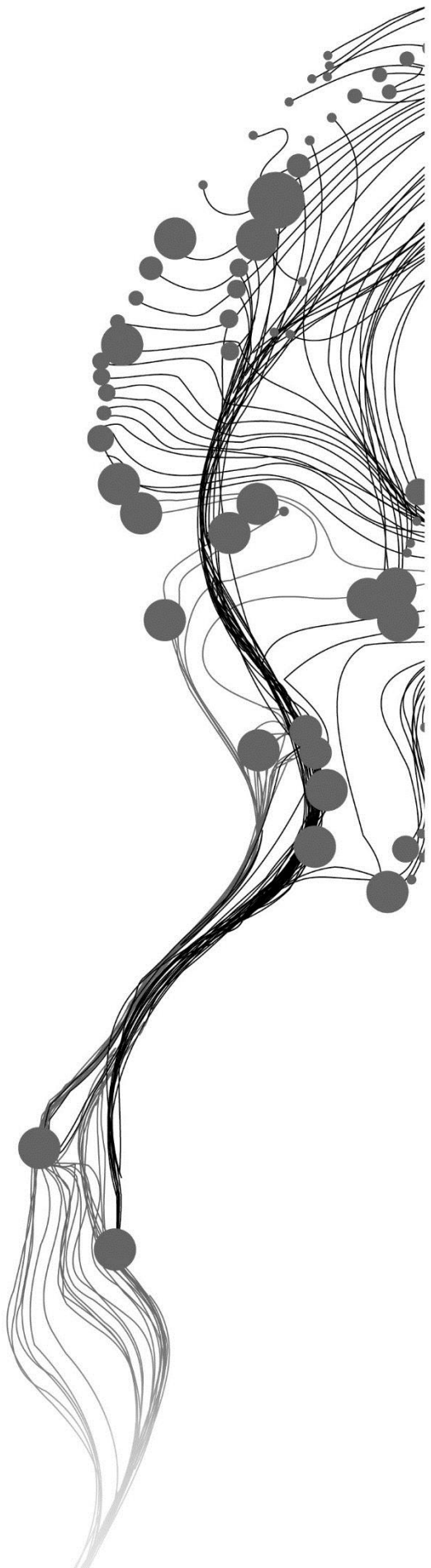


VISUALIZATION OF LINKED GEODATA ON THE WEB

JELENA MARJANOVIĆ
February, 2015

SUPERVISORS:
dr.ir. R.L.G. Lemmens
dr. C.P.J.M. van Elzakker



VISUALIZATION OF LINKED GEODATA ON THE WEB

JELENA MARJANOVIĆ

Enschede, The Netherlands, February 2015

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Geoinformatics

SUPERVISORS:

dr.ir. R.L.G. Lemmens

dr. C.P.J.M. van Elzakker

THESIS ASSESSMENT BOARD:

Chair: prof.dr. M.J. Kraak

External examiner: Ms dr. K. Ooms; University Gent, Department of Geography

Supervisor: dr.ir. R.L.G. Lemmens

Second supervisor: dr. C.P.J.M. van Elzakker

Advisor: drs. B.J. Köbben

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

Everyday, Humans are, directly or indirectly, creating vast amount of data. New methods and ideas are appearing on how to manage these data because traditional databases and other methods are no longer satisfactory. Linking the raw data on the internet is one of proposed ideas. This concept is being implemented, and the Linked Open Data (LOD) cloud is growing every day. In the LOD cloud, data with a geo component already exist, but has a need for improvement and more information is required to be added in order to improve geo-information integration and add various possibilities on how geographical data can be used.

This research serves as a contribution to geo-information integration into the cloud by defining features necessary for a good representation of linked. The notion was to achieve this through User Centered Design approach. After the research problem was defined, and existing visualisation solutions were detected and commented, three systems were chosen to be tested in order to extract the most important requirements which have to be satisfied when visualising linked data. By giving users carefully designed tasks which will force them to navigate through the graphs of existing visualisations in addition with an eye tracking machine, and video and audio recordings, the main guidelines were noted. This was done through the list of issues created before the requirements itself where main issues expected to be detected are noted. After the answers from test persons were collected and summarised, the recommendations for visualisation 1 and visualisation 2 in this research were created. JavaScript library helped in showing all the visual features detected to be needed through requirements testing. In addition, using JavaScript library increases interactivity by adding dynamic options which in return increases usability. Finally, the visualisations were commented upon and the main conclusions were drawn. The visualisations were created with varying features to allow many types of features to be evaluated with requirement testing. In this thesis, concepts on the integration of geodata in linked data and the potential benefits of adding a spatial data into the LOD cloud is discussed. This research will lead to more successful visualisations in future which will encourage more users to use LOD cloud and more scientist to improve its quality and size.

ACKNOWLEDGEMENTS

Type the text here. I would like to thank to my family who supported me in my decision to pursue higher education abroad. Thanks to my friends from all over the world and special thanks to my supervisors for their guidance and advice through the last six months. Moreover, I would like to thank all of the professors and ITC staff from who helped me in extending my knowledge.

TABLE OF CONTENTS

1.	Introduction.....	1
1.1.	Background.....	1
1.2.	Problem definition	2
1.3.	Motivation for the research	2
1.4.	Research objectives	3
1.5.	Research questions	3
2.	Linked geodata visualization systems and their users	4
2.1.	Introduction on the linked data	4
2.2.	Visualization of linked geo data	5
2.3.	Recommendations for visualisation of linked data.....	5
2.4.	Existing visualizations of linked geo data	6
2.5.	Existing evaluations of linked data visualisations	11
3.	User-Centred design and Requirement Analysis	13
3.1.	Usability.....	13
3.2.	Users	13
3.3.	Requirement Analysis	14
3.4.	Results and analysis of the of the experiment	21
4.	Conceptual design and prototype implementation	31
4.1.	Recommendations.....	31
4.2.	Describing dataset	31
4.3.	Use case.....	32
4.4.	Answers to questions from requirements analysis.....	34
4.5.	Code development and issues	36
4.6.	Why d3	37
4.7.	Design of the visualisation environment.....	37
4.8.	Conclusions	41
5.	Conclusions and recommendations.....	42
5.1.	Overview.....	42
5.2.	Discussion.....	42
5.3.	Comparison to other work	43
5.4.	Answering the research questions	44
	Bibliography.....	46
	Appendix 1: Requirements analysis experiment pre selection questionnaire.....	49
	Appendix 2: Requirements analysis experiment literal instructions	51
	Appendix 3: Requirements analysis experiment test questions.....	57
	Appendix 4: Requirements analysis experiment end questionnaire.....	58
	Appendix 5: Invitation email for requirements analysis for users	59
	Appendix 6: Example use of Python rdflib.....	60
	Appendix 7: Code for visualisation 1	66
	Appendix 8: Code for visualisation 2	75

LIST OF FIGURES

Figure 1 - Presentation of linked geodata with RDF Gravity	2
Figure 2 - Example of a triple	4
Figure 3 - Representation of RDF Gravity interface.....	7
Figure 4 - Representation of Welkin interface.....	8
Figure 5 - Representation of Lodlive interface.....	9
Figure 6 - Representation of Gephi interface	10
Figure 7 - Representation of Gruff interface	10
Figure 8 - Hardware setup in ITC lab.....	14
Figure 9 - Steps that users were following	16
Figure 10 – Initial interface look from Lodlive	18
Figure 11 – Initial interface look from Visuwords.....	18
Figure 12 – Initial interface look from RDF Gravity.....	19
Figure 13 - Organisation of codes in Atlas.ti.....	22
Figure 14 - Process of adding codes in Atlas.ti software.....	23
Figure 15 - Difference between two different types of relationships in Visuwords.....	26
Figure 16 - “heat map” during the navigation of one of the users with graph from Visuwords	27
Figure 17 – Representation of the first proposed visualisation	38
Figure 18 – Representation of what is show in the visualisation after the resource “brand” is double clicked.....	39
Figure 19 – Proposed interface for Visualisation environment 2.....	40
Figure 20 – The possibility in Visualisation 2 to fic the resource to a certain position.....	41

LIST OF TABLES

Table 1 - Background of the users.....	20
Table 2 - Predetermined issues.....	23
Table 3 - Features that should be added to future visualisation according to end questionnaire	24
Table 4 - Features that are characterized as positive according to end questionnaire.....	25

1. INTRODUCTION

1.1. Background

According to IBM (2012) “90% of the data in the world today has been created in the last two years alone”. This illustrates how data is ever emerging and becoming more pervasive. One way of tackling big data is the interlinking of data. Nowadays, the term linked data is often used, referring to a web-based network of datasets. It uses links between the data in order to obtain meaningful information and information of different origin. Linked data is based on RDF (Resource Description Framework) (RDF Working Group, 2014) which attaches meaning to content. RDF represents data as a graph and once extracted it can be used in multiple formats, allowing the users to be unattached to the format of the data. The query language SPARQL was designed to query data conforming to RDF (DuCharme, 2013). With SPARQL, it is possible to receive insight into linked geodata and to select an individual part of the cloud of data, as well as to update or create new data. This technology serves as starting point for the creation of the Semantic Web which will be able to extract meaningful data. The initiatives of big enterprises such as Google’s Knowledge graph (Google, 2014) and Facebook’s Graph search (Facebook, 2014) as well as the Wolfram Alfa search engine are already using semantic search technology in order to improve their search.

Lack of technical knowledge and understanding of complex semantic relationships restrains people from exploring and using data (Dadzie & Rowe, 2011). On the other hand, if people can have a good understanding of the relationships in data sets, they will be able to conduct an analysis and draw new conclusions from it. This refers to both professionals and non-technical users. For this reason, a user-friendly visualization solution is needed. Visualization leverages the perceptual abilities of the user to explore the data with a network structure (Bastian, Heymann, & Jacomy, 2009). However, visualizing, navigation and finding patterns in networks is inherently difficult. Nevertheless, network representation is still extremely attractive. “How the elements are put together is just as important as the elements themselves” (Perer & Shneiderman, 2006). This is another indicator of how important proper visualization is and the role of visualization in understanding relationships in the data for bringing more coherent and prompt decisions. Visualizations are useful for improving human abilities to seek for features in network structure (Bastian et al., 2009). Plaisant (2004) remarked that improving methods of evaluation is necessary to encourage a higher amount of users to adopt ontologies. Dadzie & Rowe (2011) concluded that linked data is a significant research challenge where visualization and browsing linked data play an essential role.

Even though it has many benefits, network visualization can be a chaotic, especially when the network is large, and additional problems present overlapping links and unreadable nodes (Perer & Shneiderman, 2006). Additional issues existing in visualization tools are usually created just for one particular group of users, as in Cytoscape, a tool designed for medical purposes (Cytoscape Consortium, 2014). In addition, existing tools are not user-friendly and not straightforward in installation and very often demand a very high level of understanding of RDF. Limited research has been completed in the area of visualization of linked geodata despite the extensive need for the adjustment of this visualization for different users (Stadler, Lehmann, Höffner, & Auer, 2011). In addition, research on linked geodata is even less available; however, linked data with spatial features will bring new insight into the exploration of data. Background knowledge

On the other hand, more sophisticated visualizations of concept networks are built into their applications and can only be used with their specific data. Moreover, there is a lack of usability tests on the existing tools. In this project, a visualization mechanism will be designed that will overcome these shortcomings, by making its components accessible over the web and by performing end-user tests within a usability lab with an eye tracking device.

1.4. Research objectives

The main objective of this research is to create an optimal web-based visualization environment in which linked geodata can be visualized. Such visualization environment intends to help producers/users of VGI (Volunteer Geographic Information) in understanding the semantics behind VGI applications that implement linked data, by supporting it with concept visualization. Understanding the semantics in a particular application will improve and broaden the usage of linked data based applications.

Sub-objectives:

1. To define the requirements of the visualization environment.
2. To design and produce a prototype linked geodata visualization environment.
3. To formulate recommendations for the visualization of linked geodata.

1.5. Research questions

Question related with the sub objective 1:

1. What are the needs of a user?
2. What questions do users need answers to?

Questions related with the sub objective 2:

1. What functionalities should a visualization environment have?
2. How are users finding concepts and how are they navigating through the visualization?

Questions related with the sub objective 3:

1. Which visualization variables of concepts, instances, and relationships, as well as which combination of visualization variables makes the visualization understandable?
2. Depending on different use cases, what values are appropriate for these variables?

2. LINKED GEODATA VISUALIZATION SYSTEMS AND THEIR USERS

The previous chapter discussed the obstacles ongoing in this research and provides explanation of the aim of this research. The main objectives are outlined and the research questions related to them. This chapter will dive into the existing research regarding existing evaluations of linked data as well as the existing visualisations of linked data and some recommendations regarding the visualisation of this type of data. A few visualisation systems will be briefly explained. These systems were reviewed in order to decide on the solution to be used in the requirements extraction section, which is considered in Chapter 3. Moreover, by examining existing solution, the main shortcoming can be detected.

2.1. Introduction on the linked data

Before we start it is important to outline the main idea of the linked data and its representation. RDF is a format of linked data. It is a data model where the basic unit is known as “triple” with three parts: subject, predicate and object. For example, a subject might be “girl”, a predicate might be “name” (like name of the girl) and object might be “Anna” (so the actual name of the girl). A subject girl might also have the predicate “age” and then we might have the object “13”. This example is represented in Figure 2. Since RDF is a format for presenting the information on the internet and in a machine readable format, subject, object and predicate will be URLs, so instead of a girl we might have something like: www.example.com/girl and instead of Anna: www.example.com/Anna

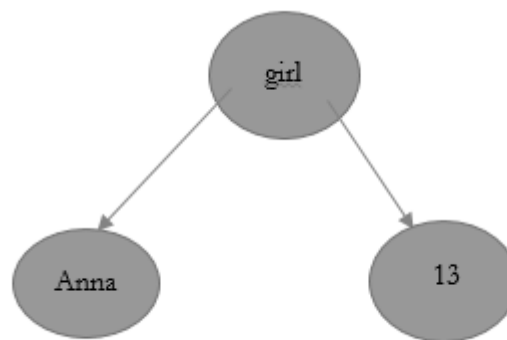


Figure 2 - Example of a triple

2.2. Visualization of linked geo data

As stated previously, one objective of this thesis is to contribute to the expansion of areas in which linked data is used and to increase the number of users. This is to be accomplished by setting rules for this type of visualization in the prospect that it will lead to the evolution of linked data. According to Card K., Jock, & Shneiderman (1999), visualization is "the use of computer-supported, interactive visual representations of data to amplify cognition". The visualization facilitates the memorization of information as well as decreasing the time for information searching. Due to the importance of creating linked data technology closer to the general users, there is a number of research efforts that attempts to develop algorithms and tool chaining. This would allow general users to visualize and utilize data without an in-depth knowledge of the applications (Sabot et al., 2014). Affective graphs evaluated interfaces of the existing graph and linked data visualisation and came to the conclusion that overall they scored very low to medium. This score refers to aesthetic properties that also contribute to the overall understanding of the visualisation. This is also an indicator of a need for development of this kind of visualisations.

Challenges pointed out by Dadzie & Rowe (2011) include the implementation of a universal method of presenting linked data to various types of users. The goal of allowing them to obtain an overview of the content and make decisions based on knowledge extracted, as well as to discover new suitable information by relating one piece of information to another one. Finding adequate visualisation that enables users to obtain valuable information for decision-making is a significant challenge in linked data. Obtainable linked data means support for the web of data by updating existing links and adding new ones. With visualisation this could be achieved, serving as a way for the presentation of abstract data that aids the recognition of patterns (Dadzie & Rowe, 2011).

2.3. Recommendations for visualisation of linked data

Even though more can be achieved if a set of specific visualization tools is designed for a specific purpose, common features can still be determined for the visualization of large amounts of data. Dadzie & Rowe (2011) wrote about high-level requirements for consuming linked data:

- Intuitive navigation through linked data structures
- Data exploration to obtain an understanding of its structure and content
- Data exploration, to identify errors, noise and other anomalies in context and syntax
- Publication/syndication and the attendant requirements for error correction, verification and validation of new data and links to existing data
- Data extraction for reuse in other application

Additionally, Dadzie & Rowe (2011) expressed some general requirements for lay or general users:

- Intuitive navigation through the large amounts of complex, multi-dimensional data
- Exploratory knowledge discovery
- Publication/syndication
- Data extraction
- Presentation of the results of the analysis of different audiences

When visualizing linked data which is almost always represented as a graph, structure rules of “Gestalt theory” can be applied. According to (Fekete, Wijk, Stasko, & North, 2008):

- Things that are close together are perceptually grouped together;
- Similar elements tend to be grouped together;
- Visual elements that are smoothly connected or continuous tend to be grouped;
- Two symmetrically arranged visual elements are more likely to be perceived as a whole;
- A closed contour tends to be seen as an object;
- Smaller components of a pattern tend to be perceived as objects whereas large ones as a background.

Mazumdar et al. (2013) base their research on another research of visualisation of UML diagrams, and they consider it to be applicable to LD application also. They listed features they considered important. They consider it important to separate hierarchical and non-hierarchical relations so that the hierarchy can be differentiated. Also, they consider that child nodes should be positioned at a median position with regards to a parent node. Additionally, semantically similar nodes should be clustered together. They also mention that it is important to avoid overlapping on edges whenever possible. Furthermore, comment of the nodes should be located close to the connected nodes, and additional specification in the graph should be standardized according to the group they belong to. In addition, special attention must be drawn to ratios, compact drawing, and symmetry.

Mazumdar et al. (2013) intended to improve the visualisation by improving the aesthetic properties of the graph. They claim that by improving aesthetic properties the understanding of the graphs by the user is higher. This is due to higher interest of users, if they are visually attracted with the tool they will try to execute a certain task multiple times even if not successful right away, furthermore users tend to memorize experience when they visually enjoy the interface. In their study, the users admitted that once they try out new ways of graphical representations with graphs it is easy to adjust and users would rather choose this kind of representation because of pleasurable and enjoyable interface. In this study authors followed four mayor requirements. They considered it important to highlight context, relation and statistics. Furthermore, they considered it important to show "data on demand", therefore not to show all the data at once but with the help of statistics to show more pleasant interface. Moreover, by testing the users, they discovered the need for easy access to data. Last requirement they followed was about providing the users with the possibility to interact and pose their queries. In addition, in this research some of the requirements for statistical information representation and visualisation of multivariate data is outlined. They also mention obstacle of ever growing data.

2.4. Existing visualizations of linked geo data

There are some existing visualisations designed for the visualization of graph data. These visualizations have many shortcomings that will be outlined in this research, and many are too complicated for lay users, both in using the software but in installing it also.

Furthermore, this ongoing research is mainly focused on the visualisation of RDF data, therefore narrowing the choice of this type of software. Moreover, many visualization tools are not focused on web representation and are desktop solutions. These types of solutions are not embracing the concept of openness and interoperability.

with either blue or red rectangle but there is no straightforward way to discover the difference between the two. In Figure 2, the interface of Welkin is represented.

