

To what extent can open geospatial resources help us overcome lack of GPS data?

Locating Raw Addresses in Emerging Cities, the Case of Dakar, Senegal

Motivation

Geospatial information is increasingly easy to collect, but is not so easily available historically, and less so for historical administrative data. Hence, **rich administrative records are often under-exploited due to a lack of geographic information**. Spatial units are required for spatial analysis and representation (GPS or fine geographic areas), and although **such data usually includes addresses, challenges arise from converting these addresses into usable geospatial information**. This is illustrated using property tax data from Senegal.

Mapping the unmapped

- **Use case:** leveraging property tax data to study the impact of public transport infrastructure on real estate
- **Requirement:** knowing if properties are near or far from the infrastructure (example BRT map)
- **Data:** Property tax data provided by the Senegal National Tax Authority (DGID) only contains text addresses

Which tools are best suited to overcome this challenge?

2 APPROACHES

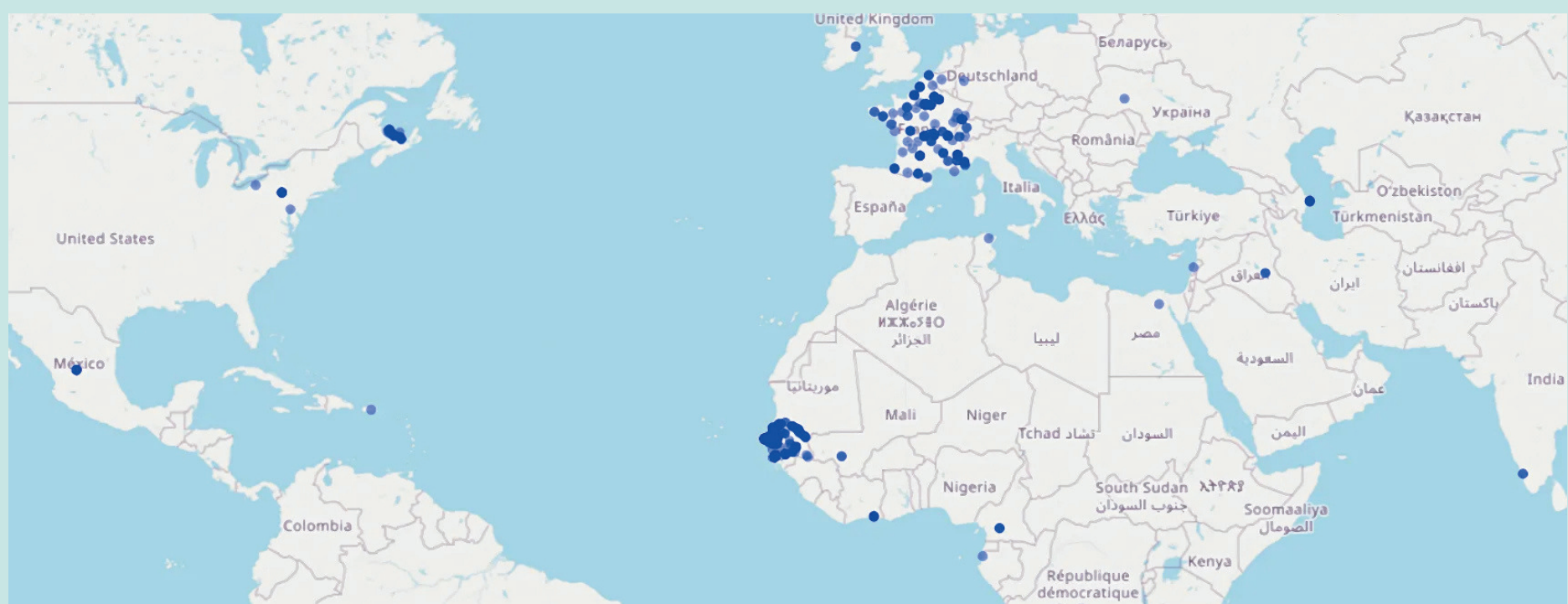
Google Maps API

Approach: Querying the Google API using an algorithm to automate and scale addresses research on Google Maps, inputting the text addresses from the data.

Result: misallocation issues (outside Senegal), only 11% of addresses matched accurately (precise geocoordinates).

Limitation: Informal address formats, landmark-based descriptions.

Raw geolocation results from Google Maps API



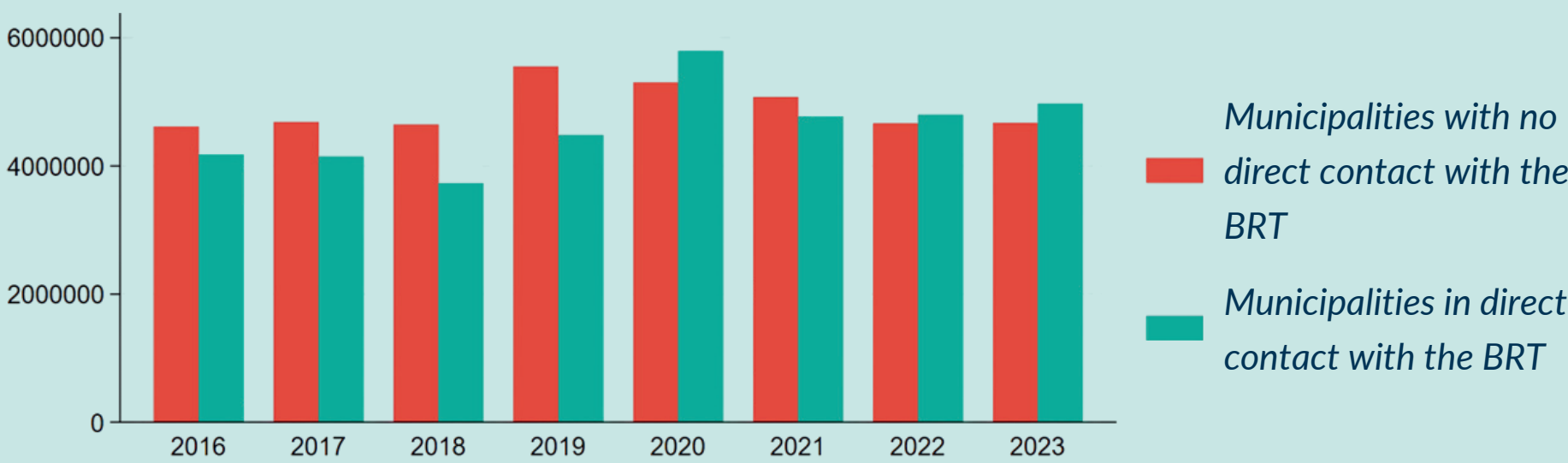
Text-area matching

Approach: Dictionary-based text analysis, where an algorithm matches the text address content with a list of predefined known and mappable locations.

Result: 62% of addresses are located, but only at the municipality area level.

Limitation: Lack of publicly available fine-grained geographic data.

Evolution of the average rental value in Dakar by year and by proximity to the BRT



Note: Raw average values after removing the lowest and highest 5% of observations

Findings

- ✓ **Address localization is difficult in cities like Dakar** due to informal and non-standardized address systems lacking detailed geographic divisions.
- ✓ **Collecting GPS data at the source is ideal**, but when unavailable, algorithms that process raw address text can help make existing datasets usable for spatial analysis.
- ✓ **Dictionary-based text matching offers a viable solution**, provided there is a comprehensive and accurate list of local area names.