

Janusz DYGASZEWICZ

Spisy powszechne jako źródło danych do analiz geoprzestrzennych

Na potrzeby przeprowadzenia spisów powszechnych GUS przygotowano przestrzenne bazy adresowe zawierające m.in. statystyczne punkty adresowe oraz wektorowe granice podziału statystycznego, czyli granice rejonów statystycznych oraz obwodów spisowych. Służby geodezyjne i kartograficzne nie dysponowały zbiorem punktów adresowych obejmujących kraj, które można by wykorzystać w spisach. Dlatego służby statystyki publicznej podjęły się opracowania „warstwy” takich punktów. Służyły one rachmistrzom w łatwym i sprawnym odnajdowaniu miejsc, w których przeprowadzano wywiad. Ponadto na etapie opracowywania wyników spisu służą one również do prezentacji przestrzennej badanych zjawisk. Głównym celem artykułu jest prezentacja portalu geostatystycznego umożliwiającego tworzenie analiz przestrzennych z danych zebranych podczas spisów powszechnych.

Statystyczne punkty adresowe określono przez atrybuty opisowe (województwo, powiat, gmina, miejscowość, ulica, numer adresowy) i graficzne (położenie wyznaczone przez współrzędne punktu adresowego). Z kolei obwody spisowe i rejony statystyczne określają przebieg granic oraz ich identyfikatory zgodne z Krajowym Rejestrem Urzędowym Podziału Terytorialnego Kraju (TERYT).

Bazy adresowe zawierają odniesienie przestrzenne (współrzędne x, y) punktów adresowych zintegrowanych z rejestrem TERYT, a zatem odzwierciedlają położenie każdego budynku mieszkalnego lub budynku, w którym znajduje się przynajmniej jedno mieszkanie. Bazy te opracowywane są w podziale na powiaty. Aktualizacja przestrzennych baz adresowych odbywa się kwartalnie. Podstawą do aktualizacji są dane rejestru TERYT.

Aktualizacja podziału statystycznego (granice obwodów spisowych i rejonów statystycznych) odbywa się podobnie. Informacja o usuniętym lub dodanym nowym obwodzie zawarta jest w pliku xml¹. Analogicznie jak przy aktualizacji punktów adresowych, na liście zadań podawany jest operatorowi numer obwodu, dla którego należy określić położenie poprzez ustalenie jego granic. Z kolei zmiany w przebiegu granic, wynikające z dostosowywania do granic obrębów ewidencyjnych bądź innych szczegółów terenowych, wykonywane są na bieżąco.

Przestrzenne bazy adresowe zawierają także zbiór danych określających przebieg ulic. Z uwagi na priorytet, jakim było przygotowanie przestrzennych baz adresowych, „warstwa” tych danych ulic nie jest kompletna, ponieważ nie wszystkie ulice występujące w rejestrze TERYT miały określoną lokalizację przestrzenną. Dane te są stale uzupełniane. Dzięki tak przygotowanym warstwom, statystyka publiczna dysponuje bardzo szczegółowymi danymi przyporządkowanymi do poszczególnych osób powiązanych z konkretnymi adresami. Dlatego też możliwe są różnego rodzaju analizy przestrzenne na różnych poziomach agregacji (przy zachowaniu tajemnicy statystycznej).

Obecnie planowane jest powiększenie przestrzennych baz adresowych przez dodanie adresów podmiotów gospodarczych znajdujących się w Bazie Jednostek Statystycznych (BJS). Rejestr statystyczny przedsiębiorstw wykorzystuje REGON, jednak w odróżnieniu od niego jest aktualizowany poprzez ciągłe uwzględnianie wyników badań statystycznych oraz informacje pochodzące ze źródeł administracyjnych, np. z Krajowego Rejestru Sądowego.

