Harmonizacja historycznych informacji przestrzennych pochodzących z różnych źródeł na podstawie "*Geograficzno-statystycznego opisania parafiów Królestwa Polskiego*" oraz map specjalnych województw koronnych Karola Perthéesa

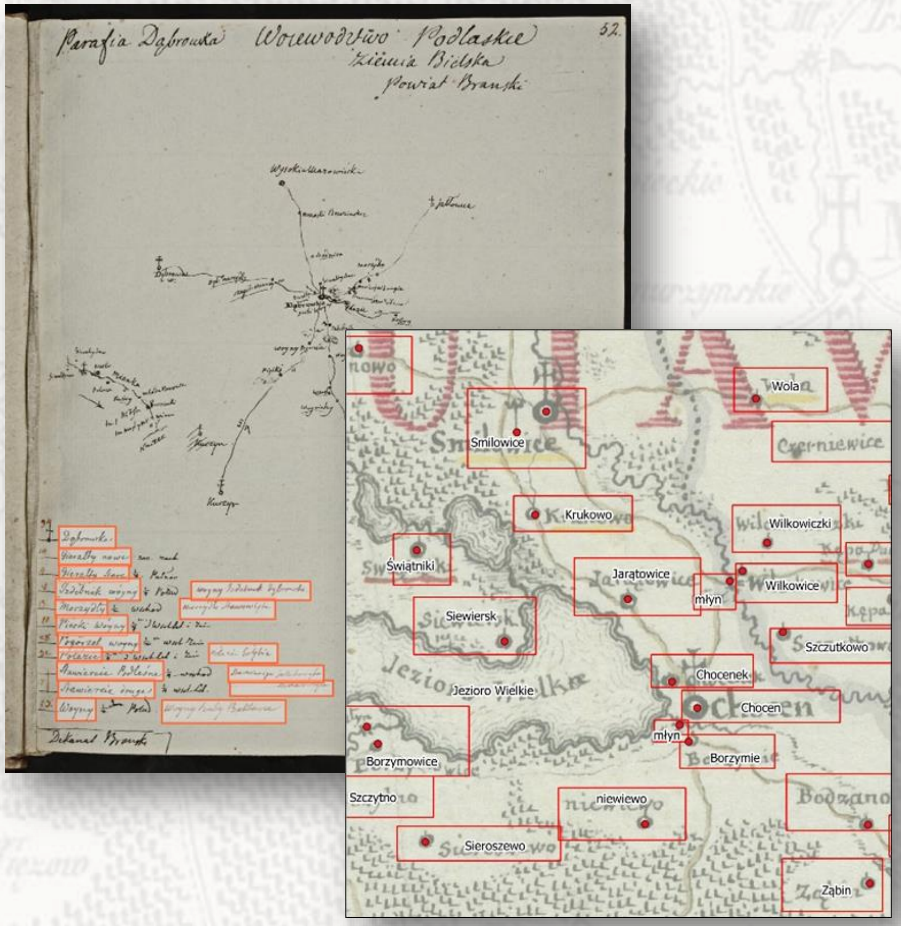
Grzegorz Myrda

gmyrda@ihpan.edu.pl

INSTYTUT HISTORII IM. TADEUSZA MANTEUFFLA POLSKIEJ AKADEMII NAUK

Projekt: Kartografia w służbie reform państwa epoki stanisławowskiej – krytyczne opracowanie »Geograficzno-statystycznego opisania parafiów Królestwa Polskiego« oraz map województw koronnych Karola Perthéesa

