

Internetowy Atlas Polski Niepodległej – metodyka opracowania bazy miejscowości i granic administracyjnych oraz analizy pochodne

Andrzej Głazewski

<https://orcid.org/0000-0003-1884-3682>
Politechnika Warszawska

Paweł J. Kowalski

<https://orcid.org/0000-0003-4228-9503>
Politechnika Warszawska

Zarys treści: Artykuł przedstawia przykłady wykorzystania zasobu historycznych danych przestrzennych, zgromadzonego w ramach Internetowego Atlasu Polski Niepodległej, opublikowanego w 2021 r. jako efekt projektu Instytutu Historii PAN i Wydziału Geodezji i Kartografii Politechniki Warszawskiej. Zasadniczym celem artykułu jest ukazanie możliwości badania struktury sieci osadniczej z wykorzystaniem analiz przestrzennych oraz propozycja sposobów odtwarzania metodami kartograficznymi struktur przestrzennych: grup obiektów osadniczych czy jednostek administracyjnych.

Słowa kluczowe: portal mapowy, atlas historyczny, Druga Rzeczpospolita, Polska niepodległa, mapa taktyczna WIG, *Słownik geograficzny Królestwa Polskiego* (SGKP), analizy przestrzenne, integracja danych przestrzennych

Abstract: The article presents examples of the use of historical spatial data collected for the Internet Atlas of Independent Poland. The Atlas was published in 2021 as a result of the project of the Institute of History of the Polish Academy of Sciences and the Faculty of Geodesy and Cartography of the Warsaw University of Technology. The article's main aim is to show the possibility of investigating the structure of the settlement network using spatial analyses and to propose ways of reconstructing spatial structures employing cartographic methods: groups of settlement objects or administrative units.

Keywords: map portal, historical atlas, Second Republic of Poland, independent Poland, WIG (Military Geographical Institute) tactical map, *Geographical Dictionary of the Kingdom of Poland* (SGKP), spatial analysis, spatial data integration

1. Wprowadzenie

W 2021 r. ukazał się Internetowy Atlas Polski Niepodległej (IAPN), którego głównym przeznaczeniem była prezentacja zintegrowanych danych historycznych i geograficznych. W georeferencyjnej bazie danych utworzonej w ramach projektu znalazły się trzy komponenty: jednostki podziału administracyjnego, obiekty osadnicze z obszaru Drugiej Rzeczypospolitej wraz z nazwami i podstawowym zestawem atrybutów oraz tabela toponimów ze *Słownika geograficznego Królestwa Polskiego*

(SGKP)¹. Granice administracyjne województw i powiatów oraz punkty reprezentujące miejscowości i mniejsze obiekty osadnicze zostały zwektoryzowane z mapy taktycznej Wojskowego Instytutu Geograficznego w skali 1:100 000 (WIG)², materiałem źródłowym była też mapa administracyjna Głównego Urzędu Statystycznego

¹ *Od „Słownika geograficznego Królestwa Polskiego” do map topograficznych Wojskowego Instytutu Geograficznego*, red. T. Epsztajn, Warszawa 2021.

² B. Krassowski, *Polska kartografia wojskowa w latach 1918–1945*, Warszawa 1974.

(GUS) 1:300 000 z 1937 r.³ Z kolei nazwy miejscowości z haseł piętnastotomowego SGKP zostały pozyskane w drodze rozpoznawania obrazowego fragmentów tekstu (OCR) oznaczonych na skanach kart Słownika. Trzy zestawy danych trafiły do bazy danych w celu powiązania odpowiadających sobie obiektów.

Powiązanie danych przestrzennych (punktów miejscowości) oraz opisowych (haseł SGKP) było kluczowym problemem technicznym i merytorycznym rozwiązaniem w projekcie w dwóch etapach: automatycznej integracji oraz manualnej weryfikacji i korekty powiązań. Integracja bazowała na nazwach własnych punktów miejscowości na mapie WIG oraz haseł SGKP. Konieczne było jednak przeprowadzenie manualnej weryfikacji wszystkich powiązań i eliminacja błędnych par, co zostało opisane w rozdziałach 3.1 oraz 3.2. W ramach drugiego etapu uzupełniano także referencję przestrzenną dla niepowiązanych automatycznie haseł SGKP.

Utworzona baza danych stała się podstawą serwisu geoinformacyjnego, który funkcjonuje jako niezależny internetowy atlas historyczny, ale może stać się również węzłem infrastruktury danych przestrzennych dla zastosowań historycznych, np. do publikacji wyników innych projektów badawczych. Przy użyciu oprogramowania GIS możliwe jest analizowanie danych w kontekście geograficznym i w wymiarze chronologicznym, co zostało przedstawione w niniejszym artykule na kilkunastu przykładach analiz przestrzennych.

2. Analizy przestrzenne w środowisku systemów informacji geograficznej

Systemy informacji przestrzennej (ang. GIS) nie są jedynie oprogramowaniem czy zestawami baz danych, ale stanowią złożone systemy geoinformacyjne, które łączą w sobie zarówno aplikacje (software), sprzęt,

zasoby i rejestry danych geoprzestrzennych, jak i metody oraz technologie z użytkownikami. Wszystkie te elementy, zintegrowane funkcjonalnie, dają możliwość zaawansowanego przetwarzania informacji przestrzennej. Informację przestrzenną stanowią wszelkie dane wraz z ich kontekstem, które posiadają odniesienie geograficzne, wyrażone w konkretnym układzie współrzędnych lub w sposób niemetryczny (systemy kodowe, nazewnicze, skorowidze, siatki regularne, regionalizacje)⁴. Dane te mogą tworzyć specyficzne struktury topologiczne: sieć linową, wypełnienie powierzchni, triangulację, sieci regularne – struktury, które modelują z pewnym przybliżeniem rzeczywistość przestrzenną. Systemy geoinformacyjne realizują wiele zadań – od pozyskiwania i strukturalizacji danych do ich udostępniania w różnych formach, lecz zawsze sercem ich funkcjonalności jest wspomniane przetwarzanie wraz z wszechstronnym analizowaniem danych. Jedynie systemy informacji geograficznej wykorzystują metody i technologie efektywnego przetwarzania danych przestrzennych, gwarantując poprawność metodyczną, szybkość obróbki danych, możliwość poprawnego wizualizowania, oraz dając pożądaną dostępność do informacji, poufność i możliwość ograniczeń w tym dostępie.

W systemach informacji przestrzennej różne rodzaje analiz mogą być wykonywane na różnych strukturach danych, można jednak zestawić najczęściej wykonywane typy analiz przestrzennych, porządkując je ze względu na stopień zaawansowania. Wśród analiz elementarnych można wymienić:

- obliczenia atrybutów i pomiary geometrii (zastosowanie operatorów tekstowych, logicznych, matematycznych oraz narzędzi kartometrii);
- selekcję obiektów wg warunków atrybutowych lub operatorów funkcyjnych, logicznych;

³ R. Cebertowicz, *Mapa Administracyjna Rzeczypospolitej Polskiej 1:300 000*, [b.m.w.] 1937.

⁴ *GIS. Teoria i praktyka*, red. P. Longley i in., Warszawa 2006.

- rekłasyfikację i agregację przestrzenną;
- generowanie pochodnej geometrii (np. ekwidystant – stref buforowych, obwiedni itp.).

Natomiast jako analizy złożone można uznać poniższe rodzaje przetworzeń danych:

- analizy zależności i wpływów na otoczenie;
- badanie intersekcji klas obiektów, selekcja wg położenia i relacji topologicznych;
- analizy sieciowe (w strukturach grafowych);
- analizy powierzchni statystycznych, interpolacja, zastosowanie algebry map;
- analizy wielokryterialne.

Analizy obliczeniowe i kartometryczne polegają na przetwarzaniu opisowej części danych przestrzennych, np. obliczeniach wartości atrybutów przy użyciu szeregu operatorów, w tym logicznych i matematycznych. Wyniki pomiarów obiektów przestrzennych są zapisywane do bazy danych jako kontynuacja opisu tych obiektów. Są to zwykle pomiary cech geometrycznych obiektów – takich jak pole powierzchni, obwód, długość, wartości współrzędnych, odległości. Odpowiednio uporządkowane stanowią dodatkową – szybko dostępną – informację, która może być np. analizowana przez obróbkę statystyczną (np. obliczenia średnich długości, sum powierzchni itp.) bądź może być podstawą źródłową do prezentacji kartograficznej.

Najczęściej wykonywanym typem analiz jest selekcja obiektów według warunków atrybutowych oraz operatorów różnego typu. Podobnie jak w przeszłości w pracy z mapą analogową często chodziło o wyszukanie interesującej informacji, tak i dziś, korzystając z odpowiednich metod i operatorów (np. języka SQL), można efektywnie wyselekcjonować zestawy danych interesujące użytkownika.

Rekłasyfikacja danych to wykorzystanie atrybutów lub innych cech obiektów

(np. geometrycznych lub sąsiedztwa) do ich klasyfikacji lub grupowania. Grupowanie to może mieć postać agregacji, czyli tworzenia nowych, uogólnionych obiektów, np. jednostek wyższego rzędu, szerszych pojęciowo kategorii itd.

Elementarnym i często występującym typem analiz przestrzennych jest generowanie pochodnej geometrii. Są to zadania prowadzące do budowy obwiedni obiektów czy izolinii (np. ekwidystant, zwanych często strefami buforowymi), dających nowe możliwości analityczne, związane z otoczeniem obiektów przestrzennych. Badając zależności zjawisk, bierze się pod uwagę zwykle różne klasy obiektów źródłowych, dlatego generowanie nowej geometrii stosuje się często w analizach wielokryterialnych.