Obecnie w przestrzennych bazach adresowych znajdują się tylko punkty, w których jest przynajmniej jedno mieszkanie, a zatem uzupełnienie baz w zakresie położenia przestrzennego adresów podmiotów gospodarczych będzie wymagało dodatkowej pracy, związanej z wyznaczeniem współrzędnych punktów adresowych pozostałych budynków. W tym celu zostaną wykorzystane dane

¹ XML — rozszerzalny język znacznikowy, ang. *extensible markup language* — uniwersalny język formalny przeznaczony do zapisywania danych w sposób strukturalizowany. XML jest niezależny od platformy, co ułatwia wymianę dokumentów pomiędzy różnymi systemami; *Internetowy leksykon geomatyczny*, Zespół Redakcyjny: Jerzy Gaździcki, Janusz Michałak, Ewa Musiał, <http://www.ptip.org.pl/>.

ze źródeł administracyjnych (z ewidencji gruntów i budynków oraz Bazy Danych Obiektów Topograficznych).

PORTAL GEOSTATYSTYCZNY

W celu prezentacji wyników Powszechnego Spisu Rolnego 2010 (PSR 2010) oraz Narodowego Spisu Powszechnego Ludności i Mieszkań 2011 (NSP 2011) GUS wykonuje rozbudowę platformy aplikacyjnej GIS przez opracowanie Portalu Geostatystycznego (PGS). Portal umożliwi udostępnianie zagregowanych danych w postaci różnego rodzaju analiz przestrzennych z zachowaniem tajemnicy statystycznej. Oto przykładowe gotowe serwisy, które znajdują się w portalu geostatystycznym:

1. W zakresie danych z PSR 2010:

- gospodarstwa z dochodami z działalności rolniczej,
- gospodarstwa korzystające ze środków wsparcia PROW² w ciągu ostatnich 3 lat,
- gospodarstwa, których użytkownicy planują przekazanie ich następcom do 2013 r.,
- gospodarstwa zużywające własne produkty rolne z rozwiniętą hierarchią,
- gospodarstwa z dochodami z działalności rolniczej zużywające własne produkty rolne z rozwiniętą hierarchią,
- gospodarstwa realizujące sprzedaż bezpośrednią z rozwiniętą hierarchią,
- gospodarstwa z ciągnikami własnymi lub wspólnymi według mocy z rozwiniętą hierarchią,
- gospodarstwa z dojarkami i krowami z rozwiniętą hierarchią,
- gospodarstwa z samochodami i przyczepami z rozwiniętą hierarchią,
- gospodarstwa według ilości posiadanych ciągników własnych lub wspólnych z rozwiniętą hierarchią,
- rodzaje urządzeń do produkcji energii odnawialnej z rozwiniętą hierarchią,
- rodzaje nawozów wykorzystanych w gospodarstwie,
- uprawy podstawowe z rozwiniętą hierarchią biorące pod uwagę liczbę gospodarstw oraz powierzchnię zasiewów,
- rodzaje gruntów z rozwiniętą hierarchią biorące pod uwagę liczbę gospodarstw oraz powierzchnię gruntów,
- skala chowu bydła z rozwiniętą hierarchią,
- rodzaje zwierząt z rozwiniętą hierarchią,
- rodzaje aktywności bieżącej osób w gospodarstwie indywidualnym;

2. Uzyskanych z NSP 2011:

- poziom wykształcenia,
- kontynuacja nauki,
- główne i dodatkowe źródło utrzymania,
- liczba świadczeń,
- niepełnosprawni biologicznie i prawnie,

² PROW — Program Rozwoju Obszarów Wiejskich, <http://prow.rolnicy.com>.

- liczba urodzonych dzieci,
- aktywność ekonomiczna ludności faktycznie zamieszkałej oraz rezydującej,
- dojazdy do pracy ludności faktycznie zamieszkałej,
- mieszkania niezamieszkane według przyczyn, liczby izb, statusu zamieszkania, powierzchni użytkowej, wyposażenia w wodociąg,
- ludność faktycznie zamieszkała i rezydująca według np. płci, wieku, stanu cywilnego, roku urodzenia itp.,
- zamieszkane budynki według rodzaju,
- zamieszkane budynki według podmiotów będących właścicielami lub współwłaścicielami tych budynków,
- zamieszkane budynki mieszkalne według wyposażenia w gaz z sieci,
- zamieszkane budynki mieszkalne według wyposażenia w kanalizację,
- zamieszkane budynki mieszkalne według liczby mieszkań.