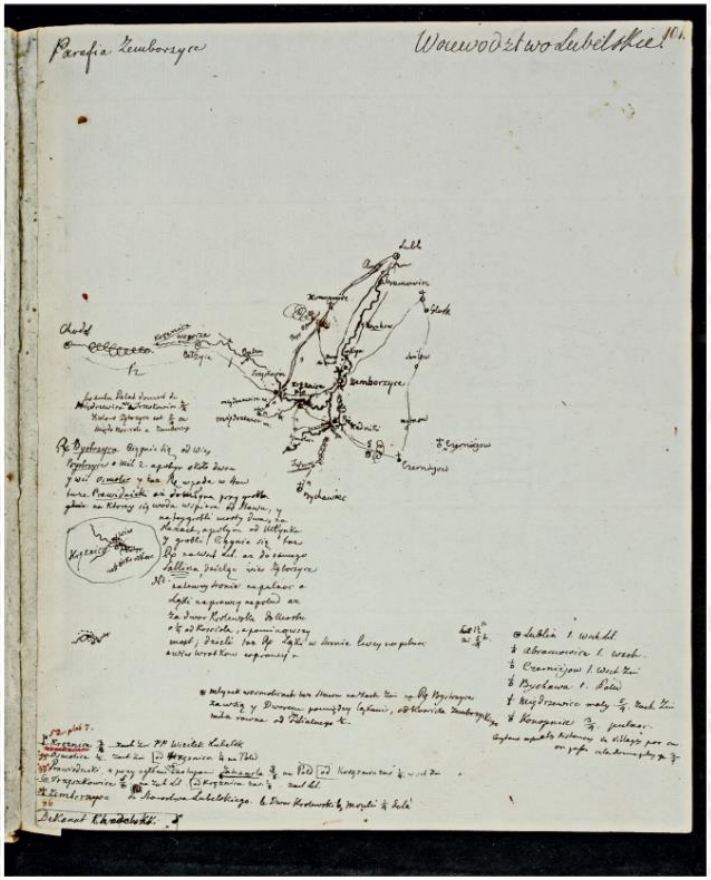
NPRH nr 11H 18 0122 87



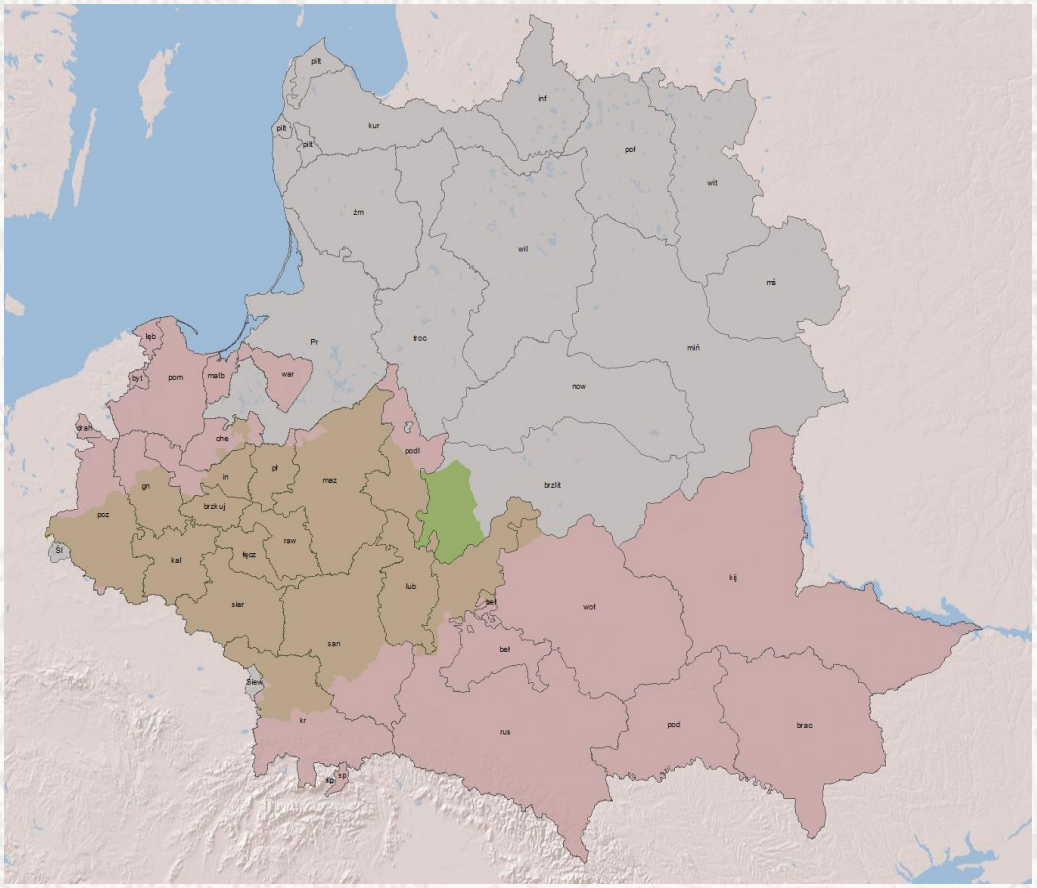
Głównym celem projektu jest krytyczne opracowanie oraz publikacja w internecie „Geo-graficzno-statystycznego opisania parafiów Królestwa Polskiego” oraz map specjalnych województw koronnych Karola Perthéesa. Karol Perthées (1739-1815) posiadał tytuł królewskiego geografa, króla Rzeczypospolitej, Stanisława Augusta Poniatowskiego. Autor szeregu map województw dawnej Rzeczypospolitej.

Cel i źródła

jest



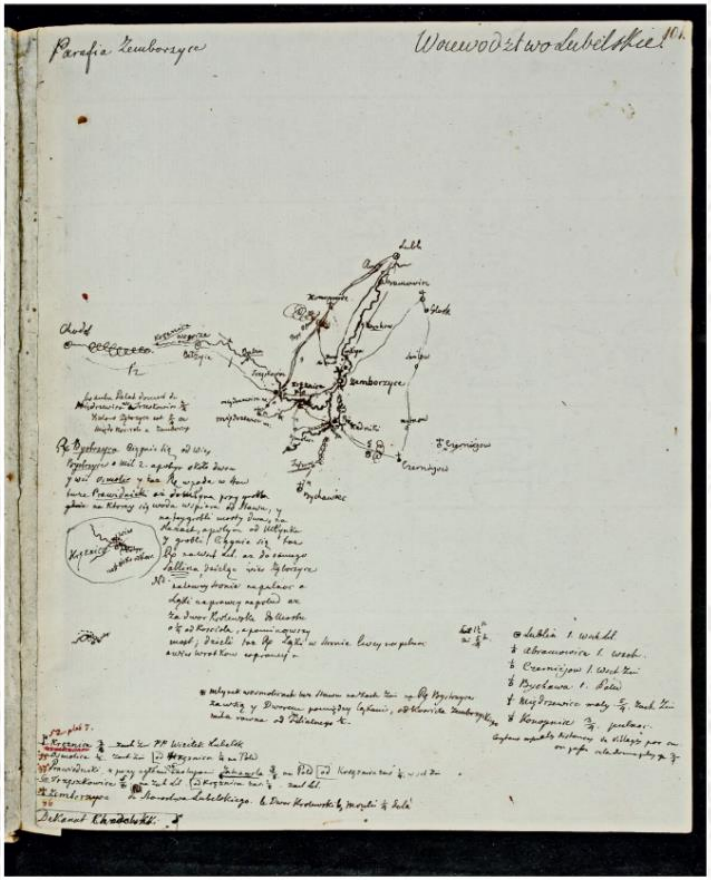
będzie



Archiwa: Litwa, Ukraina, Polska, Rosja, Białoruś

geoportal

Szkice parafialne



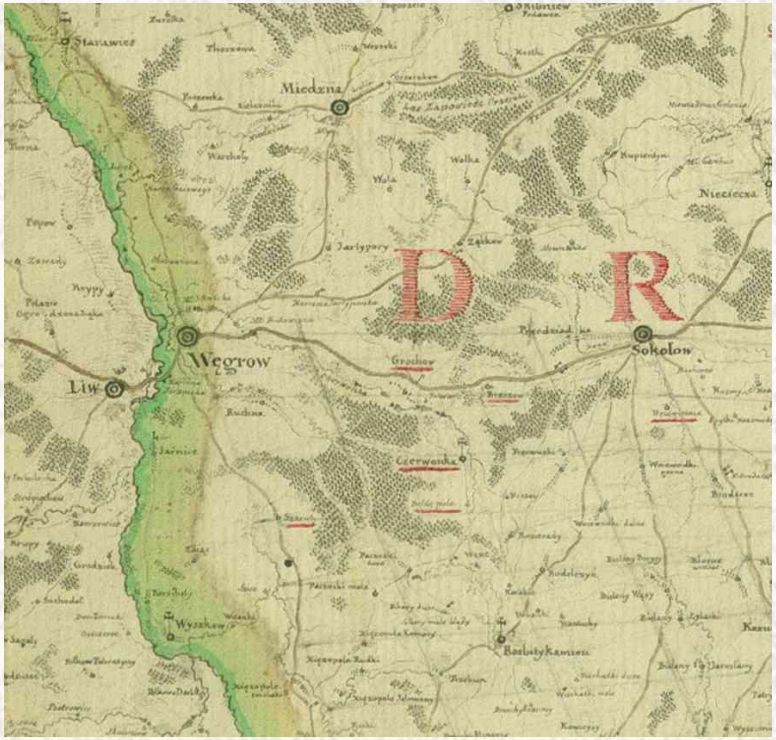
Zasięg terytorialny: Korona Królestwa Polskiego

Data powstania (orientacyjna): 1790-1796

Miejsce przechowywania: Narodowa Biblioteka Ukrainy im. W. Wernadskiego w Kijowie

12 tomów, 5500 stron, 2056 parafii

Mapy



Zasięg terytorialny: Korona Królestwa Polskiego

Data powstania (orientacyjna): 1783-1804

Miejsce przechowywania: Rozproszone egzemplarze (biblioteki i archiwa polskie, białoruskie, rosyjskie, ukraińskie)

12 map województw koronnych (poza gnieźnieńskim i sieradzkim)

Metodyka zakładająca podejście typu multimodel

- Model 1: model źródłowy, jak najbliższy pojęciowo źródłom, przy najmniejszej możliwej interpretacji treści. Inwentaryzacja toponimów z map oraz szkiców parafialnych. Przykład: fragment obrazu z napisem Zemborzyce.
- Model 2: model analityczny, obejmujący interpretacje treści. Identyfikacja miejscowości oraz jednostek administracji terytorialnej, harmonizacja z danymi z XVI wieku oraz współczesnymi. Przykład: Zemborzyce już nie tylko jako napis, ale także jako zidentyfikowana i zlokalizowana miejscowość.
- Model 3: model prezentacyjny. Połączenie danych z modelu krytyczno-analitycznego i źródłowego w celu ich prezentacji w optymalny sposób.