W metodyce modelowania baz danych geograficzno-historycznych stosuje się wiele różnych podejść – odmiennych przede wszystkim ze względu na odniesienie czasowe⁵: Model Ujęć Czasowych, Złożony Model Czasowo-Przestrzenny, Jednorodny Obiektowy Model Czasowo-Przestrzenny, Model Czasowo-Przestrzenny Zorientowany na Wydarzenie, Model Trójdomenowy czy też Geoprzestrzenny Model Wydarzeń. W projekcie IAPN podstawowy zestaw danych geometrycznych dla miejscowości był odniesiony do jednego okresu referencyjnego, a charakterystykę czasową można było określić dopiero w wyniku analiz. Dlatego we wstępnych założeniach przyjęto model struktury bazujący na dwu segmentach: źródłowym (dokumentacyjnym) i krytycznym (wynikowym)⁶. Segmenty te, spójne modelowo, zawierają odpowiednio: zapis dosłowny informacji źródłowej oraz krytyczną, zweryfikowaną przez historyka informację wynikową. Wszelkie wyniki

⁵ T. Panecki, „Koncepcja struktury bazy danych historycznych obiektów topograficznych”, rozprawa doktorska, Uniwersytet Warszawski, Warszawa 2018.

⁶ B. Szady, *Czasowo-przestrzenne bazy danych jako narzędzie w geografii historycznej*, „Acta Universitatis Lodzianensis, Folia Geographica Socio-Oeconomica”, 14, 2013, s. 17–32.

analiz przestrzennych mieszczą się w tym modelu w części krytycznej jako efekt przetwarzania źródeł i wnioskowania.

3. Procedura gromadzenia i integracji danych geograficzno-historycznych

Funkcje analiz przestrzennych zaimplementowane w aplikacjach GIS umożliwiły odpowiednią organizację zasobów na etapie realizacyjnym bazy danych historycznych IAPN, ale również były środkiem weryfikacji postępów w pracach. Wykorzystane w projekcie oprogramowanie QGIS w wersjach od 3.14 do 3.18 w zupełności zaspokajało wszelkie potrzeby geoprzetwarzania. Zaprojektowany w ramach prac koncepcyjnych schemat bazy danych objął dwa zestawy danych przestrzennych: granic administracyjnych i punktów miejscowości oraz tabelę haseł SGKP. Gromadzenie podstawowych danych przestrzennych odbyło się w drodze wektoryzacji położenia punktów miejscowości i przebiegu granic na mapach WIG. Z kolei tabela haseł powstała poprzez indeksację – oznaczenie położenia wszystkich wytluszczonych nazw na poszczególnych stronach SGKP, a następnie interpretację obrazową oznaczonych pól z nazwami⁷. Pozyskano łącznie 99 114 punktów miejscowości z obszaru Drugiej Rzeczypospolitej, w tym 61 955 w obecnych granicach państwa oraz 146 630 nazw miejscowości, dla których przypisano identyfikator jednostki administracyjnej TERYT_H, spośród 166 587 wszystkich nazw miejscowości pozyskanych z haseł SGKP.

3.1. Pozyskiwanie danych geograficznych

Lokalizację obiektów osadniczych wyznaczano na podstawie mapy taktycznej WIG w skali 1:100 000, przy czym źródłem geometrii był Państwowy Rejestr Nazw Geograficznych (PRNG) dla obszaru

⁷ O. Graszka, *Automatyzacja procesu rozpoznawania i weryfikacji nazw geograficznych ze źródeł historycznych na przykładzie Słownika geograficznego Królestwa Polskiego*, w: *Od „Słownika geograficznego Królestwa Polskiego*, s. 21–32.

współczesnej Polski oraz baza OpenStreetMap – dla pozostałych terenów Drugiej Rzeczypospolitej. Odtwarzanie sytuacji osadniczej z roku 1931 polegało na korektach położenia punktu reprezentującego miejscowość i poprawkach atrybutów nazwy i typu obiektu, jeśli istniał we wspomnianych bazach, lub wstawieniu nowego punktu centralnego i uzupełnieniu wymaganych atrybutów. Przyjęto zasady lokalizacji punktu centralnego dla miejscowości zgodnie z wytycznymi PRNG⁸. Jeśli ustalenie właściwego położenia nie było możliwe wyłącznie w oparciu o podkład rastrowy WIG, korzystano z danych topograficznych (dróg, kolei i cieków wodnych) z Bazy Danych Obiektów Topograficznych BDOT10k oraz bazy OpenStreetMap.

Lokalizowane punktowo miejscowości wymagały referencji administracyjnej, niezbędnej na kolejnych etapach integracji. W oparciu o zwektoryzowane na podkładach rastrowych map WIG linie granic powiatów – za pomocą narzędzia konwersji typów geometrycznych „Linie na poligony” z panelu „Algorytmów processingu”⁹ – wygenerowane zostały obiekty powierzchniowe. W związku z dość dużą rozciągłością czasową edycji mapy taktycznej WIG (ok. 20 lat) jako źródło uzupełniające danych administracyjnych wykorzystano wydane arkusze mapy szczegółowej WIG 1:25 000 oraz wspomnianą już mapę administracyjną GUS z 1937 r. i inne źródła¹⁰. Danymi źródłowymi w formie cyfrowej stały się współczesne zbiory Państwowego Rejestru Granic i powierzchni jednostek podziałów

⁸ A. Jagura, A. Kocot, P. J. Kowalski, *Metodyka odtworzenia sieci osadniczej z obszaru II Rzeczypospolitej*, w: *Od „Słownika geograficznego Królestwa Polskiego”*, s. 47–63.

⁹ W programie QGIS wszystkie narzędzia dostępne są w panelu „Algorytmy processingu”, a najczęściej używane – bezpośrednio w menu „Wektor”.

¹⁰ A. Głazewski, A. Jagura, A. Kocot, *Podział administracyjny – metodyka pozyskiwania historycznych danych topograficznych z obszaru II Rzeczypospolitej*, w: *Od „Słownika geograficznego Królestwa Polskiego”*, s. 33–46.

terytorialnych kraju (PRG) oraz dane dotyczące podziału państw sąsiednich na jednostki administracyjne, dostępne w ramach bazy danych EuroGlobalMap, a dla obszaru Białorusi – dane ze zbiorów DIVA-GIS, udostępnione przez www.diva-gis.org. Następnie połączonym topologicznie obszarom przypisano identyfikatory i nazwy. Odtworzono obszary powiatów na rok referencyjny 1931¹¹, ale także – w oparciu o dodatkowe źródła¹² – na rok referencyjny 1900. Zachowanie pierwotnej geometrii liniowej granic było uzasadnione technicznie i merytorycznie: łatwością ewentualnych korekt przebiegu i odwołania się do metadanych źródła (godła arkusza mapy, z której zostały pozyskane, czy np. roku jej wydania). Z kolei geometria powierzchniowa była zastosowana w procesie regionalizacji. Regionalizacja zbioru punktów miejscowości według powiatów na rok referencyjny 1900 została przeprowadzona z wykorzystaniem zapytania „Złącz atrybuty według lokalizacji” – z operatorem zawierania się obiektów (*contains*). Umożliwiła zawężenie obszarów wyszukiwania powiązań między danymi z map i ze SGKP.

3.2. Automatyczna integracja zbiorów

Powiązanie miejscowości pozyskanych z map WIG oraz haseł SGKP odbyło się w dwóch etapach: automatycznym i manualnym. W każdym z etapów procedurę poszukiwania powiązań wykonywano iteracyjnie. Zapytania SQL łączące zbiory w typowym zadaniu bazodanowym opierają się na atrybutach (kluczach) wiążących rekordy o takich samych wartościach tych atrybutów. Jednakowe wartości – zazwyczaj

identyfikatora lub nazwy własnej obiektu – są niezbędne, aby zaistniała relacja między tabelami. W przypadku danych historycznych trudno jest wskazać inny niż nazwa własna atrybut, który mógłby pełnić rolę klucza łączącego. Jednak nazwa geograficzna może zmieniać się w czasie, przyjmować różną formę ortograficzną czy zmieniać się w zależności od wpływów regionalnych, a zatem nie jest w pełni skutecznym rozwiązaniem podczas poszukiwania relacji pomiędzy obiektami o różnym datowaniu. W projekcie IAPN pomiędzy źródłowymi materiałami było kilkadziesiąt lat różnicy, co więcej oba źródła charakteryzuje spora rozpiętość czasowa: 22 lata pomiędzy pierwszym i ostatnim tomem SGKP i nawet do kilkunastu lat różnicy pomiędzy arkuszami map WIG¹³. Zmienność historyczna toponimii, ale także różnice ortograficzne, fleksyjne, a często także błędy drukarskie różnych źródeł nazw wymusiły zastosowanie iteracyjnej procedury łączenia źródeł z wykorzystaniem zapytań atrybutowo-przestrzennych.

Proces automatycznej integracji zbiorów opierał się na założeniu tożsamości odpowiadających sobie co do nazwy: miejscowości widniejącej na mapie WIG oraz hasła SGKP. Ze względu na powtarzalność nazw obiektów osadniczych na obszarze kraju konieczne było zawężenie obszarowe poszukiwań par o identycznych nazwach. Przyjęto, że ograniczenie terytorium do obszaru powiatu będzie możliwe dzięki znajdującej się w hasle słownikowym informacji o przynależności miejscowości do jednostki administracyjnej. Niejednoznaczność regionalizacji wynikająca ze zmian terytoriów powiatów w końcu XIX w. wymusiła zastosowanie rozbudowanej tabeli mapowania identyfikatorów powiatów SGKP na zdefiniowane w projekcie wartości identyfikatorów

¹¹ Rok referencyjny oznacza wybrany rok, dla którego przebiegi granic, zwektoryzowane z arkuszy map WIG o różnym datowaniu, zostały ujednolicone w oparciu o inne dostępne źródła.

¹² *Mappa Gubernij Królestwa Polskiego dla cyklistów*, M. Poturaj, Warszawa 1894; także opracowania współczesne np.: *Królestwo Polskie. Podział administracyjny 1907* (https://pl.wikipedia.org/wiki/Podzia%C5%82_administracyjny_Kr%C3%B3lestwa_Polskiego#/media/Plik:Polska_1907_adm.png, dostęp: 1 czerwca 2020).

¹³ T. Epsztajn, *Słownik geograficzny Królestwa Polskiego i innych ziem słowiańskich – zarys historii dzieła*, w: *Od „Słownika geograficznego Królestwa Polskiego”*, s. 85–115.

