DelViz: Exploration of Tagged Information Visualizations

Mandy Keck, Dietrich Kammer,
René Iwan, Rainer Groh

Technische Universität Dresden
Professur Mediengestaltung
Nöthnitzer Str. 46
01187 Dresden
mandy.keck@tu-dresden.de

Severin Taranko

queo GmbH
Tharandter Str. 13
01159 Dresden
s.taranko@queo-group.com

Abstract: Classification methods such as social tagging provide an easy way to annotate documents with metadata and describe them from various points of view. Frequently used visualization methods like tag clouds offer a limited access to the network of tags and fail at illustrating implicit relationships. Developing user interfaces which reveal these relationships and support the exploration of the resulting data structure is a considerable challenge. To address this problem, we present an interface concept, which illustrates the relationships between tags and supports different search strategies to explore a tagged item collection. The web application DelViz allows search and analysis in a collection of information visualizations.

1 Introduction

In the research field of information visualization numerous visualization and interaction techniques have emerged. Combining these techniques in an appropriate way or finding suitable information visualizations for a given context presents a significant challenge. The visualization tool DelViz (*Deep Exploration and Lookup of Visualizations)* supports searching for information visualizations from various points of view. Furthermore, it reveals hidden relationships in a tagged collection of visualization projects to allow an in-depth analysis of the connections between the involved techniques. For this purpose, the underlying data, the visualization methods, and the interaction mechanism are considered.

We first present the underlying data structure of DelViz which is affected by two classification methods: faceted classification and social tagging. In section 3 we present related work regarding our concept. Interfaces focusing search or analysis tasks are distinguished. We then present DelViz and its interface elements, which combines different approaches to search and analyze tagged information objects.

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4
weitere Artikel online:
http://informatik2011.de/519.html

2 Between Faceted Classification and Social Tagging

The data set used by DelViz currently contains 700 visualization projects which are stored with title, description, preview picture, and link to a demo of the information visualization or to a related website.

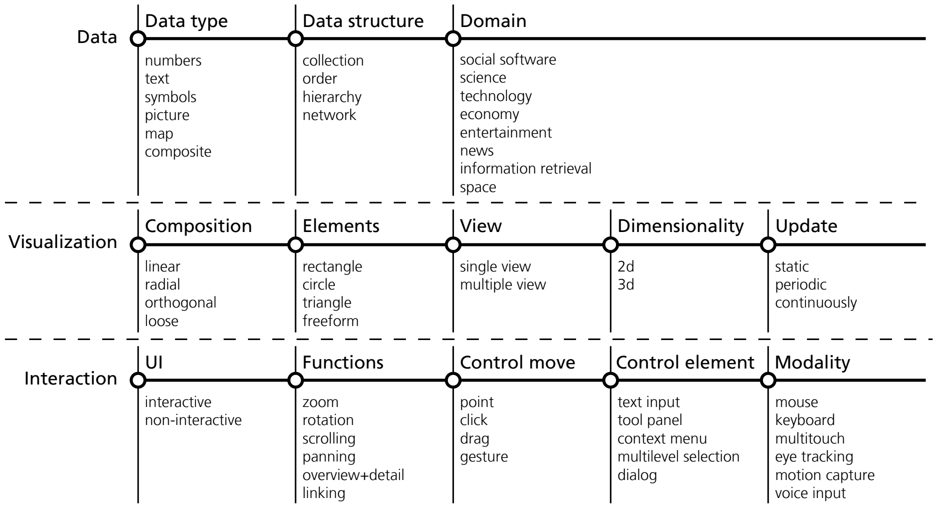

Figure 1: The DelViz classification schema contains different categories, dimensions, and tags to describe information visualizations from various points of view [Ke10]

We catalogued findings from the web and information visualizations from the "visual complexity" collection [Li10], which offers a broad range of network visualizations. For structuring the visualization collection, we considered several information visualization taxonomies (e.g. [Sh96], [Le07], [Ch00]) and visualization systems (e.g. [Vi07], [Li10]). Many of these taxonomies focus either the underlying data structure or applied visualization techniques. Our goal is to offer a more flexible classification of information visualizations that allows the analysis from various points of view. In addition, we use a simple vocabulary to support people without an academic background in using our tool (e.g. the term "collection" is used instead of nominal data, "order" instead of ordinal data, etc.). Taking these goals and the considered taxonomies into account, we created a multidimensional classification schema for structuring the visualization collection (cp. Figure 1).

