

Social Information Retrieval Experiment & Exercise

Fact Pack for Homework Assignment #4

Christoph Fuchs

2015-05-20

Source: <https://vmschlichter24.informatik.tu-muenchen.de/ha4/>

Homework Assignment #4

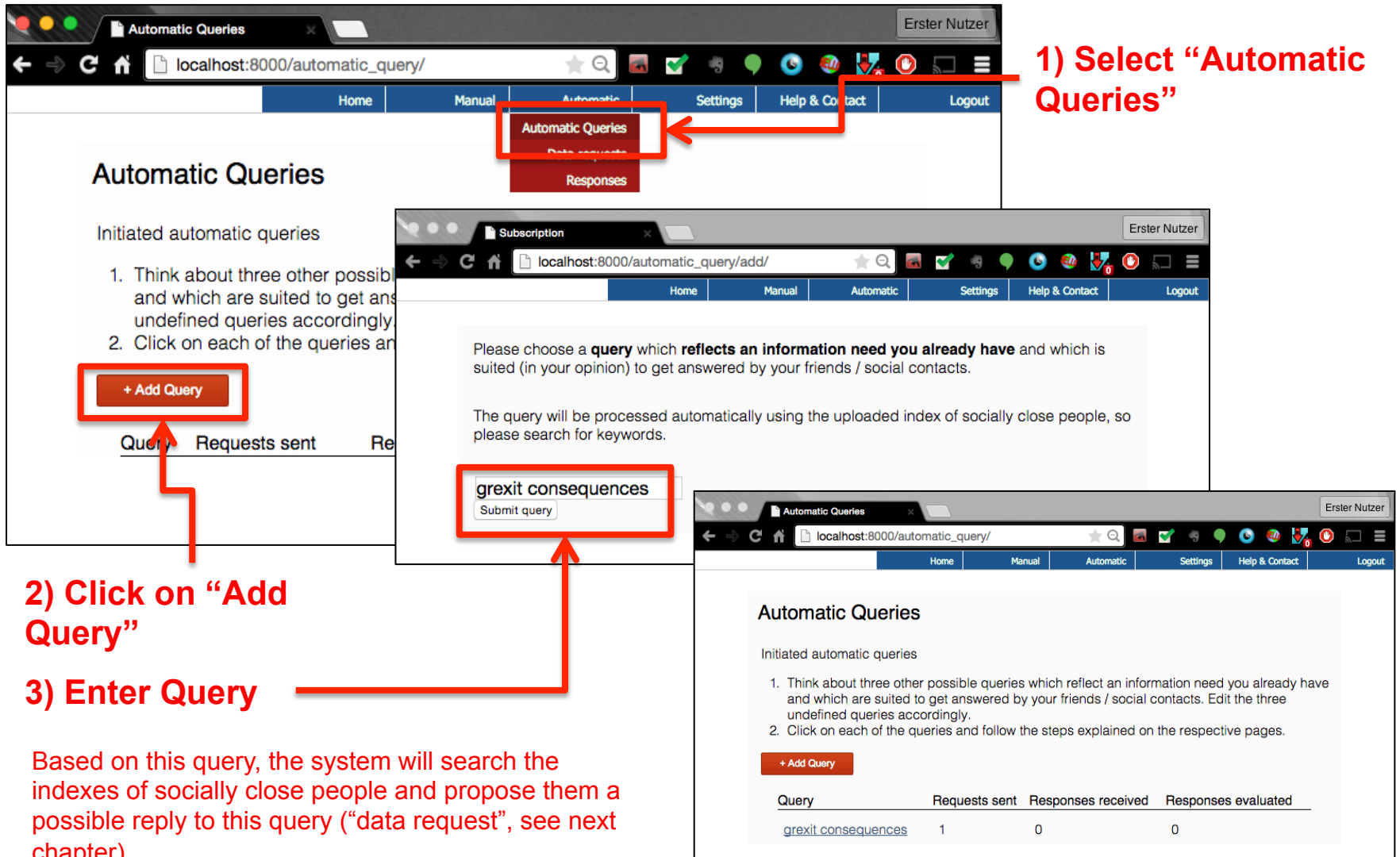
Covered area in this document

Due: **17.06. 23.59 CEST**

1. Start automatic query process - see [fact pack](#) for further details (add 6 queries to the automatic mode)
2. Maintain manual process (evaluate replies)
3. Analyze "our" anonymized social network
 - Create a Python script (using the [igraph](#) library) to
 - Identify number of components
 - Apply two different clustering algorithms
 - Calculate diameter, density, clustering coefficient of the network
 - Comment on the results in a report (~50-100 words): Which results are surprising, which not? Give reasons for your answers.
4. [Upload the Python script and your file with the comments](#)