TERYT_H¹⁴. W tabeli mapowania zapisane zostały zarówno wyjściowe pary identyfikatorów powiatów dla roku 1900, jak i warianty przynależności terytorialnej wynikające ze zmian administracyjnych zachodzących w latach wydawania *Słownika*. Były to przypadki powstających w tych latach jednostek lub zmienionych granic powiatowych. Należy dodać, że informacja o kolejnych iteracjach została wpisana w „statusie rekordu”¹⁵ tabeli integracji na wypadek wykrycia w późniejszych etapach błędów mapowania jednostek terytorialnych.

Ostatecznie zapytanie atrybutowe: „Złącz atrybuty według wartości pola”, w oparciu o „nazwę główną” z ograniczeniem do obszarów powiatów, dało wynik w postaci tabeli pośredniczącej liczącej 66 316 rekordów, zwanej dalej tabelą integracji. Tabela ta zawierała wszystkie powiązane kombinacje haseł SGKP (w liczbie 50 295, czyli 34 proc. wszystkich pozyskanych z SGKP) z punktami miejscowości WIG (w liczbie 45 376, czyli 46 proc. całego zbioru).

3.3. Etap manualnej weryfikacji i uzupełniania powiązań

W ramach drugiego etapu projektu zrealizowano przede wszystkim referencję przestrzenną dla niepowiązanych wcześniej haseł SGKP. Sprawdzono i dodano identyfikację dla ponad 52 tys. nazw pozyskanych z haseł SGKP. Brak powiązań, będący skutkiem rozbieżności nazw miejscowości, wynikał zarówno z różnej metodyki ich gromadzenia w obu źródłach, jak i ewolucji sieci osadniczej¹⁶. Wszelkie wątpliwości pojawiające się podczas poszukiwania

odpowiedników zostały wpisane w atrybucie „Niepewność” i jednocześnie taki rekord otrzymywał status powiązania niepewnego. Jednym z ciekawszych problemów, trudnych do rozwiązania bez potwierdzenia w dodatkowych źródłach, była niemożność powiązania takich obiektów osadniczych, które były wpisane w jednym haśle SGKP, a na mapach reprezentowane były przez liczne obiekty o indywidualnych nazwach. Takie sytuacje oznaczano jako „powiązanie grupowe”.

Drugim celem manualnej weryfikacji powiązań była eliminacja błędnych par – przede wszystkim spośród wielokrotnych powtórzeń popularnych nazw. Wykorzystano przy tym tabelę integracji, uzupełniając jej atrybuty o dwie wartości: „liczby powtórzeń hasła” oraz „liczby powtórzeń punktu” (ryc. 1). Do obliczenia statystyk powtórzeń wykorzystano kwerendy zliczające „Statystyki wg kategorii”. Umożliwiły one wykrycie najliczniejszych błędnych powiązań tych samych nazw w obrębie powiatów. Ponad 2300 powiązań oznaczono jako błędne i usunięto z tabeli integracji, dzięki czemu kolejna iteracja zbioru powiązań zbliżyła całość procesu integracji do zadowalającego poziomu wiarygodności.

4. Badanie struktury sieci osadniczej

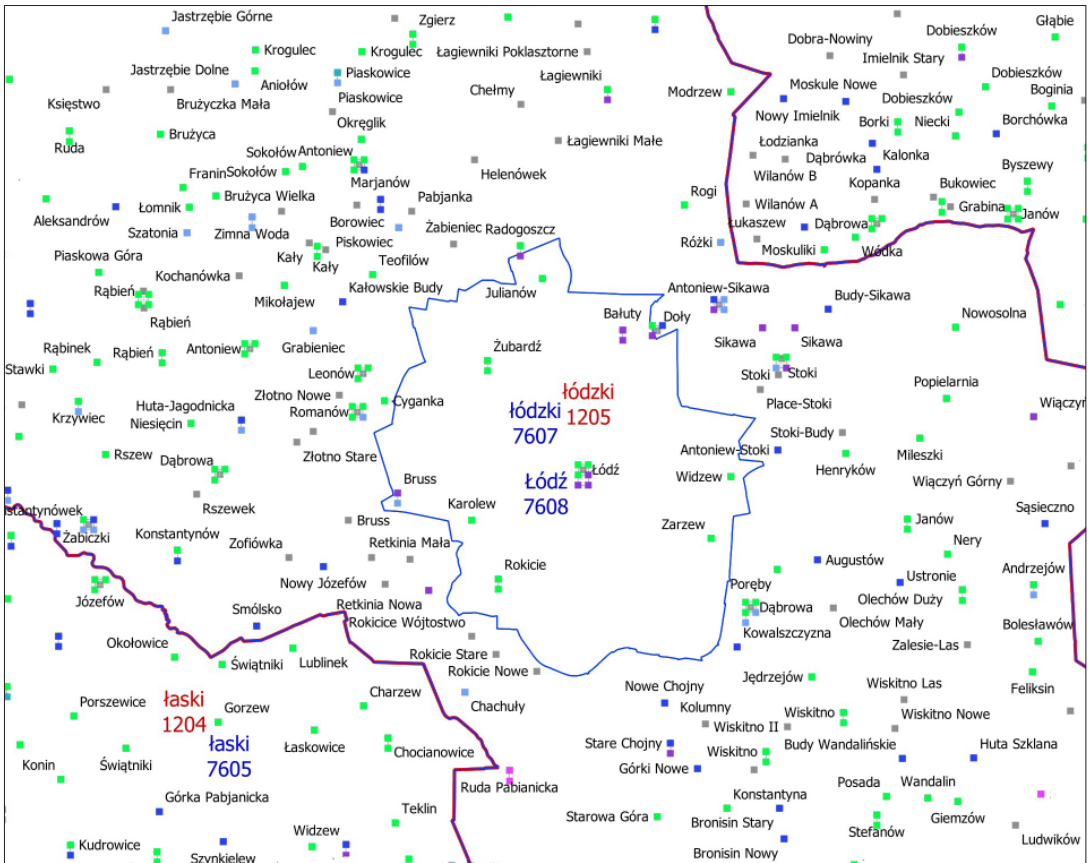
Dzięki zgromadzonym w ramach projektu danym przestrzennym możliwe jest wszechstronne analizowanie sieci osadniczej – także w kontekście regionalizacji administracyjnej. W ramach projektu zaproponowano trzy obszary badań: analizy struktur przestrzennych, analizy porównawcze i badanie dynamiki zmian oraz odtwarzanie nieskartowanych lub niezachowanych na mapach struktur przestrzennych: grup obiektów osadniczych czy jednostek administracyjnych¹⁷.

¹⁴ Identyfikator TERYT_H został zdefiniowany w oparciu o system urzędowej identyfikacji współczesnych jednostek administracyjnych, z tym że został określony dla lat referencyjnych 1900 i 1931.

¹⁵ „Status” został przypisany wszystkim rekordom tabel źródłowych i pochodnych dla oznaczenia stanu powiązania obiektów, poziomu niepewności powiązania i realizacji kolejnych etapów prac.

¹⁶ A. Jagura, A. Kocot, P.J. Kowalski, *Metodyka odtworzenia sieci osadniczej*, s. 47–63.

¹⁷ K. Osińska-Skotak i in., *Analizy przestrzenne i wizualizacja danych historycznych IAPN w środowisku systemów informacji geograficznej*, w: *Od „Słownika geograficznego Królestwa Polskiego”, s. 65–84.*



Ryc. 1. Fragment zbioru miejscowości z uwidocznieniem liczby powiązań dla danego punktu miejscowości i statusu powiązania (kolor zielony – powiązanie automatyczne, niebieski – manualne, fioletowy – niepewne, szary – punkt niepowiązany).

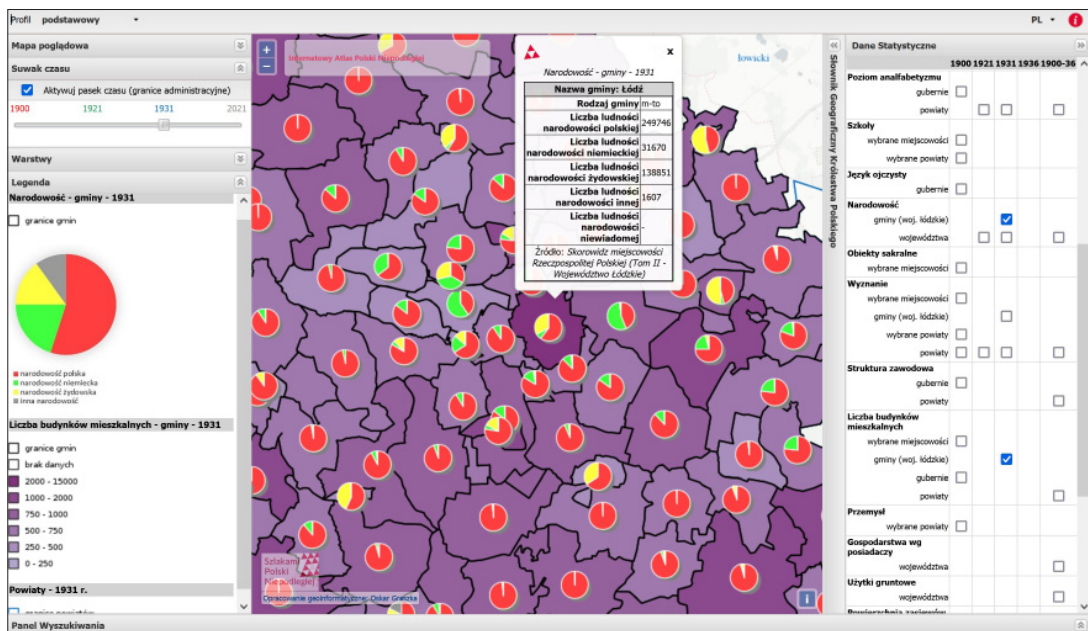
Źródło: oprac. własne

Zrealizowano tylko część z planowanych analiz przestrzennych. Obliczono takie parametry jak: liczebność według typu obiektu osadniczego, gęstość i współczynnik rozproszenia (średnia odległość pomiędzy najbliższymi sąsiadującymi ze sobą punktami). Wyniki zostały zwizualizowane z wykorzystaniem różnych metod prezentacji kartograficznej: sygnatur ilościowych, kartogramów właściwych, kartodiagramów i izolinii. Pozostałe, m.in. wspomniane w poprzednim rozdziale powiązania grupowe, wymagają dalszych prac badawczych, a przede wszystkim dodatkowych źródeł informacji mereologicznej.

