

Summary

Littar was the second-prize winning entry in an app competition. It implemented a system for visualizing places mentioned in individual literary works. Wikidata acted as the backend for the system. Here I describe the Littar system and also some of the issues I encountered while developing the system: How locations and literature can be related, what types of location-literature relations are possible within Wikidata, what limitations there are and what questions we may ask once we have enough data in Wikidata.

Literature and location

Literature science often approaches its subject with respect to time (periods of genre) or with respect to persons (the author). A newer trend, the “spatial turn”, studies literature with respect to locations^{1,2}. Locations might be:

- Real geolocatable location, e.g., Paris
- Fictitious places, e.g., Agatha Christi’s St. Mary Mead
- Generic sites, e.g., coffeehouse or police station

Literature may relate to locations through the narrative but also where the text has been composed, collected and used.²

It varies considerably how much authors use geolocatable places in their work, from none (e.g., tales of Hans Christian Andersen) to plenty (e.g., Patrick Modiano). Non-fiction works recounting past events may often refer to places.

Other Wikidata properties

For works: Main subject, place of publications, publisher, inspired by and some others. The first three are widely used.

For authors: place of birth, place of death, place of burial, educated at, residence, employer.

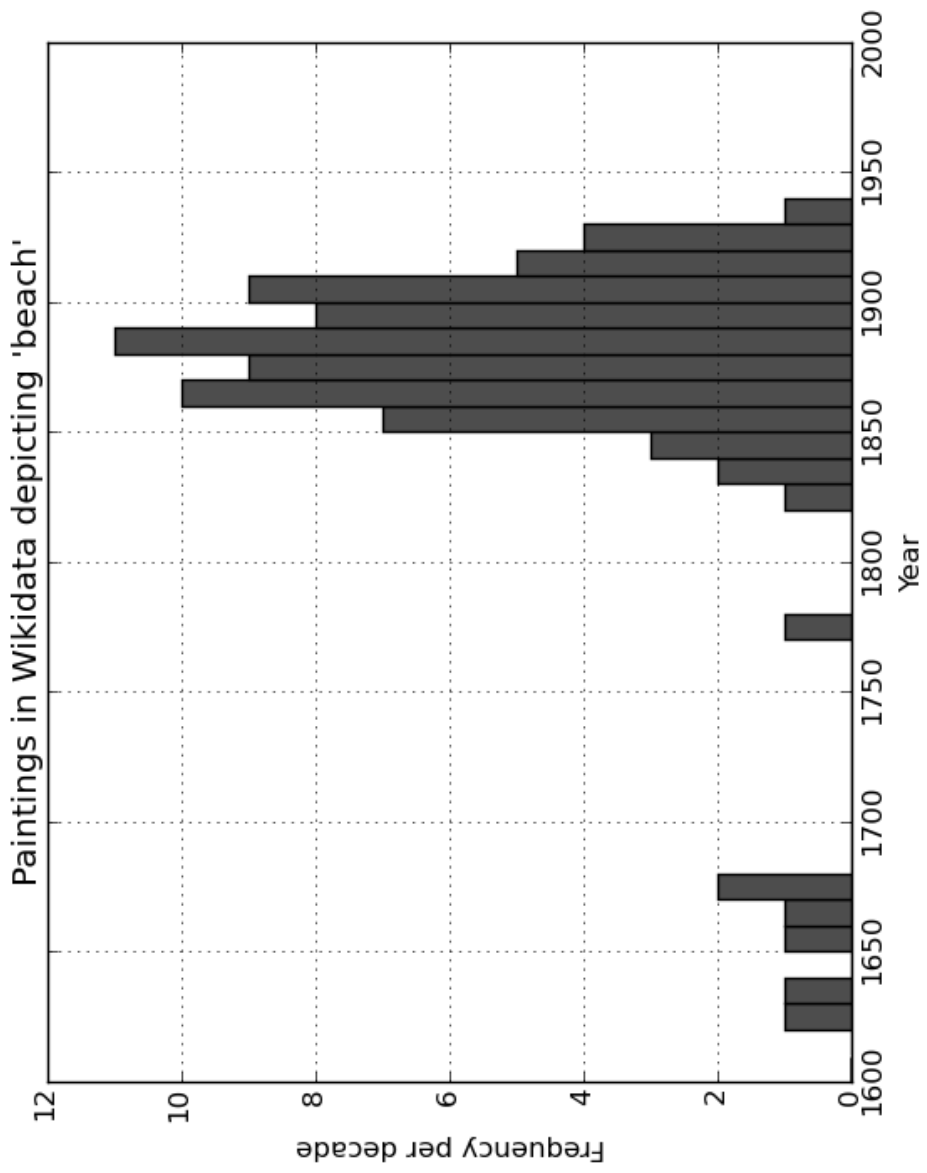
There are no appropriate properties for “described location”, i.e., locations that are not part of a narrative, but only described, e.g., in lyrical poems. Wikidata does not describe the function of the place, e.g.: home, work, murder site, travel destination.

