All Dates Lead to Rome: Extracting and Explaining Temporal References in Street Names

Rosita Andrade
Max Planck Institute for Informatics
Saarland Informatics Campus
Saarbrücken, Germany
randrade@mpi-inf.mpg.de

Jannik Strötgen
Max Planck Institute for Informatics
Saarland Informatics Campus
Saarbrücken, Germany
jannik.stroetgen@mpi-inf.mpg.de

ABSTRACT

Street names say a lot about a country's or region's identity. So far, they have mostly been analyzed manually and for very limited regions (e.g., a city), and hardly any large-scale studies have been performed automatically. A phenomenon not yet studied are street names with date references. These are of special interest as they can be used to commemorate important events in a region's history. In this paper, we present our approach to automatically extract such street names across the world. We analyze the dates' temporal and geographic distribution, and automatically gather potential explanations why specific dates occur in particular regions. We further describe DATE-ROME, a tool to interactively explore the streets, their distribution, and possible explanations.

Keywords

street names; temporal tagging; map-based exploration

1. MOTIVATION & BACKGROUND

Street names tell a lot about a country and its culture. So far, they have mainly been studied in the fields of geography, topology, and social science as they serve well for building commemorative landscapes. Typically, a manual analysis of a limited region or focusing on a particular personality is performed. For instance, [2] and [3] analyzed the renaming of streets in East Berlin after the German reunification and in Bucharest during the time of the Romanian People's Republic (1947–1965), respectively. [1] studied (difficulties of) the commemoration of Martin Luther King based on the streets in the US named after him. In contrast, an automatic, large-scale analysis tackled the distribution of streets named after male and female personalities in several major cities, showing that streets named after men are more frequent and more centrally located [4]. An interesting aspect is also why a street carries a particular name. An application providing explanations of street names (in Spanish) is the Nomenclátor de Calles, which covers the city of Mon-

Table 1: Example streets (S) with explanations (E).

S	Straße des 13. Januar (January 13), Germany (Saarland)
E	On January 13, 1935, the Saar status referendum took place; 90% voted for reunification with Germany.
S	23 Nisan Caddesi (April 23), Turkey (e.g., Ankara)
E	National sovereignty and children's day; opening of the Grand National Assembly of Turkey at Ankara in 1920.
S	Rue du 8 Mai 1945 (May 8), France (e.g., Paris)
E	May 8 is a French holiday; this day in 1945, de Gaulle announced the end of WWII in France.
S	Via XX Settembre (September 20), Italy (e.g., Rome)
E	This day in 1870, the capture of Rome ended the reign of the Papal States (754–1870).
S	Estádio 11 de Novembro (November 11), Angola (Luanda)
E	Stadium and surrounding street are named after the date of Angola's independence in 1975.

tevideo, Uruguay.[1] Given query street names, it presents manually collected explanations.

In this paper, we focus on a specific type of street names – those with references to dates – and harvest explanations for their usage automatically. Often, these names are used to refer to historic events, which are of particular importance in specific regions (e.g., a state or country). For instance, the famous *Straße des 17. Juni* in Berlin commemorates the uprising of the East Berliner workers on June 17, 1953 when several protesting workers were shot. Further examples of street names with explanations are shown in Table 1.

In contrast to most prior works on analyzing street names, we do not only consider a small region. Thus, we collect street data of the world and automatically detect temporal expressions in the names. We then apply various approaches to harvest explanations for street names. Finally, we describe DATE-ROME, a system to explore all streets with date references, their distribution, and possible explanations.

2. FINDING DATES IN STREET NAMES

Using Geofabrik[2] and Gisgraphy[3], we extracted all streets of OpenStreetMap (OSM)[4] for each country (or region, e.g., states in the US) as ⟨`osmid`, `streetname`⟩ tuples. Then, we assigned to each country or region all official languages spoken in respective areas.[5]

[1] http://www.montevideo.gub.uy/aplicacion/nomenclator
[2] https://www.geofabrik.de/geofabrik/geofabrik.html
[3] http://download.gisgraphy.com/openstreetmap
[4] https://www.openstreetmap.org/
[5] https://en.wikipedia.org/wiki/List_of_official_languages_by_country_and_territory
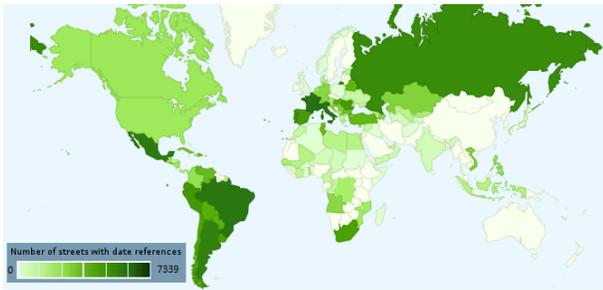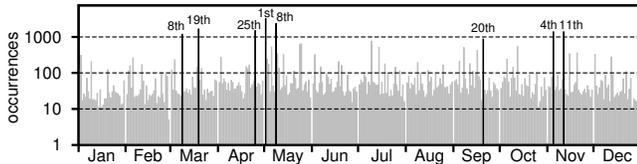
Figure 1: Geographic distribution.