4.1. Kontekst administracyjny analiz przestrzennych

Badanie struktury zjawisk rozproszonych zwykle odbywa się w referencji do obszarów stanowiących, np. administracyjnych lub sztucznych, jak geometryczne pola odniesienia. Otrzymane liczby bezwzględne (liczby miejscowości) lub pochodne parametry rozproszenia (średnich odległości) czy gęstości dają możliwość analiz porównawczych. Każda z delimitacji (administracyjna lub geometryczna) ma swoje zalety i zastosowania¹⁸. Pierwsza nadaje się do porównywania różnych wielkości

¹⁸ P. Cebrykow, *Generalizacja map statystycznych*, Lublin 2017.



Ryc. 2. Prezentacja danych statystycznych dla gmin województwa łódzkiego: porównanie struktury narodowościowej i liczby domów. Źródło: Internetowy Atlas Polski Niepodległej, 2021 (<https://atlas1918.ihpan.edu.pl>, dostęp 10 października 2022)

statystycznych, gromadzonych zwykle dla jednostek administracyjnych; w projekcie IAPN były to powiaty dla całego obszaru Drugiej Rzeczypospolitej, a dla jednego województwa – gminy (ryc. 2). Mapy tematyczne wybranych danych społecznych i gospodarczych przygotowano w oparciu o kilka źródeł: *Statystykę Polski* z 1915¹⁹, spisy powszechne z lat 1921 i 1931, *Skorowidz miejscowości Rzeczypospolitej Polskiej* z 1925²⁰, „Rocznik Statystyki Miast Polski”²¹ oraz spisy rolne z 1930/31 i 1936 r.²²

Równoległe z zadaniami realizowanymi w projekcie podjęto próby alternatywnego wykorzystania bazy danych opracowanej w ramach kilku prac dyplomowych na Wydziale Geodezji i Kartografii

Politechniki Warszawskiej. W ramach jednego z tematów opracowany został interaktywny atlas tematyczny, w którym zaprezentowano możliwość porównywania danych statystycznych z lat 1921 i 1931²³. W tym ujęciu możliwe jest zestawianie przekrojów czasowych pomimo tego, że granice administracyjne uległy pewnym zmianom. Inną zaletą tego rozwiązania jest ułatwienie odbiorcy map samodzielnego wyboru tematów prezentacji (ryc. 3).

4.2. Kontekst regionalizacji przestrzennych

Drugi sposób odniesienia, wykorzystujący sztucznie wygenerowaną referencję przestrzenną, jest użyteczny z perspektywy analiz wieloczasowych. W sytuacji zmiennego na przestrzeni lat podziału administracyjnego pola tworzące układ regularny zapewniają stabilną referencję przestrzenną

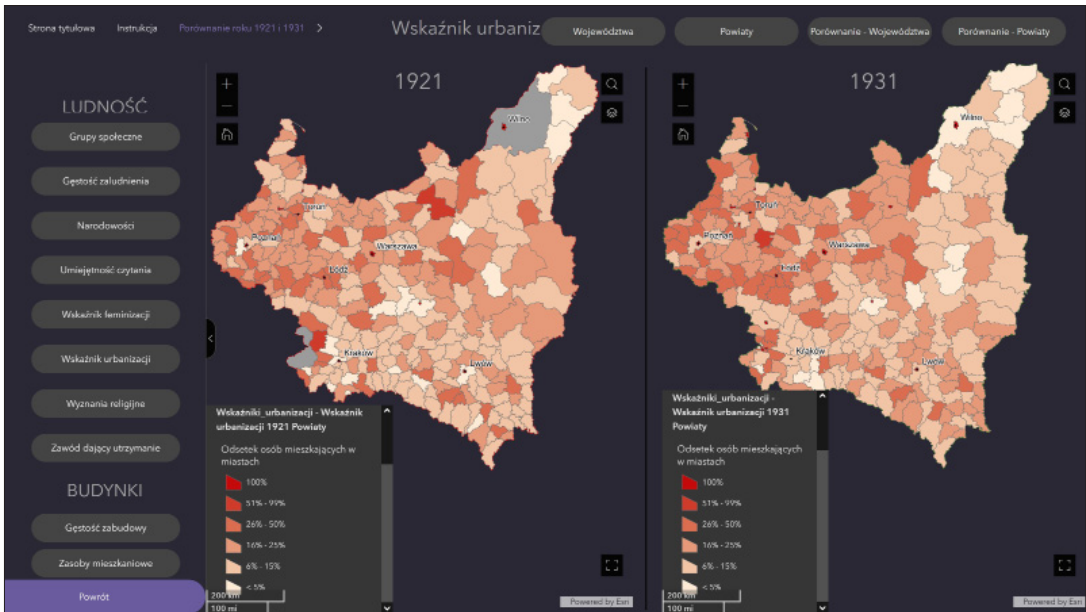
¹⁹ K.W. Kumaniecki, A. Krzyżanowski, *Statystyka Polski*, Kraków 1915.

²⁰ *Skorowidz miejscowości Rzeczypospolitej Polskiej opracowany na podstawie wyników Pierwszego Powszechnego Spisu Ludności z dn. 30 września 1921 r. i innych źródeł urzędowych*, Warszawa 1925.

²¹ „Rocznik Statystyki Miast Polski”, 1, 1928.

²² *Statystyka Rolnicza 1930/31*, Warszawa 1932; *Statystyka Rolnicza 1936*, Warszawa 1937.

²³ M. Gut, „Opracowanie atlasu tematycznego zmian demograficznych Polski po odzyskaniu niepodległości”, praca dyplomowa inżynierska, Archiwum Prac Dyplomowych Politechniki Warszawskiej 2021.



Ryc. 3. Przykład map statystycznych z lat 1921 i 1931 w internetowym atlasie tematycznym.

Źródło: M. Gut, „Opracowanie atlasu tematycznego zmian demograficznych Polski po odzyskaniu niepodległości”, praca dyplomowa inżynierska, Archiwum Prac Dyplomowych Politechniki Warszawskiej 2021

w dowolnych okresach. Dają zatem możliwość policzenia wartości zagregowanych tj. zsumowanych lub uśrednionych w każdym polu odniesienia, porównania ich pomiędzy różnymi przekrojami czasowymi czy też obliczenia wartości różnicowych.

Proces generowania pól odniesienia, zwanych zwykle polami podstawowymi (w QGIS za pomocą narzędzia „Utwórz siatkę”) obejmuje ustalenie kształtu pola (kwadrat, sześciokąt, romb lub trójkąt), wielkości pola w miarach terenowych i lokalizacji punktu początkowego sieci. Te trzy parametry dają sporą swobodę dopasowania geometrii i gęstości pól do charakteru danych źródłowych oraz odpowiedniej generalizacji obrazu rozkładu przestrzennego zjawiska²⁴. Najwygodniejszą metodą prezentacji takich układów regularnych jest kartogram geometryczny z przyjętą klasyfikacją wartości. Przykładem może

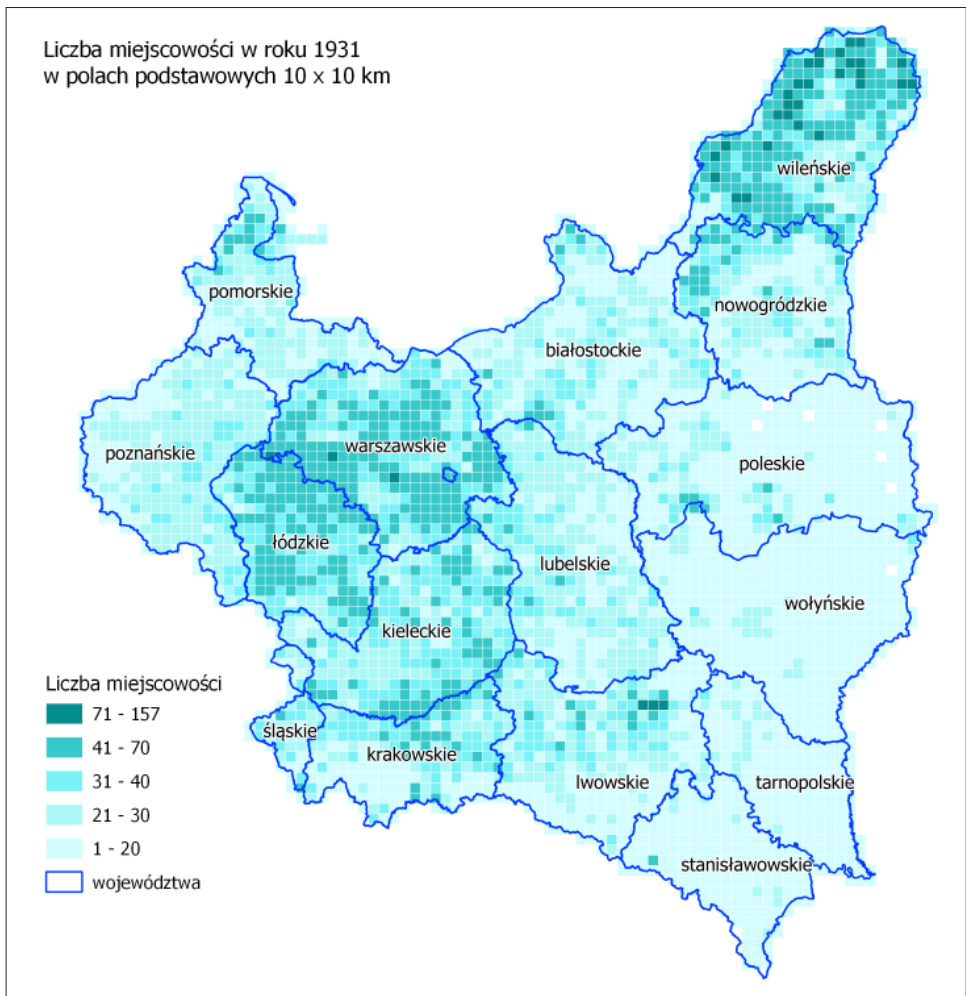
być prezentacja liczebności miejscowości (ryc. 4).

