

SeMap: A Concept for the Visualization of Semantics as Maps

Kawa Nazemi, Matthias Breyer, and Christoph Hornung

Fraunhofer Institute for Computer Graphics Research,
Fraunhoferstr. 5, 64283 Darmstadt, Germany
{Kawa.Nazemi, Matthias.Breyer,
Christoph.Hornung}@igd.fraunhofer.de

Abstract. The enhancement of the individual knowledge is a basic need that came up with changes in our society, whereas the process of learning disappears more and more. In the recent past the disappearance of a predefined learning process was named ambient learning, which came up to cope the changing need of every time and everywhere learning. Learning contents get more structure by new technologies like semantics, which specifies and defines more the semantic structure and with it the meaning of information. Users working with information system are confronted with different processes for getting the required information. The following paper introduces a new visualization technique, which uses the everyday processes of information search for imparting knowledge. The visualization technique utilizes the surplus of semantics to encourage the process of ambient learning.

Keywords: semantic visualization, ambient learning, treemap, treeview.

1 Introduction

In the information science a new paradigm of learning came up called ambient learning [9]. This paradigm tries to cope with the society change of every time and everywhere learning. Thereby the process of learning is not in the focus of the learners. Instead learning is an implicit effect for example exploring an information space for searching information. The learning effect can be encouraged by utilizing the surplus of semantic knowledge representation to fit the information visualization to the user's mental map in mind [6].

In this paper attributes are identified which solve the needs for the society change using the benefits of the technological change. A visualization of semantic annotated content had been developed which supports a fast and intuitive navigation while ambient learning is encouraged implicit.

The visualization concept had been developed in the scope of the THESEUS program [13]. THESEUS is a research program, initiated by the Federal Ministry of Economy and Technology (BMWi) [4], to develop a new internet-based infrastructure in order to better use and utilize the knowledge available on the internet.

In this paper the underlying semantic data structure will be described first. Methods for visualizing such data will be examined. In section 3 relevant aspects in semantic visualization for imparting knowledge will be identified and the concept seMap for semantic visualization using maps will be described. A case study had been build up to evaluate the benefits of the concept seMap in the field of ambient learning.

2 Visualization Techniques of Semantic Annotated Content

2.1 Semantic Structures for Information Representation

For visualizing the relations and structure of information, an elementary data representation structure had been chosen. The Resource Description Framework (RDF) [15], specified by the W3C Consortium [16], contains all relevant aspects needed to visualize the information in an adequate way. RDF is a data format for low-level representation of semantic data. It is based on XML namespaces to refer to schemas. Meaning is expressed by encoding it in sets of statements. Each statement describes a resource. RDF statements are often referred as triples that consist of a subject, predicate and object, which correspond to a resource (subject), a property (predicate), and a property value (object) [1]. The RDF Schema (RDFS) [15] extends RDF and provides basic elements for the description of ontologies. RDFS provides a higher level of abstraction than RDF, intended to structure RDF resources. Further more RDFS provides information about the interpretation of the statements given in an RDF data model.

Data elements of an RDF file should be visualized in a way the user can discern the information effortless [8]. This includes classes, instances, taxonomy (*is_a* relations), multiple inheritance, role relations and properties. The challenge for information visualization is to extract relevant information, combine it with other information sources and present it to the user adequate. The goal is a reduction of the cognitive load of the user. Even more the user should effortless realize inner data structures.