This schema represents a set of rules to assign tags to information visualizations. It contains different dimensions of an information visualization, which can be assigned to three categories according to the reference model for information visualization [CM99]: (1) the underlying Data (category "Data"), (2) the visual representation of the data (category "Visualization") and (3) the interaction techniques to explore and manipulate the visualization (category "Interaction"). The included dimensions (e.g. "Data type" in the category "Data") and values (e.g. "numbers" in the dimension "Data type") combine two classification methods.

The values correspond to tags, which are used in social tagging [Ma04]. In this classification method, an arbitrary number of tags are collaboratively created and assigned to an information resource. The dataset resulting from the tags by all involved persons is also called Folksonomy or user-generated classification [Qu05]. The collaborative knowledge is used to describe a domain from various points of view. However, this uncontrolled vocabulary leads to a number of limitations and offers restricted possibilities to compare information visualizations. Therefore, the presented classification schema is influenced by faceted classification [Ra62]. This classification method describes items though a combination of facets where each facet addresses a different conceptual dimension relevant to the collection [Qu05]. In contrast to faceted classification, the classification schema in Figure 1 is less stringent and allows multiple or no assignments in one dimension. For example, the data types “picture” and “text” can be assigned to information visualizations, if both are addressed. However, if compound objects are used as smallest units of displayed data, the “composite” tag is applied instead of the various attributes. In the same manner, no tag can be assigned in a dimension – for instance if a visualization project is classified as “non-interactive”, no other dimension in the category "interaction" is assigned.

The tagging process was conducted by a team of five researchers, providing their expert opinions on how to classify each visualization project. With the online publication of the database, we plan to integrate more users, who can add and classify visualization projects. This approach of social tagging with an initial classification should help to enhance the presented classification schema and to consider the vocabulary of a broad community of users that is part of future work (cf. section 5).

3 Related Work

In this section, related interfaces are examined with regard to their suitability for searching and analyzing a structured information collection as presented in section 2.

A common practice to represent a tag collection is the use of tag clouds. In this visual representation, the tags are normally listed alphabetically and the popularity of each tag is represented by font size or font weight. This visualization method (e.g. used at del.icio.us or flickr.com) is useful for gaining an overview of the folksonomy and quickly perceiving the most prominent tags. But they neither allow the analysis of the relationships between the tagged information objects nor do they support complex search tasks.

The visualization concept "Web Trigrams” [Ha11] supports the analysis of the relationships between different terms. The visualization in Figure 2 (left) compares a set of trigrams, starting with the words ‘I’ and ‘You’. The frequencies of the second and third word in the trigrams were combined and sorted, and rendered in decreasing order according to the frequency-of-use. The color-coded lines act like paths, connecting the words and enumerating all trigrams [Ha11].

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4
weitere Artikel online:
http://informatik2011.de/519.html

"Elastic Tag Maps" present another approach: a mapping algorithm for analyzing and illustrating emergent structures in a tag collection. The visualization uses a two-dimensional map, which places frequently co-occurring tags closely together. Highlighting one of the tags via rollover brings related tags to the front. Additionally, line thickness indicates the strength of the relation [St07].

