotowane na zamówienie, po przeprowadzeniu kontroli ujawniania danych statystycznych (tj. z wyłączeniem z udostępnianych zbiorów danych objętych pierwotnym lub wtórnym ukrywaniem). Wskazane jest poinformowanie zamawiającego już na etapie opiniowania zamówienia, że kontrola ujawniania danych statystycznych może wpłynąć na zakres udostępnianych danych oraz podanie szacunku oczekiwanego poziomu straty informacji powstałej na skutek tej kontroli. Jeżeli strata informacji została określona i przekazana zamawiającemu dla danych z wcześniejszej edycji udostępniania danych i nie zaszły żadne zmiany w zakresie udostępnianych danych ani w stosowanej metodologii kontroli ujawniania danych, wówczas można odstąpić od ponownego szacowania przewidywanego poziomu straty informacji.

Szczegółowe informacje (poza sformułowanymi w Ustawie) o parametrach metod kontroli ujawniania danych statystycznych stosowanych przez statystykę publiczną, które mogłyby być wykorzystane przez użytkowników zewnętrznych do ujawnienia tajemnicy statystycznej, nie są udostępniane użytkownikom zewnętrznym. Nie powinna być także udostępniana informacja dotycząca szacunków ryzyka ujawnienia danych podlegających ochronie. Użytkownikowi zewnętrznemu należy udostępnić informację o stracie informacji wynikającej ze stosowanej kontroli ujawniania danych statystycznych. Jeżeli ryzyko utraty informacji zostało określone dla danych wcześniejszej edycji udostępniania danych i nie zaszły żadne zmiany w zakresie udostępnianych danych ani w stosowanej metodologii kontroli ujawniania danych, wówczas można odstąpić od ponownego naliczania ryzyka utraty informacji.

Odrębne zasady postępowania z danymi statystycznymi regulują, oprócz Ustawy, dokumenty określające współpracę GUS z Narodowym Bankiem Polskim na podstawie art. 21a i art. 21b Ustawy, rozporządzenie w sprawie statystyki europejskiej i rozporządzenie w sprawie europejskiej statystyki gospodarczej (EBS) oraz inne zobowiązania wobec organizacji międzynarodowych.

4.2.2 Zasady kontroli ujawniania zagregowanych danych statystycznych opracowywanych w tablicach

Przygotowując założenia do tablic, jednostka autorska zobowiązana jest do określenia założeń do kontroli ujawniania danych statystycznych dla źródłowych danych przewidzianych przez tę jednostkę autorską do udostępniania (a przynajmniej dla opracowywanych tablic publikacyjnych oraz tych tablic wynikowych – przekazywanych innym jednostkom służb statystyki publicznej – które mają być następnie bezpośrednio opublikowane). Założenia do kontroli ujawniania danych statystycznych powinny dotyczyć wszystkich danych objętych tajemnicą statystyczną (nie tylko charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą). W szczególności, jednostka autorska:

- określa (zgodnie z informacjami wprowadzonymi w SMS) zmienne w badaniu, które są objęte szczególną ochroną przed ujawnieniem (np. wskazuje zmienne charakteryzujące wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą),
- w przypadku zmiennych charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą przyjmuje założenia do kontroli ujawniania danych statystycznych zgodne z art. 38 ust. 2 Ustawy (na tych zasadach oparto poniższe wytyczne),
- w przypadku pozostałych danych statystycznych stosuje się (jest to zalecane) do zasad zawartych w art. 38 ust. 2 Ustawy (mniej niż 3 podmioty i dominacja więcej niż 75%), jednak możliwe jest wyznaczenie przez jednostkę autorską innych zasad, właściwych dla

specyfiki tego rodzaju danych, ale jednocześnie zapewniających ochronę danych statystycznych objętych tajemnicą statystyczną przed ujawnieniem (informacja o przyjętych, w tym odmiennych zasadach powinna zostać zawarta w SMS),