2.2 Visual Techniques for Representing Semantics

Most semantic visualizations are graph-based. The key issues in graph visualization are the graph structure (oriented vs. non-oriented graphs, trees vs. cyclic graphs) and their size. A survey of basic graph visualization techniques for different graph types can be found in [7]. Existing applications realizing these techniques were collected and categorized by Katifory and Halatsis [8]. The applications were grouped in the following categories, representing their visualization type: 1. Indented list, 2. Node-link and tree, 3. Zoomable, 4. Space-filling, 5. Focus + context or distortion, 6. 3D Information landscapes. Applications of the indented list visualize semantics as tree views representing the taxonomy (*is_a* relationships). Node-link and tree visualizations present the data as a set of interconnected nodes. Users can expand or retract nodes and their sub-trees, to adjust the level of detail. Zoom-able visualizations contain methods that present nodes of hierarchically lower levels nested inside their parents. Users can zoom-in to child nodes in order to enlarge them. Space filling techniques consume the whole screen space by subdividing the space for each node. Techniques of the fifth category distort the view to combine context information

using focus. 3D Information Landscapes realize the landscape metaphor by placing information entities on a plane as color- and size-coded 3D objects.

These visualization techniques were examined in order to identify advantage aspects which may be extracted, transferred and combined to achieve a proper visualization concept. Thereby the aspect of personalization of information visualization has to be mentioned. Using personalization the visualization can be adapted to the user's preferences by filtering and context-based usage of metaphors [10].

3 Semantics as Maps

3.1 Main Criteria of Semantics Visualization for Imparting Knowledge

The main criteria of semantic visualizations are identified according to the needs for imparting knowledge. The goal of imparting knowledge in an ambient way leads to criteria, which focus on the user's preferences and the actual information context. Even more the information should be presented in a way the user can effortlessly realize the inner data structures, allocate the low-level information into the abstract structure and internalize the visualized knowledge. This supports the procedure of learning in an intuitive way due to the visualization which is adapted to the cognitive information storage and handling of the user's mind.

The following main criteria of semantics visualization for imparting knowledge had been identified:

- implicit information structure
- visual layout and order according to the user's mental map in mind
- personalized visualization
- fast and proper interaction mechanisms
- increasing level of detail
- navigation awareness and history

These criteria lead to the following characteristics of imparting knowledge;

a) visualizing the information structure implicit, not explicit by for example using edges, the visual complexity can be reduced with respect to the amount of displayed elements. This directly leads to a faster cognition of the displayed information.

b) adapting the visualization of the information structure to the user's mental map in mind, the effort of structuring and allocating the visualized information to the user's present knowledge can be minimized. So the internalization of new information for the user is encouraged [3].

c) reduces the cognitive load for the user [1]. The user can focus on the information itself instead of being deviated by inadequate visual items.

d) fast and proper interaction mechanisms supports the navigation through the data set. Here Fitts' Law [5] comes into account to decrease the time for interacting with the system and increase the precision of interactions.

e) should follow the line of a few abstract information elements to delaminated detailed low level information. An overfilling of the screen space with too much information will overcharge the cognitive load and will discourage the user. The expansion of

the level of detail has to be triggered by active interactions from the user, so the user will not be confused [11].

f) leads to orientation for the user in the information space according to his navigational interaction chain.

3.2 Using the Surplus of Semantics for Imparting Knowledge

Existing solutions for information visualization do not take the surplus of semantics into account. The benefit of the meta-information of the semantic data can be used to adapt the displayed section of the information in such an intuitive way that the user can perceive the information fast and easy.

Here the criteria for imparting knowledge will be achieved by using the surpluses of the semantic data combined with user specific data sources like a user and a community recommendation system. Especially the structure information which results from the meta-data in the semantic data can be used to organize the visualization elements in a user specific way.

The implicit information structure will be extracted out of the meta-data. The construction of the meta-data is used as basis for the progress from abstract data to a detailed level. This progress will be contained implicit in the visualization, so the user is not confronted with additional, maybe unnecessary information elements. Even more the semantic data structure can be adapted to the users mental map, so the learn process of adding new information aspects to the existing knowledge is assisted indirect.

3.3 Existing Visualization Techniques as Initials

For the new concept of semantic information visualization for imparting knowledge two existing visualization techniques were identified, which provide the opportunity of fulfilling the main criteria. Both of the visualizations have advantage aspects for this purpose. These advantages are to point out while disadvantages in the given changed case are dissolved. These two identified visualization techniques are the TreeMap concept of Shneiderman [12] and the Treeview visualization concept.