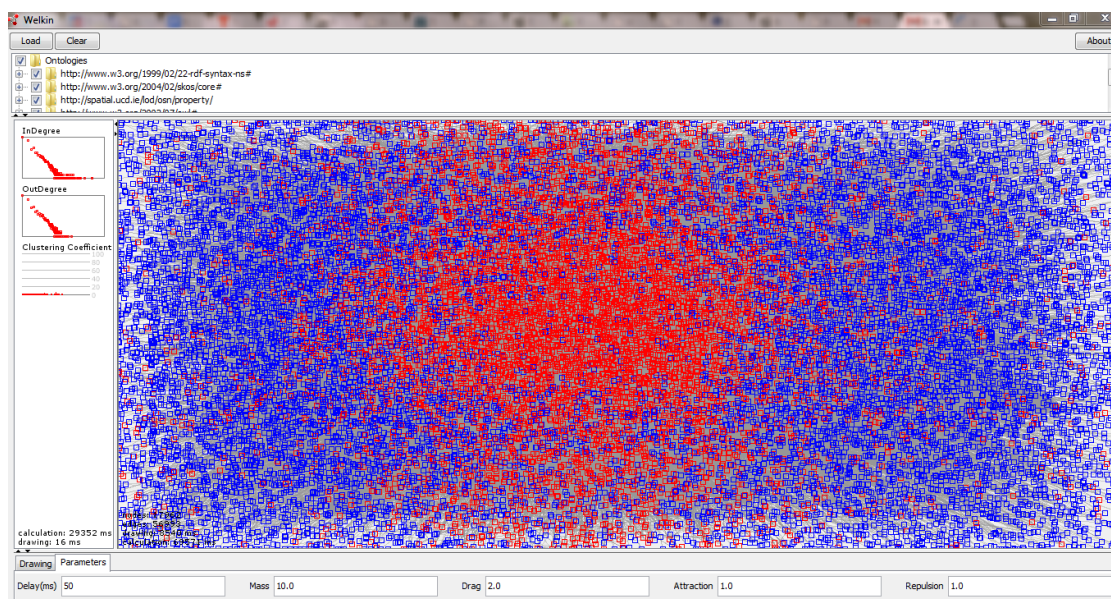


Figure 4 - Representation of Welkin interface

Another visualisation solution is Lodlive. Lodlive is mentioned as a project still in its developmental stage. The program used in the creation of Lodlive is JavaScript. It is easily available to users. It is a web solution, and it works with all up-dated browsers. Moreover, especially in comparison with other RDF visualisation systems, it is user-friendly. The user interface of Lodlive is depicted in Figure 5. In the beginning of the interaction with the Lodlive, users have to decide on the dataset, the visualisation starts with one circle on the interface surrounded with a number of smaller circles. Users have a possibility to expand this by clicking on one of the smaller circles and see with which other resource this resource is connected to. On the right top corner of the big circle, a number of small circles can be found. Here it is possible to interact further with the visualisation. So there is an option to get more information about the particular resource, to make a resource a central point, and to access the online resource in the new tab, to expand related resources or to remove specific resources. Lodlive starts showing resources by first just showing the one the user chose. After that, the user has the option to see which other concept is this concept related to using an option of showing all related resources or clicking on the smaller circles one by one and in that way revealing connected concepts. When a user in the first step shows concepts, automatically most of the concepts will be shown clearly, without significant overlap. However, if a user, on top of that, wishes to open another concept, the visualisation becomes cluttered. The example is shown in Figure 5 where the first concept was France, after that Paris (which was connected to France) was expanded and after what concepts and relationships became overlapped. It is possible to move resources quickly by just clicking the mouse. Different options such as the one to visit an online resource are easily founded. As a way to prevent the cluttering of the graph, sometimes a number of direct relations are grouped together and just after clicking them they will explode into more small circles, this happens when the name of the relationship is same. There is also a button called "additional information" which is located in the resource circle. A panel will be shown with information about the resource. On the left side of the interface, there is an option to see a map

of the resources, and where they are located. Next to this, there is an option of showing the images connected to a certain resource.

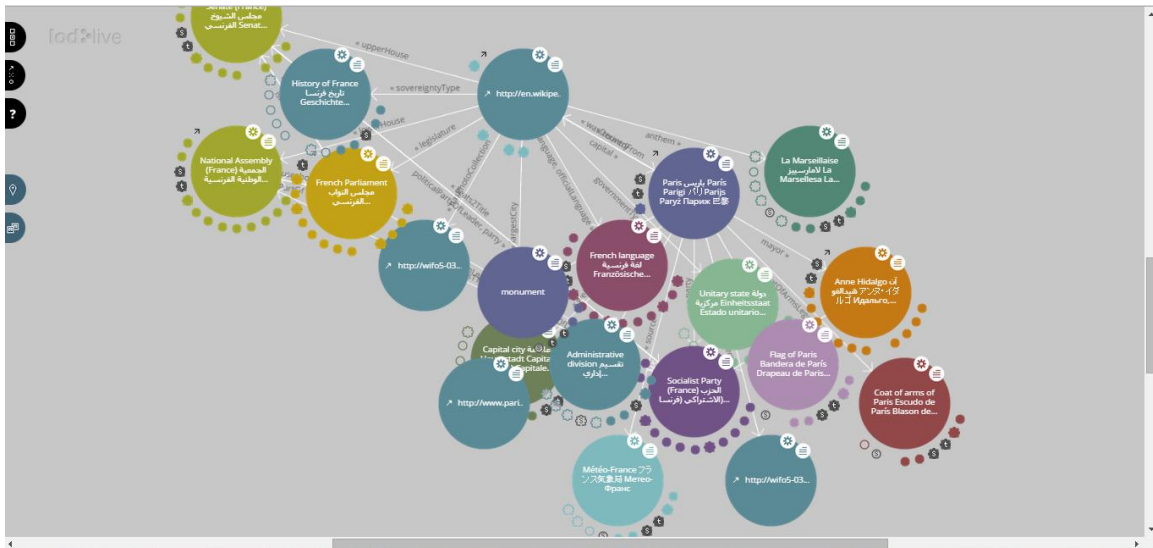


Figure 5 - Representation of Lodlive interface

Gephi (Bastian et al., 2009) software is open source software, available to anyone who wants to improve it. It is a software used for graph and network analysis. Advantage is its compatibility with different operating systems. The possibility to show large graphs is one of the reasons for this software to be used in this research. Gephi also has a plugin named “SemanticWebImport” which allows the access and visualisation of RDF files. To visualise concepts with Gephi, a user needs to explore the software in detail first, after which there are a few interesting options that can be used to create interesting visualisations. For example, there is an algorithm that measures how well a network decomposes into modular communities and can extract sub-networks after which there is an option to change the colours of the nodes according to different classes in modularity. This is a good option if the user has more concepts on the screen and wants to show different clusters in different colours but normally lay users are not interested in that, they just need an option to change the colour, and in Gephi it is not straightforward to do that but instead the user would have to know the meaning of “modularity” option and what exactly it represents. Issue encountered is that the components that need to be presented with the "semantic import" are not represented correctly. Moreover, the software is not meant for lay users, and it needs more time to be explored. Figure 6 depicts interface of Gephi.

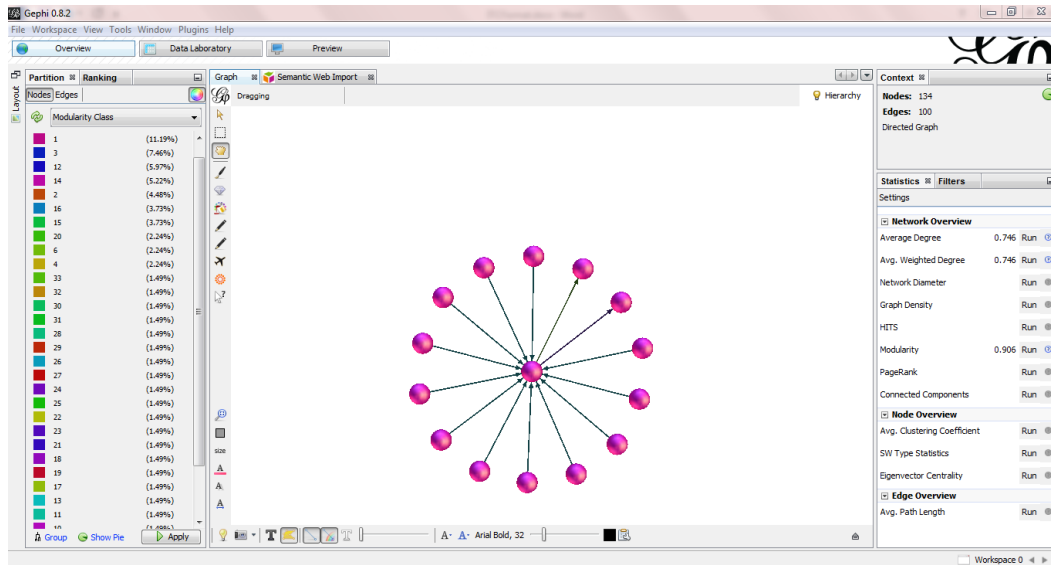


Figure 6 - Representation of Gephi interface

Gruff is a graphical triple store browser. It is an open source software available for download on the official site. When the RDF file is visualized, the software represents different relationships as different types and colours of lines. The full name of the relationships is possible to see when hovered over the relationship, on the bottom of the interface the entire URL appears. These types are represented on the left side of the interface. Additionally, the option of seeing a smaller version of the graph is also available on the left side. Resources are represented as rectangles. The colour of the rectangles differs according to the type of resource. Gruff also allows SPARQL queries. Another option, “table view”, is available and provides information on properties of a single node. The shortcoming of this system is that it shows only up to 100 triples and there is no obvious option to choose which triples one wants to visualize. Interface of the Gruff is represented in the Figure 7.

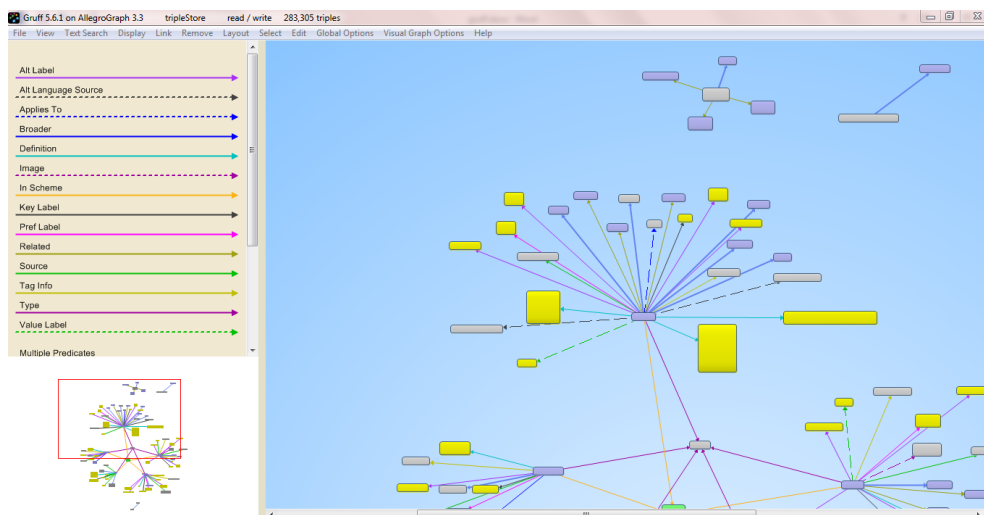


Figure 7 - Representation of Gruff interface

Mazumdar et al. (2013) implemented a system which they call "Affective Graphs". They followed the requirements found in other research as well as feedback given by users to design their solution. In their visualisation, they represented Semantic concepts as pie charts and, and they use constant values of brightness and saturation of different colours for it to make it more pleasing to the user. They put particular emphasis on the edges in the representation so they separate edges into three different types and show them in various ways, as triangles, as curves or as a satellite objects. One of the principles they followed talks about maintaining the typicality of the design, they introduced some new features but tried to save some classical presentations such as pie charts. This seems like a good solution, but the tools are not available for use, it was rather created for a proof of concepts for their research.

2.5. Existing evaluations of linked data visualisations

Fu, Noy, & Storey (2014) mentioned evaluation techniques for visualization of ontologies in two approaches: qualitative and quantitative. Qualitative methods incorporate interviews, observations and thinking aloud. Issues explained with this approach is that one can have generalizability, precision and realism at the same time. Quantitative methods include controlled experiments that measure the effects of independent variables to dependent variables. Recently, user centred evaluations have become more common. Fu et al. (2014) used eye tracking method. The eye tracking method of evaluation is popular since it can allow researchers to detect why certain features of visualization are influencing accuracy or speed. This leads to more detailed evaluations being conducted.

As found in (Fu et al., 2014), there are few approaches to ontology visualizations. For example, one pinpoints features integrated into a tool and compares with other existing tools. This is not a trustworthy process of evaluation since not all tools have coinciding options. In addition, this approach can be easily manipulated resulting in one tool appearing more advanced than the other. The next evaluation type concentrates more on interviews, observations and think aloud protocols. Limitations of this approach include the influence of subjectivity on the usability testing, particularly in the cases when the sample size is small. Another evaluation technique would be to measure the speed and accuracy. These tests are suitable for displaying the effectiveness and efficiency, however, are unable to provide an insight into the causation.

De Vocht et al. (2014) implemented visual workflow for exploration of linked data in a graph form. The central idea of their workflow is that they implemented it through a few steps in navigation. First it provides narrowing views in order to provide user overview of data, next step provides closing down to more detailed resources and finally next step is "broadening views" where users are able to discover new relations together with interaction with LOD (Linking Open Data community, 2014). Authors provided evaluation of their workflow. They have done this by giving users assignments to execute, fill in a questionnaire, and thinking aloud protocol was conducted.

In his paper, Mazumdar et al. (2013) applied a user-centred development process in shaping the final graph. There are three evaluations conducted in this paper that the author considered significant: the first one is "an objective evaluation of aesthetic properties of the system, compared with existing tools". In this evaluation author's primary concern is to answer two questions: "How does the system compare with respect to the system that was judged to be the most aesthetically pleasing tool?" and "How does the system compare with the highest score achieved by any tool for the individual categories?" This evaluation was conducted

by giving ratings to certain features of a design and comparing their solution to the maximum score in each aesthetic metrics that they defined. Second evaluation was comparing the performance of the tool. Questions that authors wanted to answer were "How does Affective Graphs perform in comparison to other systems employing different query approaches?" and "How does an aesthetically designed interface affect the users perception of the system as a whole?" In these case questionnaires, which were gathering users subjective feeling if the tool, were used. Users were supposed to formulate different types of queries and based on time that user need to formulate the query and the number of attempts required to run the query the tool was evaluated. In this evaluation, users were also asked to fill in the questionnaire as a part of qualitative analysis. A third evaluation was meant for observing of user's interaction and respond to the system as well as perception of users to new added features. Questions that need to be answered in this were: "How easy (in term of time and effort required) it is to learn how to use Affective Graphs to perform tasks of different complexity and conduct exploratory search tasks?" and "What is the effect of learning on performing tasks?" as well as "How does learning affect the aesthetic perception of the system?" Users were given with the number of search tasks. Tasks were separated into sections, and each section was happening on a different day. First session served as an introduction of users with the system; the second session was explanation about how to use system efficiently and third session was after what users were given tasks to solve. During the first and third session users were given the exploratory task, this means that they were there is no correct answer, but it is dependent on the user's perception. Mazumdar et al. (2013) extracted some features that they, based on the literature, considered as the most important for LD visualisation. Based on this they evaluated their solution. For example, one of the principles mentioned in this research is that additional graphics should be shown next to the graph itself. Well balanced interface with features such as golden ratio are mentioned, in their research in general they are pointing the importance of the pleasing interface. Another mentioned features are that types of linked should be related to values or types of data and relationships should be distanced from nodes to prevent big overlaps. Even though difficult to apply when dealing with "Big data" they propose design to be minimalistic and to show the least features at the screen as possible. Moreover, they are proposing mechanisms for navigation, filtering and access data on demand.

3. USER-CENTRED DESIGN AND REQUIREMENT ANALYSIS

In the previous chapter, a number of existing LD visualisation solutions were presented with large emphasis on open source linked data visualisation. Additionally, some existing evaluations of LD visualisations were presented in order to detect the optimal method to extract the requirements for future visualisation. This chapter introduces the UCD approach and provides the elaboration on how the user requirements were extracted. Deeper insight into the test setup itself is given and the results obtained through it.

In this research, the used centred design approach was applied. The notion behind this approach, as the name insinuates, was to put the user in the centre of the methodology to allow the designed application to be useful for the user. The user-centred design was conducted initially by executing a Requirements Analysis with a random set of potential users. The aim of this analysis was to gain a general understanding of the most important issues that users are confronted with when operating with graph visualisations. Two different open software were presented, and questions posed in such a way that users have to seek information through the graph.

3.1. Usability

UCD is an approach attempting to meet the requirements of the users. This method is progressively being used on many software applications as it is a modern and well-structured methodology for usability research (Delikostidis, 2011). According to the ISO standards, the definition of usability is: “Extent to which a product can be used by a specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” (Delikostidis, 2011).

3.2. Users

According to Gulliksen et al. (1999) :”The user is a person who will use some product or system so that it carry out his / her professional or recreational activities” (as cited in Delikostidis, 2011).

Dadzie & Rowe (2011) define two types of linked data users: namely lay-users and tech-users. Lay users, the main focus of this investigation, are described as computer literate and able to find information, but are unfamiliar with in depth with semantic web and RDF. They should be able to understand relationships even without a deeper knowledge of RDF. Even users who are familiar with RDF have issues when a large amount of data is considered. The main goal of linked data is it must be usable by both lay-users and savvy-users. The lay-users should be capable to use linked data regardless of lack of understanding of the format and structure of the document. Adequate visualization is one of the ways to help in this aim. According to Dadzie & Rowe (2011), questions that mainstream users ask are, "What is RDF?", "What is it about?", "What can I do with it?" In this research, before testing, users were presented with answers to these questions, allowing the benefits from these solutions to be apparent to the user.

The aim of this research is to discover the optimal method to visualize data in a certain format, namely RDF. RDF is a standard model for data interchange on the Web that extends the linking structure of the Web to use URIs to name the relationship between things, as well as the two ends of the link. It is shown

that there is a need for improvement of visualization of information in this format. A use case presented to a user is intended to detect the shortcomings of existing tools and detect features of visualization which influence the understanding of the user.

3.3. Requirement Analysis

To determine the user needs of the user and design recommendations for future visualisation, the first step in this research was requirement analysis. This was completed in the ITC (Faculty of Geo-information Science and Earth Observation of the University of Twente) laboratory and with an eye tracking machine. The eye tracker machine used was a Tobii X60, in case of the requirement analysis, used to detect how people are observing features that they see on the computer screen. The software used was Tobii X60 hardware and Tobii Studio software for synchronous eye tracking, screen logging, audio and video recording. The size of the screen that was used for this purpose is a 24-inch flat panel display with a resolution of 1650x1050 pixels. Hardware setup in the ITC lab can be observed in the Figure 8.