- w przypadku zmiennych charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej (oraz innych zmiennych, gdy jest to niezbędne do przeprowadzenia kontroli ujawniania danych statystycznych według przyjętych dla tych zmiennych zasad) – określa wytyczne do oznaczania agregatów zawierających dane pochodzące z mniej niż 3 jednostek (1 lub 2) – podlegające pierwotnemu ukryciu,
- w przypadku zmiennych charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej (oraz innych zmiennych, gdy jest to niezbędne do przeprowadzenia kontroli ujawniania danych statystycznych według przyjętych dla tych zmiennych zasad) – określa zmienne (jedną lub więcej), według których badane jest wystąpienie dominacji (w związku z ukryciem pierwotnym) w agregatach danego zbioru ze źródłowymi danymi oraz podaje algorytmy do badania dominacji; oznaczane są agregaty zawierające wartość pochodzącą od jednej jednostki, której udział w agregacie, wyrażony w procentach i wyliczany z dokładnością do dwóch miejsc po przecinku, jest większy niż 75,00% całego agregatu. Jeśli na podstawie takiej zmiennej/zmiennych zostanie ustalona konieczność ukrycia danego agregatu, to w ramach tego agregatu ochroną objęte są wszystkie zmienne, które (zgodnie z poniższymi zasadami ogólnymi i informacjami zawartymi w SMS) podlegają ochronie przed ujawnieniem. Możliwe jest wskazanie przez jednostkę autorską wybranej jednej lub kilku zmiennych, dla których sprawdzana jest dominacja, i ta dominacja jest podstawą do objęcia ochroną pozostałych zmiennych dla tego agregatu, przy założeniu, że informacja o dominacji dla wskazanej zmiennej/zmiennych jest właściwa dla całego zestawu zmiennych, które na tej podstawie są obejmowane ochroną. W ramach danego agregatu nie są obejmowane ochroną zmienne, które zgodnie z zasadami i informacjami w SMS nie podlegają ochronie przed ujawnieniem lub są obejmowane ochroną na innych zasadach (np. dane ilościowe o produkcji).
- w przypadku stwierdzenia, dla zmiennej będącej podstawą badania dominacji, wartości zerowych dla wszystkich podmiotów w agregacie, odstępuje od badania dominacji dla tego agregatu.

W poniższej tabeli 1. przedstawiono przykład dla agregatu obejmującego 3 podmioty:

Figure 4.2: Tabela 1. Przykład badania dominacji w przypadku wystąpienia wartości zerowych dla wszystkich podmiotów w agregacie

Podmiot	Wartość obrotu w tys. zł	Uwagi
X	0	Wartość obrotu jest dla wszystkich podmiotów w agregacie równa 0, więc przyjmuje się, że żaden z podmiotów nie jest dominujący.
Y	0	
Z	0	

- w przypadku stwierdzenia wartości ujemnej w danych jednostkowych będących podstawą badania dominacji, na potrzeby obliczenia dominacji wykorzystuje wartości bezwzględne. W poniższej tabeli 2. przedstawiono przykład dla agregatu obejmującego 3 podmioty:

Figure 4.3: Tabela 2. Przykład badania dominacji w przypadku wystąpienia wartości ujemnych w danych jednostkowych

Podmiot	Wartość obrotu w tys. zł	Uwagi
X	23	Wartość obrotu podmiotu Z jest ujemna, ale dominująca w tym agregacie. Udział = $400/(400+23+38) = 86,8\%$
Y	38	
Z	-400	

- określa ukrycie wtórne – wytyczne do wyznaczania agregatów, które należy dodatkowo objąć ochroną ze względu na możliwość odkrycia oznaczonej na pierwszym etapie tajemnicy podlegającej pierwotnemu ukryciu,
- określa zasady kontroli ujawnienia danych statystycznych po uwzględnieniu zgód uzyskanych od respondentów – podmiotów gospodarki narodowej – na publikowanie danych (art. 38 ust. 3 Ustawy), a także wprowadza odpowiednie oznaczenia komórek objętych ostatecznie ochroną przed ujawnieniem w tablicach (w tym wprowadza oznaczenia tajemnicy publikacyjnej – wynikającej z Ustawy po uwzględnieniu zgód udzielonych przez respondentów); powyższe zasady są przygotowywane w przypadku danych, dla których wykorzystywane jest pozyskiwanie zgód od respondentów; tryb postępowania w celu pozyskania zgody na opublikowanie określonych danych charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej został opisany w rozdziale 4.2.4. niniejszego *Przewodnika*,
- określa sposób symbolizowania komórek podlegających ochronie przed ujawnieniem, z uwzględnieniem przyczyn objęcia ochroną (np. wyróżnienie rodzajów ukrycia: podlegające pierwotnemu ukryciu – mniej niż 3 jednostki; podlegające pierwotnemu ukryciu – dominacja; podlegające wtórnemu ukryciu, tajemnica publikacyjna), wraz z opisem stosowanych oznaczeń. Symbole te są wykorzystywane tylko na potrzeby wewnętrzne statystyki publicznej (do oznaczania danych zagregowanych dostępnych tylko dla jednostek służb statystyki publicznej) i są stosowane tylko w tablicach, w których są widoczne wszystkie dane, łącznie z danymi objętymi ochroną (te dane objęte ochroną są oznaczone symbolami wskazanymi przedstawionymi w tabeli 3.). Przed udostępnieniem informacji z tablic dla użytkowników zewnętrznych, wszystkie dane oznaczone jako objęte ochroną powinny zostać ukryte (użytkownicy zewnętrzni nie mogą widzieć ani danych objętych ochroną, ani powyższych, wykorzystywanych wewnętrznie w statystyce symboli oznaczeń rodzajów tajemnicy statystycznej). Dostosowanie się do zamieszczonej poniżej listy symboli (tabela 3 .) jest zalecane w miarę możliwości. Jednakże w przypadku danych przekazywanych np. do instytucji międzynarodowych powinny być stosowane symbole wymagane przez te instytucje (oznaczenia tajemnicy statystycznej wykorzystywane przez Eurostat mogą być też wykorzystywane w uzasadnionych przypadkach jako podstawowe