Computational humanities?

Can we use the geographical data beyond map making? For “computation literary science”?

As an example: It has been argued that *the beach* was “discovered” between 1750-1840.¹

With Wikidata’s depict property it is possible to count the number of paintings depicting a beach.

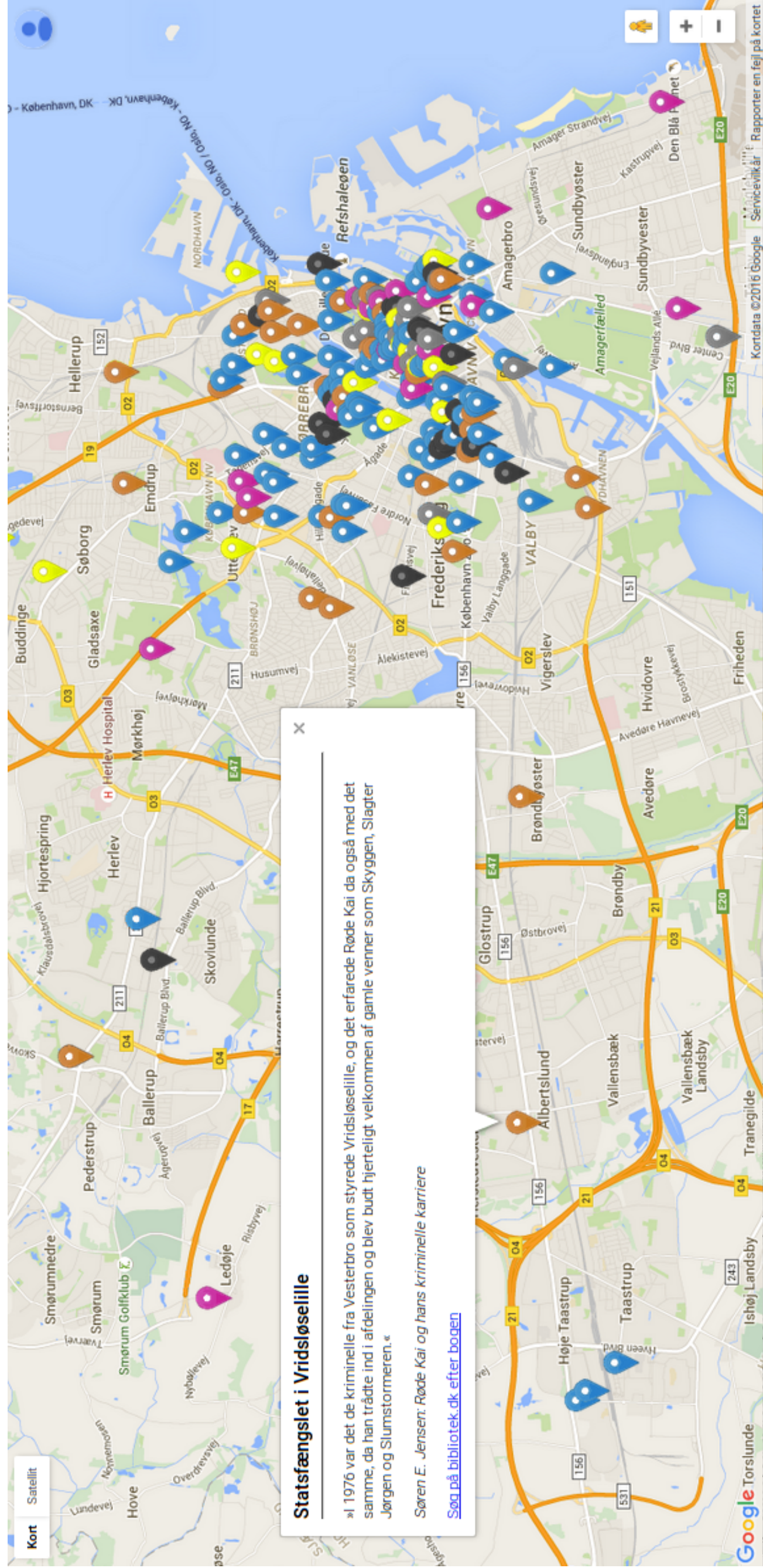


The above plot shows paintings in Wikidata depicting “beach” as a function of decade with the bulk of beach depictions after 1840.

Is it possible to do a similar analysis for literary works? There are presently too few narrative locations of this generic type in Wikidata to enable any meaningful analysis.

Littar

Littar (“literatur radar”) was an entry for an app competition set up by DBC, the Danish national library service, asking: “How can data science be used to provide library users with new and better experiences?” The map below shows part of the map for the “app”.