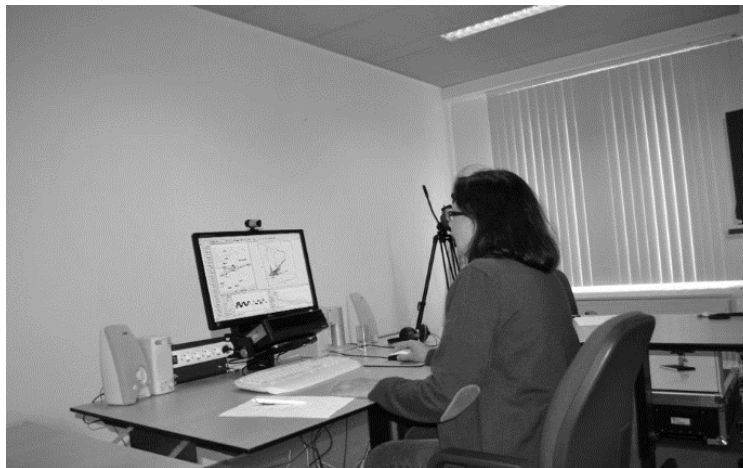


Figure 8 - Hardware setup in ITC lab

The test persons were researchers from ITC. The aim was to extract requirements so that even a person that never experienced linked data search can find the answer to a question they want to answer.

3.3.1. Test preparation

In the beginning, the initial step was to ponder which questions to pose to the users in order to extract user requirements. These were the questions that would guide through the realisation of tasks for users. The main idea before the start of test preparation was to try to create a test in such a way to obtain an understanding of: What is the first thing people notice in the linked data visualisation and what is the first thing they are seeking to find? Are users seeking for a legend to help them with navigation? When do users get confused and how often with their visualisations? Are users confused by the use of different colours in visualisations? These are general questions which served as a guidance while preparing the test set up. In Table 2, the exact issues tackled can be found. The purpose of the test was to answer questions from the Table 2. In addition, according to the user experience it was expected to detect additional issues that users

might encounter. It was expected that there would be a pattern that users would apply while searching for information. The intention of requirements analysis was to discover the main characteristics in linked data visualisations in order to later create suggestions for improvement and possible design solutions. Testing or evaluation of the existing software was not the aim here.

3.3.2. Test setup

For the purpose of requirement test, three visualisation systems were used. The reasoning for choosing these three systems was based on the difference between them, in this way users are tested with three different software solutions and that way meaningful information from different aspects is extracted. These solutions are not only different in the interface, but also in approach: two are web solutions and one is a desktop solution. The difference also comes from the purpose that these two systems were created for. The first visualisation solution is Lodlive which is more user friendly as there are limited options that the user can change in the visualisation itself. This makes it more straightforward for users to make use of the software as generally the overall picture is disregarded; the users are more interested in a method for searching for information.

Another visualisation solution that is not described in existing visualisation solutions in Chapter 2 is Visuwords (LogicalOctopus, 2014). It is not mentioned in the Chapter 2 as its purpose is not to visualize RDF software, but due to its pleasurable graphical features, it was included in the requirements testing. Visuwords is not a visualisation system meant for RDF files, but it is a dictionary that represents similar words and a meaning of the words and their relations with the help of graph visualisations. It is very simple and straightforward for a use. By entering the words into the search box the words similar to that one expand. It is possible to look them up or double-click a node to expand the graph. By clicking and dragging the background it is possible to pan around and with the mouse wheel it is possible to zoom. Hovering over nodes provides a definition of the word and in order to clarify connections the user can click and drag nodes. Used dataset is from WordNet (Princeton University, 2014) - a large lexical database of English developed by Princeton University.

Various datasets are used for the requirement analysis. For Lodlive and RDF Gravity, the DBpedia dataset (Wikipedia Community, 2014) is used. To avoid users from learning these data, different data topics have been presented. For Lodlive and RDF Gravity used data set is from DBpedia and for Visuwords, there is no option to change a dataset but they have their own dictionary that they are visualising.

With the help of different techniques used in this research, such as voice recording, video recording, thinking aloud, and eye tracking, it is possible to get insight into how people work with linked data presented on the graphs. According to Delikostidis (2011) thinking aloud explain the reasoning behind the actions of participant, especially in cases when participant is confused. Eye tracking helps analysing the behaviour of test persons better, when thinking aloud is combined with eye tracking much more information is revealed than by using these two methods separately (Delikostidis, 2011). If TPs were given only surveys, they would not be engaged, and danger of users giving the subjective opinion would exist. However, at the end of the test users filled in a short questionnaire where they were additionally asked to specify things in the visualisation that they found helpful or not. There were things that people forgot to say during the testing and also after seeing both visualisations they would have more idea about what are useful and what are not useful features in visualizations they see. The results were mostly extracted based on the comment that TPs gave during the interaction and with the system or noted in the “end questionnaire”. The steps that uses

were going through are described in diagram in Figure 9. Darker yellow represents the steps that users followed in relation with first visualisation and lighter yellow represents steps user followed in regards to second visualisation.

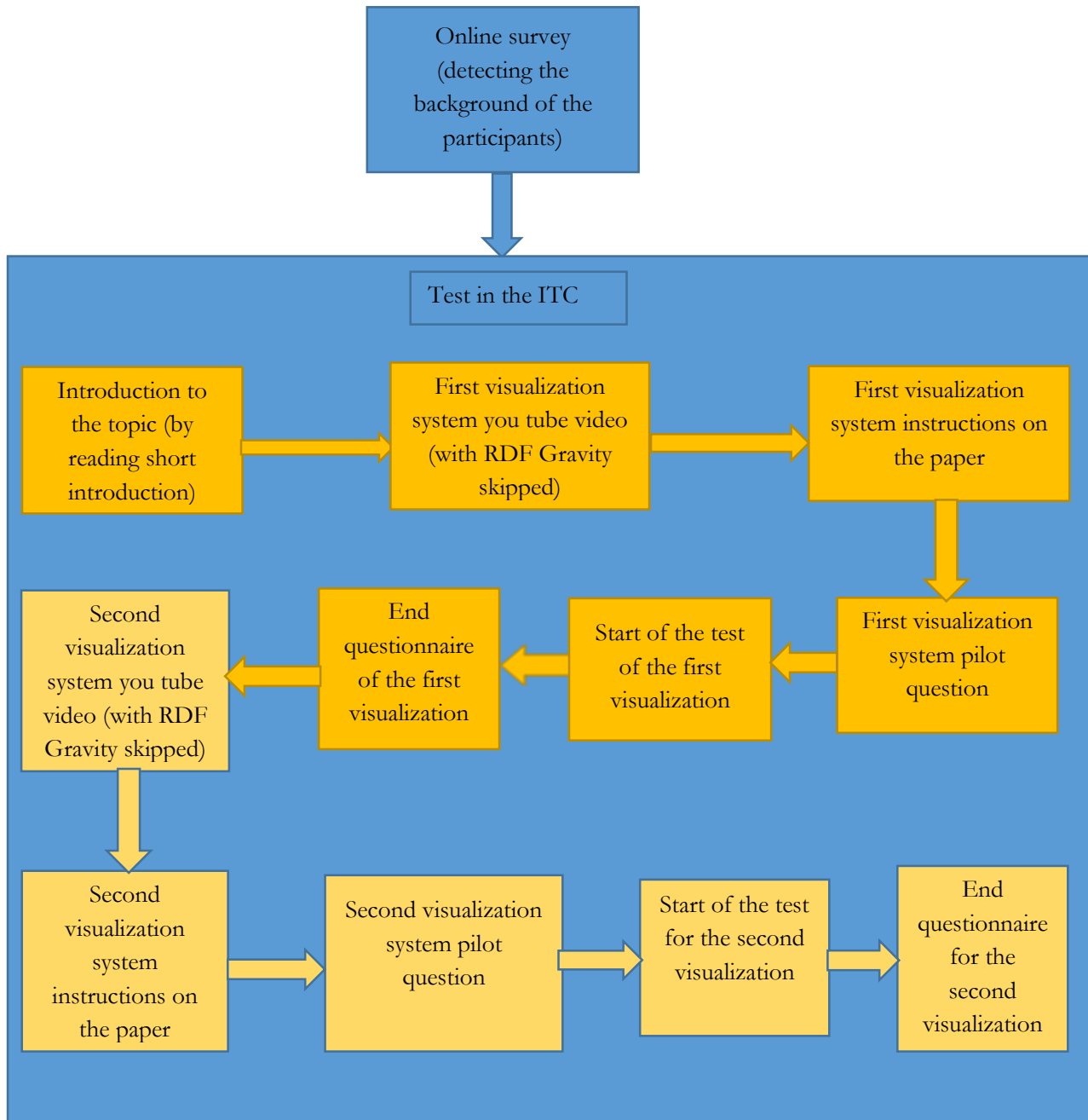


Figure 9 - Steps that users were following

However, in certain parts of the videos, "heat map" was used. For example, to answer the questions whether users concentrate more on the legend or on the graph. Even though questions from the test were prepared in such a way to help in understanding whether some features are useful or not, many significant data was already extracted in the section where users were given similar questions to those in the graph to develop a better understanding of the visualisation system they were using to allow them to feel more comfortable when using any current or future graphs. The questions could have been more detailed, as well as end questionnaire where users were submitting their overall impressions and ideas for improvement. However, this would have resulted in the test becoming too time consuming and users would not hold enough concentration to give productive answers.

There was three visualisation systems, but only two per user were tested. The test set-up allowed the three systems to be evaluated each six times, as testing all three systems per user was too time consuming for the user. Learning bias was also avoided by allowing test person to have varying order of visualisations.

3.3.3. Tasks for users

The tasks, which users were asked to complete, are in Appendix 3. The task are simple questions, written in natural language so that users can identify themselves and that it is possible to observe the approach of test persons in answering simple questions. All the answers would be found in the graph. For example, one of the questions was: "What is the name of the mayor of the Amsterdam?" In the graph, the term Amsterdam is connected to a certain name and the relationship is "leader". The users should be able to realize that the leader is indeed a mayor. Tasks are conceived to be very simple; users should not be threatened by the questions but rather intrigued and eager to find a solution.

In Figures 10, 11 and 12, respectively, the first things users saw in their visualisations in the respective systems are shown.

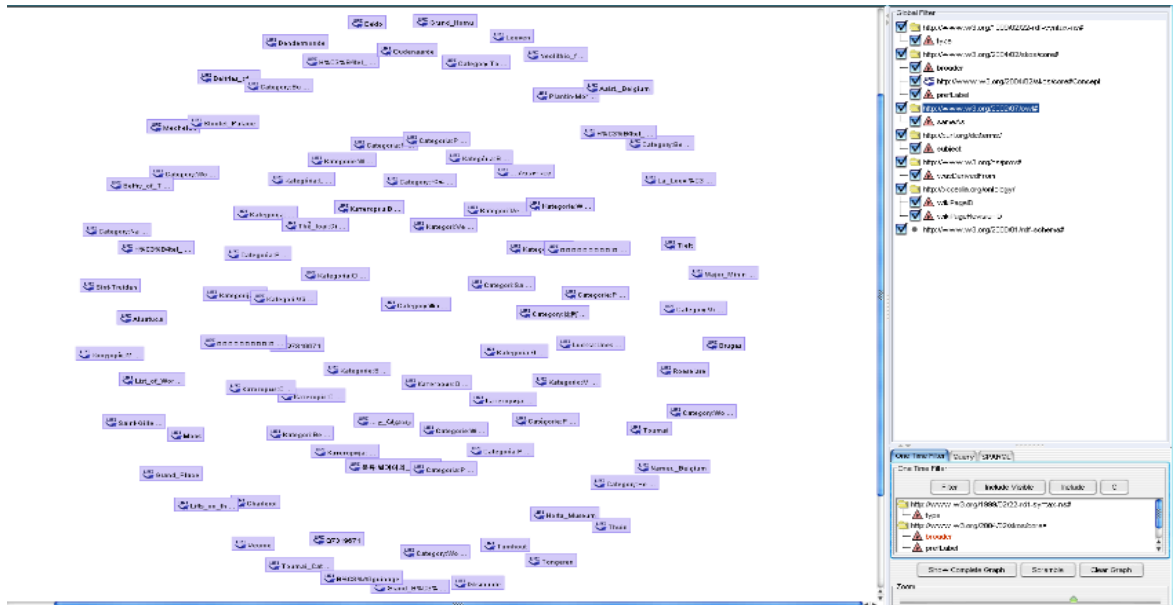



Figure 12 – Initial interface look from RDF Gravity

In the segment of the test where persons were dealing with Lodlive, four questions were asked. The overall idea of the requirements analysis was to understand the way users navigate through the graph when they want to search for the information. They were asked questions that they can find in a graph that was represented to them. The aim of the first question was to see if a certain features help users in search. With this question the idea was to discover if the users are making use of the small flower sign  that represents the concepts with same relationship name. This is the way that Lodlive deals with many resources, especially the ones with the same relationship name is that it grouped them in the small flower sign. This flower did not expand with other resources when the user chose to expand all the terms automatically. In addition, this question served to see if users are looking for an answer by reading the relationships types or resources and that will help in answering the question on what should be emphasized in the visualisation. Next question has aim of detecting the behaviour of the users when the graph becomes more cluttered. What is their strategy to deal with this and when do they become frustrated with the visualisation. The third question was to test the need for a legend and to discover the level of clumping in the graph. With fourth question, the existing way to represent relationships is tested.

The Visuwords testing was another system used for requirements analysis. Most of the questions were meant for exploration and to see how more colourful interface effects the perception of the user if certain features are helpful in the search for information. Therefore, the first question addresses most of the options available, the second question was more oriented to legend, and third one was oriented towards the option to read a meaning or certain type of relationship by hovering over the concept. Other issues are also related to legend. All the information about which issue is connected with which question are given in the Table 2 of this document. Similar counts for the questions asked when users were dealing with RDF Gravity. For this software, it was expected that users comment on what could be improved from their perspective since

it is the most basic one in the visual features. It also helps to notice whether it is helpful when visual features are emphasized with different colours and shapes or if increased the confusion of the users. Additionally, users will have to use the menu on the right side of the screen to help them in search for information so it would be beneficial to see how users responded to that.

3.3.4. Pilot test

Before the actual experiment, the pilot test had to be done. The purpose was to consider the flaws in the test questions as well as in the overall testing procedure. By noticing the shortcomings during the pilot test, possible repetition of testing or misuse of time allocated for each participant was prevented. Before the test, the questions and test setup must be designed in such a way that users easily understand what he or she is supposed to do. All the TPs have to have the same treatment; this was one of the ways to achieve this. Two pilot tests were executed for this testing. The tests were executed the same way they were expected to be executed in the real test. Minor flaws in questions and instructions were discovered which threatened to reduce the understanding of the TPs about the questions. Pilot test was very helpful in determining the number of visualisations to be showed per person since three visualisation, as initially planned showed to take more than 1 h and 30 min.

3.3.5. Test

Table 1 - Background of the users

	AGE	GENDER	COUNTRY OF ORIGIN	KNOWLEDGE OF LINKED DATA	FAMILIARITY WITH GRAPHS IN SOME OTHER FORM
TP1	25-35	F	Belgium	Understand the idea, but never tried	+
TP2	25-36	F	China	Understand the idea, but never tried	+
TP3	25-37	F	Iran	Never heard of it	+
TP4	25-38	F	Italy	Understand the idea, but never tried	+
TP5	25-39	M	Pakistan	Understand the idea, but never tried	+
TP6	25-40	M	Malaysia	Understand the idea, but never tried	+
TP7	25-41	F	China	Understand the idea, but never tried	+
TP8	25-42	F	Latvia	Understand the idea, but never tried	+
TP9	25-43	F	Spain	Understand the idea, but never tried	+

The total number of people who participated in the test was nine, not including pilot test. All test persons were researchers at ITC. They are working with computers on a daily basis as well as internet browsers. Most of the TPs heard about the concept of linked data, but never operated with it (77.78%), moreover

those who are more familiar with linked data are still introduced with it on a very low level, and only one participant stated that he never heard of this concept. Additionally, 55.56% of all participants were not familiar with any of the linked data visualisation system, while rest of have some idea. Users were willing to participate and considered operating with these graphs as a new and interesting experience. The majority of participants do not have experience in exploring information from the graph. Out of the total amount of participants, 11.11% observe graphs daily, 22.22% weekly, 44.44% rarely, and 22.22% almost never. Mind mapping is also a way of creating graphs and analysing them, therefore participants were also asked about thi;, 44.44% were using these maps and 55.56% did not.

Potential test persons were invited for analysis testing requirement by email; the body of this email can be found in Appendix 5. Before the beginning of the test, TPs were briefly introduced to the concepts of the linked data and asked to read a short document that explains this. This document is in Appendix 2. After everything was clear and no further explanation was needed, the calibration of the machine with the TPs eyes was applied. The next step was for test person to read the short tutorial written by myself about the visualisation system and to answer the pilot questions. This short tutorial for all three visualisation solutions can be found in Appendix 2. Users also had a short YouTube video in case of Lodlive and Visuwords as an introduction to what are they going to see in the upcoming tests.

3.3.6. Execution of the test

The issue encountered during the test were when participants did not talk aloud. For some people, it was not natural to think aloud to a machine but after a reminder they would usually comply, especially when they are encountering issues and detecting features that are extremely useful for them.

With two test persons, the eye tracker did not operate properly, but after one more calibration the tests were continued at the same place where they were stopped. The eye tracking machine worked properly with the majority of the people. One test person had to be replaced due to the inability to calibrate her eyes. Since there were other interested candidates this did not halt the testing.

3.4. Results and analysis of the of the experiment

People had different visions of what could be improved in this kind of visualisation. A number of participants concentrated more on the interface which was expected and also a number of people commented on the functionalities, such as “improved search”. The output was analysed with the help of Atlas.ti software(Scientific Software Development GmbH, 2015). The first step was to write comments which users gave during video recording which were later coded with Atlas.ti software. The way codes were segmented into families can be seen in the Figure 13.