oznaczenie tajemnicy statystycznej w badaniu, także do potrzeb wewnętrznych statystyki publicznej – załącznik 1). Oznaczenie, oparte na zestawie przedstawionym w tabeli 3., powinno wykorzystywać przynajmniej poziom pierwszy (oznaczany jedną cyfrą); poziomy bardziej szczegółowe (oznaczane dwiema lub trzema cyframi), mogą być wykorzystywane w zależności od potrzeby. Możliwa jest dalsza rozbudowa tej klasyfikacji, taka potrzeba powinna zostać zgłoszona do Zespołu i zostanie rozważona w trakcie prac nad aktualizacją Zasad.

Figure 4.4: Tabela 3. Przykład badania dominacji w przypadku wystąpienia wartości ujemnych w danych jednostkowych

Symbol	Status jawności / ochrony danych	
1	Dane mogą być udostępniane	
11		bez konieczności przekształcenia danych lub uzyskania zgody respondenta
12		przekształcone dane (np. metodą zakłóceńową)
13		po zgodzie respondenta
131		mimo liczby jednostek<3
132		mimo dominacji
133		mimo liczby jednostek<3 i dominacji
2	Dane nie mogą być udostępniane ze względu na tajemnicę statystyczną	
21		podlegające pierwotnemu ukryciu
211		liczba jednostek<3
212		dominacja
213		liczba jednostek<3 i dominacja
22		podlegające wtórnemu ukryciu
3	Dane nie mogą być udostępniane z innych przyczyn	
9	Brak danych (agregat zawiera 0 jednostek)	

Źródło: opracowanie własne.

Przy przygotowaniu zasad kontroli ujawniania danych statystycznych w tablicach z badania należy uwzględnić występowanie tablic łączonych, dla których powinien być wyznaczony ten sam schemat ukrywania dla wspólnych komórek. Implementacja ukrywania pierwotnego (ukrycie danych lub ich oznaczenie, w zależności od potrzeb) powinna być zawarta w SIB (jeśli dane są opracowywane w SIB badania) lub w innych systemach wykorzystywanych do opracowywania danych zagregowanych, natomiast ukrywania wtórnego – może być wykonana poza SIB, jeśli nie ma możliwości jej implementacji w SIB (więcej informacji na ten temat zawarto w rozdziale 5. *Przewodnika*).

Agregat wyliczany na podstawie badania reprezentacyjnego lub danych łączących wyniki badania reprezentacyjnego i pełnego (dla obserwacji z badania pełnego przyjmujemy wagę równą 1) należy ukryć, jeżeli występuje przynajmniej jedna z czterech poniższych sytuacji (w przypadku zmiennych charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą oraz innych zmiennych, dla których przyjęto kryteria zgodnie z art. 38 ust. 2 Ustawy):

- liczba jednostek wchodzących w skład agregatu w populacji bądź jej oszacowanie jest mniejsze niż 3,
- udział najwyższej spośród wartości jednostkowych wchodzących w skład agregatu w wartości oszacowania wielkości agregatu w populacji jest większy niż 75% – dominacja,
- w próbie występuje tylko jedna jednostka reprezentująca agregat lub występują dwie jednostki, których wagi uogólniające są identyczne lub różnią się o mniej niż 30% wielkości mniejszej z nich; warunek ten może być pominięty przy ocenie konieczności objęcia danych ochroną (możliwość nieukrywania danych, mimo wystąpienia opisywanej sytuacji) jeżeli z prawdopodobieństwem graniczącym z pewnością można przyjąć, że liczba jednostek wchodzących w skład agregatu w populacji generalnej jest większa niż 2,
- pojawia się ryzyko pośredniej identyfikacji jednostki.

Tablice wynikowe z agregatami przeznaczonymi do udostępniania, opracowywane zgodnie z założeniami przygotowanymi przez jednostkę autorską dla innych jednostek służb statystyki publicznej (w tym dla WOBOR) powinny zawierać dla poszczególnych agregatów jedynie oznaczenia konieczności pierwotnego lub wtórnego ukrycia tajemnicy statystycznej (z uwzględnieniem tajemnicy publikacyjnej, jeśli ma zastosowanie) wraz z opisem stosowanych oznaczeń – dane objęte tajemnicą statystyczną nie są w nich ukrywane. Dane oznaczone jako tajemnica statystyczna mogą być wykorzystywane tylko do wewnętrznych analiz lub przygotowania innych agregatów, które mogą być udostępniane po przeprowadzeniu kontroli ujawniania danych statystycznych. Wynikowe informacje statystyczne mogą być publikowane i udostępniane użytkownikom zewnętrznym tylko po ukryciu danych objętych ochroną oraz oznaczeń tajemnicy statystycznej wykorzystywanych wewnętrznie w statystyce publicznej. Jednostka autorska może przekazywać tablice wynikowe innym jednostkom służb statystyki publicznej (w tym WOBOR) z już ukrytymi danymi objętymi tajemnicą statystyczną, jeśli takie były ustalenia z tymi jednostkami.