Add Queries in Automatic Mode

1) Add a Query in Automatic Mode



1) Select “Automatic Queries”

2) Click on “Add Query”

3) Enter Query

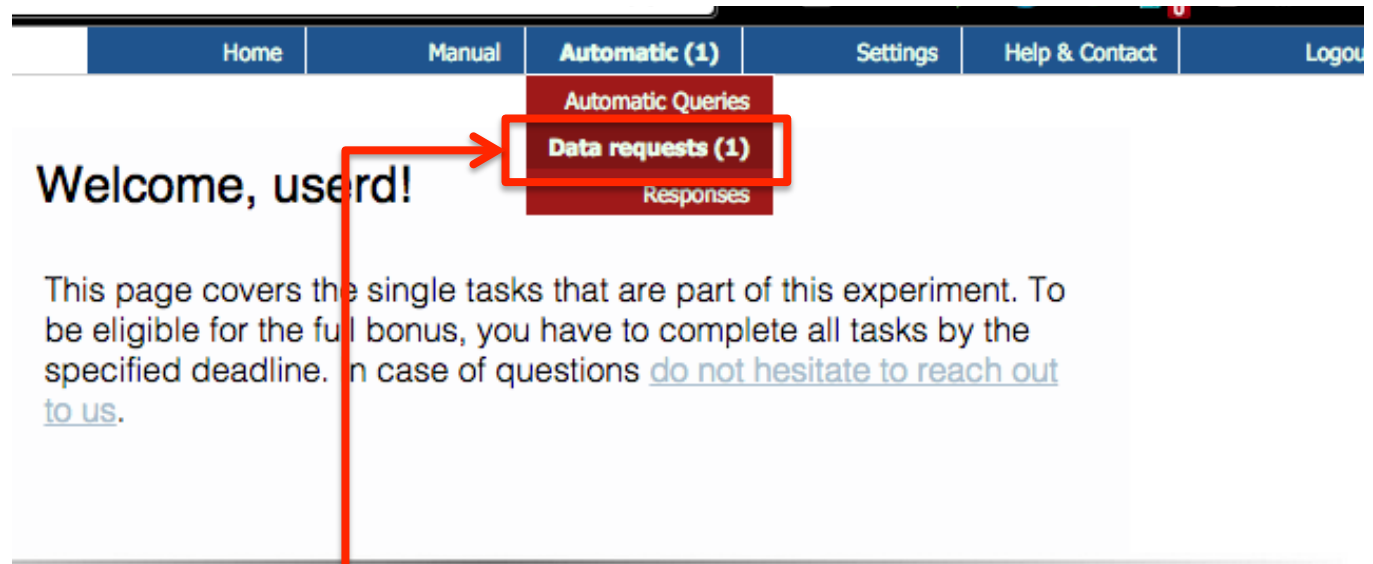
Based on this query, the system will search the indexes of socially close people and propose them a possible reply to this query (“data request”, see next chapter)

The final screenshot shows the query 'grexit consequences' has been added to the list of initiated automatic queries.

Query	Requests sent	Responses received	Responses evaluated
grexit consequences	1	0	0

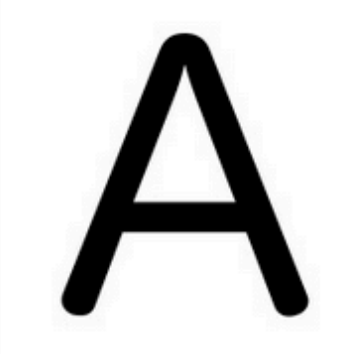
Reply to data requests in Automatic Mode