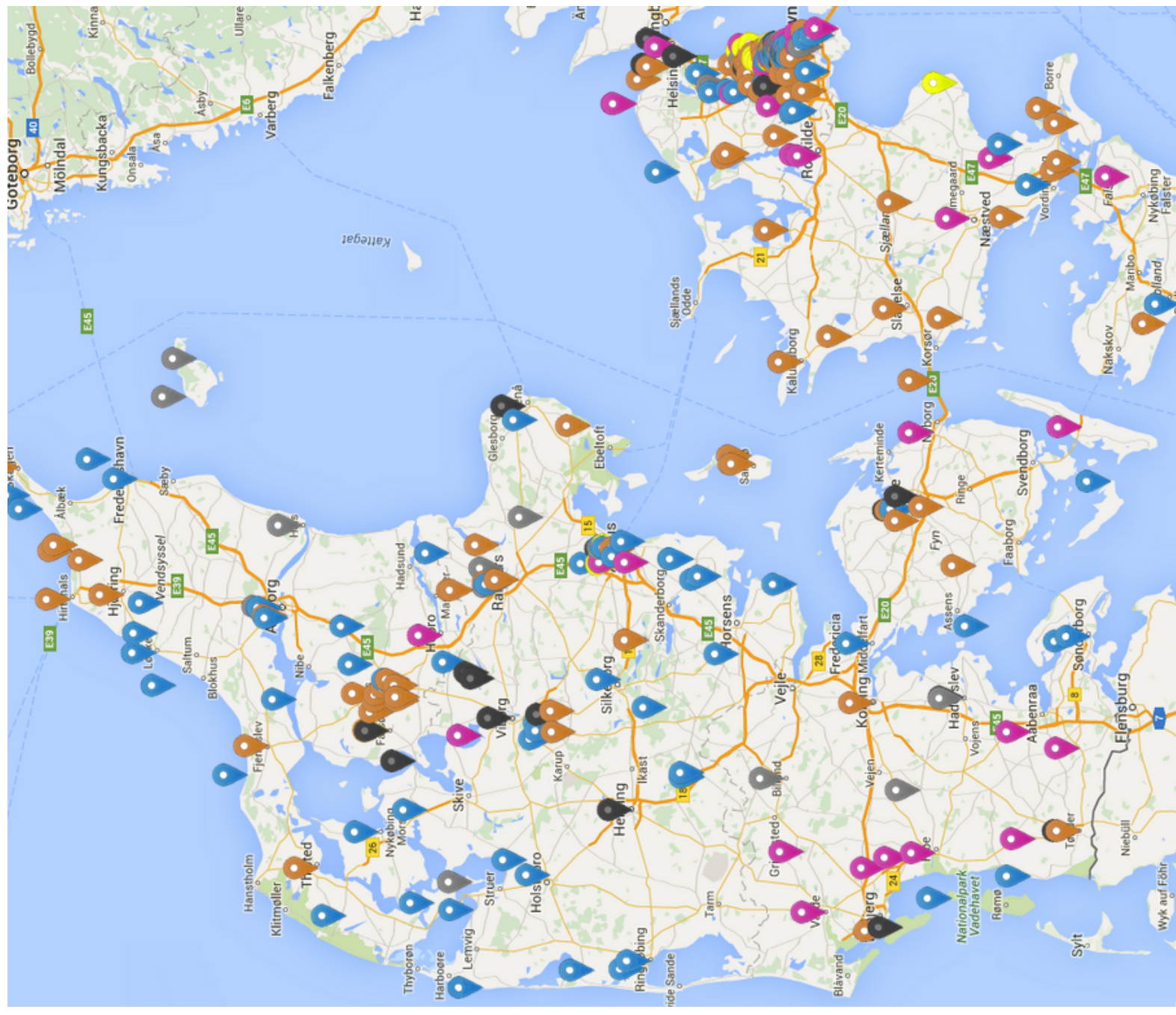


Littar map with Google Map and a quote from the book *Røde Kai og hans kriminelle karriere* by Søren E. Jensen. Here the user has navigated to the Copenhagen suburb Albertslund and pressed the placemark for The State Prison in Vridsløselille, so the popup with information about the location and work appears. Note the concentration of placemarks at the center of Copenhagen.