W wyniku kontroli ujawniania danych statystycznych może zostać ukryta wartość wynosząca zero dla zmiennej lub zmiennych wykorzystywanych do badania kryterium dominacji, ale nie może być objęty tajemnicą statystyczną wtórny agregat, który dotyczy zerowej liczby jednostek (tj. brak podmiotów, osób fizycznych, których miałyby dotyczyć dane). Ukrycie wartości wynoszącej zero może jednak wymagać przeanalizowania, czy po takim ukryciu nadal nie będzie ryzyka ujawnienia danych objętych tajemnicą statystyczną – w takim przypadku wartość zero nie powinna być wykorzystywana do wtórnego ukrycia danych, jako jedyny agregat podlegający ukryciu wtórnemu (może być wykorzystana w parze z innym agregatem o wartości różnej od 0 dla zmiennej podlegającej badaniu kryterium dominacji). Ponadto, analizując określoną zmienną w celu wykonania procedury ochrony tajemnicy statystycznej, należy pamiętać o tym, że do ukrycia wtórnego nie powinien być wykorzystywany agregat tej zmiennej oznaczony jako „zjawisko nie wystąpiło”.

Źródło: opracowanie własne.

Figure 4.5: Tabela 4. Przykład badania dominacji w przypadku wystąpienia wartości ujemnych w danych jednostkowych

PKD	Liczba przedsiębiorstw	Usługi obce w mln zł	Obrót w mln zł	Obrót jednostkowy – maksimum jednostkowe w mln zł	Wyjaśnienie
28	275	15 200	763 276	420 000	dane mogą być opublikowane
28.1	67	8 900	155 000	17 000	możliwe ukrycie wtórne (większa strata informacji)
28.2	183	2 300	532 674	420 000	konieczność ukrycia pierwotnego (dominacja)
28.3	2	4 000	2 367	1 900	konieczność ukrycia pierwotnego (liczba) i wtórnego (dominacja)
28.4	23	0	73 235	9 000	możliwe ukrycie wtórne (mniejsza strata informacji)
28.9	0	-	-	X	nie można ukryć wtórnie

Figure 4.6: Tabela 5. Przykład badania dominacji w przypadku wystąpienia wartości ujemnych w danych jednostkowych

PKD	Liczba przedsiębiorstw	Przychód w mln zł	Przychód jednostkowy – maksimum jednostkowe w mln zł	Udział maksimum jednostkowego w %	Wyjaśnienie
10	7 390	3 692 400	2 140 000	60,75	
10.1	349	23 400	350	1,50	ukrycie wtórne tylko tego agregatu nie spowoduje wyeliminowania dominacji*
10.2	884	35 700	1 200	3,36	ukrycie wtórne tylko tego agregatu nie spowoduje wyeliminowania dominacji*
10.3	3	0	0	x	ukrycie wtórne tylko tego agregatu nie spowoduje wyeliminowania dominacji*
10.4	5 995	621 700	24 900	4,01	możliwe ukrycie wtórne wyłącznie tego agregatu (duża strata informacji)
10.5	4	23 600	10 600	44,92	ukrycie wtórne tylko tego agregatu nie spowoduje wyeliminowania dominacji*
10.6	35	40 000	2 450	6,13	ukrycie wtórne tylko tego agregatu nie spowoduje wyeliminowania dominacji*
10.7	17	2 800 500	2 140 000	76,41	dominacja – konieczność ukrycia pierwotnego
10.8	5	90 800	47 300	52,09	możliwe ukrycie wtórne wyłącznie tego agregatu (średnia strata informacji)
10.9	98	56 700	3 200	5,64	możliwe ukrycie wtórne wyłącznie tego agregatu (mała strata informacji)

*Po odjęciu od wartości na poziomie działu **10** sumy odpowiednich wielkości dla nieukrytych agregatów powstanie agregat, w którym jedna jednostka będzie nadal dominować. Przykładowo, w przypadku ukrycia agregatów **10.1** wtórnie i **10.7** pierwotnie, udział maksimum jednostkowego dwóch powyższych agregatów $2140000 / (23400 + 2800500) * 100\%$ będzie wynosił 75,78%, czyli powyżej 75%.