Poza możliwością korzystania z gotowych analiz przestrzennych indywidualny użytkownik będzie mógł, w ramach portalu geostatystycznego, redagować własne mapy tematyczne (a także drukować je) w formie kartogramu, wykorzystując dowolną cechę modelu danych tematycznych. Ponadto określona funkcja portalu pozwoli na generowanie tematów analizy na podstawie buforowania wstawionego punktu, rysowanie grafiki poligonowej oraz wygenerowanie raportu sumarycznego (raport tabelaryczny może być uzupełniony o mapę obszaru i wykresy wybranych cech). Szczegółowa koncepcja struktury logicznej systemu zakłada, że rozwiązanie to będzie:

- 1) realizować usługi udostępniania danych statystycznych określonych przestrzennie;
- 2) funkcjonować w sieci Internet oraz Intranet;
- 3) implementować w środowisku serwera mapowego ArcGIS Server Standard (SE);
- 4) prowadzić geograficzną bazę danych statystycznych GeoSTAT zasilaną mikrodanymi oraz danymi zagregowanymi z Analitycznej Bazy Mikrodanych (ABM) dla danych PSR 2010 i NSP 2011 oraz innych mikrodanych z baz danych Systemu Informacyjnego Statystyki Publicznej (SISP). Baza danych GeoSTAT będzie stanowić zawsze wtórną bazę danych, czyli powstałą z przetworzenia oraz łączenia anonimizowanych danych statystycznych z danymi referencyjnymi podziału administracyjnego i statystycznego kraju, z użyciem odpowiednio baz danych: GeoTERC i GeoBREC;
- 5) wykorzystywać dane:
 - a) referencyjne: ortofotomapę i granice podziałów administracyjnych, położenie miejscowości z Państwowego Rejestru Nazw Geograficznych, sieć ulic, działki ewidencyjne,
 - b) operacyjne GUS: podział statystyczny, taki jak BREC³ i punkty adresowe NOBC⁴;

³ BREC — system rejonów statystycznych i obwodów spisowych, <http://www.stat.gov.pl/>.

⁴ NOBC — system identyfikacji adresowej ulic, nieruchomości, budynków i mieszkań, <http://www.stat.gov.pl/>.

- 6) publikować dane za pośrednictwem aplikacji oraz interfejsów dostępu osadzonych w portalu funkcjonującym w powiązaniu z systemem zarządzania treścią (CMS — Podsystem Zarządzania Treścią);
- 7) zapewniać dostęp do bazy danych za pośrednictwem serwisów serwera ArcGIS Server, będącego w posiadaniu GUS;
- 8) dostarczać dedykowane funkcje oraz usługi dla wielu różnych grup użytkowników systemu zarządzanego centralnie poprzez jego funkcje administracyjne;
- 9) zawierać komponenty aplikacyjne oparte na funkcjonalności Oprogramowania Standardowego GIS Serwera mapowego ArcGIS Server SE oraz wytworzonego przez wykonawcę oprogramowania aplikacyjnego, na które składać się będą:
 - a) Podsystem Utrzymania Bazy Danych Geostatystycznych,
 - b) Podsystem Administrowania Systemem, inaczej Podsystem Autoryzacji i Kontroli Dostępu,
 - c) Zarządzania Treścią (CMS),
 - d) Podsystem Publikowania i Aktualizacji Informacji Geostatystycznych za pomocą modułów:
 - Przeglądarki Danych Geostatystycznych, inaczej Moduł Przeglądarki Map,
 - Zaawansowanej Przeglądarki Danych Geostatystycznych, inaczej Moduł Zaawansowanej Przeglądarki Map, zapewniający również edycję danych przestrzennych,
 - Aktualizacji Gminnej (rozbudowana Aplikacja Aktualizacji Gminnej),
 - dodatkowych specjalistycznych usług, które na etapie realizacji zostaną uznane za stosowne do wprowadzenia.
 - e) Podsystem Usług INSPIRE⁵:
 - Portal Metadanych,