Metodyka zakładająca podejście typu multimodel

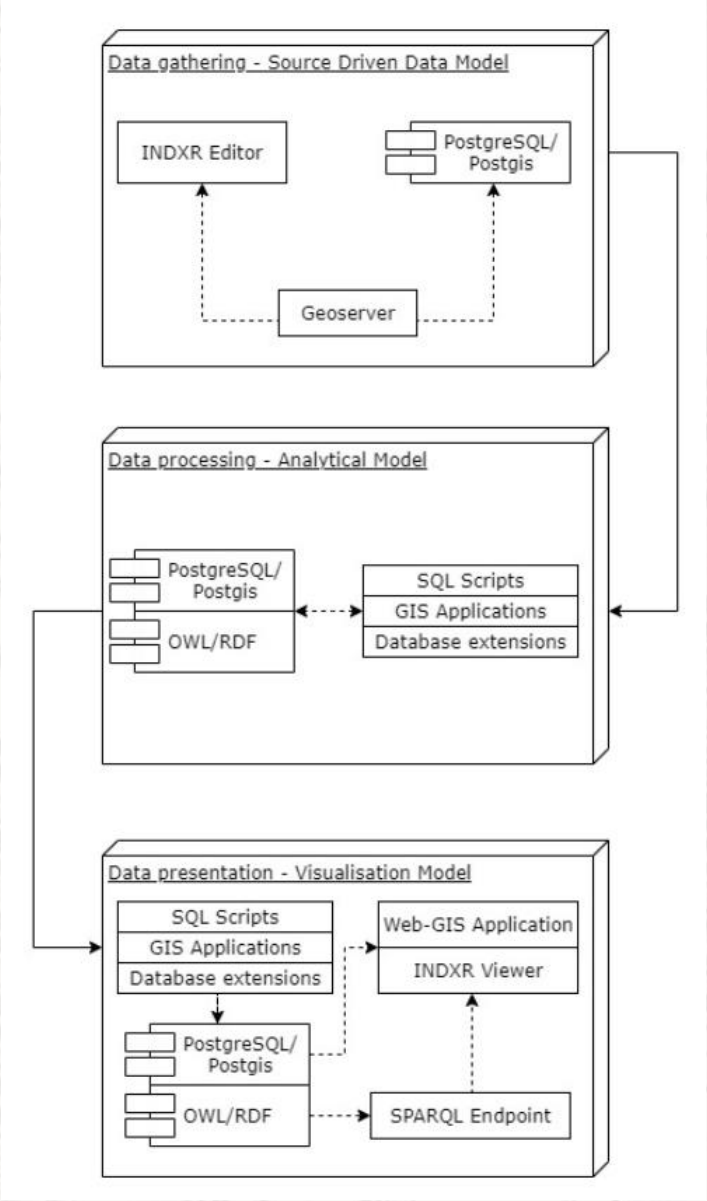
Dlaczego 3 modele?

Możliwość łatwej weryfikacji części krytycznej, nawet na etapie publikacji, z poziomu części prezentacyjnej

Możliwość prowadzenia prac z wykorzystaniem technologii zwinnych (agile) opartych o kolejne iteracje procesu identyfikacji miejscowości historycznych

Gromadzenie danych odrębnie przez dwa pod-zespoły i wprowadzenie poziomu semi-krytycznego zwiększa wiarygodność wnioskowania opartego na algorytmizacji