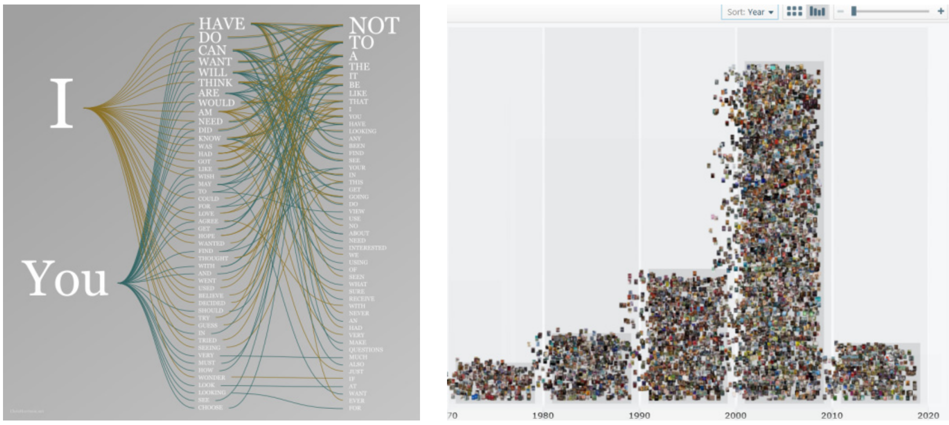

Figure 2: Web Trigrams visualize connected words with Bezier curves [Ha11] (left), Pivot View of Netflix movie database with animated state changes [Na10] (right)

Both visualization concepts support the analysis of terms or tags, but are not suitable for searching tasks, because they do not allow the exploration of the underlying data collection. The principle of Faceted Browsing is a popular interface paradigm to explore structured data collections. It is based on faceted classification (cf. section 2) and is suitable for this context due to the faceted structure of the proposed schema. Faceted Browsing supports the construction of complex search queries by selecting values of the data facets. The result set is restricted or increased iteratively with every selection or deselection. The user does not need to know exactly what she is looking for and recalling terms for searching is not required [Po09]. Thereby, Faceted Browsers are well suited for the exploration of structured data sets. Well-known examples are the Faceted Browser Flamenco [Ye03] and Elastic Lists [SM07]. Elastic Lists allow analytical tasks, too. The size of an item highlights predominant values similar to the font size of tag clouds. Additionally, a brighter color indicates that the proportional weight is significantly higher than compared to the global profile [St07].

Both Faceted Browsers allow viewing the current result set as list, whereas the approach of Pivot [Na10] allows browsing a huge result set of images. To gain an overview of the whole selection, all images are presented as thumbnails in the initial state. The user can zoom into this collage of images to see individual pieces of data more closely, or zoom out to see items grouped according to various criteria. During sorting or grouping the image collection, the change of state is animated to make the changes comprehensible to the user (cp. Figure 2, right).

ISBN 978-3-88579-286-4

4 DelViz: Deep exploration and lookup of Visualizations

As described in section 1, DelViz is intended to support different search tasks: finding suitable information visualizations for a given context, and analysis of the underlying structured data set to discover relationships between the visualizations. Different search activities are considered to support search in the information visualizations. Marchionini distinguishes lookup activities and exploratory search [Ma06]. On the one hand, Lookup is the most basic kind of search task and involves activities like "Known item search" and "Fact Retrieval". On the other hand, exploratory search describes a more complex search task such as "Knowledge acquisition", "Comparison" or "Analysis" which are associated with learning or investigating activities. Additionally, two different search strategies can be distinguished: analytical and browsing strategies. Analytical strategies depend on careful planning, the recall of query terms, and iterative query reformulations and examinations of results. Browsing strategies are heuristic and opportunistic and depend on recognizing relevant information [Ma95]. Since Lookup tasks are suited for analytical search strategies, exploratory search blends analytical and browsing strategies.

To support these different search tasks, the application offers two flexible areas: the representation of the categories, bundles and tags on the left-hand side, and the information visualizations presented as thumbnails on the right-hand side. A splitter in the middle can be used to expand one of these two areas to change the level of detail on either side. An enlarged right area (cp. Figure 3, left) automatically displays larger preview pictures of the information visualizations and supports the browsing strategy. Two equal areas support exploratory search, where both sides (tags to construct the search query and visualizations as results), are considered in the search process (cp. Figure 3, middle). An enlarged tag area automatically displays arcs between the selected and remaining tags to analyze their relationships (cp. Figure 3, right). A detailed description of these three concepts will be provided in the following sections.

