1 Welcome

Welcome, and thank you for participating in the user-study on
Association Rules vs. Learning To Rank: Which Method provides the Most Helpful RDF Vocabulary Term Recommendations in Real-World Scenarios?

General Information:
The user-study is anonymous and is part of my Ph.D. thesis on developing an RDF vocabulary term recommendation service aiding Linked Data engineers in reusing existing vocabularies.

What is the user-study about?:
By participating in this user-study, you will help us to identify which method for recommending RDF vocabulary terms, i.e., "Association Rule Mining" or "Learning To Rank", is most helpful to data providers in reusing appropriate RDF vocabulary terms. I appreciate your participation very much, as you hereby support me in my Ph.D. thesis.

Participant Information:
Participation in this user-study is voluntary and you are free to withdraw or discontinue participation at any time without any penalties.

The user-study will take you approximately **30** minutes. The principle investigator will provide you the *Informed Consent* form. If you choose to participate in this user-study, please read the Informed Consent form carefully (take as much time as you need), sign it, and press on the "Continue" button below.

Procedure:
You will be given three practical assignments, in which you are asked to pick appropriate RDF vocabulary terms to describe classes and properties between classes. The principal investigator will provide you detailed instructions for each task separately. In addition to the tasks, you are asked various questions investigating your satisfaction considering the RDF vocabulary term recommendations.

Data Collection:
We solely collect the data from the tasks and from the questionnaire. We do not collect any personal data or other data that can be used to identify you or other participants. We do not record video or audio of you or the screen. Furthermore, we do not collect any digital information in the background, such as the IP address, the web browser, or your operating system.

The user-study is ***anonymous***, and the anonymous results will be published using the GESIS *datorium* system [link].

Thank you very much for your interest and support. If you have any questions or comments, do not hesitate to email me at:
johann.schaible@gesis.org

Click "Continue" to begin with the user-study

2 Introduction

Scenario

Suppose you have some structured data and you would like to model it as Linked Open Data (LOD). You intend to reuse RDF vocabulary terms for describing the RDF types and properties in the model. Luckily, a co-worker of yours has already started to model the data, but used the vocabulary terms **owl:Thing** for some classes and **rdfs:label** for some properties, as (s)he could not find better fitting terms. These terms however are very generic or even used incorrectly. Your assignment is to find better fitting RDF types and properties to describe these classes and properties semantically correct.

In the following, the principal investigator explains how to use Karma and the provided recommendations for reusing RDF vocabulary terms based on some exaple data. The three following tasks are similar to the illustrated example, but contain data from different domains. In one of the three upcoming tasks you are provided term recommendations based on the machine learning approach “Learning To Rank”. Another task comprises the use of recommendations based on the data mining approach “Association Rule Mining”. To control the overall benefit of the recommendation, in the third task, you will also have to choose appropriate RDF vocabulary terms without any recommendations.

To continue to the first task, please specify the arbitrary ID, such that we can map the results of the task to the answer of this survey, and press "Continue". The principal investigator will provide you the task, the recommendation method, and redirects the screen to the Karma tool.

Arbitrary ID

Your arbitrary ID is:

3 Modeling Task 1

Modeling Task 1:
The principal investigator now hands you your first task. Please, read it carefully, complete it, and then answer the following questions.

How long did it take you to complete the task?

- > 6 minutes
- < 6 minutes. Please, provide needed time in format "m:ss":

What was the domain of the data in the task?

The domain of the data is specified on the task description that was given by the principal investigator.

- Data on music and musicians
- Data on museum items
- Data on offers and products

Which recommendation method were the RDF vocabulary term recommendations based on?

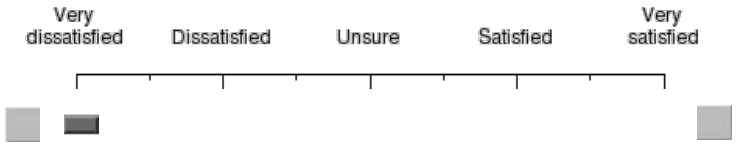
The recommendation method for the task is assigned to you by the principal investigator

- Association Rule Mining
- Learning To Rank
- None

How much did you feel supported by the tool in reusing appropriate RDF vocabulary terms?