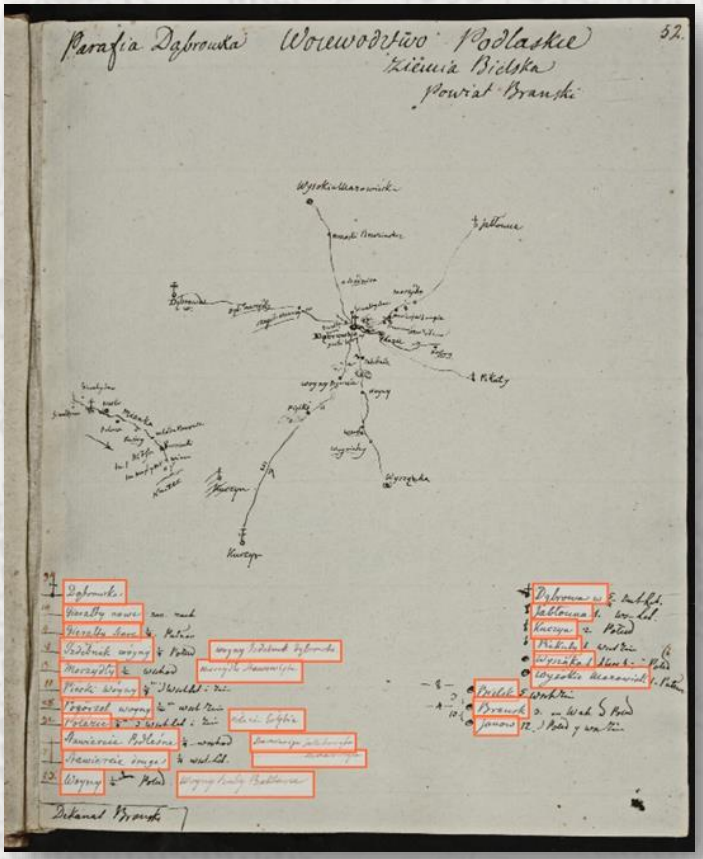
Workflow

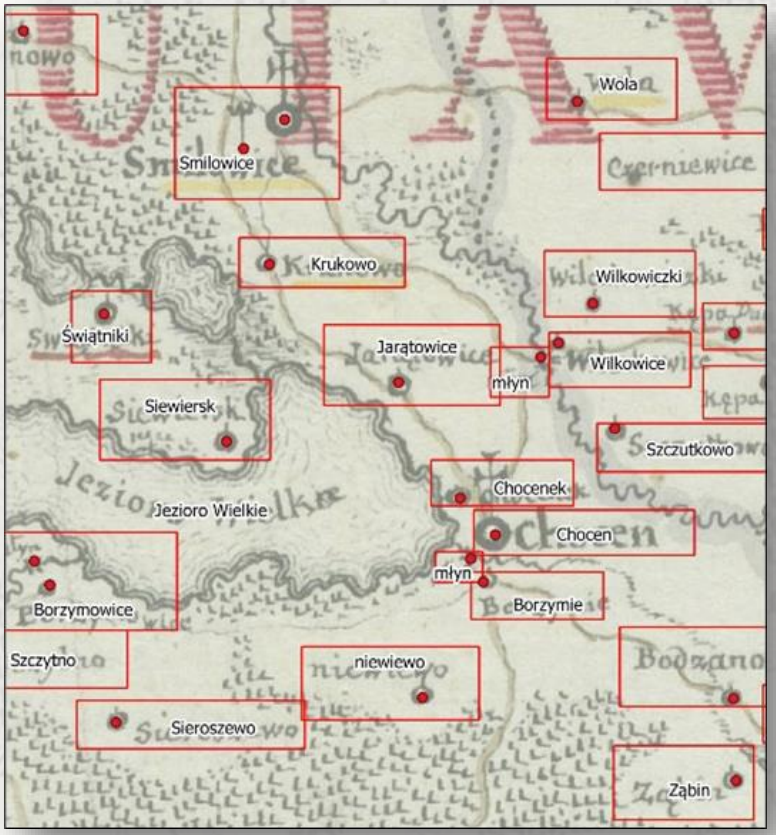

Data gathering - Source Driven Data Model
INDXR Editor
PostgreSQL/
Postgis
Geoserver
Data processing - Analytical Model
PostgreSQL/
Postgis
OWL/RDF
SQL Scripts
GIS Applications
Database extensions
Data presentation - Visualisation Model
SQL Scripts
GIS Applications
Database extensions
Web-GIS Application
INDXR Viewer
PostgreSQL/
Postgis
OWL/RDF
SPARQL Endpoint

Model źródłowy

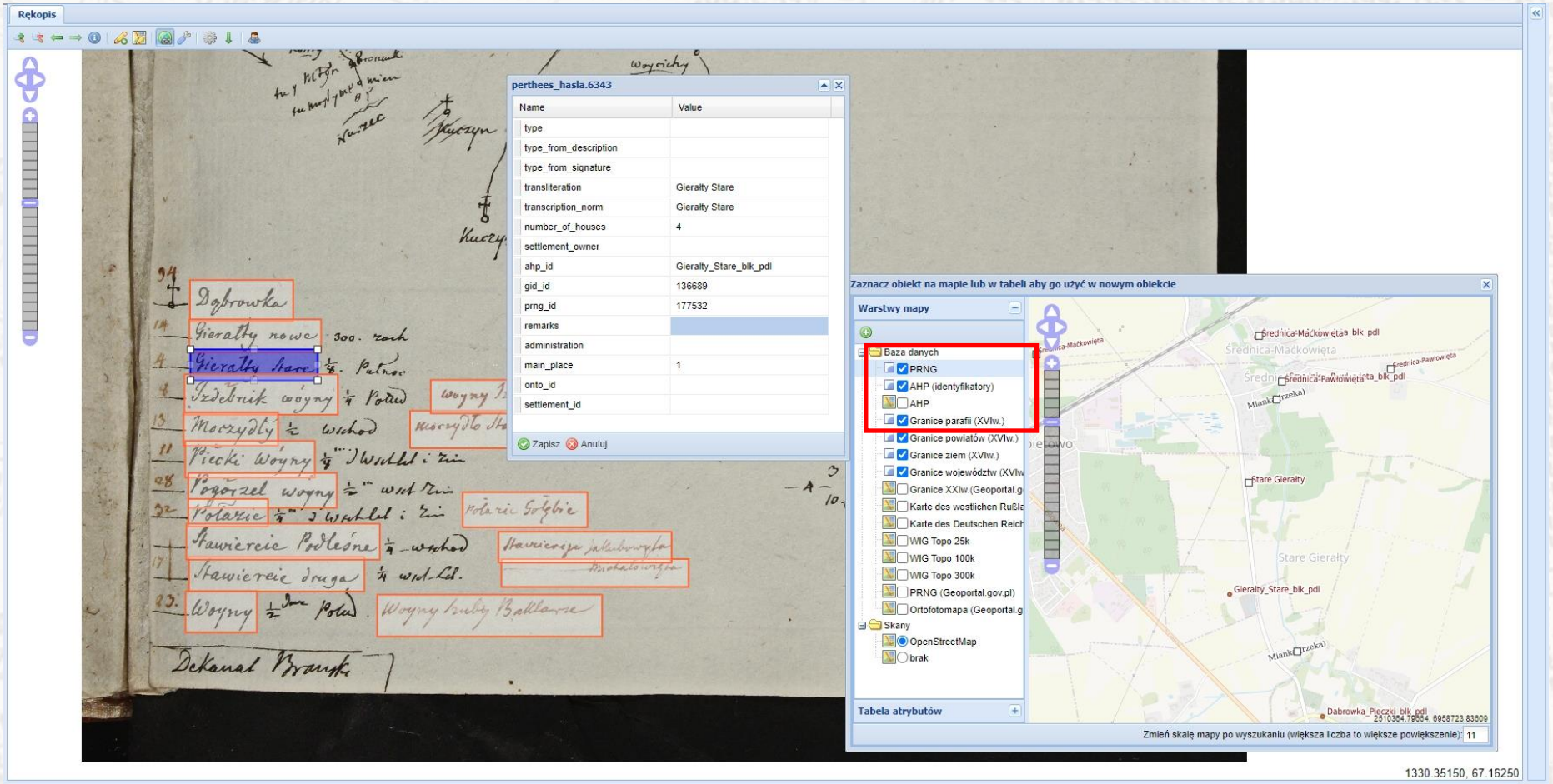
Dwa oddzielne podmodele źrodłowe – jeden dla szkiców i jeden dla map**.**

Łączy je metoda anotacyjna zapożyczona z GIS, która została wykorzystana w aplikacji INDXR.