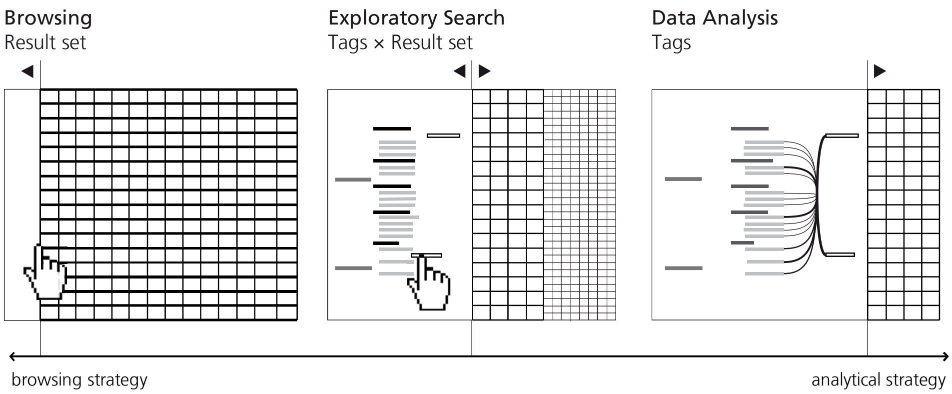

Figure 3: The application is divided into two flexible areas (the representation of tags and information visualizations) to support different search tasks

4.1 Browsing concept

According to the visual information seeking mantra [Sh96], the browsing concept initially represents a preview of all information visualizations to gain an overview of the entire collection. A zoom function is offered to examine an item or group of interest (cf. Pivot in section 3). The highest zoom level displays a gallery view, where details of the selected visualization project are presented. To support fast navigation through the visualization collection, a browsing function is available in the gallery view. The user can switch between different modes to view the result set (cp. Figure 4). In addition to gallery mode, a result set mode and a history mode are provided to explore the visualization collection depending on the current search query (cf. section 4.2).

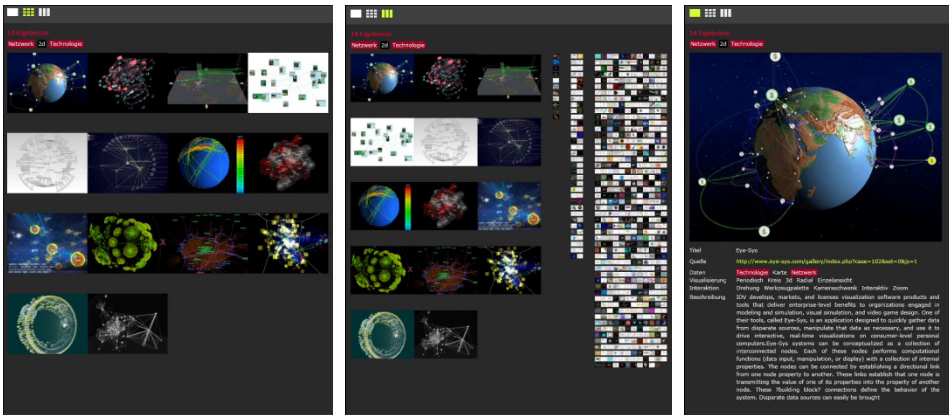

Figure 4: The result set can be regarded in three different views: the result set mode (left), the history mode (middle) and the gallery mode (right)

4.2 Exploratory search

In order to support tasks like finding a suitable visualization for an existing set of data, the user is able to filter the entire collection. Tags can be selected and removed by dragging them either to the right-hand side or left-hand side (cp. Figure 5). The facet classification is displayed on the left of the tag list to clarify the context of each tag. The combination logic within a facet is logical conjunction, i.e. the AND operator. The result set is subsequently updated to correspond with the query and hence the collection is elaborated step-by-step and convenient visualizations can be picked (cf. section 3, *principle of Faceted Browsing*). The impact of applying these filters on the collection is visualized in history mode, in which the user examines both the current result set and the subsets that are filtered out in every search step. A set of interest can be magnified by clicking on it with the mouse in order to get a closer look of the associated visualizations. The whole process of filtering the result set makes extensive use of animations as an aid to understand transitions, comprehend the effects of filters steps, and to improve the general usability [CK09].

