Delft University of Technology

User Manual

UrbanSearch

Authors:
Tom Brunner
Marko Mališ
Gijs Reichert
Piet van Agtmaal

Supervisor:
Claudia Hauff
Clients:
Evert Meijers
Antoine Peris

July 5, 2017

T̃U Delft

1

User Manual

This is the user manual for the UrbanSearch Web Interface. Here we will discuss the various interfaces that are available in the UrbanSearch system. The manual is meant for new users of the system.

1.1 General remarks

The UrbanSearch Web Interface can be accessed by visiting `http://citynetworks.bk.tudelft.nl/`. Visiting this URL will open an interactive map on which intercity relations are visualised.
The front-end of the UrbanSearch system consists of several interfaces listed below. In the coming sections we will dive deeper into the functionality offered by the different interfaces.

Interactive Map with cities and relations visualised[1]

Classification interface use to extend the training data[2]

Classifier interface to train the classifier and get info about the data sets[3]

.

1.2 Interactive Map

To display the extracted and processed data in our system, we use an interactive map(figure 1.1). The map is implemented with use of Google Maps. This means the typical map controls, such as zooming, are the same as on any Google Maps map.
Furthermore, the map contains several entities which are used to convey information to the user and to interact with or manipulate the available data.
Below we will discuss the different entities available on the map.

---

[1] `http://citynetworks.bk.tudelft.nl/`
[2] `http://citynetworks.bk.tudelft.nl/classification`
[3] `http://citynetworks.bk.tudelft.nl/classifier`

Figure 1.1: The interactive map with one relation selected

1.2.1 Cities

Cities are represented by circles, which are drawn on the map and placed in the correct position using the latitude and longitude properties belonging to the city. We scale the circle based on the size of the population of the city. An example of a city on the map is show in figure 1.2.
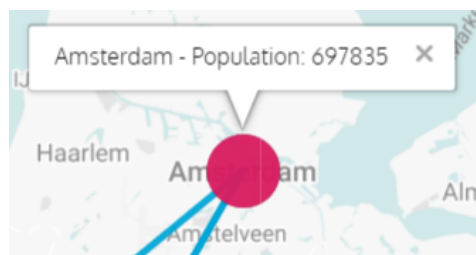

Figure 1.2: Hovering over a city will display extra information about the city

When a user hovers over a city on the map, a pop up appears with extra information about that city.

If a city is clicked, the relations to other visible cities are shown. Clicking on a selected city toggles the visibility of the previously displayed relations. To make clear for the user which cities are selected and which are not, we lower the opacity of selected cities.

1.2.2 Relations

Relations between cities are visualised by lines. Relations for a particular city are shown when a city node is selected as described in 1.2.1.

Figure 1.3: Selected city showing relations to other visible cities

Clicking on a relation opens up a new card containing information about this relation. The details about this card are discussed in 1.2.3.
The opacity of a relation indicates the relative strength of the relation compared to the relation with the highest score in the system. An example is given in figure 1.4.
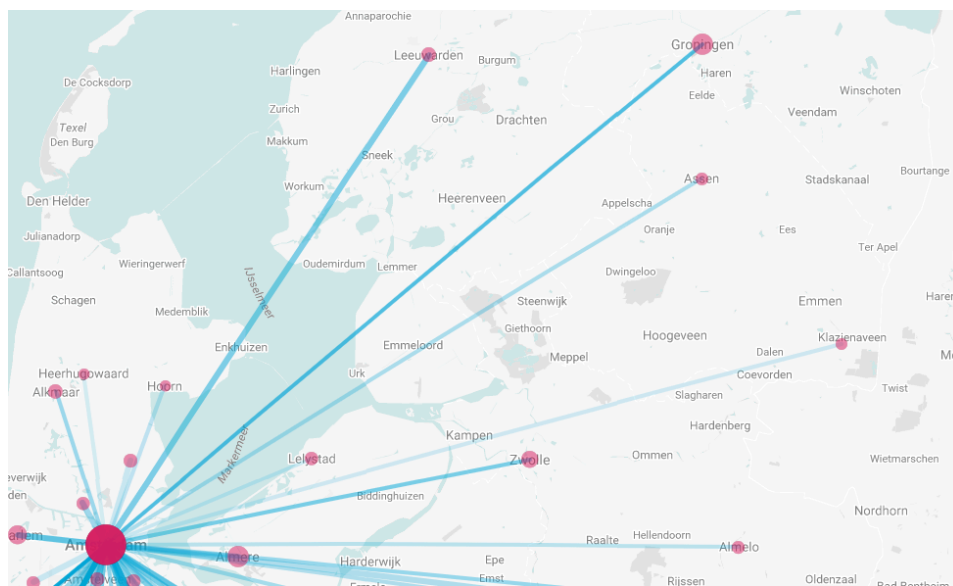