Źródło: opracowanie własne

W przypadku konieczności wyznaczenia ukrycia wtórnego i przy występowaniu kilku agregatów, które mogłyby być nim objęte, w pierwszej kolejności powinny być brane pod uwagę agregaty o najmniejszej stracie informacji, choć ukrywanie danych o zbyt małych wartościach nie jest zalecane (z uwzględnieniem zasady dot. wartości zero – opisana w akapicie powyżej dotyczącym tablic wyników); pod uwagę mogą być też brane agregaty, które są najmniej istotne z punktu widzenia badanego zjawiska.

Ochrona przed ujawnieniem danych statystycznych w tablicach jest możliwa także poprzez zmianę szczegółowości agregatów. W celu wyeliminowania agregatów o zbyt małej liczbie podmiotów lub zawierających podmiot dominujący (zgodnie z art. 38 ust. 2 Ustawy) dopuszcza się stosowanie metody zmiany szczegółowości tablicy poprzez eliminowanie niskich szczebli agregacji lub łączenie pozycji danej klasyfikacji lub nomenklatury w jedną kategorię. Odpowiednie rozwiązanie może być również zaproponowane użytkownikowi zamawiającemu dane na indywidualne zamówienie.

4.2.3 Zasady kontroli ujawniania danych statystycznych przy prezentacji danych w formie graficznej

Należy zapewnić, aby dane prezentowane za pomocą grafiki nie ujawniały danych objętych tajemnicą statystyczną.

Braki danych, będące wynikiem objęcia ochroną części kategorii/obserwacji, mogą uniemożliwiać lub znacząco ograniczać użyteczność niektórych form graficznej prezentacji danych. W związku z tym, w przypadku takiej prezentacji danych, preferowane jest stosowanie metod ochrony danych statystycznych, które nie powodują powstawania takich braków.

Wykresy – w zależności od rodzaju wykresu proponowane są następujące (przykładowe) rozwiązania:

- wykresy słupkowe – można ukryć dane statystyczne, które powinny być objęte ochroną: pominąć słupek i zamiast niego wykorzystać symbol opisany jako „konieczność zachowania tajemnicy statystycznej”. Nie uwzględniamy tak oznaczonego przekroju przy sortowaniu i zamieszczamy notkę na ten temat pod wykresem.
- wykresy strukturalne – dane statystyczne objęte ochroną można ukryć agregując dwa (lub więcej) elementy struktury.

- wykresy liniowe – dane statystyczne, które powinny być objęte ochroną, można ukryć poprzez przerwanie serii agregatów objętych ochroną przed ujawnieniem; należy wówczas zamieścić stosowną notkę pod wykresem (rozwiązanie to dotyczy sytuacji, gdy tajemnica została nałożona na dane w jakimś odcinku czasu, a nie w pełnym szeregu).

Mapy – możliwe jest ukrycie danych statystycznych objętych ochroną, czyli prezentacja takich danych po agregacji dwóch (lub więcej) elementów. Jest to możliwe w przypadku dwóch metod prezentacji kartograficznej:

- kartogramu (tylko dane względne),
- kartodiagramu skokowego (dane bezwzględne lub względne).

Ponadto na mapach wykonanych metodą kartodiagramu czasami pokazywana jest struktura diagramów. Jeżeli w którymś z elementów struktury występują dane statystyczne, które powinny być objęte ochroną, to można ukryć te dane agregując dwa (lub więcej) elementy struktury.

W przypadku grafik analizy rozkładów lub współzależności zweryfikować należy czy na grafice widoczne są obserwacje odstające (ang. *outliers*) – tj. takie, które występują w niewielkiej liczbie, ale pod względem wartości różnią się znacząco od pozostałych – oraz czy ich prezentacja nie prowadzi do ujawnienia danych statystycznych objętych ochroną. W przypadku stwierdzenia takiego przypadku należy dokonać modyfikacji grafik, uniemożliwiających ujawnienie danych statystycznych objętych ochroną (np. poprzez zastosowanie szerszych agregatów czy wydłużenie przedziałów wartości).

Należy uniemożliwić dostęp do danych jednostkowych, które zostały użyte do sporządzenia danej grafiki. Dotyczy to w szczególności elektronicznych wersji opracowań i publikacji, w których programy służące do stworzenia grafik mogą takie dane przechowywać w tle i wyświetlać je po wybraniu stosownej opcji lub utrzymywać w czytelnej postaci w źródłowym zapisie pliku.

4.2.4 Tryb pozyskiwania zgody na publikację określonych danych charakteryzujących wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej

Decyzję o wprowadzeniu do badania możliwości występowania do podmiotów gospodarki narodowej o zgodę na publikowanie danych charakteryzujących wyniki ekonomiczno-finansowe, zgodnie z Ustawą (art. 38 ust. 3), podejmuje jednostka autorska badania. Występowanie o zgodę powinno się odbywać z odpowiednim wyprzedzeniem czasowym tak, aby umożliwić zastosowanie zgody w danej edycji badania zgodnie z planami jednostki autorskiej.