Poza wspomnianymi wcześniej zaletami sztucznych pól odniesienia metodę kartogramu geometrycznego charakteryzuje bardziej precyzyjne zobrazowanie zmian przestrzennych wartości niż w kartogramie właściwym, m.in. dzięki uwzględnieniu lokalnych odchyłek od średniej otoczenia. Jeszcze inny efekt daje metoda izolunii bazująca na interpolacji uśredniającej wartości w pewnym otoczeniu – w konsekwencji wygładzająca obraz powierzchni statystycznej.

4.3. Kontekst interpolacji przestrzennych

Proces interpolacji w ogólności polega na obliczaniu wartości zmiennej dla docelowego modelu danych przestrzennych (zwykle regularnej siatki GRID) w oparciu o wartości pomiarowe w nieregularnych zbiorach punktowych. W trakcie interpolacji powstaje najpierw model sieci nieregularnej TIN (Triangulated Irregular Network),

²⁴ L. Ratajski, *Metodyka kartografii społeczno-gospodarczej*, Warszawa 1989.



Ryc. 4. Kartogram geometryczny zagęszczenia sieci osadniczej w roku 1931 dla pól podstawowych kwadratowych o boku dł. 10 km.

Źródło: oprac. własne

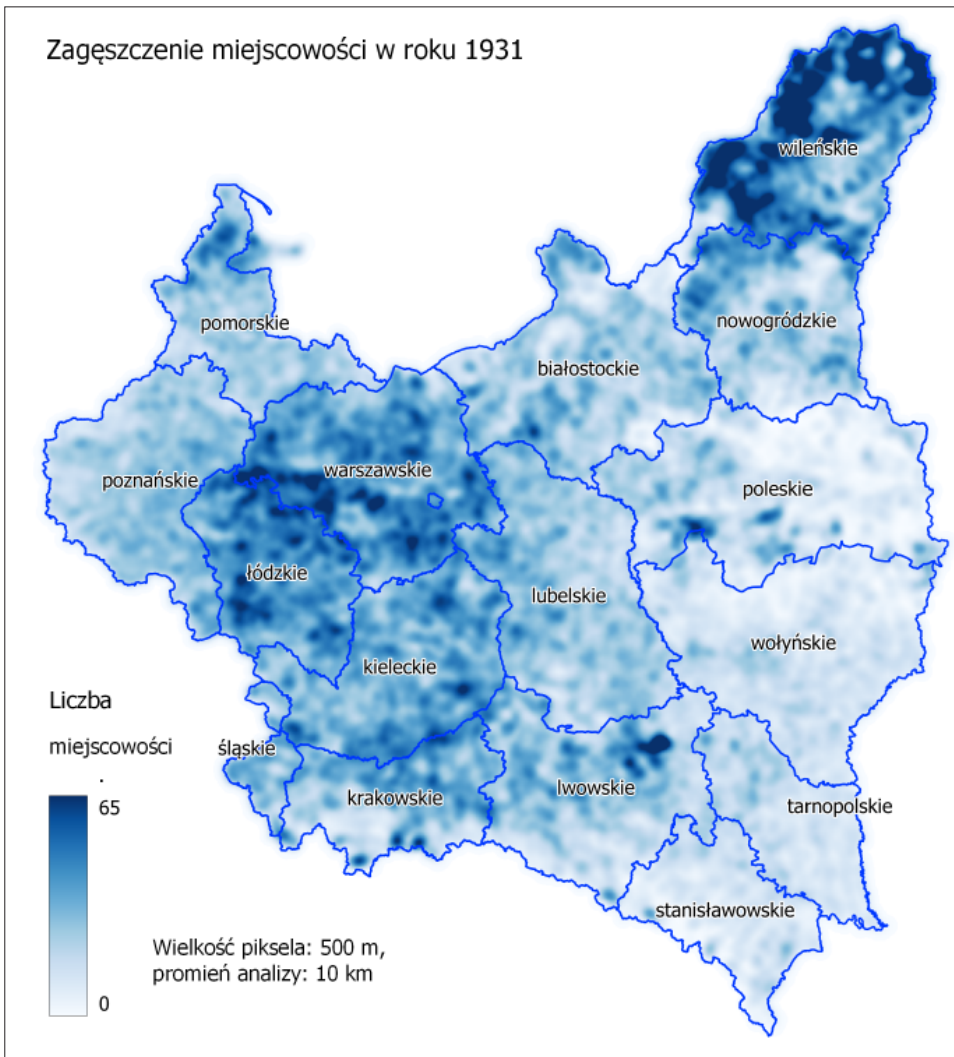
a docelowe wartości sieci regularnej wagiowane są według kryterium odległości pomiędzy punktami lub liczby punktów w założonym sąsiedztwie²⁵. Stąd bierze się silniejsze wygładzenie powierzchni statystycznej niż w modelu pól odniesienia.

Większość aplikacji GIS posiada zaimplementowane funkcje interpolacji

przestrzennych, a zakres dostępnych algorytmów interpolacji można często poszerzyć za pomocą wtyczek. Zazwyczaj dostępne są zarówno algorytmy deterministyczne: metoda odwrotnych odległości (IDW), metoda naturalnego sąsiedztwa (Natural Neighbour Interpolation) oraz metoda interpolacji funkcjami sklejanymi (Spline), jak i metody stochastyczne, np. Kriging.

Wiele algorytmów interpolacyjnych jest powiązanych z szablonami metod

²⁵ A. Fiedukowicz, J. Gąsiorowski, R. Olszewski, *Wybrane metody eksploracyjnej analizy danych przestrzennych (Spatial Data Mining)*, Warszawa 2015.



Ryc. 5. Kartogram geometryczny ciągły obrazujący zagęszczenie miejscowości w 1931 r., zredagowany za pomocą narzędzia Heatmap w aplikacji QGIS.

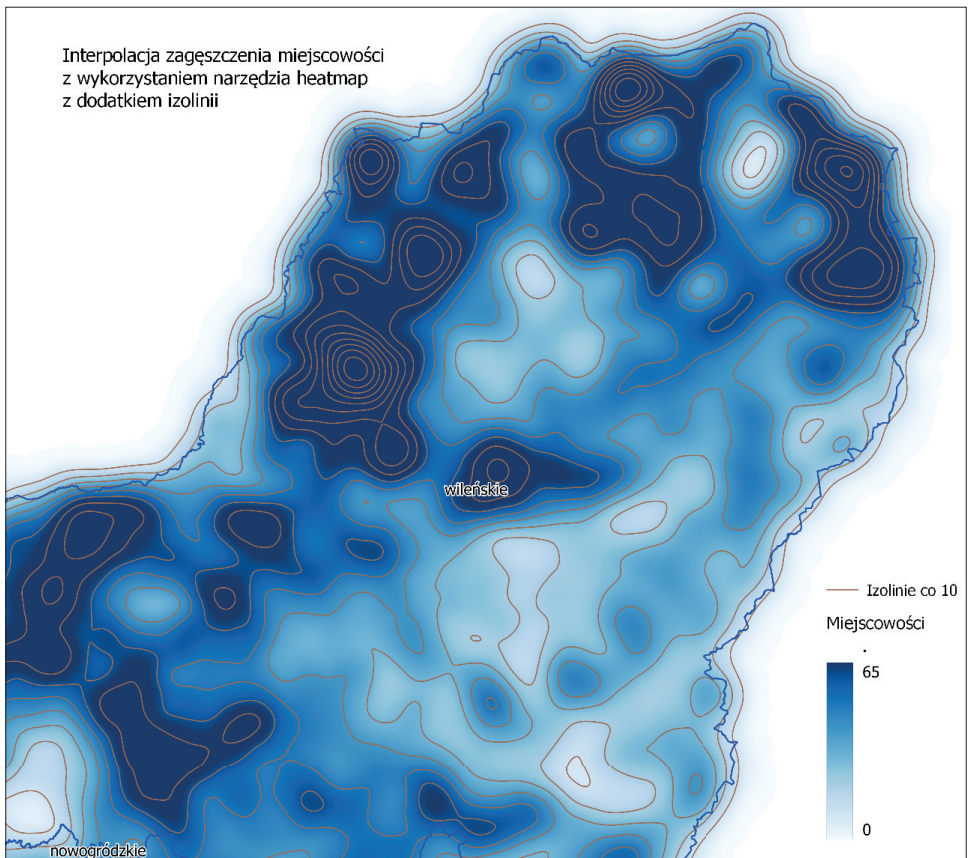
Źródło: oprac. własne

prezentacji kartograficznej w systemie GIS. Jedną z popularniejszych metod interpolacyjnych jest tzw. mapa ciepła (Heatmap), bazująca na kilku dostępnych estymatorach nieparametrycznych (Kernel Density Estimation), której wyniki mogą być zwizualizowane zarówno poprzez przypisanie ciągłej skali barw dla wygenerowanych komórek modelu regularnego, czyli z wykorzystaniem metody

kartogramicznej (ryc. 5), w formie cieniowania, jak też izol linii.

Dobór metody prezentacji kartograficznej modelu powinien być uzależniony od charakteru zjawiska²⁶. I tak dla powierzchni o dużych lokalnych skokach wartości, jak w przypadku omawianej struktury sieci

²⁶ A.H. Robinson, R. Sale, J.L. Morrison, *Podstawy kartografii*, Warszawa 1988.



Ryc. 6. Fragment mapy zagęszczenia miejscowości dla obszaru województwa wileńskiego, uzupełniony o izolinie ze skokiem 10.

