

Merging statistics and geospatial information, 2020 projects - Hungary

Statistics Explained

Integration of geospatial information and statistics; 2020 project; final report 28 February 2023

This article forms part of Eurostat's statistical report on the *Integration of statistical and geospatial information*.

Problem

Easier and quicker access is needed to a wide range of data at detailed territorial levels.

Objectives

The overall project objectives were:

- to support the integration of geospatial and statistical information during the statistical production process;
- to provide evidence-based data for policy support.

Method

A geocoded address dataset for Hungary was acquired along with map layers. The latter included polygons of individually generated blocks delimited by public areas, road/railway networks, hydrography and other landmarks. The addresses of the geocoded address dataset had unique identifiers and were classified into appropriate blocks.

The main challenge was to connect the geocoded address dataset with the existing address register of the statistical office, as the identifiers and database structures were different. For example, in the geocoded address dataset house numbers were stored in four fields rather than one; these were merged before linking.

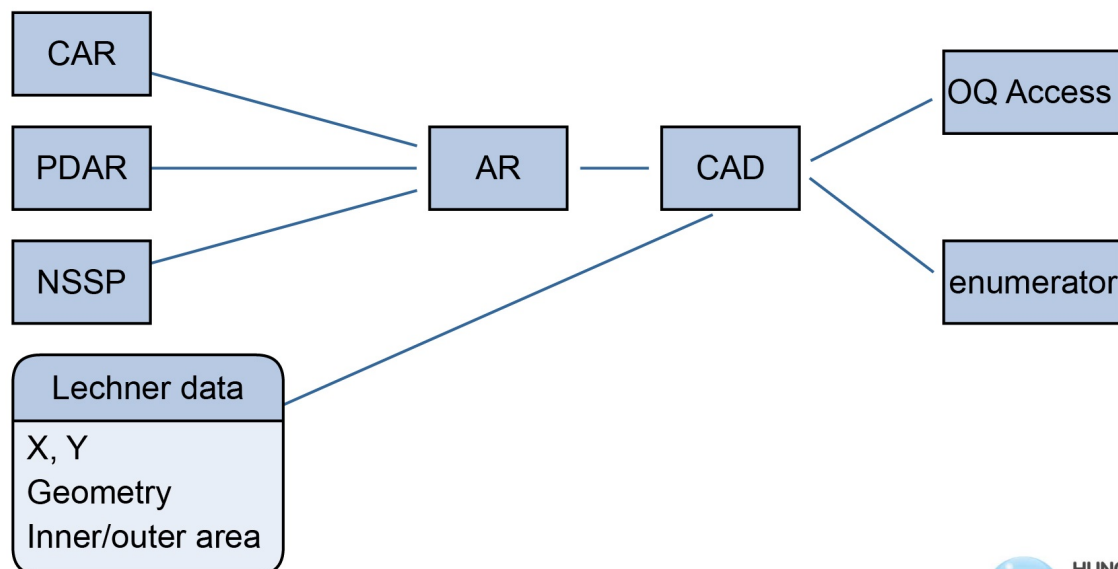
Various data sources were matched based on keys or descriptive fields. The quality of links was checked, and errors corrected using address cleaning procedures that were developed. The geocoding algorithm consists of a public area checker and three geocoder/logging functions (according to the house number, the plot number and the public area) to update the address dataset.

Results

A geocoding process was developed with addresses from external data sources. The geocoding process was automated and integrated into the maintenance and development process of various registers. The statistical office's address register was connected with the geocoded address database. It was then used during the preparation of the 2022 population census and can also be used to improve the geocoding of a range of other registers that were already linked to the address register.

- The completed geocoding application enables geocoding of various databases.
- Faster geocoding makes it easier to generate the data needed for data releases, and users can access the desired data sooner.
- Access is provided to a wider range of data at a deeper level than settlements.
- Municipalities can easily access data for settlement development projects.
- Ministries can access data more quickly for any detailed territorial level.

External Databases Internal Databases Outer Data Sources



- CAR - Central Address Register (Ministry of the Interior)
- PDAR - Personal Data Address Register (Ministry of the Interior)
- NSSP - National Statistical Survey Program (HCSO)
- Lechner data - National Mapping Agency
- AR - Address Register (HCSO)
- CAD - Census Address Database
- OQ Access - Access for the Online Questionnaire

Figure 1: Procedure to combine external data into one field in a format that can be matched to the address register

Methodology

- [Final reports \(EU Login required\)](#)

External links

- [Hungarian Central Statistical Office](#)