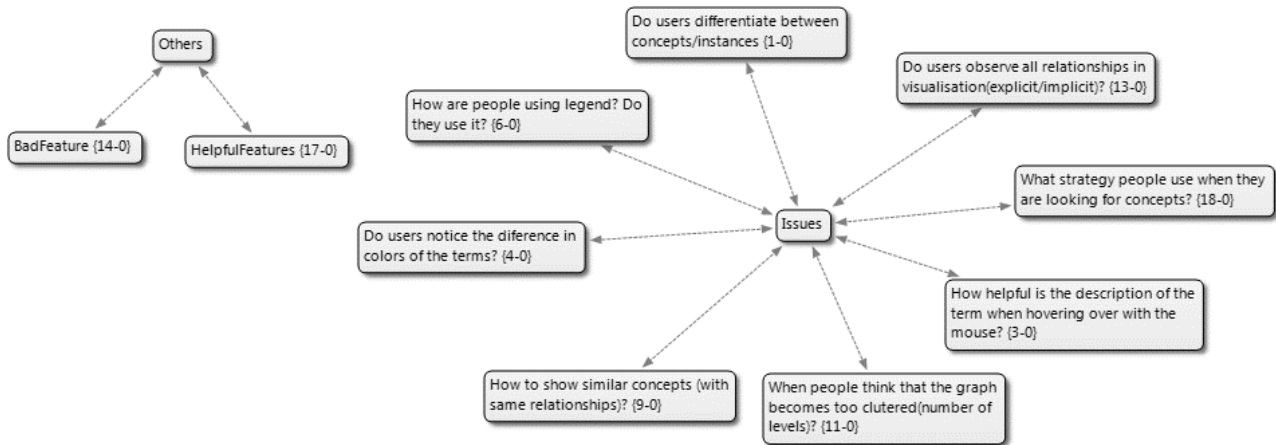


Figure 13 - Organisation of codes in Atlas.ti

From the issues defined in Table 2 codes in Atlas.ti were created. Every issue outlined in Table 2 was defined as code in the Atlas.ti software. Table 2 represents the issues that served as a starting point for the designing questions for the requirements testing. Firstly, issues were outlined and then questions were designed according to that. Every question, as presented in Table 2 had aim of tackling of one of the issues. The codes added to this software was illustrated in Figure 13. All the comments related to issues were analysed. This is described in this subchapter. Moreover TPs had some ideas of how to improve visualisations and the shortcomings about the visualization as represented in Table 2 and Table 3. The data in this tables represents summarised answers to end questionnaire from requirement analysis. First column describes feature and second column in the Table 3 and Table 4 represents the number of users which mentioned this issue. Figure 14 represents interface of Atlas.ti software and depicts the process of adding codes.

Table 2 - Predetermined issues

Issue No	Issue	Question
1	Do users observe all relationships in visualisation(explicit/implicit)?	Lodlive1, visuwords3
2	When people think that the graph becomes too cluttered(number of levels)?	Lodlive2, lodlive3,rdfgravity2
3	How are people using legend? Do they use it?	Lodlive3, visuwords1, visuwords3,visuwords4
4	How helpful is the description of the term when hovering over with the mouse?	visuwords1,visuwords2,visuwords4
5	Do users notice the difference in colours of the terms?	visuwords4,rdfgravity2
6	How to show similar concepts (with same relationships)?	lodlive1,visuwords4
7	What strategy people use when they are looking for concepts?	rdf gravity1, rdf gravity2
8	Do users differentiate between concepts/instances	visuwords4

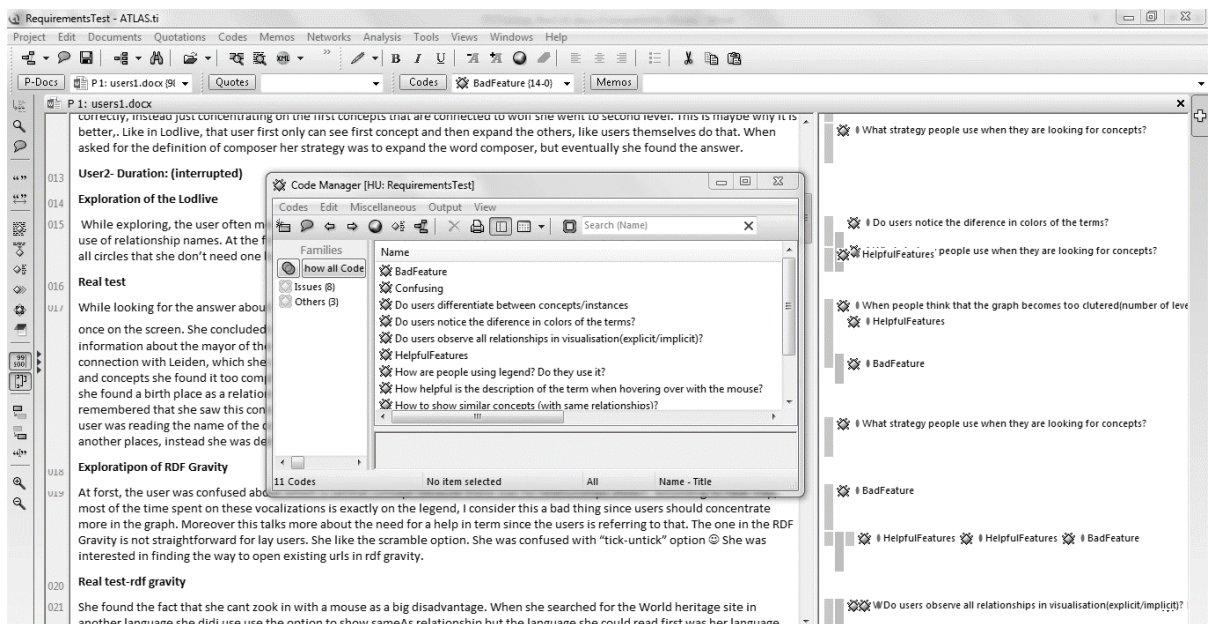


Figure 14 - Process of adding codes in Atlas.ti software

Table 3 - Features that should be added to future visualisation according to end questionnaire

should be added to existing systems:	number
fewer concepts on the screen(Lodlive)	1
friendlier fonts(than in Lodlive)	2
font size should be increased	2
zoom in and zoom out option(especially on the mouse wheel)	2
decrease number of concepts on the screen(than Lodlive)	1
With a hover of the mouse to show next concept?	1
use different colours for different types of concepts	1
URIS shouldn't appear as the name of the concepts on the circles	1
Search box	5
to show the full links	2
instead of random colours to see pictures on the circles	1
by closing one circle, all related circles to be closed	2
the relationships should be clearly represented	1
use size of circles to show the importance of concepts	1
the labels in the relationships should be visible	1
assign colours according to some rule rather than just randomly assigning	3
To alphabetize resources	1
use different shapes for different types of concepts	1
menu where you chose which relations(together with concepts) you want to show	2
larger font size	3

Table 4 - Features that are characterized as positive according to end questionnaire

good things about systems	number
simple interface(Lodlive+Visuwords)	2
colourful interface/nice graphics	5
helpful video tutorial	1
only one concept at the beginning	1
colourfulness	1
the option to move the resources according to users needs (Lodlive), easier to manipulate, it is possible to leave the node where you want	2
Legend accessible during search	2
the menu on the right side(so that you do not have to follow web page, but have instant information)	2
tooltip provides faster search	1
different colours for different concepts are helpful(example of the verb and noun In visuwords)	1
option to select concepts according to the type of the relationship	2
focus on the resource (Lodlive)option	1
different style for different relationship	1
"friendly "visualisations	1
pan option	1
arrows and circles differentiated by colours and patterns	1
possibility to expand concepts	1
force directed graph(clean)	3

Answers of TPs related to issue 1 - Do users observe all relationships in visualisation?

When it comes to the comment of the TPs on relationships during requirements analysis testing, users had to observe relationships so that they can answer the posed question. Some of the participants were confused by the use of visual features when relationships were visualized. Users found the legend in Visuwords where different types of relationships were shown helpful, but also confusing in some moments. For example, one user was confused with two types of relationships: “is a member of” and “participate” because they look similar, this is represented in the Figure 9. Another similar issue was when the legend relationships were represented as a triangle, but since there was very delicate difference, it was hard to detect the direction of the relationship.



Figure 15 - Difference between two different types of relationships in Visuwords

It was noticed that in some cases TPs were struggling to understand the meaning of the links due to not straightforward labelling of links for users who are not used to linked data exploration and different kinds of vocabularies for linked data.


Another observation is that, even though in the questions where users were supposed to answer the question by reading the name of the relationship between two concepts, they discovered the answer by opening the tab on the right side where they obtained more information about a certain resource.

It is already mentioned the option how Lodlive deals with similar concepts, however for some users this was confusing. There were some smaller circles with the letters *s* and *t* and they did not understand the purpose of these symbols. One of the TPs mentioned (but all participants did the same) that it was hard to understand in which language exactly the terms are. The way they finally concluded this is by looking for their own mother tongue which was relatively easy to observe, otherwise it would be difficult. One TP said that he does not like the option in RDF Gravity of tick and un-tick and that he prefers if different relationships are represented with different symbols.

Answers of TPs related to issue 2 - When people think that the graph becomes too cluttered?

This issue is about the maximum number of concepts that should be shown in the graph. It is not aimed at strictly defining the number but rather to specify after which number of features is frustrating for users to operate with graph. Most of the TPs commented on this issue. Four users were agitated when they saw 20 concepts on the graph. After which they had to refresh the page and start only from one concept. One also mentioned difficulties with 19. In other cases it took 12, 13 or 18 resources for users to see before they gave up on exploring before they would erase some resources. For them to start moving and organizing the graph, but by leaving the same number of concepts on the screen. On the other hand, some users even mentioned that five is too cluttered, and some mentioned this when there were 30 concepts on the screen.

Answers of TPs related to issue 3 - How are people using legend?

TPs were supposed to use the legend during exploration of Visuwords and Lodlive in order to answer questions. Only one user used legend in Lodlive to answer, users preferred to discover options by exploring or clicking the  button. In Visuwords the legend is colourful and symbols are represented with larger size symbols. It is plausible that this was the reason that they relied on the legend in this case and for this reason are more successful in their search. In Figure 8, the “heat map” is represented, which shows the accumulated processes of the navigation with the graph in Visuwords for one of the users. From Figure 10, it can be noticed that the legend was frequently used, and even though the graph was moving, the legend remained at the same location. In Figure 10, the heat map during the navigation of one TP is shown as an example. Issues that TPs had with the legend was the nuance of colours in the legend and that the graph was not the same. In addition, the space between different features in the legend should be more emphasized. Another indication, that there should be more colours on the graph, is that during the interaction with RDF Gravity, users mostly concentrated on the "legend" on the right side without many interactions with the graph.

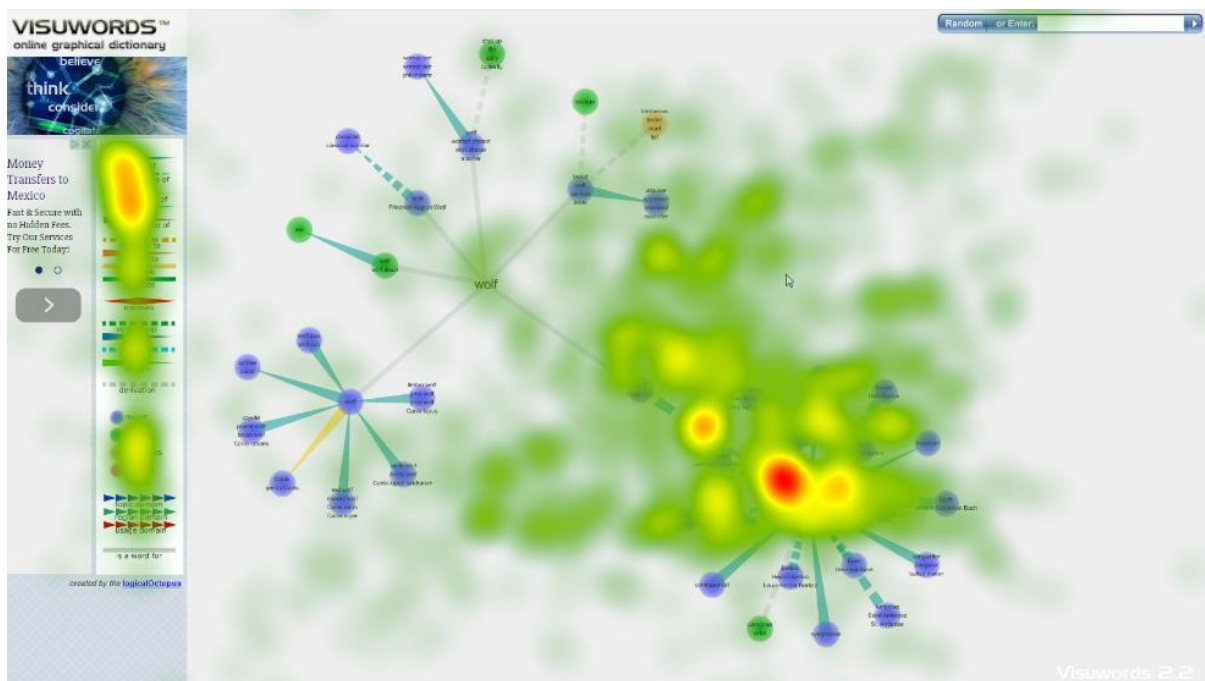


Figure 16 - “heat map” during the navigation of one of the users with graph from Visuwords

Answers of TPs related to issue 4 - How helpful is the description of the term when hovering over with the mouse?

Users did not explicitly commented on whether they find option tooltip option when hover over the node suitable for them, hover ever they made use of it. Another observation is that TPs did not make use of this

option frequently, but rather because questions demanded so. This option is available in Visuwords and RDF Gravity. In RDF Gravity it is just used to read the whole label and in Visuwords hovering over explains the meaning. Even though users did not find these options specifically interesting they suggested to implement it in Lodlive. For example when hovering over the small circles around the larger ones they would like to know the name of the relationship that the small circle is pointing to the larger one.

Answers of TPs related to issue 5 - Do users notice the difference in colours of the terms?

During the video recording, only 4 TPs commented on the use of colours in the visualisation. One of the TPs during the exploration would always mention the colour of the circle. So TP would say for example in the yellow circle I found this concept. Three users liked the feature in Visuwords that nouns were differentiated from verbs by different of circles; I believe this indicates that different types of concepts should be shown with different colours. Furthermore, many users had to rely on the colours of the circles in Visuwords to answer the questions, and they were successful in this and found the concepts very quickly.

Answers of TPs related to issue 6 - How to show similar concepts (with same relationships)?

In the previous issue, it was mentioned that colour was successfully used to differentiate types of terms (Visuwords-subject, verb etc...). Another argument for this was the fact that in RDF Gravity users only had concepts in one colour and users found it difficult to search for concepts. In Lodlive, similar concepts are grouped. For example in the sign "flower" there was hidden concepts in small circles around, however since they are not visible, one of the users suggested to add the option of naming the relationship when one hovers over the "flower" sign. Regardless of being explained in the video tutorial, users had problems with concluding what the "flower" signs actually were, therefore this option was not so straightforward. Even with a symbol legend was visible, it was too difficult for first time users meeting this kind of visualisation. A user who was working with RDF Gravity and later with Lodlive constantly tried to search for the option of turning off the relationships because it was helpful. I would propose a similar option, to user highlights all the relationships, and he or she has the option of choosing which one he or she wants to see.

Answers of TPs related to issue 7 - What strategy people use when they are looking for concepts?

In this testing environment, all users were following instructions given to them before the test. However, the main idea was to detect some usual concepts that people apply when they are confronted with the graph and if needed, to emphasize some of them. Many TPs approached the visualisations uniquely. Some first tried to observe the options they can use and some started exploring the graph and detecting the options once they need them for the task. When exploring existing concepts in most of the cases, there was no order in their search, they tried to skip and to find the concept they needed. One user searched for the concepts clockwise. Some users were expanding concepts one by one, but it was noticed this was the participant could not find the "expand all" option. When users considered the graph to be too cluttered to operate, one of the strategies was to delete concepts one by one because they were afraid if they start from the beginning they will lose previous information. Sometimes they would move the concepts around to make the interface clear. However, in both of these techniques users mentioned it as boring and annoying because there were many concepts on the screen.

Generally, when people observe graphs, especially for the first time, they will only notice concepts on the screen. This was concluded with the help of the eye tracker. Sometimes the graph is not centrally positioned on the screen, and users tried to put it in such a way that it is exactly at the centre of the screen. Another


common thing most of the users did even though it was not explained to them or the questions did not ask for that is to open the tab on the left side on the screen and search for information there.

Answers of TPs related to issue 8 - Do users differentiate between concepts and instances?

Out of 6 persons who were asked to detect the instance of the composer, only one answered incorrectly. However, it can be observed that relationship "instance" did not contribute in this search, they did not pay attention on this, and they rather tried to expand resources.

Additional comments

During the interaction with three different systems users were also talking about good and bad parts of the visualisation that did not fall in any of the category form issues, they are outlined below:

- More users stated that they liked force directed graph and the option to expand resources.
- More users considered the scramble option in RDF Gravity useful.
- Zoom in and zoom out is helpful.
- More users were interested in the option to read more about the resource without going directly to the web page by clicking on the  in Lodlive.
- Expressed interest in finding the way to open existing URLs in RDF Gravity.
- Expand button is helpful

During the testing there were also comments about the features that users did not like, the most common ones are outlined here:

- Due to overlapping nodes and concepts she found the graph too complicated to operate with
- Not being able to zoom in/out is a big disadvantage
- Comment for Visuwords given by TP3: "I cannot copy information and paste it, this is just for viewing the information". This comment was not repeated constantly but the users had similar issues where they would propose something like a search box where it would be possible to search for the name of the concept out of the existing ones presented in the interface.
- TP3 complained about short explanations for some circles (so I could conclude that the tab on the right side of Lodlive is helpful) I believe that this could be improved with something like short explanation of the URLs to which certain concept leads so that users can inform themselves about the concept without leaving the graph, in case that it is not the information they needed. Moreover, struggling to understand the meaning of the links, sometimes there are not labels on the concepts, only URIs but this can be also improved by adding more straightforward labels for the links.
- It is difficult to find the answer when all the relationships are overlapping all those names one on each other.
- Small letters on the concepts are preventing the user from seeing the label and they waste lot of time trying to read the label name.
- The size of the square is too small.
- When relations are vertically written it is difficult to read them

This chapter explained UCD approach as well as how this approach was integrated onto this research and applied to the requirements testing. The process of requirement testing was elaborated. The test setup was explained and justified and in the end the results were presented. Raw data from usability test was processed and the most important and most common comments and observations were summarised as answer to issues from Table 2. This results will build up on the next chapter which will elaborate on how the results obtained in the requirements testing segment were be integrated into the proposed visualisation environment.

4. CONCEPTUAL DESIGN AND PROTOTYPE IMPLEMENTATION

Testing how users interact with existing visualisation solutions provided insight in how users interact with the system. Users were trying to answer given questions while using various techniques. After noting the comments and impressions of the users the overall answer to the issues noted in Table 2 was given. In this chapter these issues were integrated into future visualisations. The interface of the two proposed visualisation environments were presented and features represented are elaborated, they are in accordance with the requirements extracted from previous chapter.

4.1. Recommendations

This Chapter will concentrate on a design of two produce a prototype visualization tools and the foundation of that is laid in a thorough requirement analysis executed with the help of potential users. Before the visualisation is created it is important to take into account recommendations from previous research.