Źródło: oprac. własne

osadniczej, metoda kartogramu sprawdza się dużo lepiej na mapie w małej skali (dla obszaru całej Polski), niż najczęściej stosowane izolinie, których podstawową zaletą jest precyzja zobrazowania powierzchni statystycznej (ryc. 6). Parametryzacja metody izolinii obejmuje przede wszystkim przyjęty interwał (skok wartości), a także stopień wygładzenia linii.

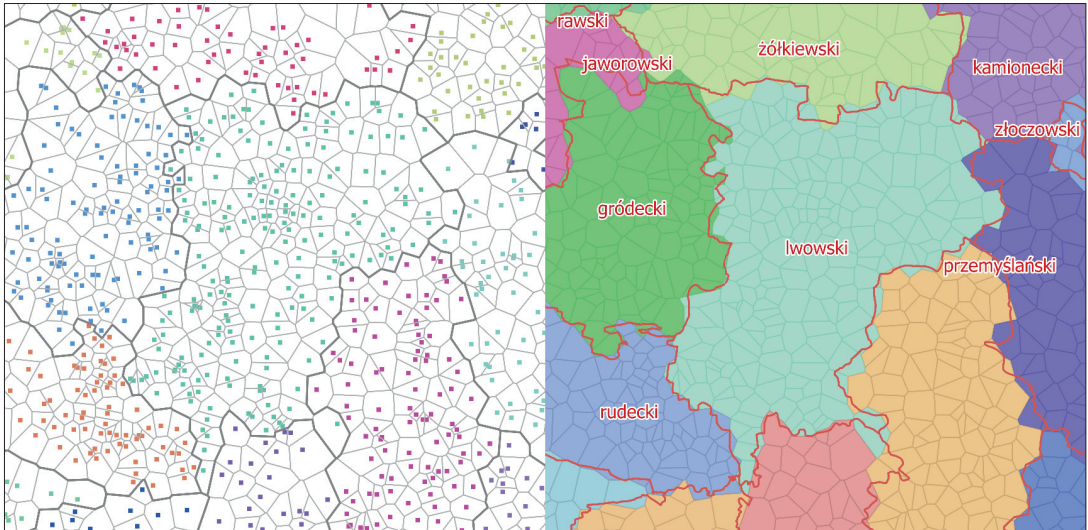
5. Odtwarzanie struktur przestrzennych

W badaniach historycznych materiały źródłowe, które zawierają informacje o lokalizacji jedynie w formie opisowej, mogą być zwizualizowane na mapie poprzez referencję do istniejących danych przestrzennych

– oczywiście jedynie na pewnym poziomie uogólnienia. Takie założenie przyświecało budowie bazy danych IAPN, w której miejscowości SGKP mogą być zwizualizowane w oparciu o lokalizację miejscowości z map WIG. Ta prosta referencja do pojedynczych punktów otwiera możliwość ekstrapolacji przestrzennej faktów, które nie posiadają źródeł mapowych.

5.1. Ekstrapolacja przestrzenna

Procedurę odtwarzania struktur przestrzennych można podzielić na dwa etapy: klasyfikacji obiektów źródłowych i ekstrapolacji przestrzennej. Klasyfikacja zbioru miejscowości może bazować na



Ryc. 7. Poligony Woronoja dla punktów miejscowości (po lewej) oraz porównanie wygenerowanych zasięgów z granicami powiatów pozyskanymi z map z przełomu XIX i XX w. (po prawej).

Źródło: oprac. własne

identyfikatorze jednostki administracyjnej, nazwie miejscowości nadrzędnej lub innych zgromadzonych informacjach, które umożliwiają regionalizację. Najprostszą formą utrwalenia kwerend atrybutowych byłby zapis zagregowanych grup punktów (typ multigeometrii) np. dla zobrazowania zależności miasto – przysiółki, wieś – folwarki itp. Taką funkcję pełnią zarejestrowane w ramach projektu IAPN „lokalizacje grupowe”, dzięki którym możliwe jest generowanie zasięgów dziewiętnastowiecznych nazw miejscowości odpowiadających kilku czy kilkunastu obiektom osadniczym zlokalizowanym na mapach topograficznych z początku XX w. Dzięki odtworzeniu takich relacji powstaje możliwość wizualizacji ekspansji terytorialnej miast poprzez wchłanianie przysiółków i sąsiednich miejscowości, które z czasem stawały się dzielnicami aglomeracji.

W przypadku przynależności administracyjnej (do powiatów, gmin, parafii) pochodna geometria powinna mieć charakter powierzchniowy. Narzędzie „Poligony Woronoja” umożliwia generowanie zasięgów obszarów zdefiniowanych za pomocą grup

punktów. W pierwszej fazie procedury wieloboki otaczające każdy punkt (zwane poligonami lub diagramami Woronoja) tworzone są indywidualnie dla każdego punktu miejscowości, a następnie agregowane terytorialnie według wybranego atrybutu. Na rycinie 7 zaprezentowano poligony Woronoja wygenerowane w oparciu o punkty miejscowości z bazy IAPN oraz porównanie zagregowanych jednostek administracyjnych na rok referencyjny 1900 z granicami powiatów zwektoryzowanymi z map, jak zostało to opisane w rozdziale 3.1.

Wichrowatość sztucznie wygenerowanych granic względem ich przebiegu utrwalonego w źródłach kartograficznych jest tym mniejsza, im więcej zlokalizowanych punktów miejscowości bierze udział w generowaniu poligonów. W zaprezentowanej symulacji granice powiatów zostały odtworzone w ogólnym przebiegu, zatem tak określona procedura regionalizacji mogłaby zostać zastosowana do wizualizacji innych delimitacji nieposiadających źródeł mapowych.

Mniej liczne zbiory punktów dają możliwość odtwarzania nieskartowanych

dotychczas struktur przestrzennych w znacznie bardziej przybliżony sposób. Przykład analizy opartej o zbiór miast – siedzib powiatów został zrealizowany w ramach pracy dyplomowej Karola Śpili²⁷. Celem eksperymentu było stwierdzenie, czy możliwa jest ekstrapolacja obszarów wyłącznie w oparciu o lokalizację siedzib jednostek. Zbiór 283 miast powiatowych z 1931 r. porównano z centroidami powiatów z tego okresu, wygenerowanymi w oparciu o zwektoryzowane na mapach granice. Narzędzie „Macierz odległości” umożliwiło obliczenie średnich odległości między odpowiadającymi sobie punktami obu zbiorów. Średnia odległość miasta od centroidu powiatu wyniosła 7,9 km przy wartości odchylenia standardowego: 5,2 km, co można uznać za niewielką odchyłkę w przypadku analiz ogólnokrajowych (w skalach poniżej 1:1 mln). Wizualizacja zasięgów w oparciu o poligony Woronoja nie odzwierciedla geometrii granic, ale może być wykorzystana do prezentacji danych statystycznych dzięki zachowaniu topologii podziału administracyjnego.

5.2. Analizy dynamiki zmian

Alternatywne reprezentacje geometryczne, takie jak wspomniane wyżej centroidy obszarów, mogą stanowić ciekawe uzupełnienie analiz w kontekście dynamiki zmian terytorialnych. We wszystkich systemach geoinformacyjnych większość narzędzi wymaga podania parametrów, a także wyboru konkretnego algorytmu realizacji zadania²⁸. Standardowe położenie centroidów jest obliczane jako wynik średniej arytmetycznej ze skrajnych współrzędnych X i Y obiektu (narzędzie „Centroidy”). W aplikacji QGIS można ponadto użyć zmodyfikowanego algorytmu wyznaczającego

punkt wyłącznie w środku obszaru (algorytm „Punkt na powierzchni”), wreszcie inny algorytm („Biegun niedostępności”) wyznacza punkt najbardziej oddalony od granic danej jednostki. Wybór zależy od potrzeb i charakterystyki terytoriów i ma znaczenie szczególnie dla obszarów nieregularnych o wydłużonych kształtach (ryc. 8).

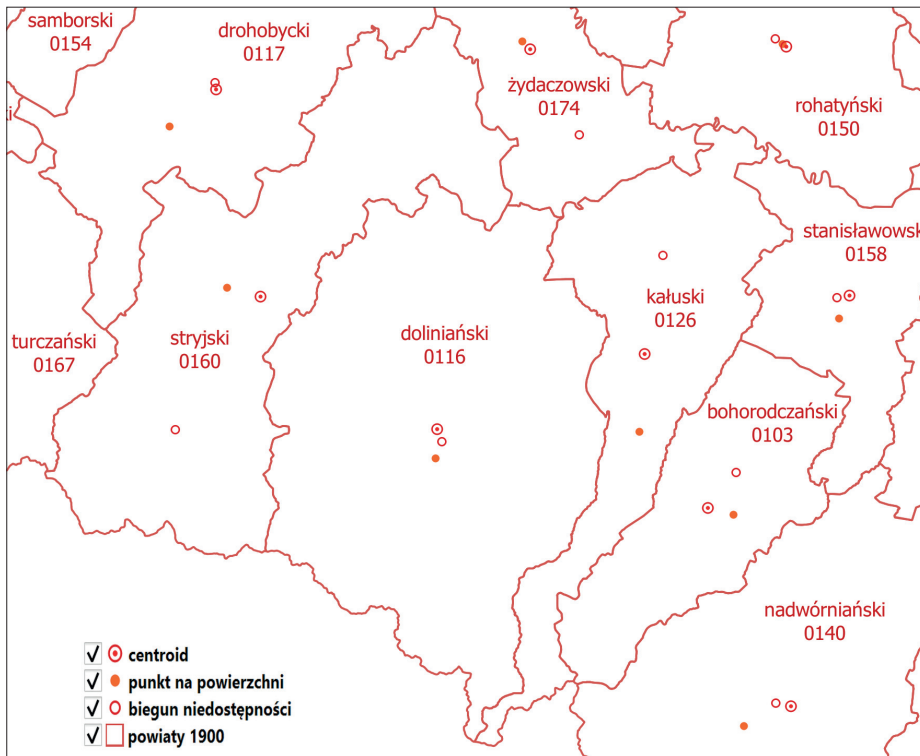
Wprowadzenie sztucznej reprezentacji obszarów w postaci centroidów zdecydowanie ułatwia analizowanie zmian zasięgów jednostek administracyjnych: pokrywające się centroidy z różnych przekrojów czasowych świadczą o stabilności granic danego terytorium, natomiast wielkość i kierunek przesunięcia obrazują ogólną tendencję zmian obszaru. Punktowa reprezentacja – w przeciwieństwie do powierzchniowej – daje możliwość zobrazowania pierwotnego położenia obiektów w więcej niż dwóch przekrojach czasowych (ryc. 9). Pewnym rozwiązaniem problemu pojemności graficznej map w klasycznej metodyce kartograficznej są pochodne reprezentacje powierzchniowe, np. mapy bilansowe czy mapy typów²⁹, jakkolwiek dostępność interaktywnych map (zarówno w systemach GIS, jak i WebGIS) pozwala na swobodniejsze stosowanie reprezentacji pierwotnych – użytkownik mapy ma możliwość samodzielnego wybrania przekrojów czasowych, które chce porównywać (pasek czasu na ryc. 2 oraz opcja podwójnego okna mapowego na ryc. 3).