Figure 1.4: A higher opacity indicates a weaker relative total strength of the relation. So Groningen has a stronger connection with Amsterdam than Assen.

1.2.3 Side-menu

The side-menu displayed in figure 1.5 allows the user to control the information that is displayed on the map. First we will discuss the elements that are always available in the side-menu. Finally we will discuss dynamic content that can be added and removed from the side-menu.
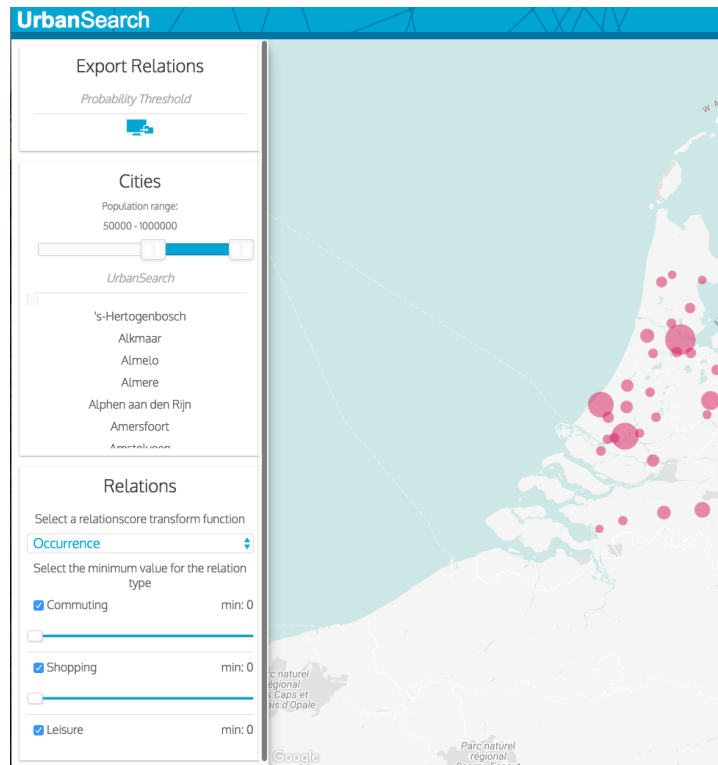
Figure 1.5: The side-menu displayed on the left side of the interface

City Controls

The city controls displayed in figure 1.6 allow users to control the visibility of cities and city relations.

The population slider can be used to set the visibility of cities. All cities with a population size within the selected range are displayed on the map.

The search input allows the user to search a list of visible cities. When a city in the list is clicked, the relations for this city are shown on the map.
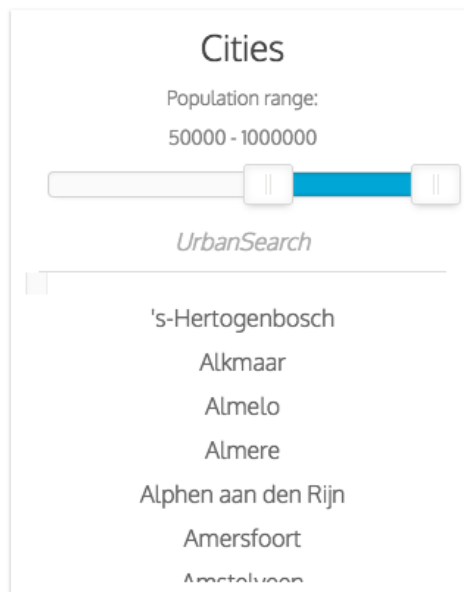
Figure 1.6: The city control card

Relation Controls

Using the relation controls displayed in figure **??**, a user can manipulate how intercity relations are displayed.

