

Warszawa, 2021-04-26

Znak sprawy: GUS-BD.573.1.2021

Znak pisma: GUS-BD.573.1.2021.357

Szanowny Pan
Bartosz Wojciechowski

Szanowny Panie,

W odpowiedzi na Pana pismo z 8 kwietnia 2021 r., wystosowane do Prezesa GUS na podstawie przepisów o dostępie do informacji publicznej, pozwalamy sobie przekazać następujące wyjaśnienia odnośnie do poruszonych w piśmie zagadnień.

W odniesieniu do poruszonej przez Pana kwestii anonimizacji (odpersonalizowaniu) danych informujemy, że zasadniczo operacja ta polega na oddzieleniu danych identyfikowalnych badanych podmiotów i zastąpieniu ich sztucznymi identyfikatorami. Anonimizacji podlegają dane osobowe, takie jak imię, nazwisko oraz numer pesel, ale również cechy identyfikujące inne rodzaje badanych podmiotów, jak np. dane adresowe.

Należy zauważyć, że indywidualne dane osobowe same w sobie nie są przedmiotem badania statystycznego, ani celem zainteresowań i dociekań badawczych, lecz pełnią tylko rolę danych pomocniczych, wykorzystywanych wyłącznie w procesie pozyskiwania i integracji danych. Są bowiem jedynym skutecznym środkiem (sposobem) do uwierzytelniania i trafnego identyfikowania (rozdzielania czy odróżniania) badanych jednostek, a przez to umożliwiają łączenie danych z wielu źródeł, tj. przyporządkowanie danych pozyskanych w różnych badaniach (względnie rejestrach) lub w różnym czasie właściwym – zawsze tym samym – jednostkom.

Dane osobowe na etapie przetwarzania są zatem zastępowane sztucznymi identyfikatorami (nierozpoznawalnymi) zapewniając integralność danych.

W kwestii wyrażonych przez Pana wątpliwości co do sposobu pozyskiwania, przetwarzania i organizowania pozyskanych w ramach spisu danych o poszczególnych badanych jednostkach, pozwalamy sobie przedstawić kilka informacji o charakterze metodologiczno-technicznym.

Jak był Pan uprzejmy zauważyć, przyjęto w badaniu spisowym takie rozwiązanie, że w ramach jednej sesji – czy inaczej rzecz ujmując – jednego elektronicznego pakietu mieszkaniowego kwestionariuszy spisowych, możliwe jest spisanie i przekazanie kompletu przewidzianych w spisie danych o budynku (w przypadku domów jednorodzinnych), mieszkaniu oraz o każdej zamieszkałej pod danym adresem osobie. Przy czym należy wyraźnie podkreślić, że takie rozwiązanie, choć uznane za modelowe i najbardziej wygodne dla respondentów, nie jest jedynym możliwym, ponieważ przewidziano również możliwość odrębnego spisywania się osób w mieszkaniu.

Pragniemy zapewnić, że niezależnie od tego, że aplikacja spisowa umożliwia płynne przechodzenie do kolejnych części kwestionariusza i udzielanie informacji o mieszkaniu i poszczególnych jego mieszkańcach, konstrukcja elektronicznych kwestionariuszy spisowych oraz sposób przetwarzania danych zapewniają to, że pozyskiwane informacje o każdej spisanej jednostce są odrębnie rejestrowane w odpowiedniej strukturze danych. Istotą i celem przetwarzania informacji pozyskanych w spisach powszechnych i innych badaniach statystycznych jest bowiem zorganizowanie danych do postaci zbiorów statystycznych, stanowiących optymalną podstawę do przygotowywania statystycznych opisów i analiz badanych zbiorowości (populacji) pod kątem cech przewidzianych w zakresie przedmiotowym badania.

W przypadku Narodowego Spisu Powszechnego Ludności i Mieszkań, odpowiednio do rodzajów spisywanych podmiotów, przygotowano osobne statystyczne zbiory danych o budynkach, mieszkaniach i osobach. Cały zestaw pozyskanych w spisie informacji, dotyczących każdego zbadanego indywidualnego podmiotu, w tym również każdej osoby spisanej w mieszkaniu, jest zapisywany w jednym osobnym rekordzie (wierszu) danych statystycznego zbioru, właściwego dla tego podmiotu.

W kwestii sprawdzania prawdziwości danych spisowych informujemy, że spójność i adekwatność danych jest już zapewniona na poziomie samych kwestionariuszy spisowych, poprzez zastosowanie w aplikacji spisowej standaryzacji pytań oraz odpowiednich mechanizmów kontroli logicznej. Niezależnie od tego dysponujemy w badaniu spisowym danymi z różnego typu administracyjnych rejestrów i systemów informacyjnych, które również są wykorzystywane przy opracowaniu wyników spisu, w tym do weryfikacji i walidacji danych.

Należy jednak zauważyć, że zakres przedmiotowy spisu powszechnego obejmuje dane o bardzo zróżnicowanej naturze: począwszy od twardych i faktograficznych, w tym również prawnie usankcjonowanych, do zupełnie subiektywnych, uznaniowych i deklaracyjnych. Stosownie to tej różnorodności cech istnieją również adekwatne możliwości ich weryfikacji.

Niezależnie od powyższego statystyka publiczna w ustawie i dokumentach spisowych, a także w aplikacji spisowej odwołuje się do autorytetu prawa po to, aby udział w badaniu spisowym był traktowany z należytą powagą, a pozyskane w toku badania spisowego informacje były możliwie najbardziej rzetelne i zgodne z rzeczywistością, a w konsekwencji – wyniki spisu umożliwiły uzyskanie miarodajnego statystycznego obrazu polskiego społeczeństwa i jego warunków mieszkaniowych.

Z poważaniem

/-/ Dorota Száltys

Dyrektor Departamentu Badań Demograficznych