Zmiana reprezentacji geometrycznej obiektów może ułatwić ocenę wizualną zmian, ale też może być utrwalona zarówno w formie tabeli (wspomniane wyżej narzędzie „Macierz odległości”), jak i w postaci zbioru liniowego połączeń między punktami (narzędzie „Połącz liniami”), których fragment pokazano na

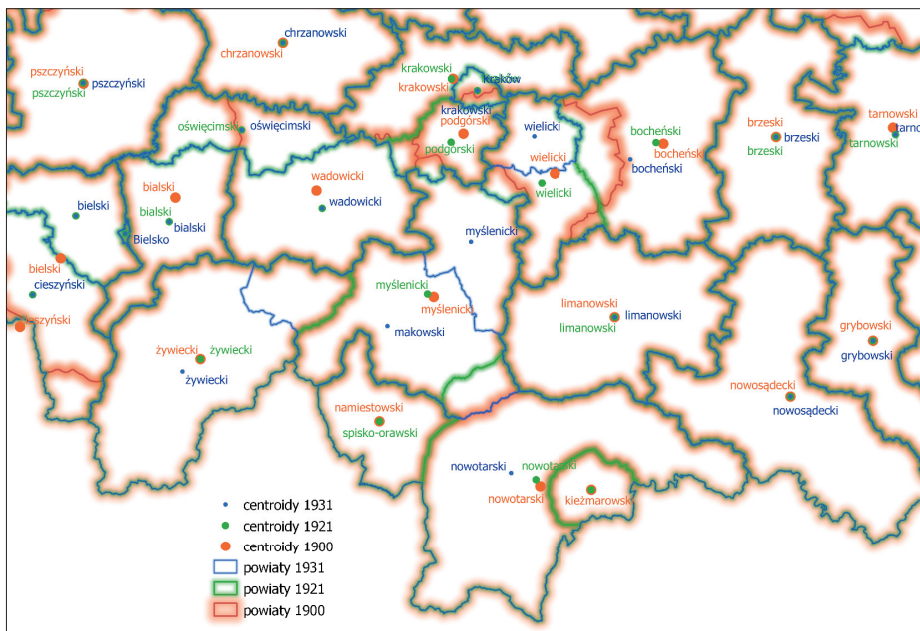
²⁷ K. Śpila, „Koncepcja historycznych analiz przestrzennych na podstawie danych topograficznych i statystycznych”, praca dyplomowa magisterska, Archiwum Prac Dyplomowych Politechniki Warszawskiej 2021.

²⁸ M.W. Meksuła, *Kartograficzna prezentacja dynamiki zjawisk*, w: *Świat mapy – świat na mapie. Główne problemy współczesnej kartografii 2002*, red. W. Pawlak, Wrocław 2002, s. 122–136.

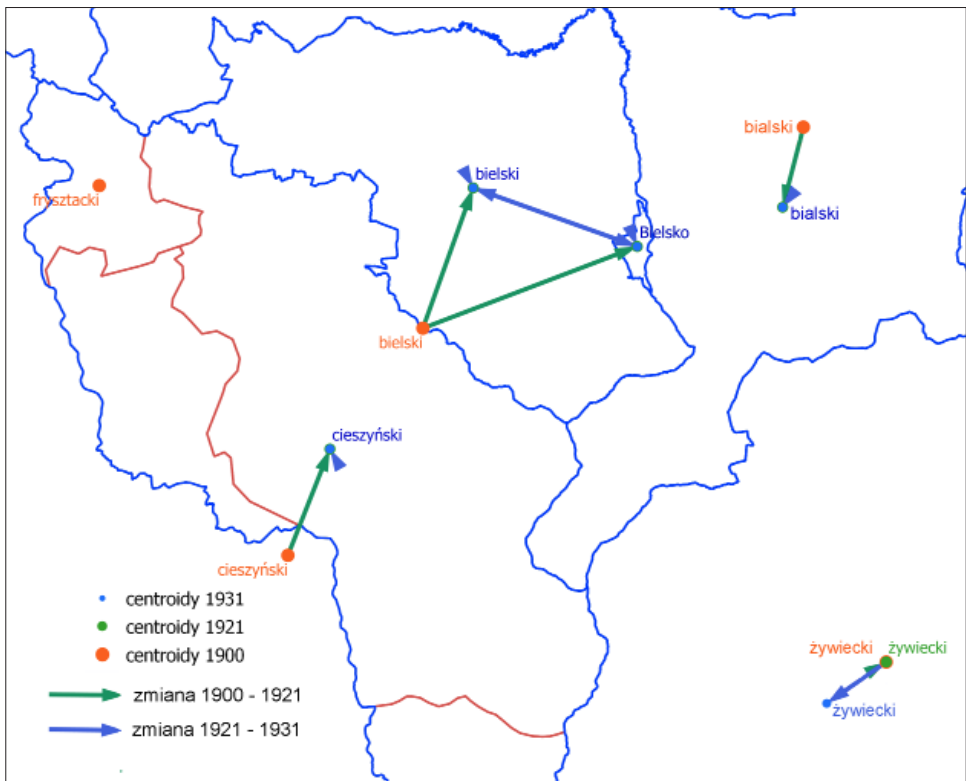
²⁹ D. Dukaczewski, „Kartograficzna prezentacja zmian użytkowania ziemi za pomocą animowanych map elektronicznych na przykładzie Gór Iżerskich”, rozprawa doktorska, Polska Akademia Nauk, Instytut Geografii i Przestrzennego Zagospodarowania im. S. Leszczyckiego, Warszawa 2000.



Ryc. 8. Porównanie lokalizacji centroidów generowanych za pomocą różnych algorytmów programu QGIS.
Źródło: oprac. własne



Ryc. 9. Zmiany terytorialne jednostek administracyjnych zobrazowane za pomocą linii granic i centroidów.
Źródło: oprac. własne



Ryc. 10. Wektorowa reprezentacja dyslokacji centroidów powiatów.

Źródło: oprac. własne

rycinie 10. W tej operacji kluczowe znaczenie mają atrybuty wspólne dla obiektów ze zbioru początkowego (*spoke*) i końcowego (*hub*) – mogą to być identyfikatory (jednostek), jak i nazwy własne (miasta lub powiatu).

Przykład analizy czasowej zmian obiektów pokazuje, że funkcje analityczne w środowisku GIS dają swobodę dostosowania poziomu generalizacji danych do założonych celów: zasięgi obszarów mogą być reprezentowane przez punkty, a te z kolei mogą być łączone wektorami zmian położenia. Dla każdej reprezentacji geometrycznej dysponujemy metryką: polami powierzchni, współrzędnymi czy też odległościami, a także parametrami pochodnymi: gęstością, natężeniem, współczynnikiem korelacji itd.

6. Podsumowanie

Celem pracy było wskazanie zastosowań analiz przestrzennych, które umożliwiają przede wszystkim integrację różnych zbiorów danych geograficzno-historycznych, ale także badania zmian charakterystyki sieci osadniczej w określonych zakresach czasowych. Zaprezentowane przykłady analiz struktur przestrzennych i analiz porównawczych pozwoliły określić dynamikę zmian rozmieszczenia i gęstości osadnictwa, ale także odtworzyć neutralne na mapach struktury przestrzenne, np. jednostek administracyjnych.

Mimo że opisane zadania bazowały tylko na dwóch podstawowych zbiorach miejscowości oraz jednostek administracyjnych, to udało się zaprezentować potencjał i wieloaspektowość możliwych

do prowadzenia badań historycznych. W oparciu o doświadczenia z realizacji projektu IAPN można przytoczyć kilka kluczowych wniosków technologicznych:

- fundamentalne znaczenie dla efektywności badań historyczno-geograficznych ma stosowanie spójnych struktur bazodanowych, dostosowanych do zamierzonego celu i zapewniających nie tylko wydajne przetwarzanie, ale też bezpieczeństwo danych;
- format zastosowanego modelu bazy danych powinien gwarantować pełną kontrolę danych i transfer pomiędzy systemami: w aplikacji QGIS jest to natywna geobaza GPKG (geopackage) oraz SQLITE, jako format wymiany;
- atrybutyzacja obiektów powinna uwzględniać niezależnie: schemat źródłowy i krytyczny, co zapewnia zarówno możliwość odwołania do pierwotnego dokumentu, jak i wielopoziomą interpretację, np. źródłowa nazwa miejscowości zapisana na kartach SGKP i nazwa pozyskana z mapy topograficznej, a w efekcie interpretacji: nazwa uzgodniona dla danego przekroju czasowego;
- metadane zbiorów danych powinny zawierać oznaczenia przekrojów czasowych, zarówno referencyjnych (w odniesieniu do całego zbioru), jak i źródłowych (na poziomie obiektów);
- metadane obiektów w zakresie geoprzetwarzania powinny przede wszystkim obejmować globalne identyfikatory,

a także znaczniki wykonanych przetworzeń i stanu obiektu.