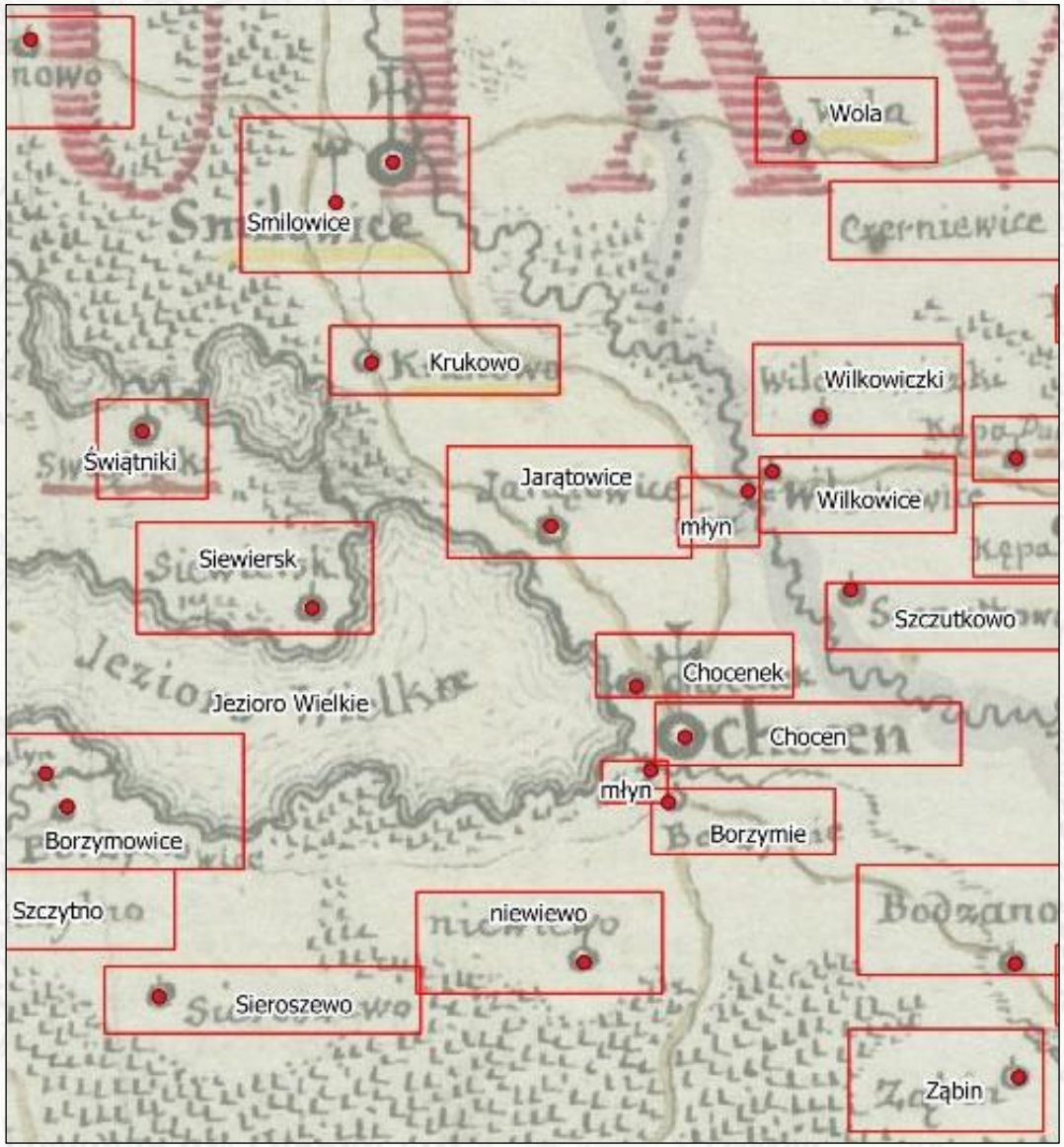


INDXR



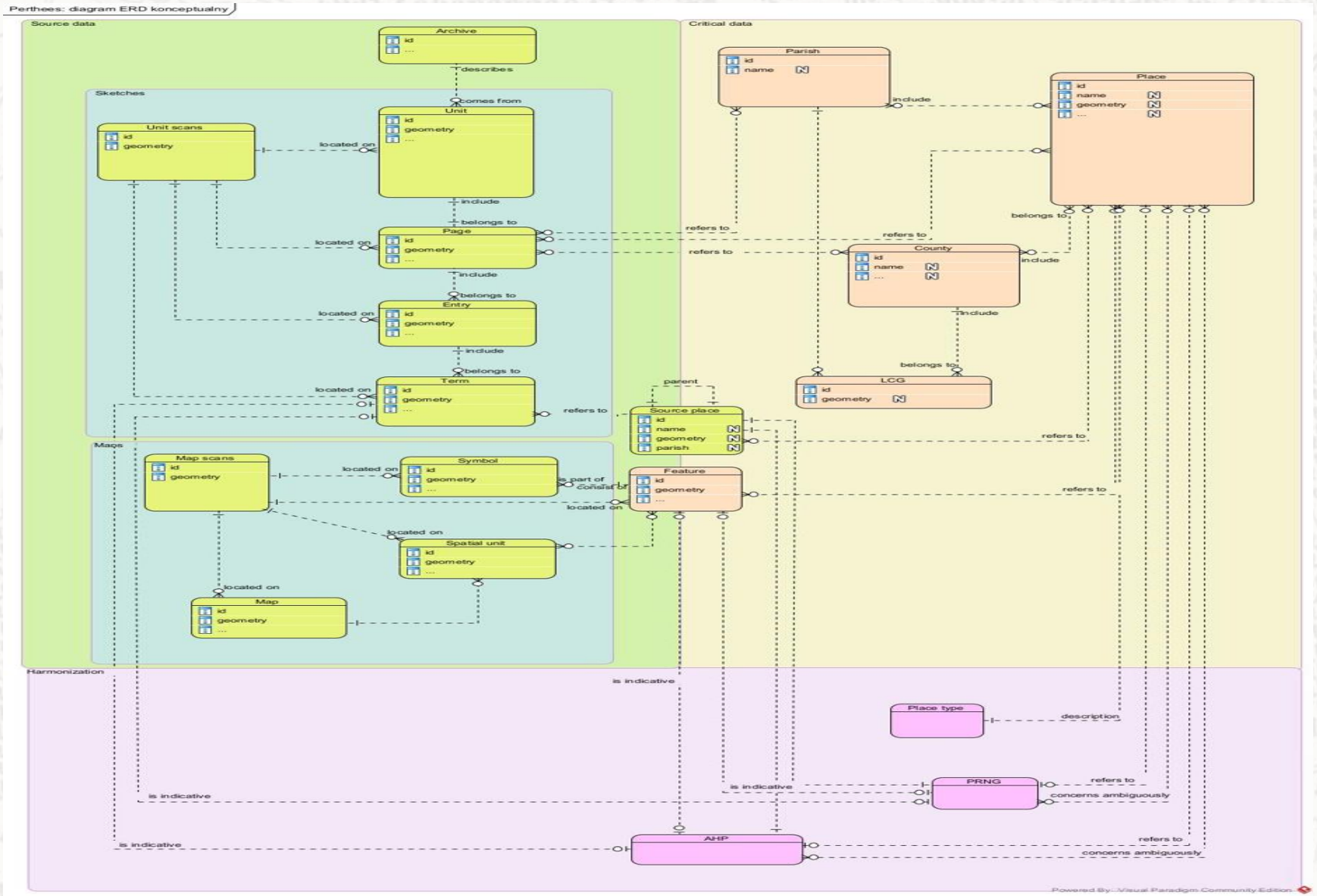
Model źródłowy dla ankiet. Indeksowanie (anotowanie) treści: Indeks miejscowości, szkic, itd.