TreeMap. The TreeMap view concept, drafted by Ben Shneiderman [12], is a 2 dimensional space-filling approach in which each node is a rectangle whose area is proportional to some attribute such as node size. Thereby the screen is divided into rectangle regions, which represent a node of the visualized tree. Child nodes of one node are placed into the region of the node by subdividing this space again according to some space filling rules. The resulting visualization is a representation of the whole information tree, arranged on the entire available display screen.

The identified advantages of this visualization which can be used to semantic knowledge representation are the concept of space filling approach, the utilization of the whole available screen space and the visually structuring as rectangles, which lead to a well arranged and clear visualization.

But for the concept of semantic visualization the space filling approach had been reduced from two dimensions to one dimension. Even more the attributes which determine the size of one element according to the used dimension for space filling is specified using external information sources containing user specific information.

Also the user will not be confronted with the whole information displayed at the initial startup. Instead the visualization displays an abstract level of detail at the startup and the user can increase this level by navigating through the visualized information space. This leads to an implicit learn effect at the user due to the learning by exploring paradigm.

Treeview. The Treeview visualization concept is well known and often used for presenting information to the user. The visualization is organized as a hierarchy. Thereby the visualization starts at a root node of a tree and displays the next and more detailed level by an interaction of the user.

The major advantage of this visualization concept is that the initial displayed information is sparse and the user has the ability to increase the displayed information. Due to these active interactions the user knows why the information comes up, where the information has to be located in the informational context and he can effortlessly realize the inner data structures.

3.4 SeMap for Semantics Visualization

The two presented visualization techniques offer different advantages as well as disadvantages for information visualization. Shneiderman's treemap uses different graphical primitives for indicating meta-information of the displayed content. With this kind of overview-visualization and the usage of graphical primitives like color, order and size, the user has the ability to gather information from this visualization in a fast and adequate way. This kind of visualization is beyond doubt real adequate to give an overview of small sets of information and indicate where the most relevant information is placed. But if the given content or knowledge is huge and the whole set of semantic information is displayed in only one screen, the user needs a longer time to identify or find the required information. Further it is more complex to gather the semantic relations of the given knowledge. The user gets information about trends, e.g. in which area are which kind of information placed, but the real content behind the given meta-information will not offer an adequate learning effect. Shneiderman's treemap does not offer an efficient way to impart knowledge from huge semantic annotated content in the way it exists.

Figure 1 shows an example of a treemap with a huge amount of data. The surpluses of this kind of visualization with the usage of the different graphical primitives offer a real adequate way to consider the meta-data in semantic structures, but how can a less complex way be used for semantic content with the same benefits? A successive buildup of the treemap reduces the complexity of the visualization. So the graphical metaphor of the treeview visualization helps to build up a map of semantics content. Semantic data offer a great value of taxonomical structure, which can be used for navigation in Semantics and got through the data in a top-down metaphor. The other surplus of a treeview is that the user is choosing the way a treemap is built up. By his navigation through the semantics the treemap only shows the relevant aspects and information for the user and his interests.

The Semantic Map (seMap) uses these two metaphors to combine the surpluses for a special case: the usage of semantic annotated data and the implicit impartation of knowledge. Figure 2 shows a screenshot of seMap.

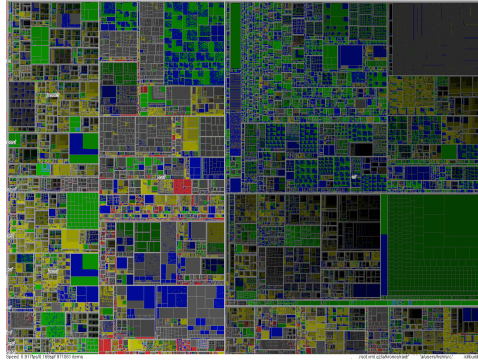


Fig. 1. Shneiderman's treemap with huge amount of data [14]

Graphical primitives like color, order and size are used to communicate relevance information in a way the user can fast and proper percept it. The relevance metrics are based on the semantics data and user specific information which are gathered from recommender systems. The primitive color indicates user specific relevance whereas the primitives order and size are determined by a combination of user- and data-based relevance. The primitive order arranges the most relevant element next to the selected element of the last row, as the arrangement of the blue elements in figure 2 indicate. Fitt's Law [5] has been considered for order and size to speed up the navigation process while reducing interaction error rate.