Wyniki projektu pokazują, że dzięki integracji rozproszonych danych w jednorodnym środowisku systemów informacji geograficznej możliwe jest rozszerzenie analizy i oceny historycznej faktów i zjawisk z perspektywy geograficznej. Głównym osiągnięciem badawczym było określenie lokalizacji miejscowości pochodzących ze *Słownika geograficznego Królestwa Polskiego*. W artykule przedstawiono dodatkowe możliwości odtwarzania hipotetycznego przebiegu granic administracyjnych, nawet jeśli nie zachowały się źródła kartograficzne wystarczająco precyzyjnie dokumentujące ich położenie. Wachlarz możliwości analitycznych jest bardzo szeroki i zależy głównie od zawartości źródeł historycznych, które posiadają referencję topograficzną wyłącznie w formie opisowej, ale którym można nadać geolokalizację z innych źródeł. Powiązanie informacji historycznej i geograficznej, nawet jeśli jest możliwe na pewnym poziomie dokładności i zaufania, otwiera nową perspektywę badawczą. Za pośrednictwem Internetowego Atlasu Polski Niepodległej możliwe staje się pozyskanie w formie zarówno obrazowej (rastrowej), jak i wektorowej danych geograficznych i historycznych od końca XIX w. do okresu dwudziestolecia międzywojennego, ale także wybranych danych statycznych w formie tabelarycznej i kartograficznej. ■

Bibliografia

- Cebertowicz R., *Mapa Administracyjna Rzeczypospolitej Polskiej 1:300 000*, [b.m.w.] 1937.
- Cebryk P., *Generalizacja map statystycznych*, Lublin 2017.
- Dukaczewski D., „Kartograficzna prezentacja zmian użytkowania ziemi za pomocą animowanych map elektronicznych na przykładzie Gór Izerskich”, rozprawa doktorska, Polska Akademia Nauk, Instytut Geografii i Przestrzennego Zagospodarowania im. S. Leszczyńskiego, Warszawa 2000.
- Epsztein T., *Słownik geograficzny Królestwa Polskiego i innych ziem słowiańskich – zarys historii dzieła*, w: *Od „Słownika geograficznego Królestwa Polskiego” do map topograficznych Wojskowego Instytutu Geograficznego*, red. T. Epsztein, Warszawa 2021, s. 85–115.
- Fiedukowicz A., Gąsiorowski J., Olszewski R., *Wybrane metody eksploracyjnej analizy danych przestrzennych (Spatial Data Mining)*, Warszawa 2015.
- GIS. Teoria i praktyka*, red. P. Longley i in., Warszawa 2006.

- Głazewski A., Jagura A., Kocot A., *Podział administracyjny – metodyka pozyskiwania historycznych danych topograficznych z obszaru II Rzeczypospolitej*, w: *Od „Słownika geograficznego Królestwa Polskiego” do map topograficznych Wojskowego Instytutu Geograficznego*, red. T. Epszstein, Warszawa 2021, s. 33–46.
- Graszka O., *Automatyzacja procesu rozpoznawania i weryfikacji nazw geograficznych ze źródeł historycznych na przykładzie Słownika geograficznego Królestwa Polskiego*, w: *Od „Słownika geograficznego Królestwa Polskiego” do map topograficznych Wojskowego Instytutu Geograficznego*, red. T. Epszstein, Warszawa 2021, s. 21–32.
- Gut M., „Opracowanie atlasu tematycznego zmian demograficznych Polski po odzyskaniu niepodległości”, praca dyplomowa inżynierska, Archiwum Prac Dyplomowych Politechniki Warszawskiej 2021.
- Internetowy Atlas Polski Niepodległej*, 2021 (<https://atlas1918.iopan.edu.pl>, dostęp: 10 października 2022).
- Jadura A., Kocot A., Kowalski P.J., *Metodyka odtworzenia sieci osadniczej z obszaru II Rzeczypospolitej*, w: *Od „Słownika geograficznego Królestwa Polskiego” do map topograficznych Wojskowego Instytutu Geograficznego*, red. T. Epszstein, Warszawa 2021, s. 47–63.
- Krassowski B., *Polska kartografia wojskowa w latach 1918–1945*, Warszawa 1974.
- Kumaniecki K.W., Krzyżanowski A., *Statystyka Polski*, Kraków 1915.
- Mappa Gubernij Królestwa Polskiego dla cyklistów*, M. Poturaj, Warszawa 1894.
- Meksuła M.W., *Kartograficzna prezentacja dynamiki zjawisk*, w: *Świat mapy – świat na mapie. Główne problemy współczesnej kartografii 2002*, red. W. Pawlak, Wrocław 2002, s. 122–136.
- Od „Słownika geograficznego Królestwa Polskiego” do map topograficznych Wojskowego Instytutu Geograficznego*, red. T. Epszstein, Warszawa 2021.
- Osińska-Skotak K. i in., *Analizy przestrzenne i wizualizacja danych historycznych IAPN w środowisku systemów informacji geograficznej*, w: *Od „Słownika geograficznego Królestwa Polskiego” do map topograficznych Wojskowego Instytutu Geograficznego*, red. T. Epszstein, Warszawa 2021, s. 65–84.
- Panecki T., „Koncepcja struktury bazy danych historycznych obiektów topograficznych”, rozprawa doktorska, Uniwersytet Warszawski, Warszawa 2018.
- Ratajski L., *Metodyka kartografii społeczno-gospodarczej*, Warszawa 1989.
- Robinson A.H., Sale R., Morrison J.L., *Podstawy kartografii*, Warszawa 1988.
- „Rocznik Statystyki Miast Polski”, 1, 1928, Warszawa, 1929.
- Skorowidz miejscowości Rzeczypospolitej Polskiej opracowany na podstawie wyników Pierwszego Powszechnego Spisu Ludności z dn. 30 września 1921 r. i innych źródeł urzędowych*, Warszawa 1925.
- Statystyka Rolnicza 1930/31*, Warszawa 1932.
- Statystyka Rolnicza 1936*, Warszawa 1937.
- Szady B., *Czasowo-przestrzenne bazy danych jako narzędzie w geografii historycznej*, „Acta Universitatis Lodzianensis, Folia Geographica Socio-Oeconomica”, 14, 2013, s. 17–32.
- Śpila K., „Koncepcja historycznych analiz przestrzennych na podstawie danych topograficznych i statystycznych”, praca dyplomowa magisterska, Archiwum Prac Dyplomowych Politechniki Warszawskiej 2021. ■

The Internet Atlas of Independent Poland – the methodology for developing a database of towns and administrative boundaries, and derivative analysis

Summary

The article presents examples of using a resource of historical spatial data collected during the preparation of the Internet Atlas of Independent Poland (IAPN). The atlas, in the form of an internet portal, was published in 2021 as a result of a project carried out by the Institute of History of the Polish Academy of Sciences and the Faculty of Geodesy and Cartography of the Warsaw University of Technology.

At the beginning of the text, the authors present the typology of spatial analyses carried out

in the environment of geographic information systems (GIS), and a set of spatial analyses performed when developing the Atlas database and the IAPN geo-information portal. The GIS analytical apparatus was also a tool for organising resources during the project and verifying the progress of the work. Geometric conversion of database objects and spatial aggregation were used. For a more complex task, i.e. for integrating spatial data with dictionary data, the set joins were employed. The latter was a multi-stage

task and required the integration of spatial data extracted from a tactical map of the Military Geographical Institute (WIG) with descriptive data contained in the Geographical Dictionary of the Kingdom of Poland (SGKP), in which location is given descriptively or by the name of an administrative unit.

A vital element of the article are chapters presenting the possibilities of investigating changes in the characteristics of the settlement network over the last century. Through analyses of spatial structures and comparative studies in various time sections, it is possible to determine the dynamics of changes in the distribution and density of settlements and reconstruct spatial structures not recorded on maps: clusters of settlement objects or administrative units.

Groups of spatial analyses carried out were described and presented in several contexts: administrative, regionalising, in which the reference field is the network of basic fields; interpolation, extrapolation, when an attempt is made to reconstruct spatial structures; and dynamic, in which the analysis and visualisation of changes in a phenomenon over time come to the fore.

The authors also emphasised the role of the spatial database, which enables the implementation of several types of analysis, as well as the importance and possibility of using spatial metadata, which describes both the data sources and the resulting resources. ■

Dr inż. Andrzej Głazewski – kartograf, specjalista z zakresu projektowania baz danych i prowadzenia analiz przestrzennych oraz wizualizacji kartograficznej. Zainteresowania naukowe wiąże z kartografią topograficzną, projektowaniem i zastosowaniem baz danych referencyjnych, semiotyką kartograficzną oraz percepcją przestrzeni. Od 1994 r. pracuje w Politechnice Warszawskiej, w Zakładzie Kartografii Wydziału Geodezji i Kartografii (andrzej.glazewski@pw.edu.pl)

Andrzej Głazewski, PhD Eng. – cartographer, expert in database design and spatial analysis and cartographic visualisation. His scientific interests include topographic cartography, design and application of reference databases, cartographic semiotics and perception of space. Since 1994, employee of the Warsaw University of Technology, Department of Cartography, Faculty of Geodesy and Cartography (andrzej.glazewski@pw.edu.pl)

Dr inż. Paweł J. Kowalski – absolwent Wydziału Geodezji i Kartografii Politechniki Warszawskiej. Pracownik naukowy i dydaktyczny w Zakładzie Kartografii od roku 1994 do chwili obecnej. Od początku pracy zawodowej głównym obszarem jego zainteresowań są systemy informacji geograficznej, bazy danych topograficznych, techniki multimedialne w zastosowaniach kartograficznych, projektowanie serwisów geoinformacyjnych i internetowe publikacje kartograficzne (pawel.kowalski@pw.edu.pl)

Paweł J. Kowalski, PhD Eng. – graduate of the Faculty of Geodesy and Cartography, Warsaw University of Technology. Since 1994, academic at the Department of Cartography. His main areas of interest are: geographic information systems, topographic databases, multimedia techniques in cartographic applications, design of geo-information services and internet cartographic publications (pawel.kowalski@pw.edu.pl)