Model źródłowy dla map



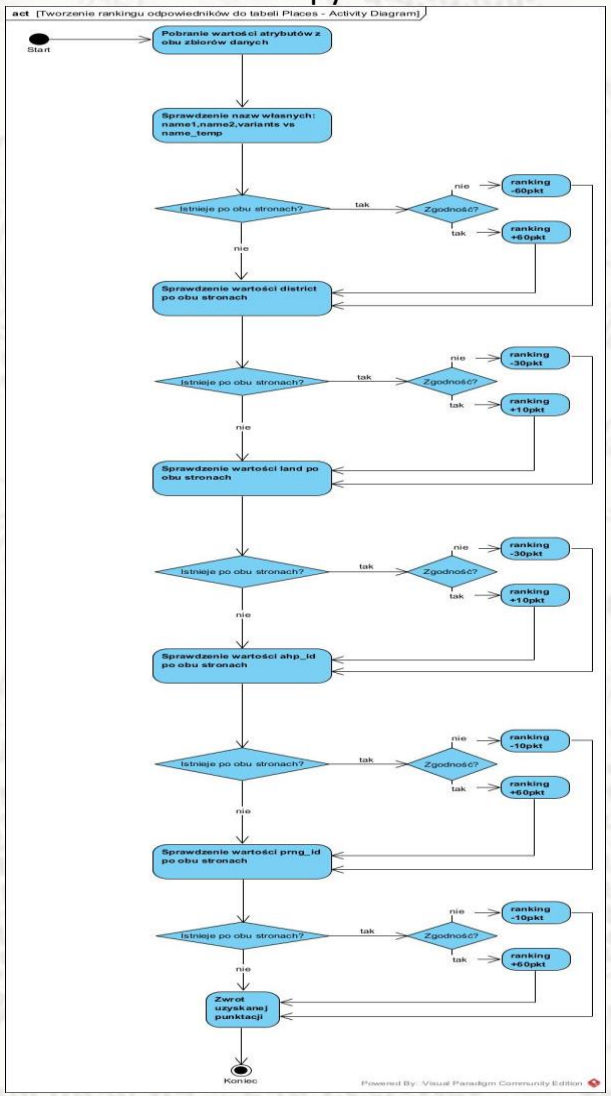
Struktura bazy danych odzwierciedla strukturę materiałów mapowych: Mapy, Jednostki, Obiekty, Symbole, Etykiety tekstowe.