First of all, the user can select several "transforms" from the drop-down that is available in the relation controls card. These transforms calculate a new value for the relation strengths based on the original values of the relations and a given property (which basically means the count of documents per category is transformed for every relation). After the transform, new totals and opacity values are calculated for every relation. This way a user can quickly assess the difference between intercity relations for several relative measures.
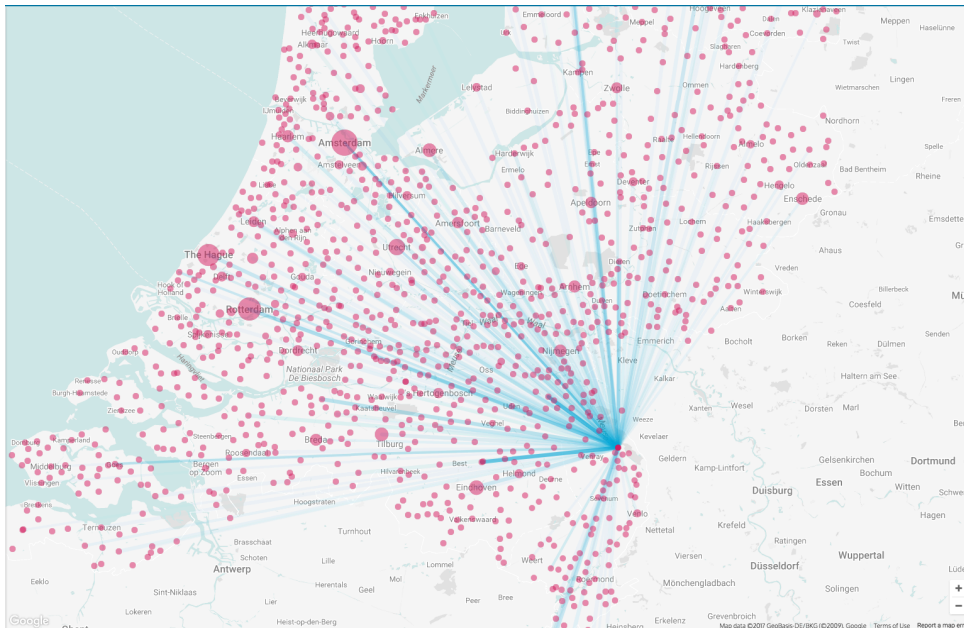

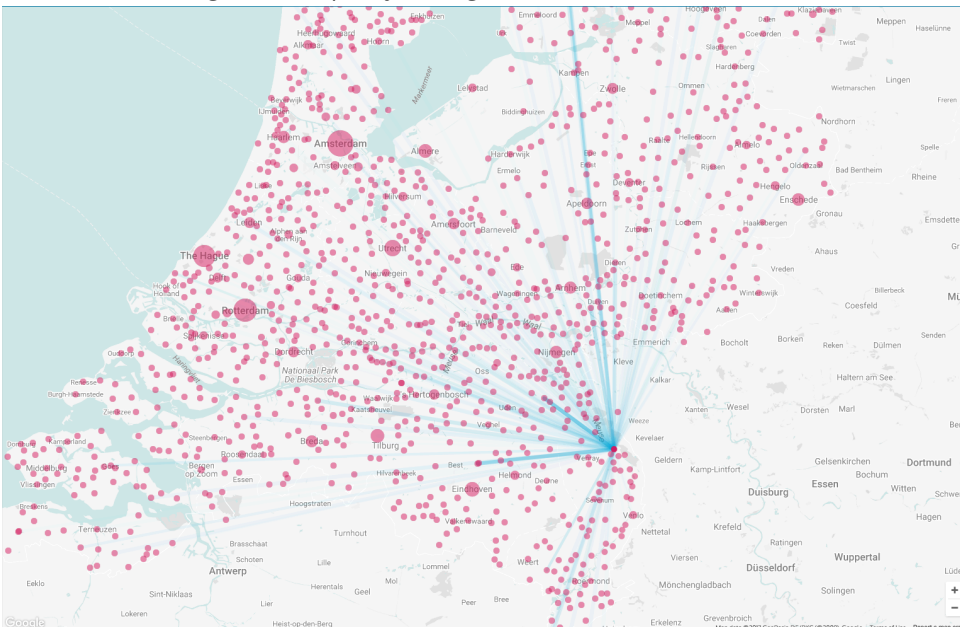Figure 1.7: Opacity scaling based on total occurrences

Figure 1.8: Opacity scaling based on total occurrences over minimum population size of a relation

   Using the sliders for each category a user can set a minimum threshold for a relation category strength. For example, if the value for the category "education" is set to 25, all relations that have a score lower than 25 on "education" are hidden. Finally, relation categories can be selected and deselected. This triggers an update of the total relation strengths, adding to or subtracting from, respectively, the total relation score the score of the toggled category.