Wystąpienie o zgodę co do zasady dotyczy jednego badania i objętej nim grupy wybranych respondentów, którzy wpływają na wzrost liczby ukryć danych ze względu na tajemnicę statystyczną. Jeśli wystąpienie ma dotyczyć respondenta, którego skala działalności może istotnie wpływać na obraz badanych zjawisk w wielu obszarach statystyki – w takim przypadku autor

badania inicjujący wystąpienie powinien ustalić, czy inne obszary statystyki są również zainteresowane takim wystąpieniem lub już wystąpiły do tego podmiotu o zgodę. Zainteresowanie innych jednostek autorskich odpowiednim wystąpieniem do istotnej jednostki powinno być, po odpowiednich konsultacjach, uwzględnione w przygotowywanym wystąpieniu.

Jednostka autorska przygotowuje notatkę (według obowiązującego schematu notatki informacyjnej) na temat potrzeby zwrócenia się o zgodę do określonych podmiotów gospodarki narodowej (respondentów) na publikowanie danych (każdorazowe przeprowadzenie akcji uzyskiwania zgód wymaga wystąpienia do Prezesa GUS). Notatka zawiera m.in. uzasadnienie takiego podejścia, informacje o podmiotach, do których planowane jest zwrócenie się o zgodę i zakres danych, o zgodę na publikowanie których występuje jednostka autorska, a także związane z tym szacowane koszty (tj. liczba podmiotów, do których zostanie skierowana prośba o zgodę oraz ile stron będzie miało pismo przewodnie i kwestionariusz zgody – jeśli jest to wysyłka do większej grupy podmiotów, powodująca istotne koszty – istotność kosztów ocenia jednostka autorska). Notatka jest przekazywana do Prezesa GUS z prośbą o akceptację, a po uzyskaniu akceptacji jest przekazywana do Gabinetu Prezesa i do wiadomości Departamentu Programowania, Koordynacji Badań i Rejestrów oraz urzędu statystycznego specjalizującego się w badaniu (jeśli jednostką autorską jest departament GUS).

Informacja o udzieleniu zgody Prezesa GUS na pozyskiwanie zgód respondentów jest umieszczana przez jednostkę autorską w Systemie Metadanych Statystycznych w modelu *Opisy badań – pozostałe informacje o badaniu*, w zakładce 5. Informacje o tajemnicy statystycznej, w punkcie 5.3. *Dodatkowe informacje, istotne dla pracowników jednostek statystyki publicznej, przydatne dla prawidłowego nakładania tajemnicy statystycznej*.

Pismo do podmiotów gospodarki narodowej (respondentów istotnych z punktu widzenia publikacji wyników informacji statystycznych) jest podpisywane przez Prezesa GUS. Pismo zawiera prośbę o wyrażenie zgody przez podmiot gospodarki narodowej na publikowanie wskazanego w dołączonym do pisma kwestionariuszu zakresu danych przez niego przekazywanych, także w przypadku, gdy w agregacie, na który się składają przekazane dane, będzie mniej niż 3 podmioty lub jeden z nich byłby jednostką dominującą (a więc zgodnie z Ustawą dane te powinny zostać ukryte; należy podkreślić, że zgoda nie dotyczy udostępniania i publikacji danych jednostkowych identyfikowalnych lub nieidentyfikowalnych). Wzór pisma do wykorzystania zamieszczono w Załączniku 2. do *Przewodnika*.

Do pisma powinien zostać dołączony spersonalizowany kwestionariusz zgody podmiotu gospodarki narodowej, który po podpisaniu przez osobę upoważnioną do reprezentowania danego podmiotu należy odesłać do GUS. W Załączniku 3. do *Przewodnika* zamieszczono szablon kwestionariusza w zakresie elementów stałych, który wymaga dostosowania przez jednostkę autorską do specyfiki badania w części dot. zakresu danych, dla których GUS występuje o zgodę na publikację. Wysyłka pisma i kwestionariusza będzie realizowana zgodnie z obowiązującymi w danym momencie ogólnymi zasadami wysyłki.

Nieudzielenie odpowiedzi przez podmiot traktuje się jak brak zgody. Jednocześnie, w przypadku braku odpowiedzi, możliwe jest nawiązanie kontaktu z jednostką (np. telefonicznie, z

dodatkowymi wyjaśnieniami w razie wątpliwości jednostki), aby zmotywować ją do udzielenia zgody. Zgoda udzielona przez osobę uprawnioną do reprezentowania podmiotu obowiązuje także po zakończeniu pełnienia funkcji w podmiocie przez tę osobę.