Model konceptualny bazy danych

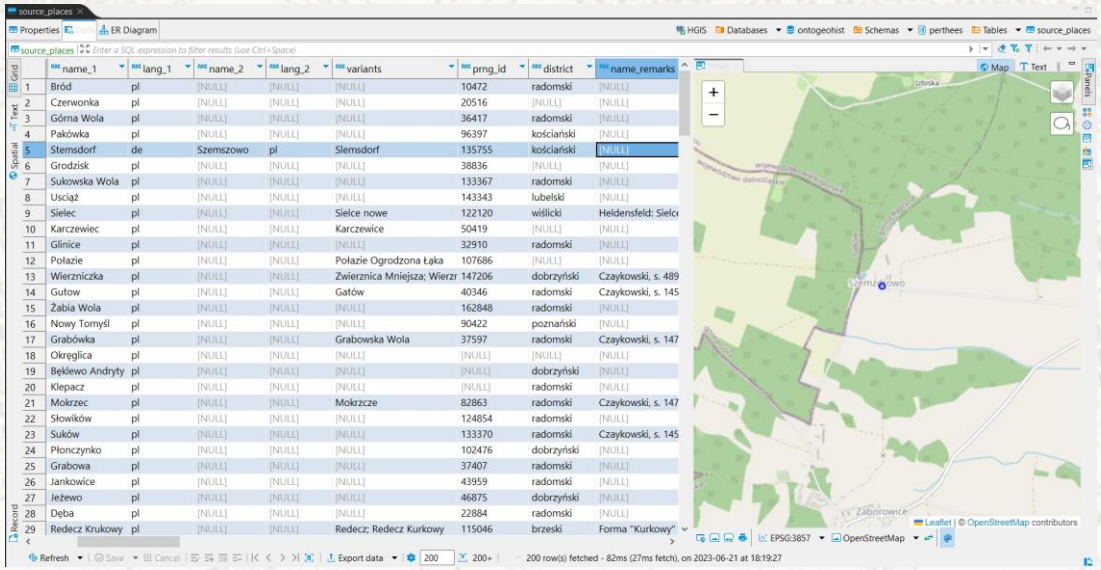


Harmonizacja wewnętrzna

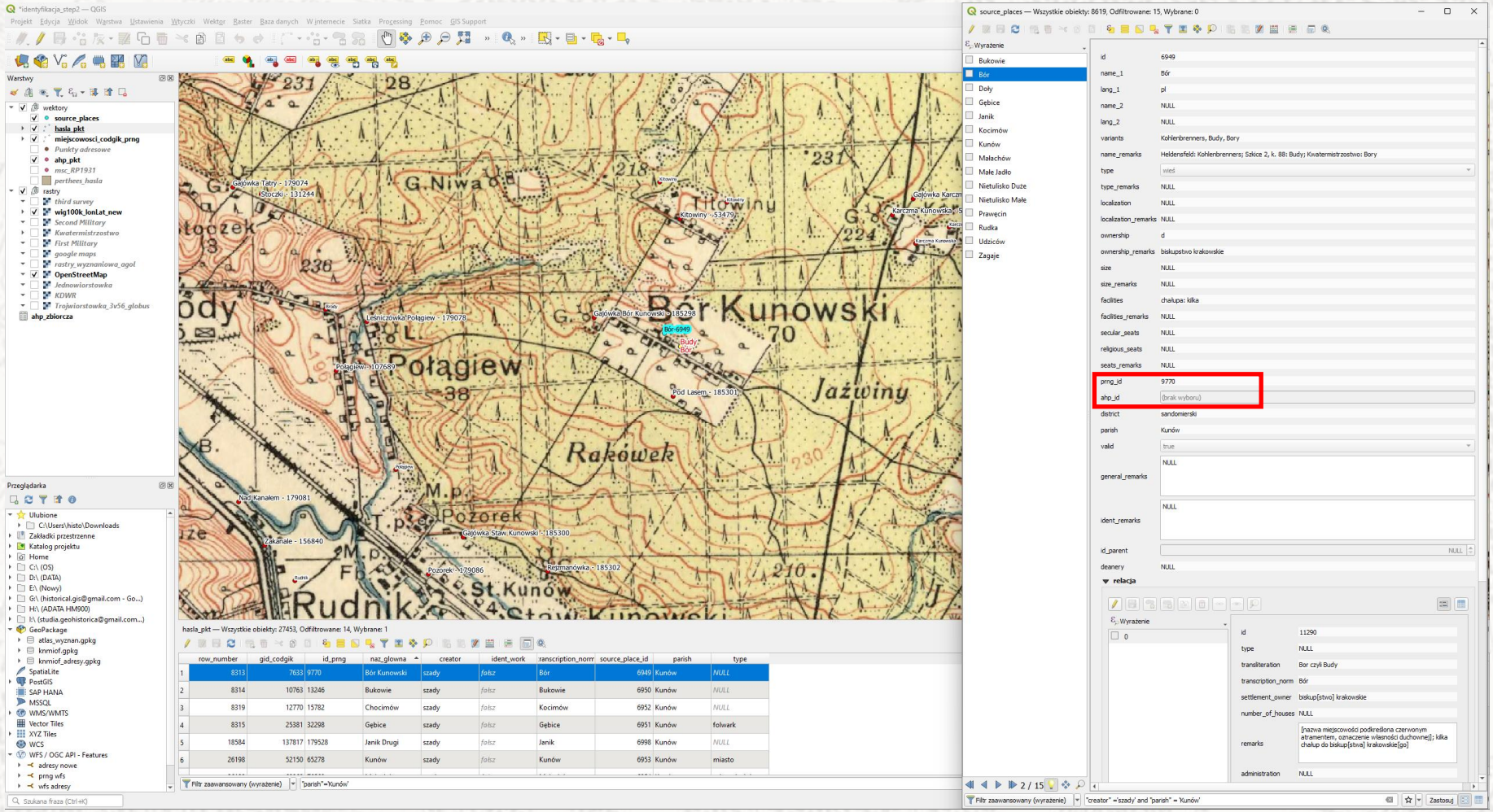
etapy



- prng id
- ahp_id
- nazwy
- przynależność adm.



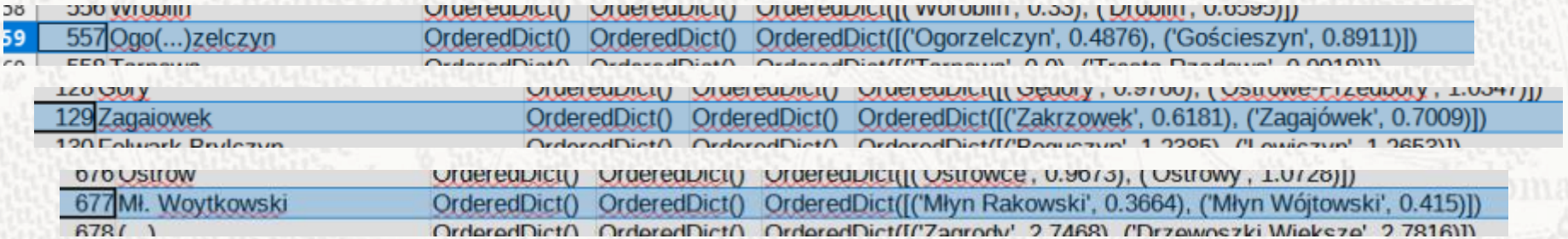
Identyfikacja, harmonizacja



Harmonizacja wewnętrzna

Porównywanie nazw:.

SQL + Levenstein + TrigramMatching + DeezyMatch



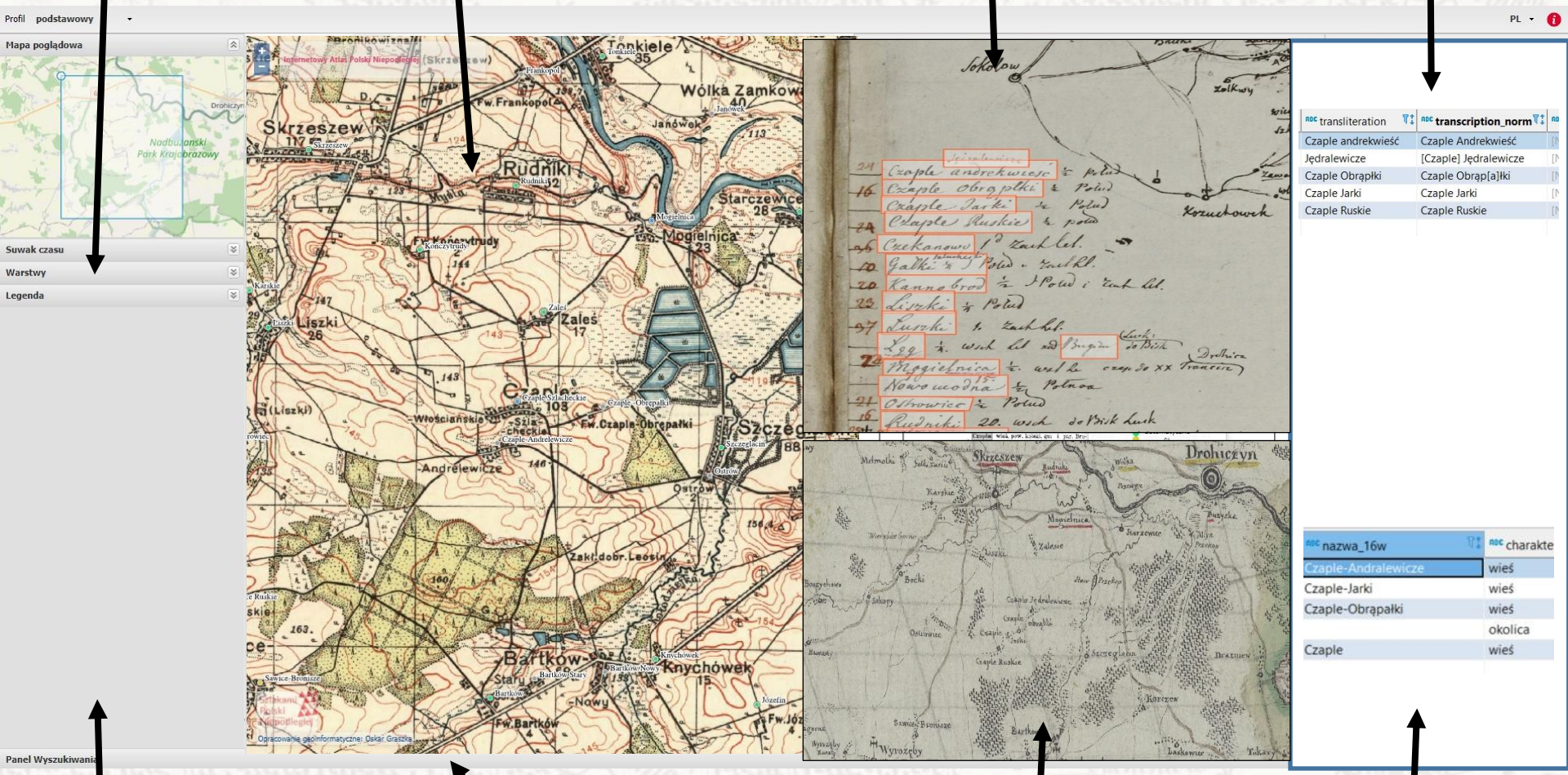
Koncepcja prezentacji danych - geoportal