ISBN 978-3-88579-286-4

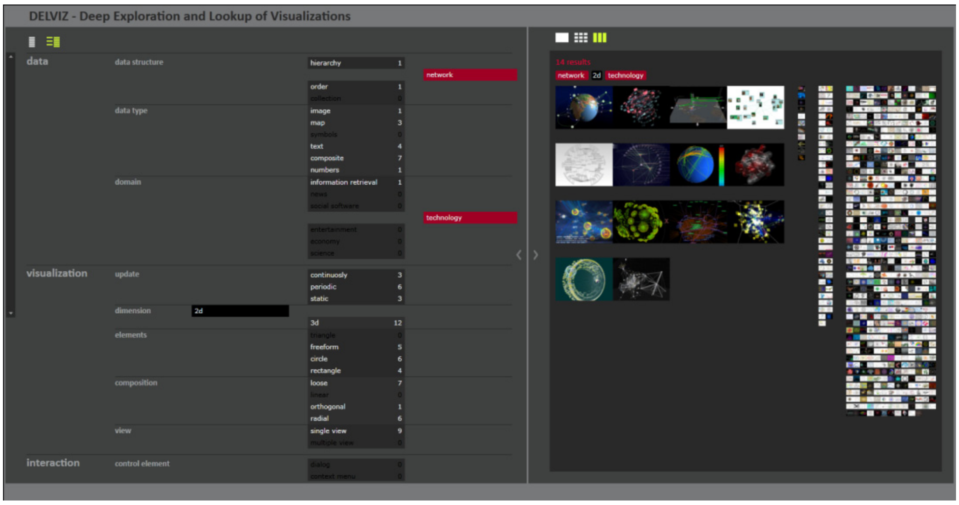

Figure 5: Exploratory search concept with history mode: the left-hand side displays the selected (red) and removed tags (black), the right-hand side is responsible for the analysis of generated subsets

4.3 Data Analysis

As soon as the left area reaches a certain size, arcs are displayed between the selected and remaining tags (cp. Figure 7). Arcs, represented as Bezier curves, are used to illustrate the relationships between tags. By using the presented classification schema in section 2, different relationships arise between the visualization projects and tags, which can be used for the analysis of the database.

These relationships are presented in Figure 6 (left) can be distinguished in direct or indirect relationship. The direct relationship is achieved by the given structure of the schema consisting of categories, dimensions and tags and by the assignment of relevant tags to a visualization project. Indirect relationships result from the use of the same combination of tags in different visualization projects. This is illustrated in Figure 6 by two different tags and visualizations. The more often two tags are assigned to a visualization project in conjunction, the stronger their relationship. To immediately identify strong and weak relationships, the weight of each relationship is mapped to the edge thickness and transparency of the Bezier curve. For the calculation of the relative weight of an edge, the Jaccard coefficient is used [HH10] that measures similarity between sample sets and is defined as the size of the intersection divided by the size of the union of the sample sets. In contrast to “Web Trigrams” (cf. section 3), not all relationships are displayed by default. Similar to “Elastic Tag Maps” (cf. section 3), specific relationships are visualized by selecting one or more tags. In Figure 6 (right), the tag “technology” is selected (red) and indicates a strong relationship to the tags ”network”, “composite”, and “3d” as shown with the comparatively thick and bright Bezier curves. In contrast, there is no relationship to the tags “triangle” and “economy”. Consequently, no arc is displayed.