3.5 Learning by Exploring

In contrast to the Shneiderman's treemap the seMap is built up by the navigation of the user. The user plays an active role in the buildup-process. He navigates through the hierarchy of the semantics for doing any process with semantic data, for example for searching a specific content. By navigating through these structures he gets more information about the searched content, e.g. he searches for a specific, which is given in different disciplines or domains. The user gets information about the higher level meaning of a term, e.g. he searches for the document containing information about usability, the user gets different higher-level information to precise his search, e.g. he is able to read documents from psychology or computer sciences. So the implicit information he gets here is that there are different disciplines, which work on this topic. The more the user knows, the more structured are the presented information and navigation abilities, so in each case a learning process is given, while the learner explores the content.

This effect can be amplified with visualizing the first or second degree of the same level semantic relations; by using graph-layout algorithms. This is the most detailed way of presenting implicit knowledge, without focusing on imparting knowledge. The user gets the first semantic relations of a given content to other contents and is able to view similarities and differences.

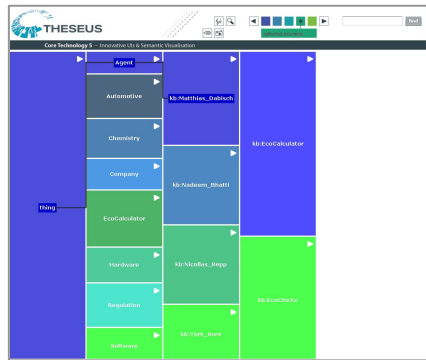


Fig. 2. Screenshot of seMap at the fourth level of detail

The following figure 3 shows a screenshot of seMap with the combination with graph-layout algorithms.

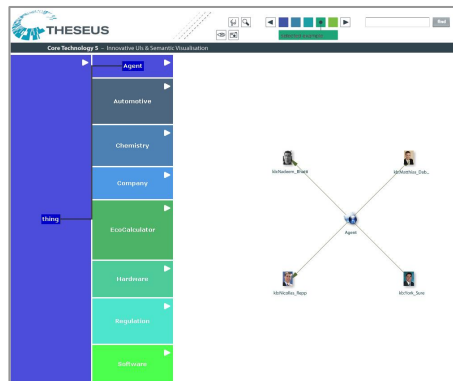


Fig. 3. Screenshot of seMap combined with graph-layout algorithm

4 SeMap for Learning

SeMap had been developed within the program THESEUS. This program consists of six Use Cases developing approaches for different scenarios. As central package the Core Technology Cluster develops general solutions which can be integrated into the applications of the Use Cases and be adapted to the specific requirements. Alexandria is one of these Use Cases creating a consumer oriented knowledge database wherein users can publish, process and search. Using seMap in Alexandria the consumers implicitly internalize knowledge while interacting with the database.

A case study of using seMap in ALEXANDRIA had been the visualization of a dataset about the person Napoleon. The dataset also included associated information like for example friendships and liaisons and had been enriched with timestamps. Furthermore actual sources like books or movies about the person Napoleon are integrated two. Based on this case study we are actually performing an evaluation.

5 Evaluation

Based on the case study of visualizing semantic database of the Use Case ALEXANDRIA by seMap we are actually performing an evaluation. The evaluation measures the learn effect of the users who interacted with the visualization. Even so the evaluation process is not finished yet the intermediary results predict well effects of remembering detailed information the user had not known before.