⁵ INSPIRE — dyrektywa 2007/2/WE Parlamentu Europejskiego i Rady z 14 marca 2007 r. ustanawiająca infrastrukturę informacji przestrzennej we Wspólnocie Europejskiej (INSPIRE), określająca legalne ramy do ustanowienia i działania Infrastruktury Informacji Przestrzennej w Europie. Celem dyrektywy jest wspieranie działań dotyczących polityki wspólnotowej w zakresie ochrony środowiska oraz polityk lub działań mogących oddziaływać na środowisko. W szczególności działania te wiążą się z implementacją, monitorowaniem aktywności i oceną strategii zaangażowanych społeczności na różnych poziomach, od poziomu lokalnego, przez krajowy aż po poziom europejski. W swoich założeniach INSPIRE łączy infrastruktury informacji przestrzennej tworzone w różnych krajach członkowskich w jeden spójny system. Elementami tej infrastruktury są: metadane, zbiory danych przestrzennych oraz usługi danych przestrzennych, usługi i technologie sieciowe, porozumienia w sprawie wspólnego korzystania, dostępu i użytkowania oraz mechanizmy kontroli i monitorowania, procesy i procedury ustanowione, stosowane lub udostępniane zgodnie z dyrektywą. Jej zakres tematyczny określono w załącznikach nr I, II i III dyrektywy; <http://www.geoportal.gov.pl/>. Transpozycją dyrektywy na grunt polski jest ustawa z 4 marca 2010 r. o infrastrukturze informacji przestrzennej (Dz. U. z 2010 r. Nr 76, poz. 489).

- Komponenty — usługi sieciowe OGC/ISO: WMS⁶, WFS⁷, WCS⁸, CSW⁹.

10) posiadać strukturę scentralizowaną, funkcjonującą w infrastrukturze technicznej GUS.

W przyszłości planowane jest powiększenie portalu geostatystycznego o prezentację wyników pozostałych badań statystycznych, a także zharmonizowanie prezentacji wyników badań pomiędzy portalem i dotychczasowymi bazami, np. Bankiem Danych Lokalnych (BDL). Portal będzie również umożliwiał podłączenie innych danych, np. sieć dróg, czy sieć kolei bądź obszary chronione poprzez usługi WMS z zewnętrznych portali, np. <http://maps.geoportal.gov.pl/webclient/>.

Planowane jest powiększenie portalu o węzeł wojewódzki, co zostało zapisane w planie informatyzacji państwa. W każdym urzędzie statystycznym system GIS¹⁰ składać się będzie z dwóch kluczowych elementów: profesjonalnej bazy danych wraz ze środowiskiem zarządzania danymi przestrzennymi oraz ze środowiska edycji i przetwarzania danych. Dane lokalne po przesłaniu na poziom wojewódzki utworzą spójną przestrzenną bazę danych. W ten sposób w województwach powstaną (w ograniczonym zakresie informacyjnym) bazy statystyczne punktów adresowych. Będą one uzupełnione o zbiór precyzyjnych obiektów wektorowych, będących łącznikiem pomiędzy danymi opisowymi rejestru TERYT i BJS a położeniem przestrzennym wyrażonym w postaci współrzędnych. Nastąpi zatem integracja statystycznych danych przestrzennych

⁶ WMS — *Web Map Service* — sposób udostępniania danych przez serwer map, a także określania przez serwer swoich możliwości w tym zakresie. Mapy te są obrazowane na podstawie danych geograficznych zawartych w bazie i wyświetlane po stronie klienta w takich formatach rastrowych, jak GIF, JPEG, PNG, <http://ikar.pgi.gov.pl/>. Przykładem usług WMS są serwisy zawarte w <http://maps.geoportal.gov.pl/webclient/>.

⁷ WFS — *Web Feature Service* — międzynarodowy standard pozwalający na pobieranie i wykonywanie operacji (edycję) na danych zakodowanych w GML, dostępnych na serwerach typu *Web Feature*. Dane przesyłane są w postaci wektorowej, *Słownik IKAR*, <http://ikar.pgi.gov.pl/>.

⁸ WCS — *Web Coverage Service* — międzynarodowy standard umożliwiający interoperacyjny dostęp do danych geoprzestrzennych zapisanych w postaci pokryć macierzowych. Termin „pokrycie” zwykle odnosi się do zdjęć satelitarnych, cyfrowych zdjęć lotniczych, numerycznych modeli terenu i innych zjawisk, których zmienność przestrzenna jest reprezentowana za pomocą pokryć rastrowych, *Słownik IKAR*, <http://ikar.pgi.gov.pl/>.