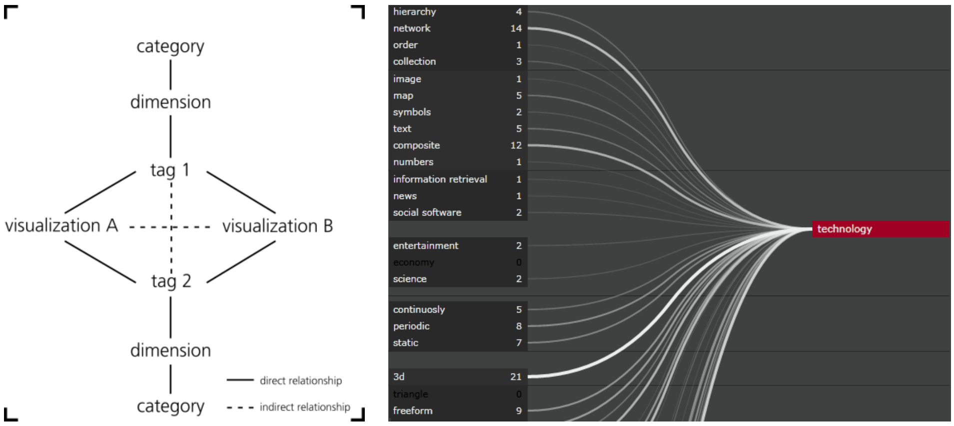

Figure 6: Relationships between the tag data (left), Bezier curves between the selected tag "technology" and other tags – tags without a relationship to the selected tag are presented with black font color and no Bezier curve is displayed

If multiple tags are selected, the Bezier curves are merged to bundled relationships. In this case, the individual relationships between two tags can be displayed using mouse over. In Figure 7, the tags "network" and "technology" are selected (red) and indicate a strong relationship to the tag "3D". The excluded tag "2D" (black at the left side) shows the relationship to the rest of the tag collection when hovered with the mouse. The relations between tags can help users to identify interesting tags for the next filter step. Selecting tags with a strong connection generates more results and thus shows commonly employed combinations of tags in visualization projects. Adding less connected tags produces fewer results, representing exceptional examples regarding the combination of data, interaction, and visualization techniques. Apart from supporting the search, the Bezier curves can be used to show suitable combinations of data, visualization, and interaction techniques, which can support the designer of information visualizations.

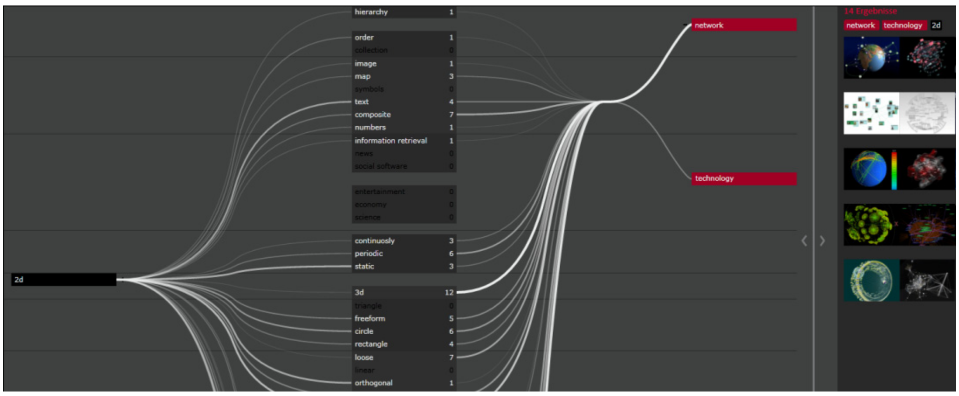

Figure 7: Bezier curves are used to illustrate the relationships between selected tags (red), removed tags (black), and remaining tags (gray).

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4

5 Conclusion and Future Work

In this paper, we presented a concept to browse, search, and analyze information visualizations. The concept is implemented as web application in Microsoft Silverlight and uses the Silverlight control "deep zoom", as shown in the Pivot example in section 3. A live demo is available online: http://www.delviz.com. DelViz is based on a client-server architecture with a Microsoft SQL Server as relational database. The Windows Communication Foundation (WCF) provides the web services for information exchange between the database and the Silverlight client. Inviting a larger community can enhance the collection of information visualizations. This is projected for the future.

The design and implementation of additional views onto the data set is part of future work as well. Also intended is the evaluation and enhancement of the presented classification schema. The approach of social tagging with a predefined classification schema used for the DelViz database also needs to be evaluated in a user study with an adequate number of participants.

Acknowledgements

This work has been supported by the European Union and the Free State Saxony through the European Regional Development Fund (ERDF).

