

# Frankenplace: An Application for Similarity-Based Place Search

Benjamin Adams<sup>†</sup> and Grant McKenzie<sup>\*</sup>

<sup>†</sup>Department of Computer Science and <sup>\*</sup>Department of Geography  
University of California, Santa Barbara  
Santa Barbara, CA, USA 93101

<sup>†</sup>badams@cs.ucsb.edu and <sup>\*</sup>grant.mckenzie@geog.ucsb.edu

## Abstract

When experiencing or describing a new place people will often compare it against other places that they already know. However, this human attention to the simultaneous similarities and differences between places is not reflected in the design of user interfaces of current place search technologies. In this demo, we present Frankenplace, an application for doing similarity-based place search that allows users to interactively find new places based on mixtures of features drawn from different places. The features of places are derived from a combination of authoritative data sources and unstructured observation data from social media, and organized into an extensible set of layers. We demonstrate the Frankenplace interface, which lets a user build a profile of a target place by selecting the most relevant of the properties shared by known places.

## Introduction

We present Frankenplace, a web-based application designed for performing semantic similarity-based place search. In contrast to traditional place search in digital gazetteers, the goal of this application is not to locate named places on a map (Hill 2000). Instead, Frankenplace is designed to enable users to find new places that are analogous to places that they already know. Any computational representation of a *place* will depend on the kinds of source data used to construct the representation. Frankenplace is designed to allow for multiple layers of data to represent place properties that are derived from multiple sources. In our demo we illustrate how data of climate observations from the WorldClim dataset and topic models derived from travel blog entries can be combined by the user to find similar places depending on context (Hijmans et al. 2005).

Although places have spatial properties including a location and extent, they also have many non-spatial properties. The operationalization of place is a difficult problem, because many properties of places (including the spatial properties) are subjective, socially-constructed, and can vary from person to person (Winter, Kuhn, and Krüger 2009). Work has been done

on specifying how the spatial properties can be learned from observations (Montello et al. 2003). Additionally, multidimensional measures of ‘sense of place’ have been explored in human geography research, but they tend toward psychological experimental studies of very specific geographic settings, e.g., lakeshore properties (Jorgensen and Stedman 2006). Rather than impose one representation for a place the approach taken here is to create a system that allows users to explore the similarities and differences between places based on heterogeneous observation data from multiple sources.

## Frankenplace application

As described in the introduction, for our prototype implementation we have two data sets derived from travel blog entries and climate observations that are represented as layers in the Frankenplace application. Over 270,000 blog entries were downloaded from TravelBlog<sup>1</sup>, stemmed, and processed through a standard English stop word list. The climate data were originally organized by month but we re-encoded them by season, so that places compared across the northern and southern hemispheres would not be artificially dissimilar. The travel blog entries were analyzed using Latent Dirichlet allocation to identify latent topics (Blei, Ng, and Jordan 2003). In order to get a set of topic property values for a given place, the average topic value for each latent topic was calculated from all the entries at that place.

Each data layer in Frankenplace is modeled as a conceptual space that allows for context-dependent semantic similarity measurement (Adams and Raubal 2009). A unique similarity function can be defined for each layer as appropriate for the data. The similarity values for all the layers are then aggregated in a weighted sum. In our implementation we used a Euclidean distance metric for the climate data and the Kullback-Liebler divergence to measure the distance between topic signatures derived from travel blogs.

Upon launching the site, the user selects a location through either a mouse click on the map or a GeoNames<sup>2</sup> based place name search. The resulting location

<sup>1</sup><http://www.travelblog.org>

<sup>2</sup><http://geonames.org>

is highlighted on the map and the user is presented with a series of word clouds representing the topics that constitute the place. From this point the user has a number of options.

Each of the topics generated from the TravelBlog data for a specific location can be added to the *Laboratory* section of the website. The *Laboratory* is essentially a bucket of *Parts* (properties) that can be used to build the profile of an ideal place. These properties become the basis of a new place search. Not only can the user select from the TravelBlog topics, but other sources of data such as climate may be used as search parameters. Figure 1 shows an example of the construction process. The user has taken a property (in this case either a *topic* or a *temperature value*) from each of three different place name searches and added them to the *Laboratory*. A search based on these properties returns a set of locations, ranked in order of similarity (to the search properties).