⁹ CSW — *Catalogue Service for Web* — serwisy katalogowe umożliwiające publikowanie i przeszukiwanie zbiorów informacji opisowej (metadanych) dla danych, serwisów i obiektów powiązanych; *Słownik IKAR*, <http://ikar.pgi.gov.pl/>.

¹⁰ GIS — ang. *Geographical Information System* (GIS) — system informacji przestrzennej dotyczący danych geograficznych; system pozyskiwania, gromadzenia, weryfikowania, integrowania, analizowania, transferowania i udostępniania danych przestrzennych, w szerokim rozumieniu obejmuje on metody, środki techniczne, w tym sprzęt i oprogramowanie, bazę danych przestrzennych, organizację, zasoby finansowe oraz ludzi zainteresowanych jego funkcjonowaniem; termin ten, w liczbie mnogiej — systemy informacji geograficznej, stosowany jest również jako nazwa dziedziny zajmującej się geoinformacją oraz metodami i technikami GIS; *Internetowy leksykon geomatyczny*, Zespół Redakcyjny: Jerzy Gaździcki, Janusz Michałak, Ewa Musiał, <http://www.ptip.org.pl/>.

z danego województwa, która będzie podlegała kontroli i weryfikacji danych operacyjnych.

WYKORZYSTANIE DOŚWIADCZEŃ Z REALIZACJI PROJEKTU „PSR 2010 I NSP 2011”

W celu wykorzystania doświadczeń zdobytych podczas realizacji spisów powszechnych w GUS powołano zespół do spraw przygotowania wdrożenia nowej koncepcji organizacji badań statystycznych oraz określenia niezbędnej infrastruktury informatycznej przy wykorzystaniu doświadczeń w realizacji projektu „PSR 2010 i NSP 2011”. Zadaniem zespołu jest sprecyzowanie koncepcji organizacji badań statystycznych oraz opracowanie sposobu i harmonogramu wdrożenia tej koncepcji, a także analiza możliwości wykorzystania systemów wdrożonych w spisach powszechnych w kolejnych badaniach prowadzonych przez statystykę publiczną. W skład zespołu wchodzi podzespół do spraw:

- 1) metodyki, standaryzacji i koordynacji badań,
- 2) infrastruktury teleinformatycznej,
- 3) zbierania danych,
- 4) standaryzacji statystycznych struktur danych wejściowych,
- 5) przetwarzania danych,
- 6) udostępniania danych.

Dla odbiorców danych statystycznych najważniejsze znaczenie ma podzespół ds. udostępniania danych. Jest on odpowiedzialny za przygotowanie projektów i kierunków działania w zakresie zasad udostępniania i rozpowszechniania informacji statystycznych i metadanych oraz ustalenie zasad współpracy międzynarodowej przy wymianie danych i metadanych statystycznych, w tym z instytucjami Unii Europejskiej i innymi organizacjami międzynarodowymi. Prace podzespołu skupiają się także na opracowaniu założeń do systemów informatycznych wspierających udostępnienie danych.

mgr inż. Janusz Dygaszewicz — GUS

SUMMARY

This paper describes the spatial address databases which on the basis of data obtained from censuses will allow to prepare analyzes at different levels of aggregation. The main part of the paper is devoted to characteristics of being implemented by the Geostatistical Portal of the Polish CSO (PGS). The creation of the PGS allows to present results of the Agricultural Census 2010 and the National Census of Population and Housing 2011 and disseminate the aggregated data in the form of spatial analyses. The article gives some examples of "ready" services on the PGS, and brought closer the possibility of generating a variety of own analyzes using the site functions.

РЕЗЮМЕ

В статье представляются пространственные адресные базы подготовленные для всеобщих переписей, которые на основе данных полученных из переписей позволят разрабатывать анализ на разных уровнях обобщения. Основная часть статьи является характеристикой осуществляемого ЦСУ Геостатистического портала (PGS). Целью разработки PGS является презентация результатов Всеобщей сельскохозяйственной переписи 2010 и Национальной всеобщей переписи населения и квартир 2011, а также предоставление обобщенных данных в виде пространственного анализа. В статье приведены примерные, «разработанные», информации находящиеся в PGS, а также представляются возможности самостоятельного анализа с использованием предоставленных функций портала.