Norman (1988) (as cited in Delikostidis, 2011) extracted seven fundamental design rules that a designer should apply when following user centred design:

“1. Use the knowledge in the world and knowledge in the head. By building conceptual models, write manuals that are easily understood and that are written before the design is implemented.2. Simplify the structure of the tasks, make sure not to overload the short term memory or the long term memory of the user. On average the user is able to remember five things at a time. 3. Make sure tasks are consistent and provide mental aids for easy retrieval of information from long-term memory. Make sure the user has control over the task.3. Make things visible: bridge the gulfs of execution and evaluation. The user should be able to figure out the use of an object by seeing the right buttons or devices for executing an operation.4. Get the mappings right. One way to make things understandable is to use graphics.5. Exploit the power of constraints, both natural and artificial, in order to give the user the feel that there is one thing to do.6. Design for error. Plan for any possible error that can be made. That way users will be allowed option of recovery from any possible error made.7. When all else fails, standardize. Create an international standard if something cannot be designed without arbitrary mappings.”

4.2. Describing dataset

In order to conduct this research, the data was explored with help of Google search and the Open Knowledge Foundation (2014) which contains data in RDF format. Most of the data containing geographic content was not easy to interpret once visualised with the help of Gephi or RDF Gravity.

OpenStreetMap (OSM) is a leading VGI initiative, and the goal of this initiative is to create the map of the world that will be open to everyone for using and editing. Open issues of the OSM map are its reliance and the lack of semantic quality of classes. In OSM maps, geometries are described through tags. Tags are there to denote the meaning of the object and to define the role of that object in the real world. Tags consists of a key and a value. The key is there to represent a broad class of features such as aerial way or amenity and

the value describes the key more detailed. So in that case if we have amenity as a key, possible values might be kindergarten, library, school, etc. The following examples are used from OSM maps (OpenStreetMap community, 2014b) to better describe this tagging structure:

- `highway=residential` is a tag with a key of 'highway' and a value of 'residential' which should be used in a way to indicate a road along which people live.
- `Name=` is a tag for which the value field is used to convey the name of the particular street
- `maxspeed=` is a tag whose value is a numeric speed in km/h (or in miles per hour if the suffix 'mph' is provided). Metric units are the default (and do not need to be mentioned explicitly). Other units, such as miles per hour, knots, yard or pounds must be stated after the value. Where a regulation is specified in a particular unit then that unit should be used within the value field.
- `maxspeed:winter=` is a key that includes a Namespace for 'maxspeed' identifies a different value for maxspeed that applies only in winter.

The planned dataset for this research is in RDF format and represents a semantic web resource extracted from the OSM Wiki website and encoded as SKOS vocabulary (W3C, 2009). It contains a machine-readable representation of OSM tags, as well as few semantic relations among them. This semantic network is extracted with the help of the OSM Wiki Crawler (Ballatore, Bertolotto, & Wilson, 2012). The idea behind the extraction of these tags into a semantic network is to aid the usage of OSM data in different applications, as well as in geographic information retrieval, spatial recommendation systems, data mining, and geo-information integration. The OSM Semantic Network has implemented a method to connect tags where it is possible to find a more substantial association between certain tags than that is the case when simply following current tag structure in OSM. Even though according to the OSM maps fountains, schools, and bookshops are all values for the key amenity in this data set, it would be recognized that a school and a bookshop have stronger relationships than a school and a fountain. This is achieved by creating an open source web crawler and a study on cognitive plausibility. The paths between tags are short due to the shallow semantic structure of the OSM, and most of the concepts will be connected with two edges. This RDF is not the only attempt to extract RDF from OSM tags, however, it is the only one that provided a semantic similarity measure for OSM geographic concepts (Ballatore et al., 2012). By including general hyperlinks between pages, the OSM Semantic network provides a computation of similarity. The OSM Semantic network also contains links to Wikipedia and LinkedGeoData (Linked Geo Data Community, 2014).

4.3. Use case

By adding an option of observing the way the features in OSM maps are connected it might be simpler for future users to add and retrieve the information from the OSM maps. In the OSM (OpenStreetMap community, 2014) it is described how currently the OSM features are being presented. This information is extracted in a form of RDF data.

There are additional motivations for choosing OSM features to test linked data. Firstly, it represents linked geo data and the aim of this research is to improve the presentation of linked geo data, therefore, popularize the extraction of geo data into RDF format. Secondly, most of the OSM maps users are not familiar with

RDF format. The definition and the aim of this research is to test the understanding of lay-users and to improve their understanding through visualization, observing how it can aid in tackling real-world problems. Through more convivial interfaces, more OSM mappers will be welcomed to contribute to this OSM map community. Test persons will include someone that has no previous experience with linked data. However, the knowledge on RDF is not crucial for extracting basic user requirements. With the OSM Semantic network data two use cases were developed in this research:

Use case 1

Jothi is an MSc student in the Netherlands. He is an Indian coming to Europe for the first time and needs to stay there for a year and a half. He is coming to Enschede - Holland. Jothi needs to buy something, and the first thing to turn to is OpenStreetMap, the free map of the world. He has a list of the things that he needs to buy for everyday life that he couldn't bring from his country. He needs to buy cooking ingredients, sheets for sleeping, cooking vessels etc. Before he starts exploring OSM Maps for specific shops that he wants to find, the first thing he will need is to identify different kinds of shops that are a lot different than in India. In India for example "kirana" stores are common in India where people can buy rice, vegetables, spices, etc. and they are scattered all over the city while in Europe most of the shops are in the centre, and there are different kind of shops. By looking at the Graph, positioned on one side of the OSM interface Jothi can make a conclusion about how the shops are organized and what kind of shops he can expect: optician, supermarket, tobacco, toys etc. If he wants to find furniture for his home from the graph he can be informed about second-hand furniture shops that are not available in India and therefore he wouldn't know about their existence. On the graph he can see what concepts are related with shops, so he can notice that there exist shops of organic food of which he didn't know before, in case that he is not familiar with what they are selling the advantage of linked data is that it can be connected with the other terms so by clicking on the link in the graph he can find out more

Use case 2

Steven is a GIS analyst and for his project he needs land use information. He needs to map the land use for his purpose but his dataset is not complete, and he wants to add information from OSM to his dataset. However, he is not familiar with the tagging system in OSM and has never added features. For this reason he does not know if the attributes he uses are similar to the ones that exist in OSM maps. Additionally, he wants to know what land use classes exist in order to see if they can be merged with his dataset. After he had explored this, he can choose to extract data where $k: \text{landuse } v: \text{residential}$, where k represents key and v represents value, as explained in 4.1, or other land use types that he is interested into.

4.4. Answers to questions from requirements analysis

A visualisation of linked geo-data which will improve search for users will have to satisfy certain requirements. These requirements are tested and here main conclusions will be drawn. According to the result and analysis of the experiment from both the video recording and end questionnaire. End questionnaire is presented in Appendix 4.

4.4.1. Improvement of representation of relationships

The first issue analysed in order to extract the main requirement was whether users observe all relationships in visualisation. The aim was to suggest the improvement of representation of relationships because they are an integral part in understanding linked geo data.

Relationships have to be clear and significantly different from each other, since the similarity of various types of relationships is what confused users the most during testing. It is proposed to use two types of emphasizing relationships to be tested to decide on different types of lines, like in Visuwords, however I would propose more discrete lines, dashed for example, rather than very emphasized as they are in Visuwords. On the other hand, especially in the cases where there are many other resources that current resource is connected to, it is not enough to show relations in a visually different way. For this reason additional functionalities to be added to visualisation. For example, give an option for users to click on the relationships they want to see and group them according to relationship name. When they click on the relationship they see they will see all concepts connected with this relationship.

4.4.2. How many concepts to represent on the screen

Another issue that is tackled by this research is how many resources can a user handle before the visualisation becomes too cluttered. By analysing the results obtained through requirements testing it is noticed that numbers around 17 are when users become frustrated and not able to concentrate on the resource. Of course, the number should be much less in order for users to start research. Ideally, it would be good to, in a final usability test show to the user the most appropriate number of different concepts to test .. For visualisations that implement force directed graphs and where resources do not overlap, users can handle even more resources at once.

4.4.3. How to represent a legend for linked geo data visualisation

A legend where symbols for relation and resources will be represented is one more issue dealt with. The aim was to discover how people are using the legend, and to learn from that experience and how it can be improved. After the testing, it is concluded that an “easy to understand” legend can be very helpful in search for information.

The primary requirements when representing a legend are:

- Symbols have to be clearly separated
- With different colours for symbols clearly differentiated
- The nuance of colour from the symbols in the graph and in the legend has to be same

4.4.4. How helpful is the description of the term when hovering over with the mouse

A hovering-over option users did not find very interesting since it disappears after the mouse is not on the resource/relationship anymore. Therefore, the suggestion is to, with hover over option, only display additional information. For example, entire URLs to the resource can be shown by hovering over, while without hover over option only labels can be represented.

4.4.5. How to use colour when representing linked data visualisations

Due to limited number of users and a limited number of existing visualisations to be tested it cannot be specified which features of the visualisation exactly should be differentiated by the colour. The general impression, however, is that a colourful interface helps users to feel more pleased to use the interface. In addition, when types of terms are in different colours, users find concepts easily.

In visualisation 1 (see 4.6) a solution is represented where resources with the same relationships are in the same colour and to see how effective that is and if this is a good way to represent similar resources (with same relationship). Since concepts represent nominal information the most logical thing would be to differentiate them with different colours.

Even though colours might seem a good idea, it would be good to test whether users really find it as a good search tool, or they just enjoy colourful interface. So I would suggest a visualisation where colours do not play a role.

4.4.6. Representation of similar concepts

Another way, besides colour, would be to group similar concepts with different shapes and symbols, like it is implemented in Lodlive. This is shown to be confusing to users, moreover, it can be confusing because typically literals are represented with rectangles and URLs with ellipses.

Lastly, this is also mentioned in previous section, relationships should be presented in a table and users should be given opportunity to tick the relations they want to see.

4.4.7. Strategy for graph exploration

One more issue was to detect a strategy that people apply so that their search can be more effective after the creation of visualisation environment. One of the observations is that when people go through the instructions carefully they don't have issues during the exploration. A recommendation is to show the users

a short tutorial about the main option that the system provides, an interactive tutorial or short video, as it was shown during to the TPs and they even mentioned it as a very helpful feature. In order to engage users it would be even better to guide the users through small example, as in computer games is usually the case. This is not mentioned multiple times, but it would be beneficial to test, whether it would make search easier for persons if the resources are listed in the graph according to alphabetical order as one of the users mentioned. One more feature that should be implemented is force directed graph, which helps resources not to overlap, since many users had problem with that and spend lot of time on scattering resources themselves. Moreover, central term should be positioned in the centre of the screen. Another thing would be to implement the option to explore about the resource without having to go directly to the web page.

One of the features that will be implemented with this visualisation is the option to, when double clicked on the resource all the resources connected to that one will be seen and others will be more transparent. This option is suitable for this kind of visualisation because many different connection are shown. Another way of emphasizing would be to colour the concepts which are connected with the same colour.

4.4.8. How to differentiate between concepts and instances

In the requirements test, only systems where instances were represented was Visuwords, however instances are here marked with relationship names. It is noticed that users did not make use of this to answer the question. Since, in the examples that explain RDF instances are represented as rectangles, and according to principles of good interface, the best would be to show the users what they are accustomed to, this would be the best option. However, another idea might be to represent different datasets with different shapes so that contradicts, therefore it is best to be tested.

4.4.9. Additional features to be implemented in the final visualisation environment

Additional features, not specified in Table 2, but detected during the requirements testing are outlined here. One of the option is to have summarized information about the web page that link is referring to, without having to click on it. Moreover, option which allows expansion of the whole graph at once with all related resources is helpful. Another thing that would be useful for this kind of representation is search box which will search the resources, but only according to label name. Additional feature to be implemented is force directed graph. Labels should be used in the visualisation, instead of representing the whole links to the user, moreover the label names should be written horizontally. Finally, font size should be larger.

4.5. Code development and issues

In this research Python rdflib, a Python library for working with RDF, served for querying the OSM Semantic Network in order to test the use cases and add resources to the visualisation. Some example queries used are represented in Appendix 6.

When designing an interface for application it is best if it is something users are already accustomed to, according to is Sabol et al.(2014) When searching things on the semantic web should be similar to non-semantic search, one box where user can type the term he or she wants to know more about.

4.6. Why d3

The aim of this research is to visualize the linked data on the web, therefore JavaScript was used. D3 is JavaScript library used to bind data to document elements and apply transformations in order to change the content. Its main distinction is manipulation of documents based on data (Michael Bostock, Ogievetsky, & Heer, 2011). A vast amount of examples of use of this library can be found on the internet. According to Stuhr et al. (2011) the weakness of this library is that the performance is not optimal when large data sets are visualised and that nodes are not clickable. In their evaluation d3 library has a score of 57, while library with the highest score has 106 and library with least score has 15, in total 15 libraries were evaluated. Moreover, as Murray (2013) wrote, D3 library is meant for explanatory rather exploratory analysis, which is why in this research will be used only as a proof of concepts. After comparing D3.JS with other JavaScript libraries D3.js had the highest number of online support and examples. D3.js library was sufficient to create the visualisation environment as planned in this research. Other reviewed options were Sigma.js (Jacomy, 2015) and Arbor.js (Samizdat Drafting Co., 2011) which are both JavaScript libraries for data with networked structure, however D3.js, as already mentioned, have big online repository of codes which were very useful for this research.

4.7. Design of the visualisation environment

Conforming to requirement test results the main requirements are signalled. As a result two visualisations with d3.js library are designed. The intention is to create the visualisation which will be tested as a part of UCD approach. The intention is that these two visualisations differ from each other. The two visualisations will satisfy two use cases. First visualisation aims to satisfy smaller data sets, aimed at specific terms to be explained. Normally, in smaller data set it will be less relationships, therefore less colours to be visualised, and then this visualisation can use colours for different colours, since too many colours can confuse user. Second visualisation, due to the fact that all resources will have the same colour, leaves more space for additional data sets in the visualisation to be shown with different colours since second use case needs to be merged with other datasets. Moreover, into second visualisation, functionalities such as option to choose which resources will be show is implemented since it is supposed to work with more data. In the visualisation for first use case only option to search for resources according to its label name is provided.

Visualisation environment 1

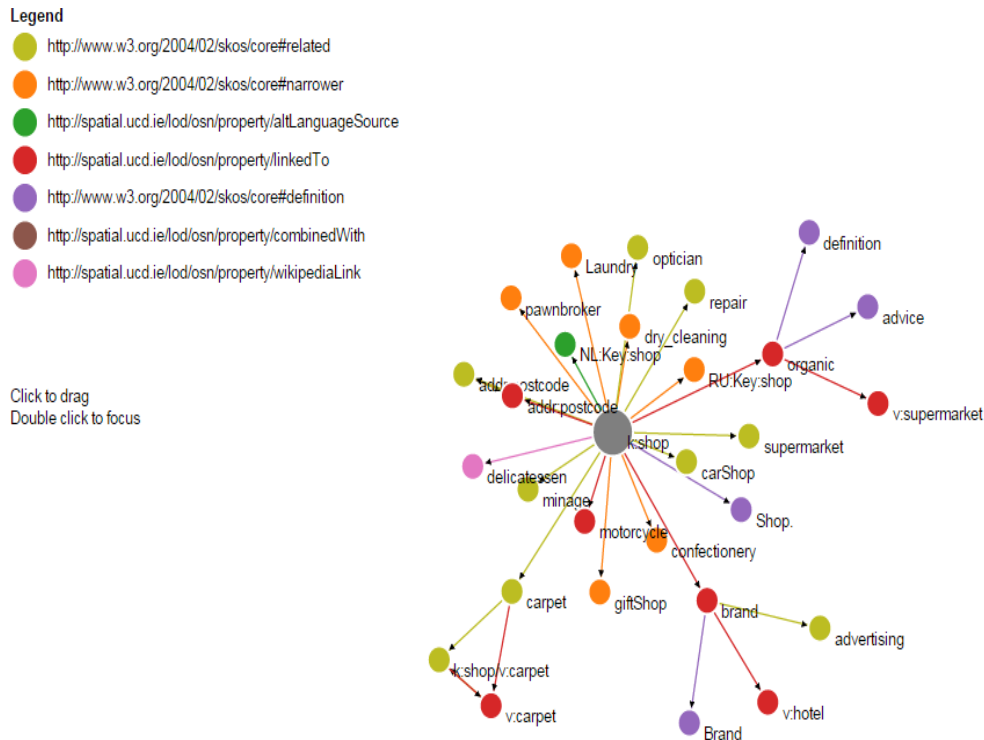


Figure 17 – Representation of the first proposed visualisation

Visualisation 1 was created through existing code from Mence (2014), which was edited and Raper (2014) was helpful as well for attaching additional features. The code for this HTML can be found in Appendix 7. The graph as seen in Figure 17 is centralised on the screen. Colours in this visualisation are used to depict similar concepts, i.e., the concepts with same relationships. Every relationship was assigned with a different colour. The line representing this relationship is coloured in the assigned colour, as well as the circle representing the particular concept. Other visualizations such as different line types and line direction could also be implemented. However, with emphasis solely on colour, the visualization and the distinct relationships as well as the overall layout are more simply defined. For example, in Figure 17, colours are only implemented, and the importance of a relationship between concepts is not emphasize more so than any other particular concept. With this method it is possible to visually group similar concepts.

A force directed algorithm was implemented, making visualisation favourable for use. Nodes are repelled by each other giving no overlap of concepts. This also allows for additional concepts on the screen if needed. To avoid crowded visualisation from abundant text, only the terminal section of the URLs are showed. The

full name can be read when the cursor hovers over a particular concept. This would allow for an additional option to be implemented for users to follow or copy individual URLs. The legend is an additional feature that has been implemented, the space, as well as the colours were optimized to allow clear and obvious differentiation between concepts in the legend. Additionally, the names of the labels in this visualisation are always positioned horizontally.

It can be observed in the legend that different relationships are visualised with different colours. In RDF this is predicate, and they are also represented with URLs. User can make a conclusion about what relation is there represented. The names of this URLs are standardized for RDF, which is machine readable format. Another aspect of usability test would be to see what kind of labels would be most appropriate for users. However, even with URLs, it is easy to conclude what kind of relationships is presented by just reading the last part of the URL.

In the OSM Semantic Network, instances are not represented, therefore they do not appear in the visualisation. However, they could be possibly added from an additional dataset as examples for certain features where the key is shop and the value is supermarket, perhaps represented by rectangles. However, if one visualisation represents two datasets, how to represent second data set? Example Additional functionality added to help users in case of too many resources displayed at once is the double click option. Double clicking on the resource emphasizes only the resources connected by making other non-connected resources transparent as can be seen in Figure 18.