Please provide your level of satisfaction on a 5-point Likert scale.

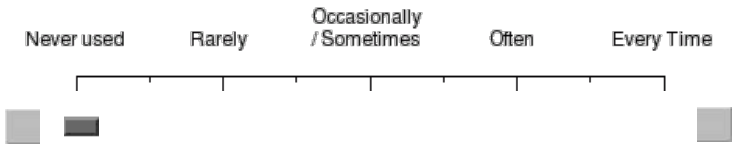
Level of satisfaction



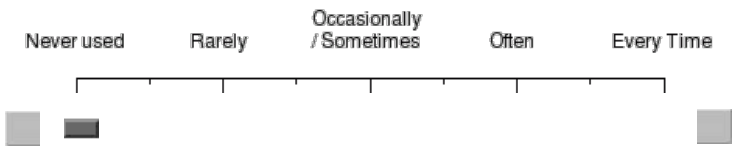
How often did you use other methods to find the appropriate vocabulary terms?

Please indicate the amount of use on a 5-point Likert scale.

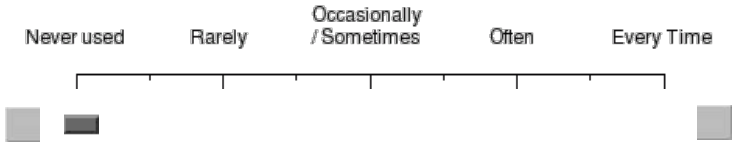
How often did you use the "search bar" to type in the desired vocabulary term?



How often did you click on "more" and browsed through all available vocabulary terms?



How often did you use external servicese outside Karma, e.g. Google?



If you used other external services, please state which services you have used.

Please name one service per line

4 Modeling Task 2

Modeling Task 2:
The principal investigator now hands you your second task. Please, read it carefully, complete it, and then answer the following questions.

How long did it take you to complete the task?

- > 6 minutes
- < 6 minutes. Please, provide needed time in format "m:ss":

What was the domain of the data in the task?

The domain of the data is specified on the task description that was given by the principal investigator.

- Data on music and musicians
- Data on museum items
- Data on offers and products

Which recommendation method were the RDF vocabulary term recommendations based on?

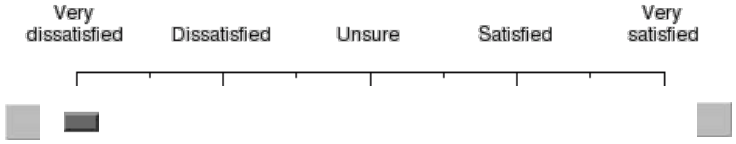
The recommendation method for the task is assigned to you by the principal investigator

- Association Rule Mining
- Learning To Rank
- None

How much did you feel supported by the tool in reusing appropriate RDF vocabulary terms?

Please provide your level of satisfaction on a 5-point Likert scale.

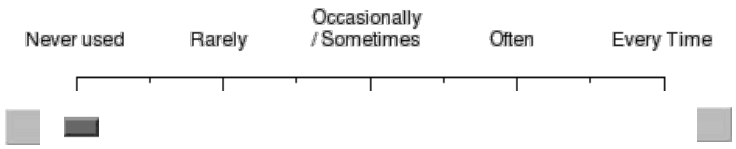
Level of satisfaction



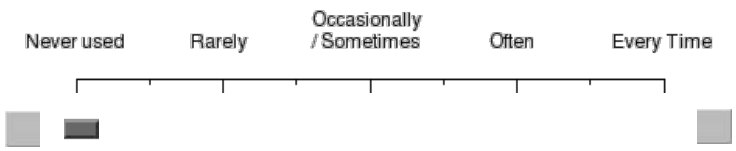
How often did you use other methods to find the appropriate vocabulary terms?

Please indicate the amount of use on a 5-point Likert scale.

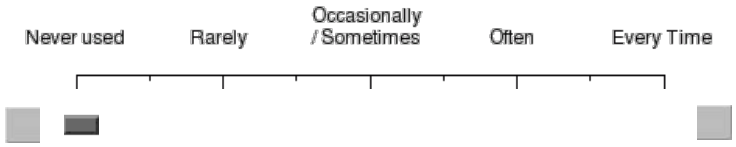
How often did you use the "search bar" to type in the desired vocabulary term?