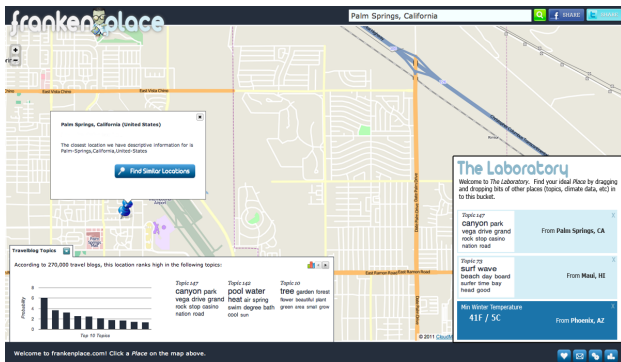


Figure 1: Using the Frankenplace laboratory.

The *Find Similar Locations* button queries the database for places with the highest overall semantic similarity in terms of a weighted sum of the similarities in each layer. The top five resulting locations are marked on the map and the most *similar* properties are displayed to the user. For example, Figure 2 shows a similar locations search for Fairfield, California in terms of the travel blog topics. After clicking *Find Similar Locations*, five markers are added to the map. In this example, the second most similar place to Fairfield, California is Strasbourg, France. The info bubble informs the user that these locations are most similar on the topics of Wines and Desserts. Note that the third most *similar* topic is not the third *highest* topic for Fairfield, California. Again, topics resulting from this query may be added to the *Laboratory*.

## Conclusion

In this paper we presented Frankenplace, a novel web-based place search application. Frankenplace offers an alternative to place name search engines currently available by representing places as unique mixtures of properties extracted from both authoritative and social me-

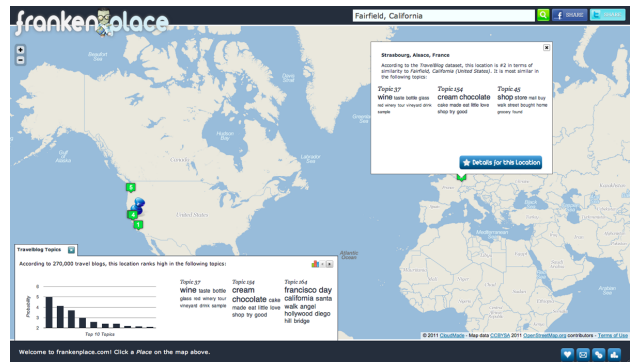


Figure 2: Sample similarity search for places.

dia sources. By allowing a user to construct a place query through interactive exploration of known places, new places can be found through their geographical analogs with respect to specific properties. We described using Frankenplace with a sample data set derived from topic modeling analysis of travel blog entries in combination with climate data.

## References

- Adams, B., and Raubal, M. 2009. A metric conceptual space algebra. In Hornsby, K.; Claramunt, C.; Denis, M.; and Ligozat, G., eds., *Spatial Information Theory*, volume 5756 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg. 51–68.
- Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3:993–1022.
- Hijmans, R.; Cameron, S.; Parra, J.; Jones, P.; and Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25(15):1965–1978.
- Hill, L. 2000. Core elements of digital gazetteers: Placenames, categories, and footprints. In Borbinha, J., and Baker, T., eds., *Research and Advanced Technology for Digital Libraries*, volume 1923 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg. 280–290.
- Jorgensen, B. S., and Stedman, R. C. 2006. A comparative analysis of predictors of sense of place dimensions: Attachment to, dependence on, and identification with lakeshore properties. *Journal of Environmental Management* 79:316–327.
- Montello, D. R.; Goodchild, M. F.; Gottsegen, J.; and Fohl, P. 2003. Where’s downtown?: Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition and Computation* 3(2&3):185–204.
- Winter, S.; Kuhn, W.; and Krüger, A. 2009. Guest editorial: Does place have a place in geographic information science? *Spatial Cognition and Computation* 9:171–173.