Export

In the export card a user can download all the relations in the system in CSV format. If the user provides a threshold, relation scores are calculated on the fly, using only the data that has been classified with a minimum certainty of the given threshold. For example, a document that has a maximum probability of 0.45 (say, for "leisure"), will be counted in the "other" category if the provided threshold is larger than 0.45.

Figure 1.9: the export card

Relation Information

When clicking on a relation an extra window in the menu will be opened. This contains all information about the relation between the two cities. The strength of the relations (amount of documents found for the relation between those cities) will be displayed, as well as the strength of the relations per category (such as commuting, shopping, leisure, etc.).
To provide the user with the option to validate the results, a "Get Documents" button, at the bottom of this card, fetches all the documents that belong to this relation.

Figure 1.10: Relations scores between two cities

Relation Documents

In this card all documents, that are assigned to the selected relation, will be displayed. The documents are listed as shown figure 1.12. Each listed item shows the category with which it has been labelled using our classifier. It also shows the probability with which this document was assigned to this category.
If a user wants to inspect the contents of a certain file a list item can be clicked, which will initiate the download of the selected document.
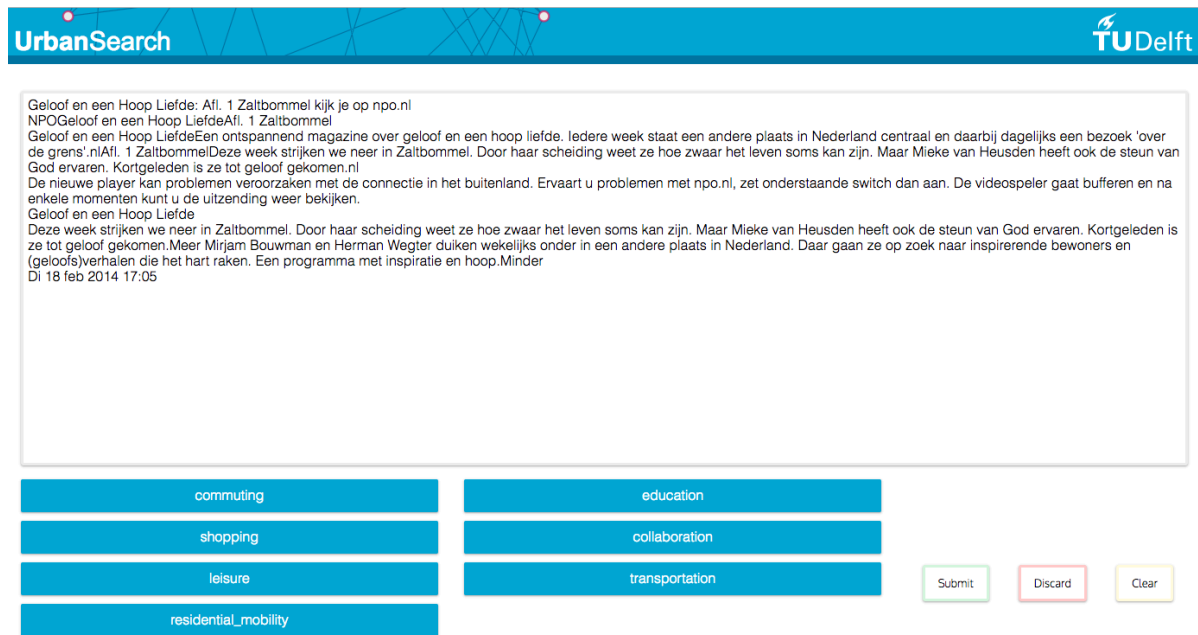
Extracted Documents for Relation
Amsterdam - Rotterdam