How often did you click on "more" and browsed through all available vocabulary terms?



How often did you use external servicese outside Karma, e.g. Google?



If you used other external services, please state which services you have used.

Please name one service per line

5 Modeling Task 3

Modeling Task 3:
The principal investigator now hands you your third task. Please, read it carefully, complete it, and then answer the following questions.

How long did it take you to complete the task?

- ○ > 6 minutes
- ○ < 6 minutes. Please, provide needed time in format "m:ss":

What was the domain of the data in the task?

The domain of the data is specified on the task description that was given by the principal investigator.

- ○ Data on music and musicians
- ○ Data on museum items
- ○ Data on offers and products

Which recommendation method were the RDF vocabulary term recommendations based on?

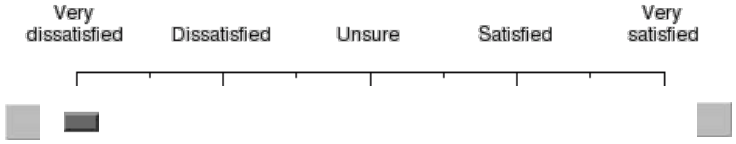
The recommendation method for the task is assigned to you by the principal investigator

- ○ Association Rule Mining
- ○ Learning To Rank
- ○ None

How much did you feel supported by the tool in reusing appropriate RDF vocabulary terms?

Please provide your level of satisfaction on a 5-point Likert scale.

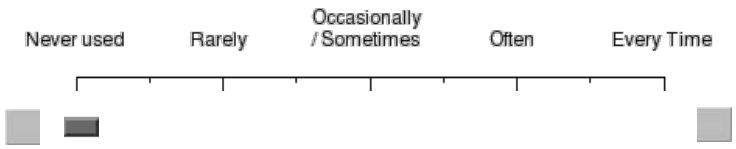
Level of satisfaction



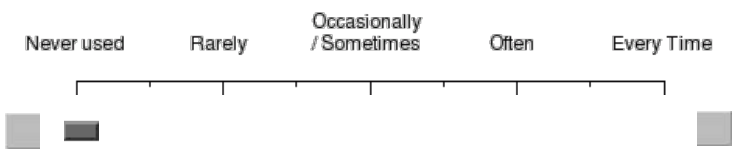
How often did you use other methods to find the appropriate vocabulary terms?

Please indicate the amount of use on a 5-point Likert scale.

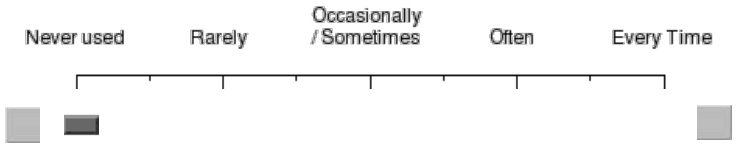
How often did you use the "search bar" to type in the desired vocabulary term?



How often did you click on "more" and browsed through all available vocabulary terms?



How often did you use external servicese outside Karma, e.g. Google?



If you used other external services, please state which services you have used.

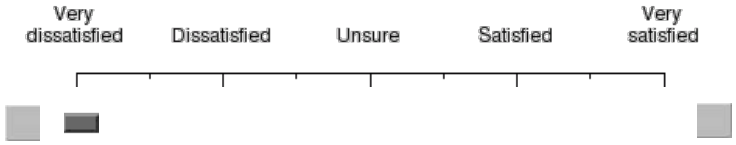
Please name one service per line

6 Comparison

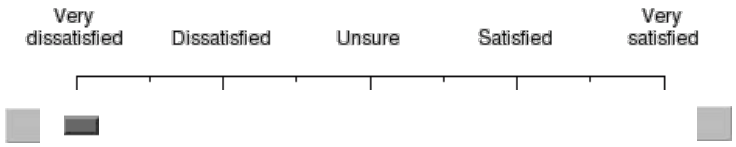
Overall, how satisfied were you with the support for reusing PRDF vocabulary terms provided by the two recommendation methods?