Figure 2: Temporal distribution (log scale).

To detect and normalize temporal expressions in natural language text (e.g., *June 17, 1953* can be normalized to `1953-06-17`), so-called temporal taggers can be used [6]. While there are some systems publicly available, we are faced with highly multilingual data so that we used HeidelTime, which contains manually developed language resources for 13 languages, and was recently automatically extended to cover more than 200 languages [5]. For languages with manually created HeidelTime resources, we used those, for all other languages, we used the automatically developed resources.

Note that our particular interest lies in street names containing references to particular dates with or without year information (e.g., *Fourth of July Road*) so that HeidelTime's automatically created resources, which are less sophisticated than the manually developed ones, can be expected to work fine for our purpose independent of the language. However, to make sure that we do not miss many date references, we also applied a post-processing step on all street names in which HeidelTime detected only a month reference and checked the rest of the street name for any numbers that might refer to a day. For some regions, this increased the number of street names with date references significantly.

Finally, we removed all duplicate street names that occur in the same postal area, suburb, or district (depending on which information was available in OSM using the `osmid`).

3. EXPLAINING DATE REFERENCES

Given a street in a particular region that refers to a particular date, our goal is to find possible explanations (cf. Table 1). For this, we run the following strategies:

1. For each country, we collect its national holiday(s).

2. For each country, we crawl its English Wikipedia page and extract all sentences with date references.

3. For ⟨date,region⟩ tuples without explanations from (1) and (2), we query our full temporally annotated Wikipedia dump for co-occurrences of the region name and date (with region ∈ {city, state, country}) on sentence level.

4. For each street name, we determine the distance to the next street with another date reference, and the number of streets with date references within a distance below a specified threshold.
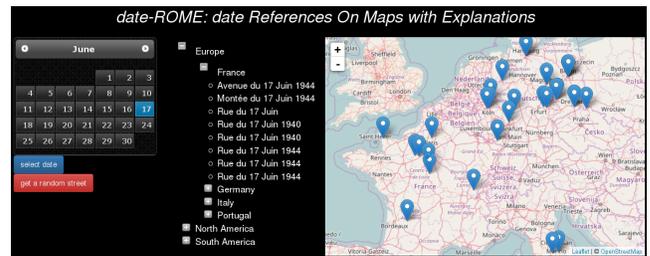

Figure 3: Screenshot of date-Rome for *June 17*.

While the first three strategies are used to determine references to important dates in a region, the last one helps to identify regions in which many street names refer to dates – as they might just be used for naming streets in a particular region instead of referring to real events in the past.

4. EXPLORING THE WORLD

In Figure 1 and Figure 2, we show the geographic and temporal distribution of street names with date references, respectively. Overall, we extracted more than 38,000 of such streets. May 1, May 8, and March 19 and France, Italy, and Brazil are the most frequent days and the countries with most streets with date references.[6]

In Table 1, we showed examples with street names in several languages. Using our system DATE-ROME[7] (date References On Maps with Explanations), all streets with date references, their geographic distribution, and explanations can be explored interactively. In Figure 3, we show a screenshot in which streets referring to *June 17* are marked (zoomed into Europe). Each street can be selected via the geographically ordered list or the map to retrieve further information and potential explanations, and to zoom into the map at respective locations. Note that streets referring to the same date might have different explanations depending on the region. Closest streets referring to the previous and next day are also linked to allow traveling the world chronologically.

5. REFERENCES

[1] D. H. Alderman. Naming Streets for Martin Luther King, Jr.: No Easy Road. *Landscape and race in the United States*, pages 213–236, 2006.

[2] M. Azaryahu. German Reunification and the Politics of Street Names: The Case of East Berlin. *Political Geography*, 16(6):479–493, 1997.

[3] D. Light, I. Nicolae, and B. Suditu. Toponymy and the Communist City: Street Names in Bucharest, 1948–1965. *GeoJournal*, 2(56):135–144, 2002.

[4] A. Sankaranarayanan. Mapping Female versus Male Street Names. https://www.mapbox.com/blog/streets-and-gender/, 2015.

[5] J. Strötgen and M. Gertz. A Baseline Temporal Tagger for All Languages. In *EMNLP '15*, pages 541–547, 2015.

[6] J. Strötgen and M. Gertz. *Domain-sensitive Temporal Tagging*. Morgan & Claypool Publishers, 2016.

---

[6] All findings rely on automatic processes. Regions might exist, for which the extraction failed or was less successful. As future work, we plan an in-depth analysis and evaluation of the extraction and explanation harvesting approaches.

[7] `http://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/TimeSEA`