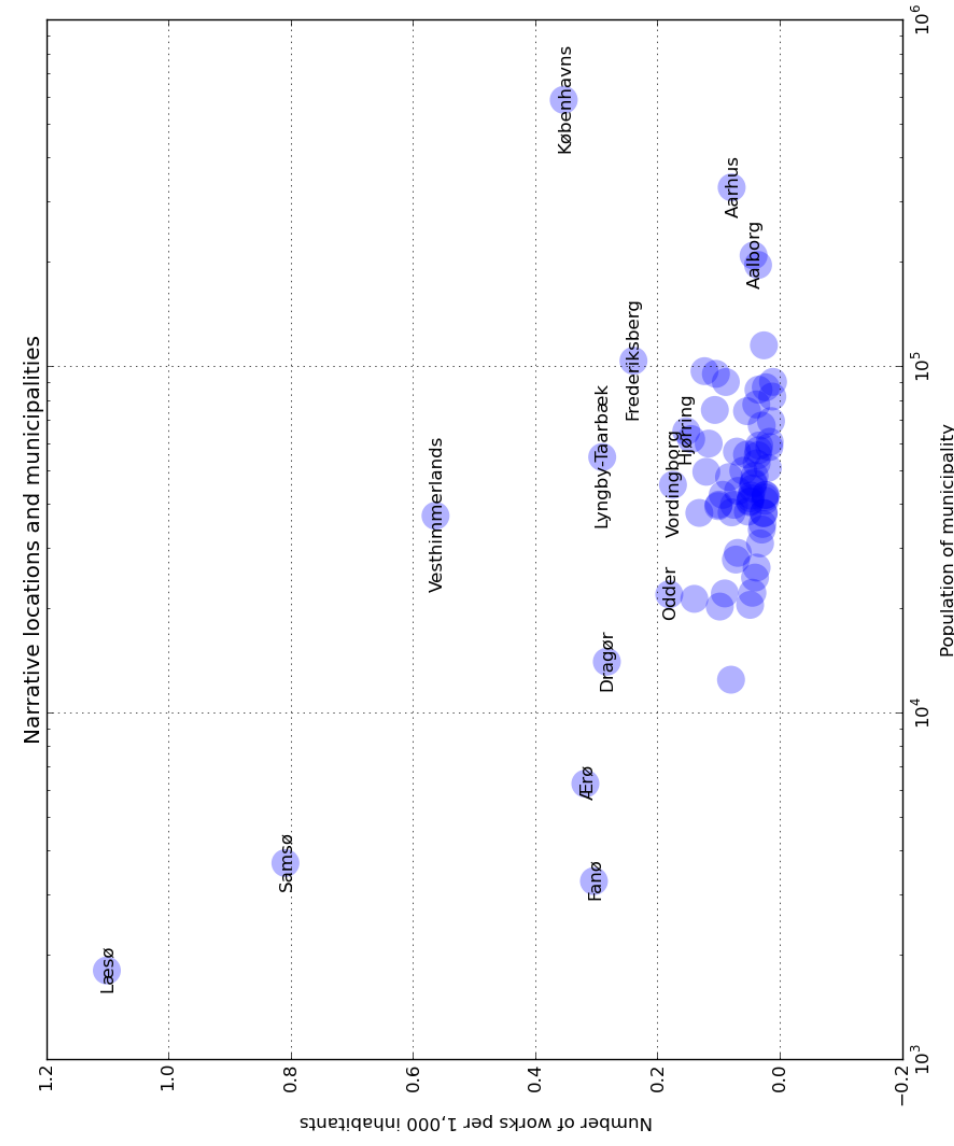
Components

- Placemarks are colored according to the kind of work: Novel, non-fiction, poetry, ...
- Black placemarks indicate multiple works use the same narrative location.
- Placemarks associated with a popup window with name of location and one or more works with title of work, a quote and a link to a library service.
- Simple implementation with HTML, Javascript and KML file generated with Python and a SPARQL call.

Literary geographical bias?



Is there a literary bias in the concentration of narrative locations? There seems to be a concentration in the capital, Copenhagen. Recent research has addressed possible geographical biases in peer-produced works.^{3,4}



Can the Copenhagen concentration be explained by a per capita effect? The plot above attempts to answer this question showing the ratio between number of narrative locations of works and the population in Danish municipalities.

Narrative locations in Wikidata

Littar uses standard Wikidata as the “backend” for the data with pages for literary works using the “narrative location” property to items that have a geographical coordinate.

 Littar Littar Littar Littar Littar Littar Littar Littar Littar Littar Littar Littar Littar Littar
--

Above is an example from Wikidata with information about narrative locations for *The Fall of the King*, a novel by Johannes V. Jensen. Here with a quote from the Copenhagen street *Pilestræde*. The 400-characters *quote* property in the references is used for the quote. (From the Wikidata page <https://www.wikidata.org/wiki/Q1758876>)