At this time 6 test persons had been interviewed before and after the usage of seMap. Most of them were able to remember 3 of 4 knowledge items they explored implicit while looking for a desired element in the data set. All further persons were able to answer 2 of 4 questions adequate. The questioned facts had not been part of the item which they were searching for, but these facts had been passed while looking for the desired element.

These intermediary results are to be verified in the ongoing evaluation process. In addition the learn effect has to be differentiated into short-term and long-term memory recognition. But at the moment the intermediary results of the implicit learn effect while exploring knowledge using seMap show great promises.

6 Discussion

SeMap is a visualization concept for visualizing semantic knowledge in a way the user can fast and proper percept it. Even more the visualization structure is adapted to the user's mental map in mind to support the process of imparting knowledge. Combined with other visualization techniques the information space can be visually structured into different regions which visualize a specific abstraction level. So the internalization of the knowledge is supported more. This visualization concept combined with the usage of graphical primitives in a way the constraints of Fitts' Law are considered lets the user percept information aspects implicit.

We are performing an evaluation to identify the benefits of seMap in the field of imparting knowledge. In this evaluation the knowledge internalization using existing visualization techniques is be compared to the knowledge internalization using seMap. Intermediary results of this evaluation regarding the implicit learn effect show great promises.

7 Conclusion

In this paper a visualization technique had been introduced which uses the surplus of semantics for imparting knowledge. For the purpose of imparting knowledge main criteria had been identified. Existing visualization techniques were examined to fulfill these main criteria. The visualization techniques Treeview and TreeMap, drafted by Ben Shneiderman, contain the most adequate aspects to fulfill these criteria. Hence the advantages of these techniques were combined to a new visualization concept of semantics for imparting knowledge. This concept is called seMap which visualizes semantics as maps.

References

1. Artino Jr., A.R.: Cognitive Load Theory and the Role of Learner Experience: An Abbreviated Review for Educational Practitioners. *AACE Journal* 16(4), 425–439 (2008)
2. Carol, J.J., Klyne, G.: Resource Description Framework: Concepts and Abstract Syntax. W3C Recommendation (2004), <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>
3. Ware, C.: Information Visualization: Perception for Design. Morgan Kaufmann, San Francisco (2000)
4. Federal Ministry of Economy and Technology, Bundesministerium für Wirtschaft und Technologie (BMWi), <http://www.bmwi.de/>
5. Fitts, P.M.: The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology* 47, 381–391 (1954)
6. Freire, M., Rodríguez, P.: Preserving the mental map in interactive graph interfaces. In: Proceedings of Advanced Visual Interfaces. ACM, New York (2006)
7. Herman, I., Melancon, G., Marshall, M.S.: Graph Visualization and Navigation in Information Visualization: A Survey. *IEEE Transactions on Visualization and Computer Graphics* 6(1), 24–43 (2000)
8. Katifori, A., Halatsis, C., Lepouras, G., Vassilakis, C., Giannopoulou, E.: Ontology Visualization Methods – A Survey. *ACM Computing Surveys* 39(4) (2007)
9. Paraskakis, I.: Ambient Learning: a new paradigm for e-learning. In: 3rd International Conference on Multimedia and Information & Communication Technologies in Education (2005)
10. The Personalization Consortium, <http://www.personalization.org/personalization.html>
11. Shneiderman, B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: Proceedings 1996 IEEE, Visual Languages, pp. 336–343. IEEE, Washington (1996)
12. Shneiderman, B.: Tree Visualization with Tree-Maps: 2-d Space-Filling Approach. *ACM Transactions on Graphics* 11(1), 92–99 (1992)
13. THESEUS Program, <http://www.theseus-programm.de>
14. UDLA Universidad de las Americas Puebla, Image Source: <http://ict.udlap.mx/people/carlos/is215/images/visualization/>
15. W3C RDF Specification, <http://www.w3.org/RDF/>
16. World Wide Web Consortium, <http://www.w3.org/>