1) Visit the web system



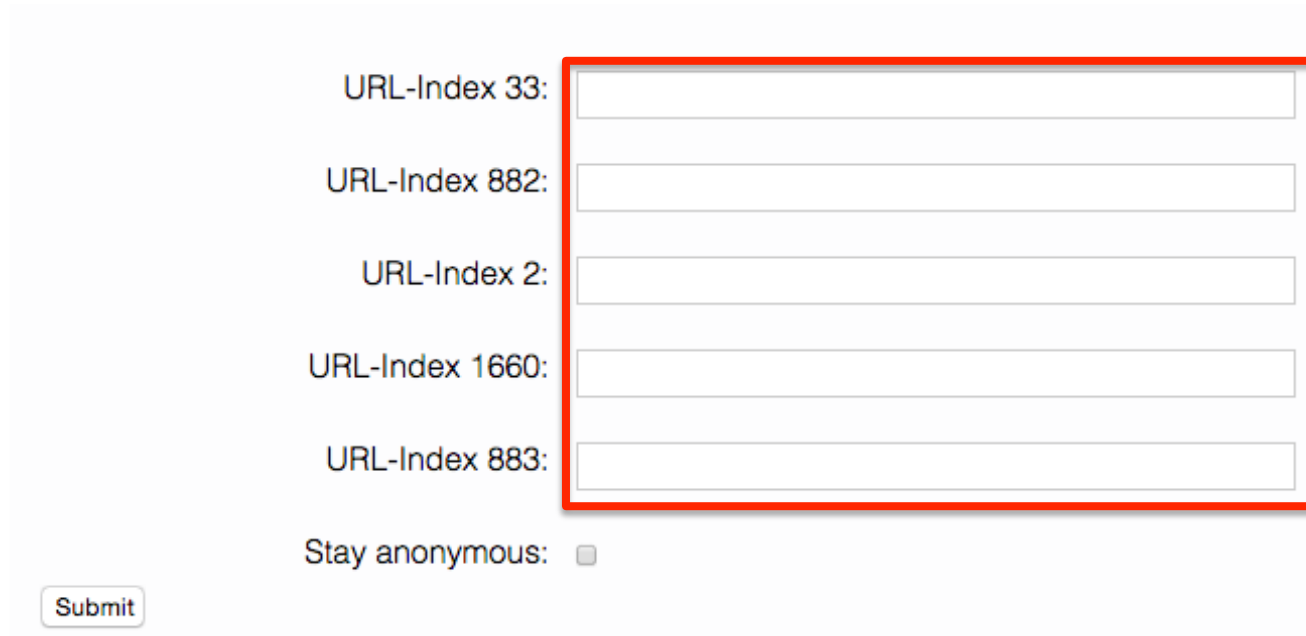
Select “Data requests” in the “Automatic” folder

2) Select the data request to reply to

			Home	Manual	Automatic (1)	Settings	Help & Contact
Data requests							
Open requests		Closed requests					
Information seeker	Query	Give a hand					
 usera	dokortitel kaufen 4 minutes ago	Reply to request					

Reply to request

3) Enter the URLs of your history (1/2)



URL-Index 33:

URL-Index 882:

URL-Index 2:

URL-Index 1660:

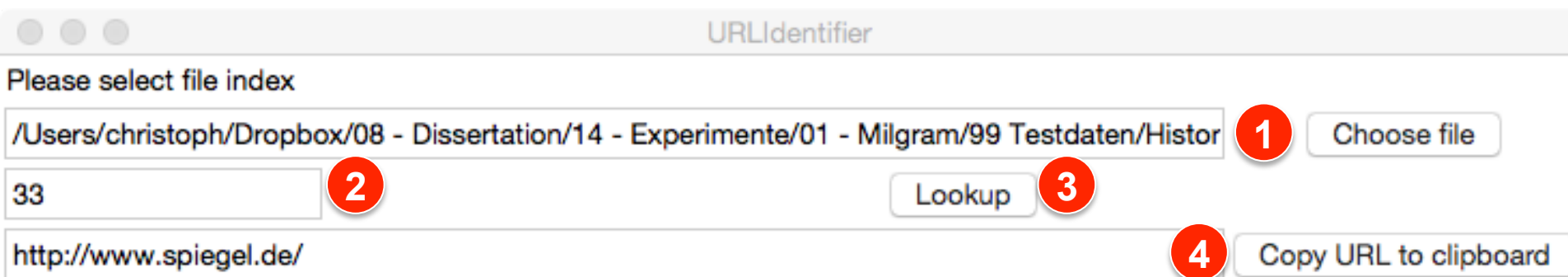
URL-Index 883:

Stay anonymous: ☐

**Fill those fields using
URLIdentifier (see
next slide)**

3) Enter the URLs of your history (2/2)

- Run the URLIdentifier tool from the software kit



The screenshot shows the URLIdentifier application window. It has a title bar with three window control buttons and the text 'URLIdentifier'. Below the title bar, it says 'Please select file index'. There is a text input field containing a long file path: '/Users/christoph/Dropbox/08 - Dissertation/14 - Experimente/01 - Milgram/99 Testdaten/Histor'. To the right of this field is a red circle with the number '1' and a button labeled 'Choose file'. Below the first input field is another input field containing the number '33'. To its right is a red circle with the number '2'. To the right of the '33' field is a button labeled 'Lookup' with a red circle containing the number '3' next to it. Below the '33' field is a third input field containing the URL 'http://www.spiegel.de/'. To its right is a red circle with the number '4' and a button labeled 'Copy URL to clipboard'.