Podmiot gospodarki narodowej ma prawo odwołać zgodę. W Załączniku 4. do *Przewodnika* zamieszczono szablon kwestionariusza odwołania zgody. Odwołanie zgody na publikację danych powinno być wyrażone przez osobę upoważnioną do reprezentowania danego podmiotu pisemnie lub za pośrednictwem Portalu Sprawozdawczego.

Odwołanie przez podmiot gospodarki narodowej zgody na publikowanie danych nie dotyczy danych za okresy, w których obowiązywała zgoda, a jedynie okresów przyszłych od momentu odwołania zgody. Odwołanie zgody obowiązuje od danej edycji badania, jeśli wpłynęło w trakcie etapu zbierania danych od respondentów w tej edycji badania, lub od kolejnej edycji badania, jeśli wpłynęło po zamknięciu etapu zbierania danych.

Informacje o zgodzie lub jej odwołaniu przechowywane są przez autora badania lub na wniosek autora badania przez urząd specjalizujący się w badaniu. Zaleca się przechowywanie informacji o udzielonej zgodzie lub jej odwołaniu w SIB oraz oznaczenie obowiązujących zgód w archiwizowanych zbiorach danych jednostkowych. Ponadto przy wewnętrznej wymianie danych zagregowanych w ramach statystyki publicznej należy zadbać o odpowiednie oznaczanie danych przeznaczonych do udostępnienia na podstawie uzyskanych zgód, np. zgodnie z zalecaną tablicą 3.

Ponowne wystąpienia o zgodę na opublikowanie danych mogą być również kierowane do podmiotów, które wcześniej odmówiły wyrażenia zgody lub odwołały zgodę, a możliwość publikacji ich wyników jest nadal istotna z punktu widzenia użytkowników statystyki, jednak nie wcześniej niż dwa lata od daty otrzymania odmowy.

Kwestionariusze zgody na opublikowanie określonych danych oraz ich odwołania powinny być przechowywane zgodnie z zasadami ochrony danych osobowych i zasadami archiwizacji.

Dane wynikowe z badania są objęte kontrolą ujawniania danych zgodnie z zasadami ogólnymi oraz zasadami szczegółowymi przyjętymi dla danego badania. Przy każdej edycji badania, przy utajnianiu danych uwzględniane są aktualne zgody udzielone przez podmioty gospodarki narodowej na publikowanie danych, z uwzględnieniem odwołanych zgód.

4.3 Zakres informacji dotyczących tajemnicy statystycznej wprowadzany w SMS

W Systemie Metadanych Statystycznych wprowadzane są przez autorów badań szczegółowe informacje spójne z zasadami ogólnymi (zob. rozdział 4.2.1), specyficzne dla danego badania / zestawu danych / pakietu zestawu danych, które są istotne dla prawidłowego ukrycia przekrojów objętych tajemnicą statystyczną. Informacje te wprowadzane są przez jednostkę autorską na potrzeby innych jednostek statystyki publicznej, ale też jako metadane przydatne

dla samej jednostki autorskiej. Jednostka autorska określa zmienne w badaniu, które są objęte ochroną przed ujawnieniem, w szczególności wskazuje zmienne, które w poszczególnych badaniach charakteryzują wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą w rozumieniu art. 38 ust. 2 Ustawy.

W SMS zaimplementowany został model *Opisy badań – pozostałe informacje o badaniu*. Model ten zlokalizowany jest w *Metadanych badań* → części *Projektowanie* → folderze *Program badań*. Po przejściu do lokalizacji modelu, wybraniu kontenera Wszystkie, wyszukaniu właściwego badania na liście metadanych w oknie filtrowania i wskazaniu wyszukanej metadanej, wyświetla się formularz edycji metadanej. Formularz składa się z pięciu zakładek, na których ulokowane są atrybuty lub powiązania. Przy określaniu tajemnicy statystycznej istotne są dwie, tj. zakładka 1. *Pozostałe informacje o badaniu* oraz zakładka 5. *Informacje o tajemnicy statystycznej*.

W zakładce 1. *Pozostałe informacje o badaniu* znajdują się dwa powiązania:

- 1.1. *Opis badania* – wskazany jest tu opis badania z modelu *Opis badania*, który jest przygotowywany w ramach Pbssp,
- 1.2. *Zmienne wykorzystywane w badaniu* – wskazane są tu zmienne (jeżeli jednostka odpowiedzialna za badanie je wypełniła), które wykorzystywane są w tym badaniu do opracowywania wyników. Zmienne, które można dodać do powiązania za pomocą przycisku *Dodaj z listy*, pochodzą z zestawów danych wskazanych w polu 4.1 w modelu *Opis badania*, zlokalizowanym w *Metadanych badań* → części *Projektowanie* → folderze *Program badań* podfolderze *PR opisy badań*.