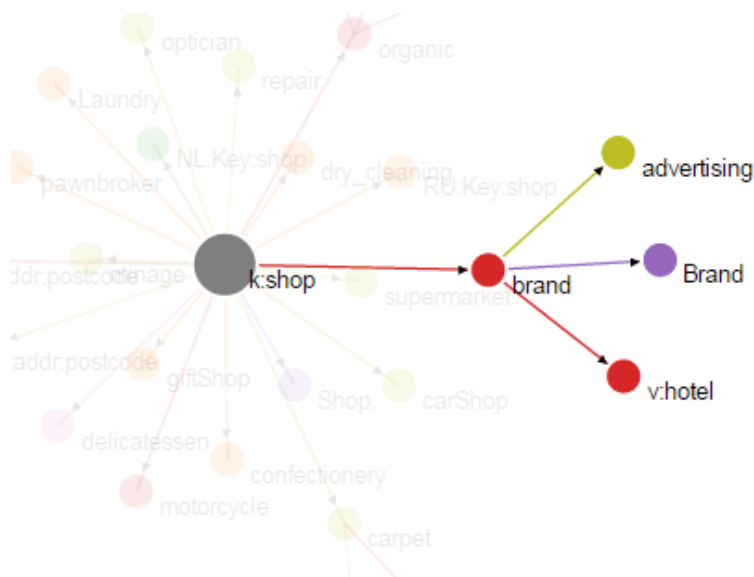


Figure 18 – Representation of what is show in the visualisation after the resource “brand” is double clicked

Visualisation environment 2

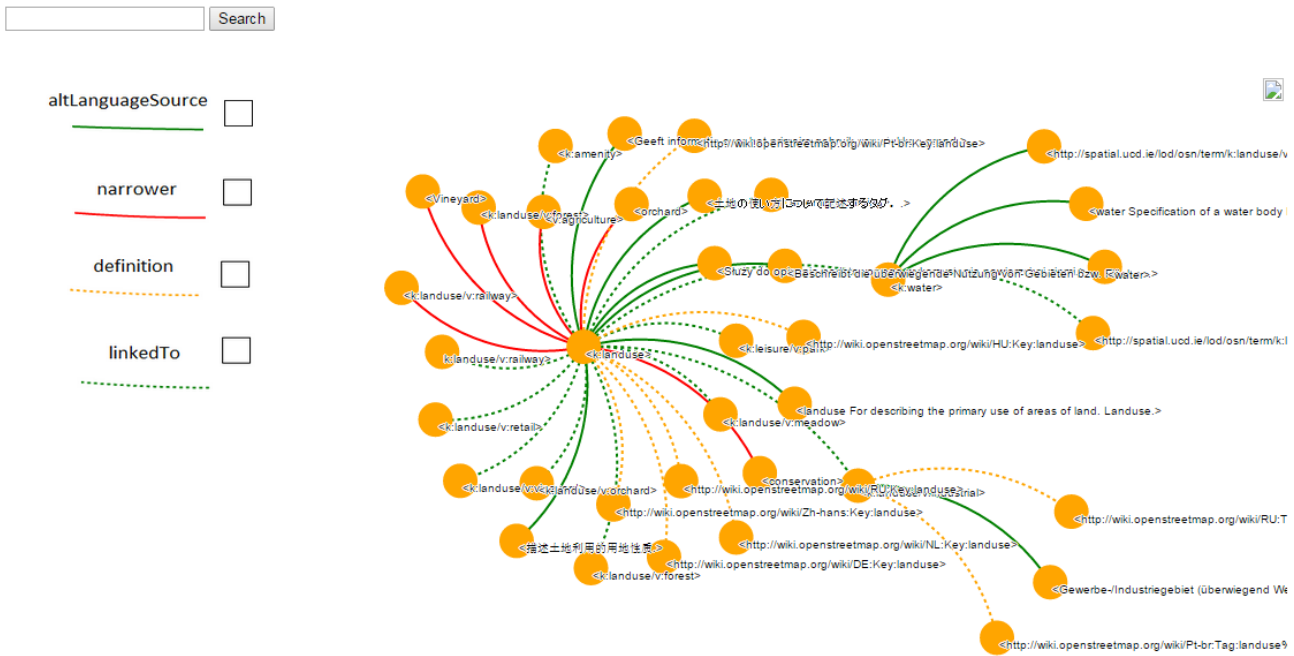


Figure 19 – Proposed interface for Visualisation environment 2

The second proposed visualisation shown in figure 20 was created with the help of existing visualisation which was adapted from Bostock (2011). The code for this HTML can be found in Appendix 8.

The second visualisation has a different approach than Visualization 1. One shortcoming of the previous visualisation was the doubling of concepts due to different relationships and the ability to only assign one colour per concept. However, in the second visualisation, all concepts are mono-coloured with the relationships lines displaying different colours and line types. In this case of visualisation 2, it is less obvious for the user which concepts are similar, but on the other hand concepts are not doubled as in visualisation 1. Relationships are visualised with varying types and colours of lines. They are not overly emphasized to avoid a crowded graph, but differentiated enough to allow concepts with the same relationship to be distinct. The menu has been functionally implemented on the left side of the interface into this visualisation based according to user performance during the requirements testing. This menu provides the option to select which relationships are shown in the visualisation. The relationships in this menu are shown by the name and the symbol with which they are depicted in the visualisation. Another functionality allows users to move the concepts to the position they want, like shown in Figure 21.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Overview

Visualisation of linked data is a potential support in search of information with semantic background, besides this, according to Dadzie & Rowe (2011), it is anticipated that visualisation will help to discover "false links" and "broken links" created due to automatized extraction of linked data. The aim of this research was to, with a UCD approach, identify the most important features when linked data is represented in order to support linked geodata visualisation. Linked geodata already exists in the cloud, but it is anticipated that after its visualisation is improved, more users will be encouraged to use linked geodata. This will empower more scientists to utilize linked geodata in the cloud which will inherently lead to an improvement of quality of linked geo data and integration of geodata information.

In Chapter 1, the background of the research and research questions are outlined. In Chapter 2, some existing linked data visualisation systems are depicted and advices from different research regarding the principles of visualisation that can be integrated in linked-data graphs are summarised. Chapter 3 summarizes the UCD approach and explains how the requirements test was conducted. Chapter 4 gives recommendations about the linked geo data and proposes potential solution.

The recommendations for this document were developed in several stages. After exploring existing systems, linked data concepts and representation of graphs in general, the list of the most important features that needed improvement was created. Those are issues that were expected to be detected. After detecting the most important issues, the question to tackle these issues were created. These are the questions that helped in extracting the main requirements and are outlined in Table 2. This served as a starting point, however users added additional opinions as expected.

5.2. Discussion

In this work, some of the most important features were outlined, and they should be used as a guidance for future visualisations. There are multiple visualisation systems, as described in the first part of this document, and each one has its advantages and disadvantages. RDF is a relatively new concept and introducing it to new users is difficult. Satisfying two main user requirements in visualisation – to be simple and to be understandable – is not an easy task. Straightforward visualisations where everything is shown the way it is and connected are easy to follow if the relationship names are outlined properly or if there is an option of turning the relations according to their names. However, the problem exists when there are many concepts connected to one node. Features for grouping concepts then have to be applied. This is already getting more complicated for lay users, especially with the relations such as "type" and "sameAs" that are not straightforward. The reason for this is that RDF is a machine readable format and not fully adapted for the visualisation.

Another issue that was encountered while operating with some systems is when representing larger RDF files, filtering is applied rather than the whole dataset being shown and it is difficult to detect what method was applied in the filtering and which part of the data is visualised.

In the requirements testing section, users were not shown random data sets, but rather datasets that are adequate for certain visualisation system. Therefore, for RDF Gravity for example, smaller data set had to be shown because RDF Gravity does not have a grouping system as in Lodlive, therefore users very often concluded that RDF Gravity is easier. However, this was not able to be taken into account, especially since the idea of requirements testing was not to test existing systems, but rather to detect the needs of the users.

Lodlive applied some grouping to its visualisation, therefore when person clicks on the “expand all resources” option it is less crowded than Gravity because Lodlive does not show all relationships at once. This is favourable, however a user does not have the overview of the relationships. An option would be to combine options from RDF Gravity and Lodlive and to show all relationships on the side where a user would be able to choose which relationships they want to show and that those relationships can appear and disappear (together with assigned concept) as the user defines. In the questionnaire given to the users after the test, many people noticed the friendly interface of Lodlive, regardless of the not so straightforward way to search for information, in comparison with that RDF Gravity which is visually not the most pleasant system, but the search was easier.

In conclusion, when improving the system for linked data visualisation in such a way that is more pleasing for users, it is not enough just to decide on visual variables. An obstacle is also present on big amounts of data and how to operate with it, as well as the format of the data, while creating RDFs creator to make user adequate labels. When having graphs, one of the principles is to filter the data and it is always challenging to do so without losing valuable information. This research tries to provide subtle contribution to the expansion of linked geodata. There are many more aspect that needs to be improved for this technology to grow, and one of them is government and other public and private bodies to open their data. This is not easy task since publishing data as linked data is creating additional expense of for companies, and significant expenses. In addition to economic aspect, intellectual property rights of particular companies are jeopardized.

5.3. Comparison to other work

Chapter 2 mentioned some relevant recommendations regarding linked data visualisation, one example being (Mazumdar et al., 2013). They compared visualisation of UML diagrams, arguing that it can be applied to a graph visualisation of linked data. Outcomes that are mentioned in this work are confirmed through this research as well. An example is the preventing of overlap of relationships and grouping similar concepts together. In their work they also mention the importance of emphasizing the statistics in the data, which in proposed visualisations in Chapter 4 was not implemented.

5.4. Answering the research questions

In the end, in this section, a summarised version of answers to research questions mentioned in section 1.5 is given.

What are the needs of a user?

The answer to this research question was answered as a result of requirements testing. Elaborated version of the users needs can be found in section 4.4 of this research. A summarised version of the most important needs are:

- A clear way of differentiating relationships
- Representing not more than 17 concepts on the screen
- Not to show significant information with tooltip
- Information on the graph should be represented with colours (one of the solutions is to represent similar concepts with different colour and second is to represent different data sets)

What questions do users need answers to?

These are questions that users will need answer to while operating with given data set. These question were drawn from two use cases elaborated in this research. Therefore, the most important questions that users needs answer to are:

- What is the tagging structure in OSM?
- What kind of shops can I find in OSM?
- What is a proper tagging structure for land use?
- What tagging structure does the user use to represent residential areas?
- Can I find instances of particular land use?

What functionalities should a visualization environment have?

Functionalities that visualisation environment should have are elaborated in more detail in section 4.4 of this research. They are derived as outcome of the requirements testing. A summarised version of the most important functionalities are:

- Interactive tutorial or short video on how to use visualization environment
- Align concepts according to alphabetical order
- Implement the option to explore about the resource without having to go directly to the web page.
- Implement the option that users themselves choose which resources, relationship, that is, they want to choose.
- Option to search for resources thought their labels

How are users finding concepts and how they navigate through the visualization?

Answer to this research question is more elaborated in 4.4.7. The research question was answered by observing users while navigating through the graph.

Users usually first examine circles, concepts that is. They do not pay attention to relationships unless directly asked to. If the legend is colourful and emphasized, they will pay attention to it also. In addition, users also consider summarised information about resources helpful.

Which visualization variables of concepts, instances, and relationships, as well as which combination of visualization variables makes the visualization understandable?

With the help of the basic requirements extracted as a result of the test in the ITC usability lab, two potential visualizations were created. Henceforth, the intention for further research was to test their usability in order to produce final recommendations and an improved visualisation environment. The two visualisations varied highly from each other and this characteristic aided in detecting which features are appropriate for linked geodata visualisation. All test persons in a final usability test should test both visualisations. Moreover, there should be three groups of users since it is important to differentiate between users who are expert in linked data, users who are new to these concepts and those who are new to linked data, but expert in the area of interest related to the dataset(s). This is important because these three types of users have different points of view on how to present the same information. This kind of test was expected to answer exactly which variables will make the visualisation understandable. However, from the visualisation designed in Chapter 4 recommendations were implemented.

Depending on different use cases, what values are appropriate for these variables?

This research question is derived from the section 4.3. and 4.4.

Depending on different use cases, users will also differ as well as their approach. Therefore it is important to adapt the visualisation according to this condition. The difference in the two use cases described in this thesis can explain this problem. While the first use case is represented as a graph on the side of an already developed application to assist users in using this application in their search, a simple interface with only the search box for labels should be applied. Since the user is expected to choose a specific term and to orient towards this term, the assumption is that simple differentiation with colours like presented in visualisation 1 of this thesis will be sufficient because the expectation is that not many different relationships will be presented. Moreover, search of the labels for concepts should be provided, but features such as an option to turn on and turn off the relationships would not be implemented due to lack of space.

On the other hand, use cases such as the use case 2 presented in this research would need a merge of multiple datasets. If different types of concepts are represented with different colours, it will be very difficult to understand the visualisation. In addition to this, it is possible to add the option of choosing which resources will be shown according to relationship name since the second use case is elaborated more as an independent visualisation. Visualisation 2 is more adequate for this use case.

BIBLIOGRAPHY

- Auer, S., Lehmann, J., & Hellmann, S. (2009) LinkedGeoData – Adding a Spatial Dimension to the Web of Data. In *International Semantic Web Conference (ISWC)*.
- Ballatore, A., Bertolotto, M., & Wilson, D. C. (2012) Geographic Knowledge Extraction and Semantic Similarity in OpenStreetMap.
- Bastian, M., Heymann, S., & Jacomy, M. (2009) Gephi: an open source software for exploring and manipulating networks. In *Third International ICWSM Conference* (pp. 361–362). Retrieved from <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154/1009>
- Bostock, M. (2011) Mobile Patent Suits. Retrieved February 09, 2015, from <http://bl.ocks.org/mbostock/1153292>
- Bostock, M., Ogievetsky, V., & Heer, J. (2011) D³: Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12), 2301–9. doi:10.1109/TVCG.2011.185
- Card K., S., Jock, M., & Shneiderman, B. (1999) Readings in Information Visualization: Using Vision to Think.
- Cytoscape Consortium. (2014) Cytoscape. Retrieved from <http://www.cytoscape.org/>
- Dadzie, A., & Rowe, M. (2011) Approaches to Visualising Linked Data : A Survey. *Semantic Web Journal*, 1(2), 34. doi:10.3233/SW-2011-0037
- De Vocht, L., Dimou, A., Breuer, J., Van Compernelle, M., Verborgh, R., Mannens, E., ... Van de Walle, R. (2014) A Visual Exploration Workflow as Enabler for the Exploitation of Linked Open Data. *Proceedings of the 3rd Workshop Intelligent Exploration of Semantic Data*. Retrieved from http://ruben.verborgh.org/publications/devocht_iesd_2014/
- Delikostidis, I. (2011) *Improving the usability of pedestrian navigatoon systems*. University of Twente Faculty of Geo-Information and Earth Observation (ITC).
- DuCharme, B. (2013) *Learning SPARQL*. (S. St. Laurent & M. Blanchette, Eds.) (2nd ed., p. 357). O'Reilly Media.
- Facebook. (2014) What Is Graph Search? Retrieved May 30, 2014, from <http://whatisgraphsearch.com/>
- Fekete, J.-D., Wijk, J. J., Stasko, J. T., & North, C. (2008) *Information Visualization*. (A. Kerren, J. T. Stasko, J.-D. Fekete, & C. North, Eds.) (Vol. 4950, pp. 1–18). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-70956-5
- Freie Universitat Berlin. (2011) Pubby – A Linked Data Frontend for SPARQL Endpoints. Retrieved November 15, 2014, from <http://wifo5-03.informatik.uni-mannheim.de/pubby/>
- Fu, B., Noy, N. F., & Storey, M. (2014) Eye Tracking the User Experience - An Evaluation of Ontology Visualization Techniques. *Semantic Web Journal*.

- Google. (2014) Knowledge – Inside Search – Google. Retrieved August 07, 2014, from <http://www.google.com/intl/en/insidesearch/features/search/knowledge.html>
- Goyal, S., & Westenthaler, R. (2009) RDF-Gravity. Retrieved August 21, 2014, from <http://semweb.salzburgresearch.at/apps/rdf-gravity/>
- Gulliksen, J., Lantz, A., & Boivie, I. (1999) *User Centered Design in Practice - Problems and Possibilities*.
- IBM. (2012, February 28) IBM Analytics – IT Business Intelligence - United States. Retrieved June 03, 2014, from http://www.ibm.com/smarterplanet/us/en/business_analytics/article/it_business_intelligence.html
- Jacomy, A. (2015) Sigma.js. Retrieved February 16, 2015, from <http://sigma.js.org/>
- Linked Geo Data Community. (2014) linkedgeo.org : About. Retrieved November 15, 2014, from <http://linkedgeo.org/About>
- Linking Open Data community. (2014) The Linking Open Data cloud diagram. Retrieved November 15, 2014, from <http://lod-cloud.net/>
- LogicalOctopus. (2014) Visuwords™ online • Visual Dictionary, Visual Thesaurus. Retrieved January 11, 2015, from <http://www.visuwords.com/>
- Massachusetts Institute of Technology. (2008) SIMILE | Welkin. Retrieved November 14, 2014, from <http://simile.mit.edu/welkin/>
- Mazumdar, S., Petrelli, D., Elbedweihy, K., Landfranchi, V., & Ciravegna, F. (2013) Affective Graphs: The Visual Appeal of Linked Data. Retrieved November 27, 2014, from <http://www.semantic-web-journal.net/content/affective-graphs-visual-appeal-linked-data>
- Mence, C. (2014) Interactive network visualisation of murders in The Wire — sub-subroutine. Retrieved February 09, 2015, from <http://www.subsubroutine.com/sub-subroutine/2014/11/28/interactive-network-visualisation-of-murders-in-the-wire>
- Metaweb Technologies. (2014) Freebase. Retrieved November 15, 2014, from <https://www.freebase.com/>
- Murray, S. (2013) *Interactive Data Visualization for the Web*. (M. Blanchette, Ed.) (p. 255). O'Reilly Media.
- Norman, D. (1988) *The Design of Everyday Things*. Retrieved from <http://www.amazon.com/Design-Everyday-Things-Donald-Norman/dp/0465067107>
- Open Knowledge Foundation. (2014) Datahub. Retrieved December 29, 2014, from <http://datahub.io/>
- OpenStreetMap community. (2014a) Map Features - OpenStreetMap Wiki. Retrieved November 13, 2014, from http://wiki.openstreetmap.org/wiki/Map_Features
- OpenStreetMap community. (2014b) Tags - OpenStreetMap Wiki. Retrieved November 12, 2014, from <http://wiki.openstreetmap.org/wiki/Tags>
- Perer, A., & Shneiderman, B. (2006) Balancing systematic and flexible exploration of social networks. *IEEE Transactions on Visualization and Computer Graphics*, 12(5), 693–700. doi:10.1109/TVCG.2006.122