- 1 Choose the URL_Mapping.pkl file generated while calculating your topic models (is in the same directory as the lda_model.model file)
- 2 Enter the requested index (taken from the web interface), click on “Lookup” (or press return)
- 3 Enter the requested index (taken from the web interface), click on “Lookup” (or press return)
- 4 Copy resulting URL to clipboard
 - Paste URL in respective form in web system (see previous slide)

Evaluate Replies to Automatic Queries

1) Select Response to assess



The screenshot shows the TUM system interface. At the top, there is a navigation bar with the following items: Home, Manual, Automatic (2), Settings, Help & Contact, and Logout. Below the navigation bar, there is a dropdown menu for 'Automatic (2)' which contains 'Automatic Queries', 'Data requests', and 'Responses (2)'. The 'Responses (2)' item is highlighted with a red box. To the right of the dropdown menu, there is a red arrow pointing to the 'Responses (2)' item, with the text '1) Select "Responses" from the "Automatic" menu' next to it. Below the dropdown menu, there is a section titled 'Responses to data requests'. This section has two tabs: 'Open responses' and 'Evaluated responses'. The 'Evaluated responses' tab is selected. Below the tabs, there is a table with the following columns: 'Information Provider', 'Your query', and 'View & evaluate response'. The table contains two rows. The first row shows a profile picture of a person named 'fridolin' with the query 'dokortitel kaufen 2 days, 8 hours ago' and a 'View & evaluate' link. The second row shows a profile picture of a person named 'userd' with the query 'dokortitel kaufen 2 hours ago' and a 'View & evaluate' link. A red arrow points to the 'View & evaluate' link in the first row, with the text '2) Select "View & evaluate" to see (and assess) the response' next to it.

Automatic (2)

- Automatic Queries
- Data requests
- Responses (2)**

Responses to data requests

Open responses Evaluated responses

Information Provider	Your query	View & evaluate response
 fridolin	dokortitel kaufen 2 days, 8 hours ago	View & evaluate
 userd	dokortitel kaufen 2 hours ago	View & evaluate

1) Select "Responses" from the "Automatic" menu

2) Select "View & evaluate" to see (and assess) the response


2) Fill Evaluation Form

dokortitel kaufen - evaluat


localhost:8000/automatic_query_responses/9/


Home Manual Automatic (2) Settings Help & Cont

dokortitel kaufen

 fridolin

fridolin replied to your request. Please **take some time** and **visit the given resources**. After reviewing them, please **fill out the respective forms thoroughly** (there is no chance to edit them later).

External urls are marked with an arrow  and will open in a new window/tab. All information you enter is kept confidential. The data will be processed in anonymized form, your data won't be linked to your account/name.

URL:  <http://www.dokortitel-kaufen.net/>

Is this link relevant to answer your query?	Not relevant at all	<input type="range"/>	Highly relevant
Did the website contain information you did not expect or that was not obvious; did the content surprise you?	Content did not surprise me	<input type="range"/>	Content was highly unexpected
Was the link you received personalized to you as a specific person? (e.g. did the	Not personalized at all	<input type="range"/>	Highly personalized

1) Review provided result

2) Answer questions (e.g. by using sliders)

PREPARATION – NOTES & REMARKS

(1/2)

We take Data Privacy seriously: We don't do anything without your permission & we are transparent on what we do:

- 1. We don't know the names of your Facebook friends:** We only ask you to upload the hashed version of the network file¹
- 2. None of your data** (evaluation of ties, names, products, etc.) **will ever be related to you**
- 3. We don't get to know your browser history** – the uploaded topic models only contain document IDs and word vectors

¹ The social network built from all individual ego-networks will get hashed again and randomly changed (edges and nodes changed/removed) before used in exercise to ensure that it is not possible to identify nodes

PREPARATION – NOTES & REMARKS

(2/2)

4. The data will only be used to **conduct sound scientific research** (and student exercises)
5. You will use a (**obfuscated**) version of the dataset to **solve homework assignments**
6. Your **name, matriculation number, email address, etc. is only used to grade the homework** – it **will be deleted afterwards**
7. We will **never give** the data to anyone
8. Publications based on this dataset will only contain **highly aggregated information**
9. We will **hash the Amazon products after the experiment** and delete all copies of the original dataset