W zakładce 5. *Informacje o tajemnicy statystycznej* znajdują się dwa powiązania oraz atrybut:

- 5.1. *Zmienne charakteryzujące wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą* – po kliknięciu *Dodaj z listy* wyświetla się lista zmiennych wskazanych w polu 1.2. Spośród nich należy wybrać te, które charakteryzują wyniki ekonomiczno-finansowe. Dane statystyczne charakteryzujące wyniki ekonomiczno-finansowe podmiotów gospodarki narodowej prowadzących działalność gospodarczą to dane pozyskane przez statystykę publiczną z różnych źródeł na potrzeby prowadzonych badań, które zarówno bezpośrednio jak i pośrednio mogą informować o wynikach ekonomiczno-finansowych osiąganych przez podmiot. Są to zarówno wszystkie pozycje ze sprawozdania finansowego i wyliczane na ich podstawie wskaźniki oraz inne istotne finansowe informacje, które mogą stanowić tajemnicę przedsiębiorstwa, jak i informacje, które mogą na przykład pokazywać pozycję podmiotu na rynku. Przykładowo tajemnicą powinna być objęta produkcja sprzedana w ujęciu wartościowym i jej dynamika, informacje dotyczące podatków. Przy klasyfikowaniu danych istotna jest też wrażliwość danej grupy respondentów na upublicznianie tego typu danych.

- 5.2. *Zmienne wykorzystywane do badania kryterium dominacji (udział jednego podmiotu w określonym zestawieniu jest większy niż trzy czwarte całości)* – po kliknięciu *Dodaj z listy* wyświetla się lista zmiennych, która wynika ze zmiennych wskazanych w polu 1.2. Spośród nich należy wybrać te, które wykorzystywane są do badania kryterium dominacji.

Nie należy wypełniać punktu 5.2., jeżeli podjęto decyzję o sprawdzaniu kryterium dominacji dla każdej zmiennej z punktu 5.1. osobno po kolei. Punkt 5.2. należy natomiast wypełnić, w przypadku gdy do badania dominacji stosowane są wybrane zmienne spośród tych wymienionych w punkcie 5.1. Z praktyki wiadomo bowiem, że nie zawsze zasadne jest sprawdzanie kryterium dominacji dla wszystkich zmiennych z punktu 5.1. Dominacja może być sprawdzana dla kilku wybranych zmiennych i rozciągana na pozostałe zmienne wymienione w punkcie 5.1., w którym określamy jaki zestaw danych chronimy. Jest to uproszczone badanie dominacji: bada się tylko niektóre zmienne, najbardziej charakterystyczne i istotne dla danego zestawu i przenosi schemat ukrywania na pozostałe zmienne, wyszczególnione w punkcie 5.1. Sprawdzenie wybranych kilku zmiennych przyspiesza i usprawnia proces badania kryterium dominacji.

- 5.3. *Dodatkowe informacje, istotne dla pracowników jednostek statystyki publicznej, przydatne dla prawidłowego nakładania tajemnicy statystycznej* – w polu tym wpisywane są wszelkie istotne informacje dotyczące ukrywania przekrojów objętych tajemnicą statystyczną.

W punkcie 5.3. *Dodatkowe informacje, istotne dla pracowników jednostek statystyki publicznej, przydatne dla prawidłowego nakładania tajemnicy statystycznej* należy wpisać takie informacje, które nie zostały zawarte w punktach 1-5, a są istotne dla spójnego ukrycia przekrojów objętych tajemnicą statystyczną w jednostkach statystyki publicznej dla danych pochodzących z tego badania/zestawu danych, np. specyficzne dla badania / zestawu danych zasady przy wtórnym ukrywaniu przekrojów, informacja o pozyskiwaniu zgód na publikację danych zgodnie z art. 38 ust. 3 i uwzględnianiu ich przy wyznaczaniu tajemnicy publikacyjnej, czy też prośba jednostki autorskiej o konsultowanie z nią wtórnego ukrywania komórek w przygotowywanych nowych przekrojach. Nie trzeba tu powtarzać na przykład ogólnych zasad ukrywania przekrojów objętych tajemnicą statystyczną zawartych w Ustawie.

Hundepool, Anco, Josep Domingo-Ferrer, Luisa Franconi, Sarah Giessing, Rainer Lenz, Jane Longhurst, E Schulte Nordholt, Giovanni Seri, and P Wolf. 2010. "Handbook on Statistical Disclosure Control." *ESSnet on Statistical Disclosure Control*.

Hundepool, Anco, Josep Domingo-Ferrer, Luisa Franconi, Sarah Giessing, Eric Schulte Nordholt, Keith Spicer, and Peter-Paul Wolf. 2012. *Handbook on Statistical Disclosure Control*. John Wiley & Sons.

"The Eurostat Concepts and Definitions Database." 2001. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/coded-circa>.