- Plaisant, C. (2004) The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces - AVI '04* (p. 109). New York, New York, USA: ACM Press. doi:10.1145/989863.989880
- Princeton University. (2014) About WordNet - WordNet - About WordNet. Retrieved January 11, 2015, from <http://wordnet.princeton.edu/>
- Raper, S. (2014) An A to Z of extra features for the D3 force layout. Retrieved February 09, 2015, from <http://www.coppelia.io/2014/07/an-a-to-z-of-extra-features-for-the-d3-force-layout/>
- RDF Working Group. (2014) RDF - Semantic Web Standards. Retrieved August 07, 2014, from <http://www.w3.org/RDF/>
- Sabol, V., Tschinkel, G., Veas, E., Hoefler, P., Mutly, B., & Granitzer, M. (2014) Discovery and Visual Analysis of Linked Data for Humans. In P. Mika, T. Tudorache, A. Bernstein, C. Welty, C. Knoblock, D. Vrandečić, ... C. Goble (Eds.), *The Semantic Web – ISWC 2014* (Vol. 8796). Cham: Springer International Publishing. doi:10.1007/978-3-319-11964-9
- Samizdat Drafting Co. (2011) Arbor.js. Retrieved February 16, 2015, from <http://arborjs.org/>
- Scientific Software Development GmbH. (2015) ATLAS.ti: The Qualitative Data Analysis & Research Software. Retrieved January 31, 2015, from <http://atlasti.com/>
- Stadler, C., Lehmann, J., Höffner, K., & Auer, S. (2011) LinkedGeoData : A Core for a Web of Spatial Open Data. *SemanticWeb Journal*, 20.
- Stuhr, M., Roman, D., & Norheim, D. (2011) LODWheel - JavaScript based Visualization of RDF Data. In *Second International Workshop on Consuming Linked Data* (p. 12).
- W3C. (2009) SKOS Core Vocabulary Specification. Retrieved November 15, 2014, from <http://www.w3.org/TR/swbp-skos-core-spec>
- Wikipedia Community. (2014) DBpedia. Retrieved November 12, 2014, from <http://dbpedia.org/About>

APPENDIX 1: REQUIREMENTS ANALYSIS EXPERIMENT PRE SELECTION QUESTIONNAIRE

Requirements analysis questionnaire:

1. Address

Name

Country

Email Address

Phone Number

2. What is your age?

- 18 to 24
- 25 to 34
- 35 to 44
- 45 to 54
- 55 to 64

3. What is your gender?

- Male
- Female

4. Previous knowledge about linked data:

- Never heard of it
- I understand the idea behind it, but never tried it out
- I operated with it and I understand RDF in depth
- Other (please specify)

5. Are you familiar with some visualization systems about linked data?

- No
- Yes (please specify which one)

6. How often do you analyse information by observing graphs?

- Daily
- Weekly
- Rarely
- Almost never

Prev

Next

7. Do you use mind mapping technique in studying or in some other case?

8. If yes, do you draw your own maps or do you use mind mapping software?

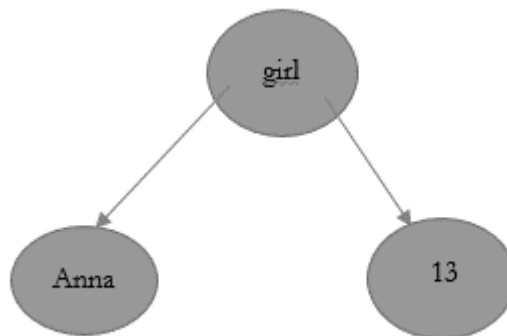
Prev

Next

APPENDIX 2: REQUIREMENTS ANALYSIS EXPERIMENT LITERAL INSTRUCTIONS

General instructions:

The aim of this research is to discover suitable ways to visualize linked geodata. The idea behind the use of linked data is to link data on the internet so that they lead to each other. This will reduce costs of maintenance of databases and also reduce time for searching information. RDF is a format of linked data. It is a data model where the basic unit is known as “triple” with three parts: subject, predicate and object. For example, a subject might be “girl”, a predicate might be “name” (like name of the girl) and object might be “Anna” (so the actual name of the girl). A subject girl might also have the predicate “age” and then we might have the object “13”. Since RDF is a format for presenting the information on the internet and in a machine readable format, subject, object and predicate will be URLs, so instead of a girl we might have something like: www.example.com/girl and instead of Anna: www.example.com/Anna



RDF helps to improve the interoperability between devices and systems. It helps to connect information from various sources. Introducing RDF in the semantic web will improve web search due to the fact that data will be interlinked.

There is a need for the improvement of the visualization of linked data. Linked data is a format created for machines but eventually humans are the final users and that is why it is important to make it more available and understandable for users like you.

During this testing you will be shown three visualisations with two different pieces of software for visualisation of linked data. The graphs with URLs will be presented to you. You will also have the option

of manipulating the graphs so that you can find the answers more easily. For each of the visualisations presented you will be asked to answer specific questions related to only that visualisation. You will be given small tasks so that we can detect how well a user who is not very familiar with linked data can search for information and find answers to possible questions. In the end you will be asked to evaluate and decide on the visualisation that was the easiest for you to use to answer the questions.

DBpedia represents an effort to extract data from Wikipedia to a structured format, namely RDF. During this testing you will be introduced to a source from DBpedia. It means that you will be presented with a data from Wikipedia in a graph representation.


For the test, it is necessary that **you speak aloud** about the reasoning behind your actions and possible issues you may encounter. Your voice and video will be recorded, moreover, eye tracking machine will be used. That will be used for this research since my main objective is to detect the main issues with the visualisations you will see, so that I can create further recommendations for improvement. In case you already know some of the answers please ignore that fact and try to search in the graph for answers.

Good luck in your search!

Lodlive

Before we begin short tutorial on use of Lodlive, the web visualisation system that you will use now, will be presented.

Lodlive is a web based visualisation system for Linked data, data that will be presented to you is linked data from DBpedia.

During the search central term or concept shown will be in the centre of the screen and there will be concepts related to it. Circles that you will see can be moved with the mouse, and if you scatter them a bit it will be easier to answer the question. With a left click of the mouse on the little  you will expand the number of options and by hovering over them you can learn about the meaning of them. Moreover, most of the options are explained in the symbol legend on the left side of the screen

While searching for the answers on the given questions, talk out loud about your experience and the way you are thinking. Your voice will be recorded and will be used to help to improve future solutions. The idea of this is to identify reasoning of the person when searching for the information in the graph. According to this, features that people were more interested in and which helped them in answering the question will be improved during the creation of an improved visualisation environment.

Pilot question

In front of you is the example of the graph when the Belgium is the central concept.

You can expand resources one by one, by clicking each circle around the big one, or by clicking the option "expand all visible relations". Keep in mind that if visualisation becomes too cluttered you can always refresh the page by using "reload" option in Chrome or Mozilla browser and start from the beginning. Notice that in the legend above you have the option to select any of the resources as a central one.

Expand related resources on in the visualisation you see and answer next questions:


1. What is the relationship between "Belgium" and "Senat (Belgium)"?

In this question you should concentrate on the relationship names (predicates, as explained in the instructions you have gone through before)

2. What is "Brabanconne"?

You should be able to notice the concept “Brabanconne” in the graph. Then, according to relations with the concept “Belgium” or other concepts that Brabanconne is connected to, you should be able to discover what it is.

3. Who is the author of “Brabanconne”?

You will need to expand more resources in order to answer this question. The easiest way to do this is to use the option “focus on the resource” after clicking the . More resources will expand and you will be able to answer the question.

4. Name 3 concepts that are connected to the concept “Belgian Federal Parliament”.

Visuwords

Before we begin short tutorial on use of Visuwords, the web visualisation system that you will use now, will be presented.

This is a relatively simple search engine for searching the meaning of different words in a dictionary. Your task will be to search for specific terms and then you will have to navigate through the graph in order to answer questions.

You can try out these options by searching certain term, let's try options described below on Denmark example. Type Denmark in the search box

- When you search for the term, different kinds of linked terms and symbols might appear. The explanation for each one of these you can find in the legend on the left side of the screen.
- A pan option is there with a left click of the mouse.
- By hovering over some of the terms you will notice the explanation for that term appearing.
- By double-clicking the circle, the term will expand to show whether there are any related items.

Pilot test:

Answering this question will help you in answering the questions from the real test.

- According to the graph, what is Scandinavia?
- Can you point out all terms related to the term "Jutland"?
- What is the difference between relationships Denmark - Aalborg and Denmark -European Union?

RDF Gravity

The dataset which is presented to you is from DBpedia. It is RDF data.

We will go through an example of a smaller dataset to show you how the questions will look like and how to use the interface. Try the options described on the example of World heritage sites in Moldova.

- In the interface you see there are no relationships, just terms (concepts). The central term is representing the URL to the DBpedia page about Moldova heritage sites, and all the other concepts are URLs somehow related (in various ways) to the central one, usually referring to another DBpedia page.
- The menu on the right side of the screen might be helpful. In this case you will only make use of the upper window. The red triangles are representing what is, in RDF, a “type”. However, that is not relevant for you to answer the questions.
- Notice how, when the relations in the right upper window are ticked, the connections between the concepts disappear: e.g. untick the term “broader” in the upper window. You will notice a few relationships appearing. So, for example, broader refers to concepts such as Europe Heritage Sites or World Heritage Sites, since the central one is Heritage Sites in Moldova. Untick the “sameAs” and you will notice new links appearing. These are related to URLs that are referring to Wikipedia pages in different languages. Finally, if you untick ”subject” you will be able to see links to URLs to wiki pages with heritages in Moldova.
- In the menu at the top of the page, you will notice the option: ‘Configure’ where you can choose whether to show or hide relationships names and whether to show the whole label name on the screen by hovering over the term in the visualisation.
- At the bottom of the screen is “scramble” option which re-initialize layout of the graph and “zoom” for zooming the graph. For a better understanding you can try this options on the given example.

Pilot test:

Navigating through the graph will help you in answering the next questions:

- Find “World heritage site in Moldova” in another language besides English.
- Name the broader concepts of ”World heritage sites in Moldova”.

APPENDIX 3: REQUIREMENTS ANALYSIS EXPERIMENT TEST QUESTIONS

Lodlive

1. What is the government type in the Netherlands?
2. Who is the mayor of the Amsterdam and how is he connected with Leiden? (For this question you will need to expand more resources).
3. What concepts are connected with the concept: “Unitary state”?
4. What is the relationship between “Netherlands” and “States General of the Netherlands”?

Visuwords

Type “wolf” in the search box. You will notice more terms which are connected with the term “wolf”. This will help you answer next questions. Do not type other terms in the search box, all answers can be found in the graph or by expanding graph.

1. There are different kinds of meanings for this word. Besides the meaning that the wolf is an animal, what else you can find as a meaning from the visualisation on the screen? Name one noun and one verb and write their meaning.
2. What is the relation between the term “wolf” (with the predator meaning) and “Canis” (genus canis)?
3. According to the graph, what is the description (definition) of the word “composer”?
4. In the graph you can notice that “Hugo Wolf” is given as an example of a composer. Find another composer and write his name.

RDF Gravity

Example for the test is about World heritage sites in Belgium.

Navigating through the graph will help you in answering the next questions:

1. Find “World heritage site in Belgium” in another language besides English.
2. Name the broader concepts of “World heritage sites in Belgium”.

APPENDIX 4: REQUIREMENTS ANALYSIS EXPERIMENT END QUESTIONNAIRE

Which one of the existing visualisations was for you the easiest one to find concepts and to orient yourself? Why?

Regarding the interface and visualisation parameters: is there something that would significantly help you in the search for information?

What is the main disadvantage and advantage you found in these visualisations?

Do you have any additional ideas for improving these kinds of visualizations?

APPENDIX 5: INVITATION EMAIL FOR REQUIREMENTS ANALYSIS FOR USERS

Dear ITC researcher,

I am GFM MSc student. I am working on a visualisation of linked geo-data on the web under the supervision of Rob Lemmens and Corné van Elzakker. As part of my research, I am doing a requirements analysis of linked data visualizations in order to discover features that need to be improved in existing visualization systems. If you decide to participate in this test you will be given some examples of linked data visualisations (graphs with different concepts connected together) and you will be asked to search for the information. This will be done in the usability lab and your search actions will be recorded. Do not worry: I guarantee your anonymity and everything will be explained carefully beforehand.

Aim of the research is to improve this visualisation in a way that it is understandable for any person that might want to operate with it. Therefore, the target user of these visualisations can be any person, no matter the background or knowledge about the linked data, as long as it has basic computer literacy and knows how to use browser.

If you are willing to participate, I would like to ask you first to complete an online survey by clicking on the link: <https://www.surveymonkey.com/s/K9ZVJKB> in which you will be asked questions about your background. Be informed that in reporting the research results your name and identity will not be revealed and you will remain anonymous.

Thank you in advance for your time and consideration.

Kind regards,

Jelena Marjanovic

APPENDIX 6: EXAMPLE USE OF PYTHON RDFLIB.

- This is query used for selecting triples where subject is `http://spatial.ucd.ie/lod/osn/term/k:landuse` and predicate is `http://www.w3.org/2004/02/skos/core#narrower`

```
query1 = p.query(
"""SELECT DISTINCT *
WHERE {
<http://spatial.ucd.ie/lod/osn/term/k:landuse
<http://www.w3.org/2004/02/skos/core#narrower> ?o
}""")
```

- Then the query below was created to print the triples in .nt format

for row in query1:

```
print("http://spatial.ucd.ie/lod/osn/term/k:landuse
<http://www.w3.org/2004/02/skos/core#narrower> <%s>" % row)
```

-
- The output below shows the result of the abovementioned query:

```
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:forest>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:military>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/proposed_term/Surface_Mining>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/proposed_term/Vineyard>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:greenfield>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:garages>
```

<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/tag:landuse/v:garages>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<[http://spatial.ucd.ie/lod/osn/term/k:landuse/v:wood_\(Don%27t_use\)](http://spatial.ucd.ie/lod/osn/term/k:landuse/v:wood_(Don%27t_use))>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/Depot>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/Misc._urban_open_space>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:allotments>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/village_green>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/highway>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:brownfield>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:quarry>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:commercial>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:harbour>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:user_defined>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:tillage>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:reservoir>>

<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/proposed_term/landuse/v:trees>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/proposed_term/greenhouse_horticulture>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:farm>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:industrial>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:wellsite>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/proposed_term/Community_food_growing>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:apiary>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:fishfarm>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:meadow>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:well>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:basin>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:pasture>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:cemetery>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:salt_pond>

<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:peat_cutting>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<[http://spatial.ucd.ie/lod/osn/term/k:landuse/v:wood_\(Don't_use\)](http://spatial.ucd.ie/lod/osn/term/k:landuse/v:wood_(Don't_use))>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/Wasteland>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/centre_zone>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:plant_nursery>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:construction>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:vineyard>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:grass>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/orchard>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:railway>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:retail>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:observatory>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:farmyard>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/Traffic_island>

<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:residential>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:pond>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:aquaculture>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:farmland>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/proposed_term/apiary>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:utility>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/proposed_term/Building_site>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/proposed_term/Plant_nursery>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:agriculture>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:field>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:landfill>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:traffic_island>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:breeding>
<http://spatial.ucd.ie/lod/osn/term/k:landuse>
<http://www.w3.org/2004/02/skos/core#narrower>
<http://spatial.ucd.ie/lod/osn/proposed_term/landuse/v:animal_keeping>

<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:orchard>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:street>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/fishfarm>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/plantage>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/Aquaculture>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:depot>>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:recreation_ground>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:greenhouse_horticulture>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/farmland>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:village_green>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/conservation>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/Salt_Pond>
<<http://spatial.ucd.ie/lod/osn/term/k:landuse>>
<<http://www.w3.org/2004/02/skos/core#narrower>>
<http://spatial.ucd.ie/lod/osn/proposed_term/farmyard>

APPENDIX 7: CODE FOR VISUALISATION 1

```
<!DOCTYPE html>
<meta charset="utf-8">

<style>

body {
  background-color: #FFF;
}

text {
  fill: #000;
  font: 10px sans-serif;
  pointer-events: none;
}

circle {
  fill: #ccc;
  stroke: #fff;
  stroke-width: 1.5px;
}

#network {
  min-height:500px;
  position:absolute;
  margin:auto;
  top: 0; left: 0; bottom: 0; right: 0;
  float:left;
}

#network rect:hover{
  fill: black;
}

.link {
  stroke: #999
  stroke-opacity: .30; // kada ovde povecam linkovi se produze i postane citljivije
  stroke-width: 1.5px;
}

</style>
```

```

<body>
<script src="http://d3js.org/d3.v3.js"></script>
<script src="d3.v3.min.js"></script>

<svg id='network'></svg>
<script>

// note that nodes must be listed in id order corresponding to links
// so first node corresponds to 0, second to 1, etc.
var nodes = [
  {node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop', name:'k:shop', group:7,value:14} //shop
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:repair', name:'repair', group:0,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:pawnbroker', name:'pawnbroker',
group:1,value:8}
    ,{node_name:'http://wiki.openstreetmap.org/wiki/RU:Key:shop', name:'RU:Key:shop',
group:1,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:brand', name:'brand', group:3,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:addr:postcode', name:'addr:postcode',
group:0,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:minage', name:'minage', group:0,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:furniture/v:carpet', name:'carpet',
group:0,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:confectionery', name:'confectionery',
group:1,value:8}
/*9*/ ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:organic', name:'organic',
group:3,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:dry_cleaning', name:'dry_cleaning',
group:1,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:craft/v:optician', name:'optician',
group:0,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/proposed_term/Laundry', name:'Laundry',
group:1,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:gift', name:'giftShop',
group:1,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:motorcycle', name:'motorcycle',
group:3,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:car', name:'carShop',
group:0,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:supermarket', name:'supermarket',
group:0,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:addr:postcode', name:'addr:postcode',
group:3,value:8}
    ,{node_name:'A place selling a retail product or service. ', name:'Shop.', group:4,value:8}
    ,{node_name:'http://en.wikipedia.org/wiki/Delicatessen', name:'delicatessen', group:6,value:8}
/*20*/ ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:carpet', name:'k:shop/v:carpet',
group:0,value:8}

```

```

    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:carpet', name:'v:carpet',
group:3,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:tourism/v:hotel', name:'v:hotel',
group:3,value:8}
    ,{node_name:'brand the identity of a specific product, service, or business. Often trademarks
Properties.', name:'Brand', group:4,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/proposed_term/advertising', name:'advertising',
group:0,value:8}
    ,{node_name:'Please note that shop=organic is seen as depreciated. You should use the more
versatile organic=* in combination with a general shop=* key. Related terms: organic store.',
name:'definition', group:4,value:8}
    ,{node_name:'http://spatial.ucd.ie/lod/osn/term/k:shop/v:supermarket', name:'v:supermarket',
group:3,value:8}
/*27*/,{node_name:'A shop selling organic food.Alternatively you can use the more versatile organic=*
in combination with a general shop=* key (In this case: shop=supermarket or shop=convenience). A
shop selling organic food. Alternatively you can use the more versatile organic=* in combination with a
general shop=* key (In this case: shop=supermarket or shop=convenience). Please note that
shop=organic is seen as depreciated. You should use the more versatile organic=* in combination with a
general shop=* key. Related terms: organic store. A shop selling organic food. Alternatively you can use
the more versatile organic=* in combination with a general shop=* key (In this case: shop=supermarket
or shop=convenience).', name:'advice', group:4,value:8}
    ,{node_name:'http://wiki.openstreetmap.org/wiki/NL:Key:shop', name:'NL:Key:shop', group:2,value:8}