Please specify your level of satisfaction on a 5-point Likert scale

Satisfaction with recommendations based on "Learning To Rank"



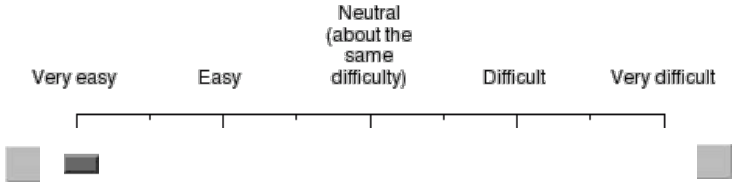
Satisfaction with recommendations based on "Association Rule Mining"



How difficult was it fo you to find appropriate RDF vocabulary terms without any recommendations?

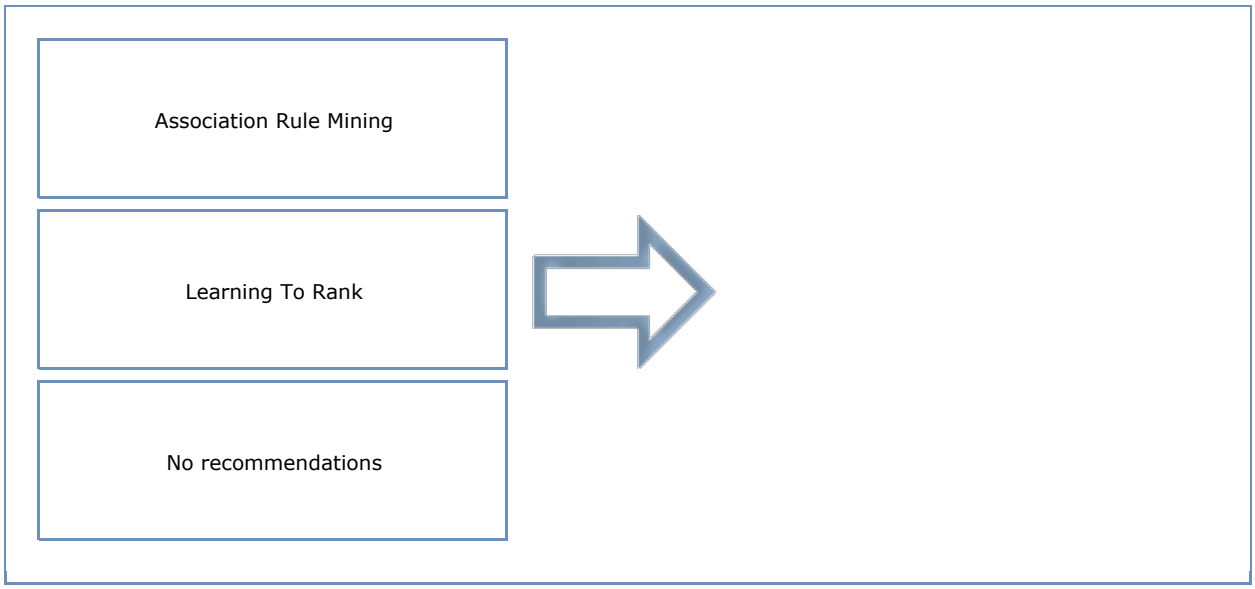
Please specify the level of difficulty on a 5-point Likert scale compared to the tasks comrising RDF vocabulary term recommendatios

Level of difficulty



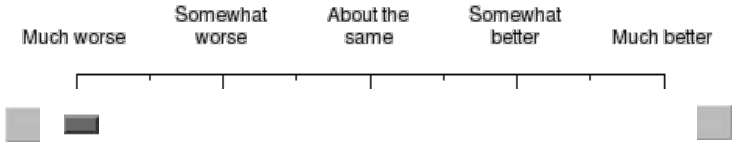
How would you rank the provided RDF vocabulary term recommendations?

Please rank the recommendation approaches "Association Rule Mining", "Learning to Rank", as well as the "no recommendations"-option from good to bad in descending order, i.e., the best approach should be at the top. (To do so, drag and drop the items to the right side)



How did you encounter the recommendations based on "Association Rule Mining" compared to the recommendations based on "Learning to Rank"?