leisure	0.617
leisure	0.555
collaboration	0.464
leisure	0.523
education	0.76

Figure 1.11: Documents included in relations

1.3 Classification Interface

The classification interface is meant as an easy to use tool for extending the systems training-set. From our document set, which we have filtered from all the CommonCrawl pages, we fetch a random document. This document is displayed in the interface shown in figure 1.12. The user can select one or more categories to tag this document, after which the user can hit the submit button to save the document to our training set.
If a user deems the document to be of no use for any of the categories, the document can be discarded by clicking the "discard" button.
Finally if the user wants to add his or her own documents this can be done by hitting the "clear" button and copy-pasting the contents of the documents into the document field, selecting one or more matching categories and submitting the document.

Figure 1.12: The classification interface

1.4 Classifier Interface

Finally we have the classifier interface. This interface provides us with a way to train a new default classifier which is used by the system for the API and for classification.

A new classifier can be trained by clicking the "Train!" button. If all goes well, the "Train!" button will change to a "Done!" button, indicating the successful creation of a new default classifier for the system.

Although we can train a new default classifier, this classifier will not be used by the system until the complete system is restarted.

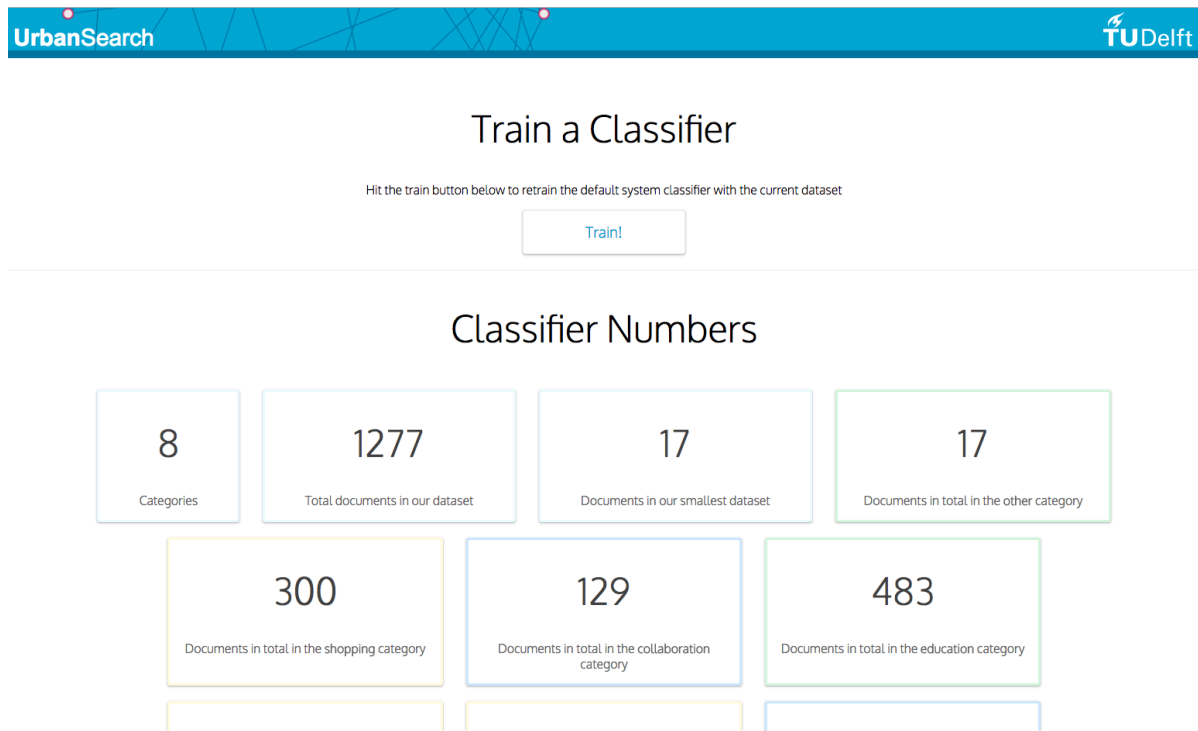The bottom part of the classifier page consists of tiles providing the user with information about the size of the training set.

Figure 1.13: The classifier interface