```

```
];
```

```

var links = [
  {source:0, target:1}
  ,{source:0, target:2}
    ,{source:0, target:3}
    ,{source:0, target:4}
    ,{source:0, target:5}
    ,{source:0, target:6}
    ,{source:0, target:7}
    ,{source:0, target:8}
    ,{source:0, target:9}
    ,{source:0, target:10}
    ,{source:0, target:11}
    ,{source:0, target:12}
    ,{source:0, target:13}
    ,{source:0, target:14}
    ,{source:0, target:15}
    ,{source:0, target:16}
    ,{source:0, target:17}
    ,{source:0, target:18}
    ,{source:0, target:19}
    ,{source:5, target:17}
    ,{source:17, target:5}

```

```

    ,{source:7, target:20}
    ,{source:7, target:21}
    ,{source:21, target:20}
    ,{source:20, target:21}
    ,{source:4, target:22}
    ,{source:4, target:23}
    ,{source:4, target:24}
    ,{source:9, target:25}
    ,{source:9, target:26}
    ,{source:9, target:27}
    ,{source:0, target:28}
];

// dimensions of plot
var width = 900,
    height = 600;

// Colour scheme for nodes by cause of death
var color = d3.scale.category10().domain(['8','1','2','3','4','5','6','7','0']);

var grayscale = d3.scale.linear()
    .domain([0, 10])
    .range(["white", "black"]);

// opacityscale = scale min-max link values between 0 and 1
var opacityscale = d3.scale.linear()
    .domain([d3.min(links, function(d) { return d.value; }),
        d3.max(links, function(d) { return d.value; })])
    .range([0.2,1]);

/**for scaling node sizes based on # of murders
var hitscale = d3.scale.linear()
    .domain([d3.min(nodes, function(d) { return d.value; }),
        d3.max(nodes, function(d) { return d.value; })])
    .range([5,10]);**/

///// create network plot /////
var force = d3.layout.force()
    .nodes(d3.values(nodes))
    .links(links)
    .size([width, height])
    .linkDistance(30)
    .linkStrength(1)
    .charge(-1200) //voila, kad ovaj broj smanjim(tj. povecam njegovu apsolutnu vrednost) krugovi se
razvoje, nisu slepljeni onako
    .friction(0.7)
    .gravity(0.5)

```



```
.on("tick", tick)
.start();

var svg = d3.select("#network")
  .attr("width", width)
  .attr("height", height);

// build the arrow
// Thanks to Mike Bostock for this code (and much more below)
svg.append("svg:defs").selectAll("marker")
  .data([ "end" ])
  .enter()
  .append("svg:marker")
  .attr("id", "end")
  .attr("viewBox", "0 -5 10 10")
  .attr("refX", 32) // ovde kliknes da se strelice promene
  .attr("refY", -0.5)
  .attr("markerWidth", 4)
  .attr("markerHeight", 4)
  .attr("orient", "auto")
  .append("svg:path")
  .attr("d", "M0,-5L10,0L0,5")
;

var path = svg.selectAll(".link")
  .data(force.links())
  .enter().append("line")
  .attr("class", "link")
  .attr("marker-end", "url(#end)")
  .style("stroke", function(d) { return color(d.target.group); })
;

var node = svg.selectAll(".node")
  .data(force.nodes())
  .enter().append("g")
  .attr("class", "node")
  .call(force.drag)
  .on('dblclick', connectedNodes)
;

node.append("circle")
  .style("fill", function(d) { return color(d.group); })
  .attr("r", function(d) { return d.value; }) // ovde kada stavim * 3 povecaju se krugovi
  .attr("opacity", 1.0)
  .attr('class', function(d) { return 'node' + d.node_name; })
;
```

```
node.append("text")
  .attr("x", 12)
  .attr("dy", ".35em")
  .text(function(d) { return d.node_name; })// kada je d.node onda prikazuje kracu verziju, d.node_name
  duzu
  .attr('class', function(d) { return 'label' + d.node_name; } )
  .style({opacity:'0.0'})
;
```

```
node.append("text");// kada sam ceo ovaj deo dodala pojevio se text, i bez hover overopcije
  .attr("dx", 10)
  .attr("dy", 10)//a bio atribut ".35em" na pocetku
  .text(function(d) { return d.name });
  // .style("stroke", "gray");
```

```
function tick() {
  path.attr("x1", function(d) { return d.source.x; })
  .attr("y1", function(d) { return d.source.y; })
  .attr("x2", function(d) { return d.target.x; })
  .attr("y2", function(d) { return d.target.y; });

  node
    .attr("transform", function(d) {
      return "translate(" + d.x + "," + d.y + ")"; });
}
```

```
// INTERACTIVITY functions for network
node.on("mouseover", function(d) {
  d3.select(this.parentNode.appendChild(this))
    .select("text")
    .transition()
    .duration(400)
    .style('opacity',1);

  d3.selectAll("." + d.node_name)
    .attr('fill', '#000000');

  d3.selectAll("." + d.node_name)
    .transition()
    .duration(400)
    .attr("x", function(d) { return (width-x(d.value))/2 ; })
    .transition()
    .duration(400)
    .attr("x", function(d) { return width - x(d.value) ; })
  ;
}
```

```
});
```

```
node.on("mouseout", function(d) {  
  d3.select(this).select("text").transition()  
    .duration(750)  
    .style('opacity', 0);  
  d3.selectAll("." + d.node_name)  
    .attr('fill', function(d) { return color(d.group); });  
});
```

```
/// CODE TO ALLOW DOUBLE-CLICK TO FOCUS ON LOCAL NETWORK ///
```

```
//Toggle stores whether the highlighting is on  
var toggle = 0;
```

```
//Create an array logging what is connected to what  
var linkedByIndex = {};  
for (i = 0; i < nodes.length; i++) {  
  linkedByIndex[i + "," + i] = 1;  
};
```

```
links.forEach(function (d) {  
  linkedByIndex[d.source.index + "," + d.target.index] = 1;  
});
```

```
//This function looks up whether a pair are neighbours  
function neighboring(a, b) {  
  return linkedByIndex[a.index + "," + b.index];  
}
```

```
function connectedNodes() {  
  if (toggle == 0) {  
    console.log(toggle);  
    //Reduce the op  
    toggle = 1;  
    //Reduce the opacity of all but the neighbouring nodes  
    d = d3.select(this).node().__data__;  
    node.style("opacity", function (o) {  
      return neighboring(d, o) | neighboring(o, d) ? 1 : 0.1;  
    });  
    path.style("opacity", function (o) {  
      return d.index == o.source.index | d.index == o.target.index ? 1 : 0.1;  
    });  
  } else {  
    console.log(toggle);
```

```

toggle = 0;
//Put them back to opacity=1
node.style("opacity", 1);
path.style("opacity", 1);
}
}

/// Add legend to LHS ///

var legend = [
  {group:0, target:'http://www.w3.org/2004/02/skos/core#related'}
  , {group:1, target:'http://www.w3.org/2004/02/skos/core#narrower'}
  , {group:2, target:'http://spatial.ucd.ie/lod/osn/property/altLanguageSource'}
  , {group:3, target:'http://spatial.ucd.ie/lod/osn/property/linkedTo'}
  , {group:4, target:'http://www.w3.org/2004/02/skos/core#definition'}
  , {group:5, target:'http://spatial.ucd.ie/lod/osn/property/combinedWith'}
  , {group:6, target:'http://spatial.ucd.ie/lod/osn/property/wikipediaLink'}
  // , {group:7, target:'CentralTerm-http://spatial.ucd.ie/lod/osn/term/k:shop'}
  // , {group:8, target:'Survived'}
];

var circle = svg.selectAll("circle2")
  .data(legend)
  ;

svg.selectAll("text2")
  .data([0], function(d) { return d; })
  .enter()
  .append("text")
  .text("Legend")
  .attr("x", 5).attr("y", 50)
  .style("font-weight", "bold");

circle.enter().append("circle")
  .attr("cy", function(d) { return 22 * d.group + 65; })
  .attr("cx", 15)
  .attr("r", 9)
  .style("fill", function(d) { return color(d.group); })
  ;

circle.enter().append("text")
  .attr("dy", function(d) { return 22 * d.group + 68; })
  .attr("dx", 30)
  .text(function(d) { return d.target; })
  ;

```

```
svg.selectAll("text2")
  .data([0], function(d) { return d; })
  .enter()
  .append("text")
  .text("Click to drag")
  .attr("x", 5).attr("y", 272)
;
```

```
svg.selectAll("text2")
  .data([0], function(d) { return d; })
  .enter()
  .append("text")
  .text("Double click to focus")
  .attr("x", 5).attr("y", 285)
;
```

```
function mouseover() {
  d3.select(this).select("circle").transition()
    .duration(750)
    .attr("r", 40);
}
```

```
function mouseout() {
  d3.select(this).select("circle").transition()
    .duration(750)
    .attr("r", 8);
}
```

```
</script>
```

```
</body>
```

```
</html>
```

APPENDIX 8: CODE FOR VISUALISATION 2

```
<!DOCTYPE html>
<meta charset="utf-8">

<div class="ui-widget">
  <input id="search">
  <button type="button" onclick="searchNode()">Search</button>
</div>

<style>

.link {
  fill: none;
  stroke: green;
  stroke-width: 2px;
}

#narrower {
  fill: red;//<ovo je da samo strelice posanu crvene>
}

.link.narrower {
  stroke: red;//<a ovo je da cele linije postanu crvene>
}

.link.linkedTo {
  stroke-dasharray: 0,2 1;
}

.link.altLanguageSource {
  stroke-dasharray: 0,2 1;
  stroke:orange
}

.link.definition {
  stroke: green;//<a ovo je da cele linije postanu crvene>
}
```

```

circle {
  fill: orange; //<ovo je za boju krugova>
  stroke: #555;
  stroke-width: 1.5px;
}

text {
  font: 10px sans-serif;
  pointer-events: none;
  text-shadow: 0 1px 0 #fff, 1px 0 0 #fff, 0 -1px 0 #fff, -1px 0 0 #fff;
}

</style>
<body>
<script src="http://d3js.org/d3.v3.min.js"></script>


//<IMG STYLE="position:absolute; TOP:70px; LEFT:1100px;/* WIDTH:200px; HEIGHT:600px*/"
SRC="legendVis2.png">

<script>

// http://blog.thomsonreuters.com/index.php/mobile-patent-suits-graphic-of-the-day/
var links = [

{source:'<k:landuse>',target:'<k:leisure/v:park>' , type:'linkedTo'},//sa zelenom isprekidanom
{source:'<k:landuse>',target:'<k:water>', type:'linkedTo'},
{source:'<k:landuse>',target:'<k:landuse/v:forest>' , type:'linkedTo'},
{source:'<k:landuse>',target:'<k:amenity/v:parking>' , type:'linkedTo'},
{source:'<k:landuse>',target:'<k:landuse/v:meadow>' , type:'linkedTo'},
{source:'<k:landuse>',target:'<k:landuse/v:vineyard>' , type:'linkedTo'},
{source:'<k:landuse>',target:'<k:landuse/v:industrial>', type:'linkedTo'},
{source:'<k:landuse>',target:'<k:amenity>' , type:'linkedTo'},
{source:'<k:landuse>',target:'<k:landuse/v:orchard>' , type:'linkedTo'},
{source:'<k:landuse>',target:'<k:landuse/v:railway>' , type:'linkedTo'},
{source:'<k:landuse>',target:'<k:landuse/v:retail>' , type:'linkedTo'},

{source:'<k:landuse>',target:'<k:landuse/v:forest>', type:'narrower'},//sa crvenom obicnom linijom
predastavljene
{source:'<k:landuse>',target:'<v:agriculture>', type:'narrower'},
{source:'<k:landuse>',target:'<conservation>', type:'narrower'},
{source:'<k:landuse>',target:'<k:landuse/v:railway>', type:'narrower'},
{source:'<k:landuse>',target:'<Vineyard>', type:'narrower'},
{source:'<k:landuse>',target:'<orchard>', type:'narrower'},

{source:'<k:landuse>',target:'<Sluży do opisanja podstawowego wykorzystania powierzchni ziemi.>',
type:'definition'},//orange dashed

```

```

{source:'<k:landuse>',target:'<土地の使い方について記述するタグ。.>', type:'definition'},
{source:'<k:landuse>',target:'<landuse For describing the primary use of areas of land. Landuse.>',
type:'definition'},
{source:'<k:landuse>',target:'<Beschreibt die überwiegende Nutzung von Gebieten bzw. Flächen.>',
type:'definition'},
{source:'<k:landuse>',target:'<描述土地利用的用地性质.>', type:'definition'},
{source:'<k:landuse>',target:'<Geeft informatie over het primaire gebruik van stukken grond.>',
type:'definition'},
/*{source:'<k:landuse>',target:'<Informe sur la nature des terrains exploités par l&#39;homme.>',
type:'definition'},
{source:'<k:landuse>',target:'<Destinazione d&#39;uso del terreno.>', type:'definition'},
{source:'<k:landuse>',target:'<Para descrever a principal utilização de áreas de terra.>', type:'definition'},
{source:'<k:landuse>',target:'<Földterületek elsődleges használatának megadására szolgál.>',
type:'definition'},
{source:'<k:landuse>',target:'<Для описания назначения земель.>', type:'definition'},*/

{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/Pt-br:Key:landuse>',
type:'altLanguageSource'},
{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/HU:Key:landuse>',
type:'altLanguageSource'},
{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/NL:Key:landuse>',
type:'altLanguageSource'},
{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/Zh-hans:Key:landuse>',
type:'altLanguageSource'},
{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/RU:Key:landuse>',
type:'altLanguageSource'},
{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/DE:Key:landuse>',
type:'altLanguageSource'},
/*{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/FR:Key:landuse>',
type:'altLanguageSource'},
{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/JA:Key:landuse>',
type:'altLanguageSource'},
{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/IT:Key:landuse>',
type:'altLanguageSource'},
{source:'<k:landuse>',target:'<http://wiki.openstreetmap.org/wiki/Pl:Key:landuse>',
type:'altLanguageSource'},*/

{source:'<k:landuse/v:industrial>',target:'<Gewerbe-/Industriegebiet (überwiegend Werkstätten, Fabriken
oder Lagerhallen).> ', type:'definition'},
{source:'<k:landuse/v:industrial>',target:'<http://wiki.openstreetmap.org/wiki/Pt-
br:Tag:landuse%3Dindustrial> ', type:'altLanguageSource'},
{source:'<k:landuse/v:industrial>',target:'<http://wiki.openstreetmap.org/wiki/RU:Tag:landuse%3Dindu
strial> ', type:'altLanguageSource'},

{source:'<k:water>',target:'<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:salt_pond> ',
type:'<linkedTo>'},

```



```
{source:'<k:water>',target:'<water Specification of a water body Natural This tag is used to specify what type of water body is tagged with natural=water, and must be used only in conjunction with it. When the latter tag is omitted, it is up to a renderer to decide whether natural=water is assumed.> ',
type:'definition'},
```

```
{source:'<k:water>',target:'<water> ', type:'<http://spatial.ucd.ie/lod/osn/property/keyLabel>'},
```

```
{source:'<k:water>',target:'<http://spatial.ucd.ie/lod/osn/term/k:landuse/v:reservoir> ',
type:'linkedTo'},
```

```
];
```

```
var nodes = {};
```

```
// Compute the distinct nodes from the links.
```

```
links.forEach(function(link) {
  link.source = nodes[link.source] || (nodes[link.source] = {name: link.source});
  link.target = nodes[link.target] || (nodes[link.target] = {name: link.target});
});
```

```
var width = 1200,
    height = 600;
```

```
var force = d3.layout.force()
  .nodes(d3.values(nodes))
  .links(links)
  .size([width, height])
  .linkDistance(100) //ovde menjas duzinu linkova izmedju tih tacaka
  .charge(-1000)
  .on("tick", tick)
  .start();

  var node_drag = d3.behavior.drag()
    .on("dragstart", dragstart)
    .on("drag", dragmove)
    .on("dragend", dragend);
function dragstart(d, i) {
  force.stop() // stops the force auto positioning before you start dragging
}
function dragmove(d, i) {
  d.px += d3.event.dx;
  d.py += d3.event.dy;
```

```

    d.x += d3.event.dx;
    d.y += d3.event.dy;
  }
  function dragend(d, i) {
    d.fixed = true; // of course set the node to fixed so the force doesn't include the node in its auto
positioning stuff
    force.resume();
  }
  function releasenode(d) {
    d.fixed = false; // of course set the node to fixed so the force doesn't include the node in its auto
positioning stuff
    //force.resume();
  }
}

var svg = d3.select("body").append("svg")
  .attr("width", 800)//u originalu je umesto 800 bilo samo width ali ja promenila jer kad dodam sliku sve
se poremeti
  .attr("height", 600);

// Per-type markers, as they don't inherit styles.
svg.append("defs").selectAll("marker")
  .data([{"linked to"}])
  .enter().append("marker")
  .attr("id", function(d) { return d; })
  .attr("viewBox", "0 -5 10 10")
  .attr("refX", 5)
  .attr("refY", -1.5)
  .attr("markerWidth", 5)
  .attr("markerHeight", 6)
  .attr("orient", "auto")
  .append("path")
  .attr("d", "M0,-5L10,0L0,5");

var path = svg.append("g").selectAll("path")
  .data(force.links())
  .enter().append("path")
  .attr("class", function(d) { return "link " + d.type; })
  .attr("marker-end", function(d) { return "url(#" + d.type + ")"; });

var circle = svg.append("g").selectAll("circle")
  .data(force.nodes())
  .enter().append("circle")
  .attr("r", 15)
  .call(force.drag)
  .call(node_drag); //Added to pin down nodes, and we defined this node_drag what it does in this
code also;

```

```
var text = svg.append("g").selectAll("text")
  .data(force.nodes())
  .enter().append("text")
  .attr("x", 2)
  .attr("y", "1em")
  .text(function(d) { return d.name; });

// Use elliptical arc path segments to doubly-encode directionality.
function tick() {
  path.attr("d", linkArc);
  circle.attr("transform", transform);
  text.attr("transform", transform);
}

function linkArc(d) {
  var dx = d.target.x - d.source.x,
      dy = d.target.y - d.source.y,
      dr = Math.sqrt(dx * dx + dy * dy);
  return "M" + d.source.x + "," + d.source.y + "A" + dr + "," + dr + " 0 0,1 " + d.target.x + "," +
d.target.y;
}

function transform(d) {
  return "translate(" + d.x + "," + d.y + ")";
}

</script>
```