Please provide a level of comparison



Do you have any other comments considering the recommendations, the recommendation services, the Karma tool, or the tasks themselves?

Please provide any feedback and comments (positive and negative)

7 Perosnal Information

Thank you very much for participating in the user-study :-)

Last but not least, please provide some (demographic) information
(and, if you like, also some feedback on the user-study).

Please specify your gender

- ○ Male
- ○ Female

Please specify your age

18 - 25
26 - 30
31 - 35
36 - 40
41 - 45
56 - 60
46 - 50
51 - 55
61 - 65
> 65

Please specify your highest academic degree

- ○ Bachelor's degree
- ○ Master's degree (Dimploma)
- ○ Doctoral degree (Ph.D.)
- ○ Other:

Please specify your current proffession (position at work)

- ○ Master's Student (Research Assistant)
- ○ Programmer
- ○ Ph. D. Student (Graduated Research Associate)
- ○ Postdoctoral Research Associate
- ○ Project Leader
- ○ Proffessor
- ○ Other:

Do you work in academia or industry?

☐ Academia ☐ Industry

Please specify the domain of the data you typically work with

If you work with data from several different domains, please specify one domain per line

How many years have you worked in the field of or with Linked Open Data (LOD)

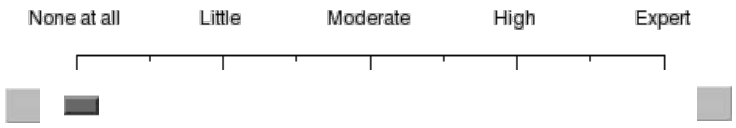
- not at all
- less than 1 year
- 1-2 years
- 3-4 years
- 5-6 years
- > 6 years

Do you consider yourself more a "LOD consumer" or more a "LOD publisher"?

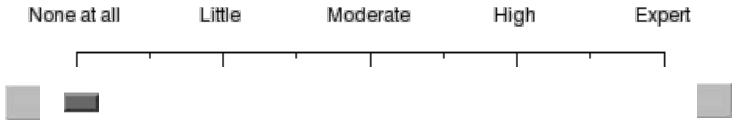
- LOD consumer
- LOD publisher
- Evenly spread
- None

Please specify the degree of your knowledge / experience in the following areas

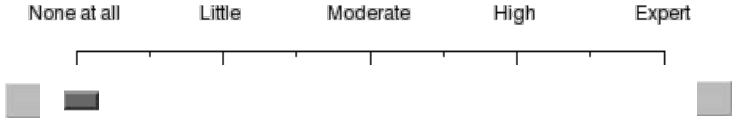
Experience in consuming Linked Open Data



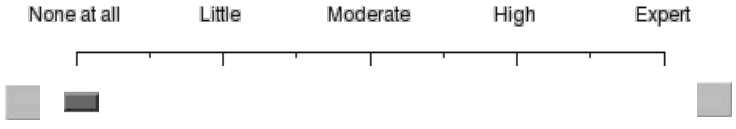
Experience in publishing Linked Open Data



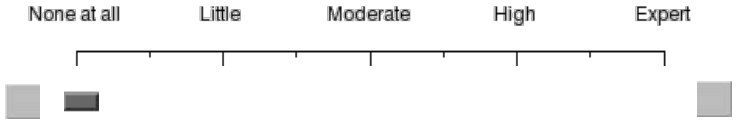
Experience in working with Karma



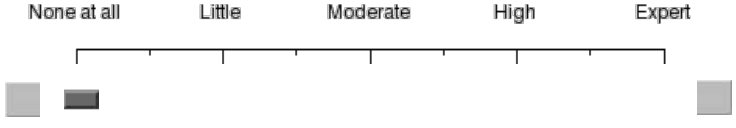
Knowledge about data from the domain "Music and musicians"



Knowledge about data from the domain "Museum items"



Knowledge about data from the domain "Offers And Products"



Do you have any comments or feedback about the user-study?

What did you miss in this study? What did you like in this study? What did you not like?

8 End

Thank you very much for participating in this user-study and for helping me with my Ph.D. thesis. I appreciate it very much.

You may close the window now.

Best Regards,
Johann Wanja Schaible
johann.schaible@gesis.org