legenda

mapa wynikowa

szkice parafialne

dane źródłowe



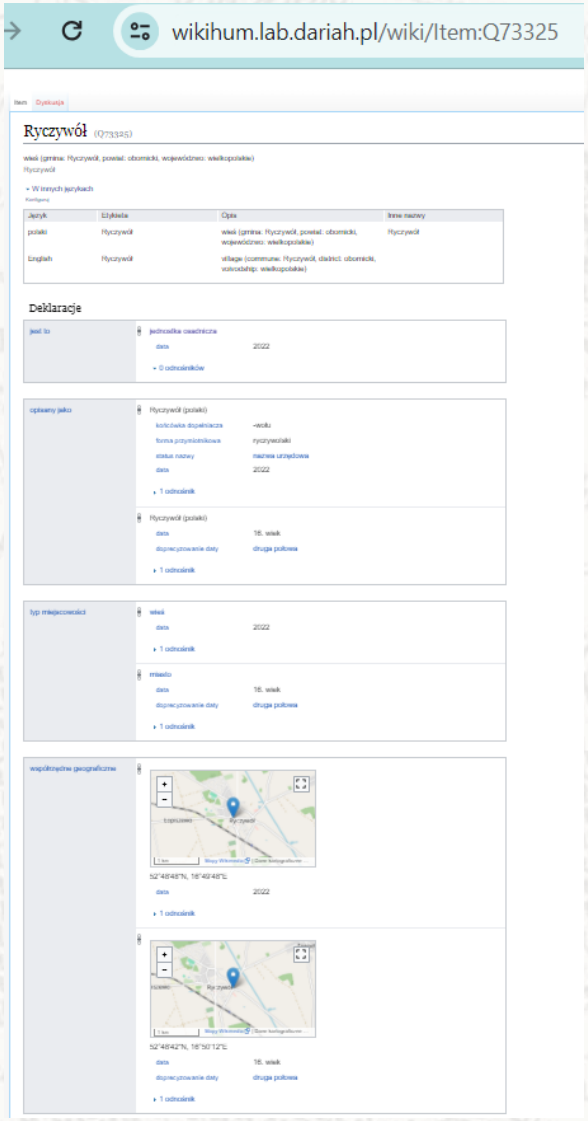
referencje

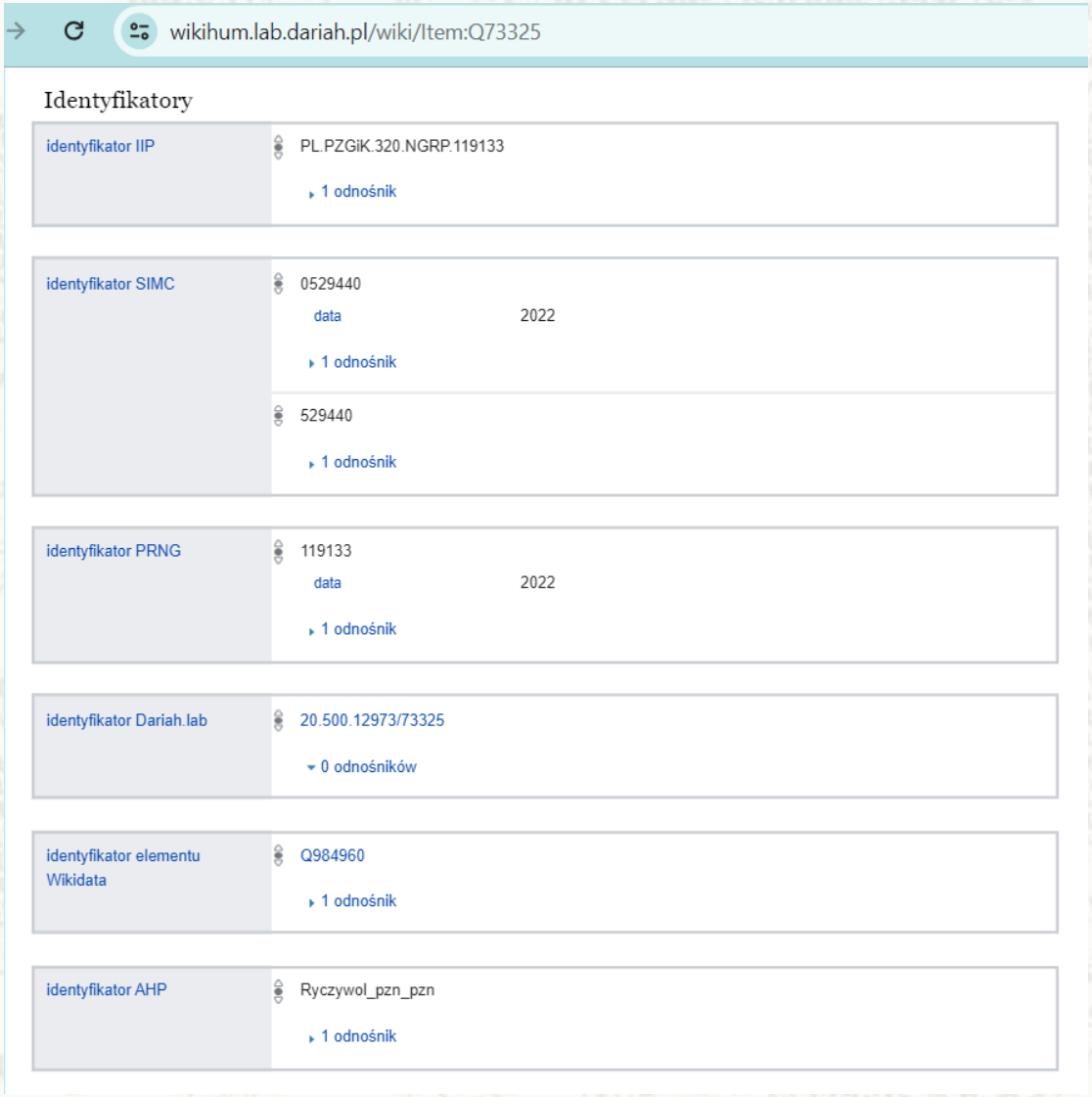
wyszukiwanie

mapa by Perthées

dane krytyczne

Manifestacje czasowe, trwałe identyfikatory, ontologia HGIS





Dziękuję za uwagę!