- In Wikidata there has been a tendency to use broad geographical items (cities, counties) for narrative locations. If the data is to generate an “interesting” map then narrative locations must be associated with geographical items that are more narrow: streets, buildings, small towns, etc.
- Much too few Danish narrative locations were available at the beginning of the project. I entered much of the data myself.

Discussion

- Interesting opportunities to use Wikidata for geographically-oriented literary science.
- Interesting information retrieval opportunities, e.g., “give me nearby poems”, or identification of areas with a low density of narrative locations.
- Yet fairly few literary works are tagged in Wikidata wrt. location.
- Narrative locations are often used for geographically broad items, such as cities and countries. Tagging should be more narrow to generate “interesting” maps.

Availability

The prototype for the Littar web service is presently available from <https://fnielsen.github.io/littar/>.

The code to query Wikidata and generate the KML file is available from <https://github.com/fnielsen/littar/>

Acknowledgment

Reinholdt W. Jorck og Hustrus Fond is thanked for their generous support.

References

[1] Rosiek J. *Danmark, Gurre, stranden. Steder i dansk litteratur*. U Press, 2015.
[2] Mai AM. *Hvor litteraturen finder sted*. Gyldendal, 2010.
[3] Johnson L et al. Not at home on the range: Peer production and the urban/rural divide. *Information Geographies*, 2014.
[4] Straumann R and Graham M. The geographically uneven coverage of Wikipedia. *Information Geographies* at the Oxford Internet Institute, 2014.