References

[Ch00] Chi., E.H.: A Taxonomy of Visualization Techniques Using the Data State Reference Model. In Proceedings of the IEEE Symposium on Information Vizualization 2000 (INFOVIS '00). IEEE Computer Society, Washington, DC, USA, 2000

[CK09] Cockburn, A., Karlson, A., Bederson, B. B. A review of overview+detail, zooming, and focus+context interfaces. ACM Comput. Surv. 41, 1, Article 2 (January 2009).

[CM99] Card, S.K., Mackinlay, J., Shneiderman, B. Readings in Information Visualization: Using Vision to Think. San Franciso: Morgan Kaufmann, 1999, S. 7ff.

[Ha11] Harrison, C.: Web Trigrams: Visualizing Google's Tri-Gram Data, http://chrisharrison.net/index.php/Visualizations/WebTrigrams, [24.06.2011]

[HH10] Hassan-Monteroa, Y.; Herrero-Solanaa, V.: Improving Tag-Clouds as Visual Information Retrieval Interfaces. International Conference on Multidisciplinary Information Sciences and Technologies. Mérida, Spain: InSciT2006, 2006, S. 1-6

[Ke10] Keck, M. et.al.: DelViz: Untersuchen von Visualisierungsformen durch eine Klassifizierung beruhend auf Social Tagging, In: Virtual Enterprises, Communities & Social Networks: Workshop GeNeMe '10 , TUDpress , Dresden , 2010 , S. 69-79

[Le07] Lengler, Ralph und Eppler, Martin J.: Towards a periodic table of visualization methods. Florida: In Proceedings of the 2007 IASTED Conference on Graphics and Visualization in Engineering, 2007

[Li10] Lima, M.: Visual Complexity. 2010. www.visualcomplexity.com, [20. 10 2010].

ISBN 978-3-88579-286-4

[Ma95] Marchionini, G.: Information seeking in electronic environments. New York, USA: Cambridge University Press, 1995

[Ma04] Mathes, A.: Folksonomies - Cooperative Classification and Communication Through Shared Metadata. University of Illinois Urbana-Champaign: Computer Mediated Communication - LIS590CMC, 2004

[Ma06] Marchionini, G.: Exploratory search: from finding to understanding. Communications of the ACM. New York, USA: ACM, 2006, S. 41-46

[Na10] Naone, E.: Making Sense of Mountains of Data - A new tool explores large sets of data--and might help organize the Web. 2010. http://www.technologyreview.com/web/24645/page1/ [24. 06 2011]

[Po09] Polowinski, J.: Widgets for Faceted Browsing. Human Interface and the Management of Information. Designing Information Environments, LNCS 5617 proceedings. Heidelberg: Springer-Verlag, 2009, S. 601-610

[Qu05] Quintarelli, E.: Folksonomies: power to the people. 2005 www.iskoi.org/doc/folksonomies.htm [13.10 2010]

[Ra62] Ranganathan, S.R.: Elements of Library Classification. Bombay, Calcutta, New Delhi, Madras, London and New York: Asian Publishing House, 1962.

[Sh96] Shneiderman, B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization. Los Alamitos, CA: IEEE Symposium on Visual Language, IEEE Computer Societ, 1996, S. 336-343

[SM07] Stefaner, M.; Müller, B.: Elastic lists for facet browsers. 18th International Conference on Database and Expert Systems Applications (DEXA 2007). Regensburg, 2007, S. 217-221

[St07] Stefaner, S.: Visual tools for the socio–semantic web. Master's Thesis. University of Applied Sciences Potsdam, 2007

[Vi07] Viegas, Fernanda B., et al.: ManyEyes: a Site for Visualization at Internet Scale. Los Alamitos, CA: IEEE Transactions on Visualization and Computer Graphics, IEEE Computer Society, 2007. S. 1121-1128.

[Ye03] Yee, K.-P.; Swearingen, K.; Li, K.; Hearst, M.: Faceted Metadata for Image Search and Browsing. ACM Press, 2003, S. 401-408

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4