Visual Innovations for Product Search Interfaces

Mandy Keck[1], Martin Herrmann[1], Dana Henkens[2], Severin Taranko[2], Viet Nguyen[3], Fred Funke[3], Steffi Schattenberg[3], Andreas Both[3], Rainer Groh[1]

[1]Chair of Media Design, Technische Universität Dresden, 01062 Dresden
{mandy.keck, martin.herrmann, rainer.groh}@tu-dresden.de
[2]queo GmbH, Tharandter Str. 13, 01159 Dresden
{d.henkens, s.taranko}@queo-group.com
[3]Unister GmbH, Barfußgäßchen 11, 04109 Leipzig
{viet.nguyen, fred.funke, steffi.schattenberg, andreas.both}@unister-gmbh.de

Abstract: As the 'gatekeepers' of the internet, search engines provide access to a huge amount of data. Nowadays, many search engines are able to interpret user queries and automatically suggest additions or corrections. Despite these technological advances, the search paradigm itself remains stable for most engines. Especially for users with little knowledge of the search domain or a very vague idea of what they are looking for, it is difficult to formulate a query with the established keyword search paradigm. Faceted Search is one of the techniques that facilitate the process of query formulation by offering a structured and classified overview of the domain. In this paper, we give an overview of search paradigms besides keyword-based techniques and will present alternative visual approaches for product search.

1 Introduction

In the last decade search engines have become the 'gatekeepers' for the contents of the internet. Since the first search engine 'Archie' in early 1990[1], which was able to search files in remote public FTP directories with an 8 character search string, the scope of functions supporting search activities have grown enormously to compete with the tremendously rising amount of data. Every day, 2.5 quintillion bytes of data are created, so much that 90 percent of the data in the world today has been created in the last two years alone[2]. Mobile devices, social networks and real-time information are driving forces for this developement. Hence, the size of items, the variety of sources as well as the overall amount of data will continue to grow. With this development, accessing relevant information is an increasingly difficult problem [CHJR+13].

To address this problem, many websources (e.g. DBpedia[3]) have started to publish structured data on the web according to Linked Data principles. The so called 'Web of Data' contains billions of RDF triples describing relations between entities and connecting dif-

[1]http://people.lis.illinois.edu/~chip/projects/timeline/1990archie.htm, last accessed: 23.06.2014
[2]http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html, last accessed: 23.06.2014
[3]http://dbpedia.org, last accessed: 23.06.2014

ferent data sources [BBBE11]. This large knowledge base (open linked data) can be used to structure and contextualize unstructured data like emails or images and facilitate real time decision making. However, without an appropriate interface, it is hard to formulate queries with complex RDF query languages like SPARQL.

Appropriate visual query systems (VQS), focusing on expressiveness and usability, are needed to transform a request of a non-technical user into a valid query for the underlying database [SGJR+13]. Especially when the user does not have a precise idea of his desired search result and is not familiar with the search domain, the system should provide a vocabulary to prevent incomplete or ambiguous queries.

In this paper, we focus on users with a vague information need in product search scenarios. We discuss different visual approaches and interaction strategies to enable the user to put his desires in concrete terms. Furthermore, we classify these approaches and emphasize the use of contextualized and structured data to improve the search experience and the quality of results.

2 Search Paradigms

Lookup and Exploratory Search are search activities that can be distinguished in the context of web search [Mar06]. Lookup describes the most basic kind of search tasks such as fact retrieval, known item search, and question answering. Exploratory Search is a more complex process that usually starts with a vague information need and therefore requires multiple iterations of learning, investigation, and reformulation of the search query.

Our goal to support product search based on vague ideas of the possible result can also be described as Motive-based Search [KHB+13]. Motive-based Search refines Explorative Search by specifying the motive and the aim of the process to find a suitable product. A motive can be defined as the reason for a search as well as particular conditions like how much a product should cost.

Thus, search interfaces are required that support the user to express his information need based on these reasons and conditions. Therefore, two search paradigms are well established in the web search world: Direct Search and Navigational Search [Tun06]. Direct Search allows users to simply write their queries in a text box and became enormously popular with web search engines, such as Google and Yahoo! Search. Text boxes and search forms are well-suited for lookup-scenarios, in which the user has a concrete idea of the desired product (e.g. looking for a flight to London). However, in motive-based search scenarios, it is difficult to transform a possibly vague information need into a specific query. In contrast, Navigational search systems provide guidance through the use of a taxonomy [Tun06]. Common web interfaces often use hierarchical structures to browse the information space in a predetermined order. To allow multiple access points for the search, the principle of Faceted Browsing provides an alternative that is based on a multidimensional taxonomy [ST09]. It enables users to explore the information space by progressively refining their choices in each dimension [RR11].

The described search functionalities are based on the principle of reduction where the results from the set of all resources are progressivley narrowed down to the desired set of answers. Clemmer and Davies define the term 'general-to-specific' (GtS) to refer to query interfaces that allow the user to specify abstract criteria for a result set [CD11]. By contrast, they define 'specific-to-general' (StG) to refer to interfaces that explicitely support starting with an example and generalzing it to find similar examples. This interface paradigm is based on expansion where aspects of a concrete example are progressively generalized to find other result that match the pattern [CD11].

In the following, some examples for these search paradigms that match to our motive-based search scenario are described. For instance, the Faceted Browsers Elastic Lists [SM07] and Flamenco [YSLH03] can be classified according to the GtS-Paradigm and Navigational Search. Both support the construction of complex search queries by selecting values of the data facets, thereby restricting the result set iteratively and allowing to browse through the information collection. Elastic lists offer a visual interface that allows additional analytical tasks (e.g. the size of each facet value highlights predominant values) [SM07]. Compared with this, Flamenco provide a textual interface with additional search functionalities such as Keyword Search and Similarity Search. Hence, it combines all of the mentioned paradigms: GtS, StG, Navigational Search, and Direct Search. Faceted Browsers are well suited for the exploration of structured data. For semi-structured data, Similarity Search is an effective way to start or continue a search. Similarity searches or search by example approaches aim to find objects with attributes as close as possible to a reference object and can be classified according to the StG paradigm. Whereas Flamenco and the fotosharing service Flickr[4] allow to find items with similar properties based on the selection of tags (Navigational Search), other search by example approaches solve query formulation problems with multimedia queries such as last.fm[5], retrievr[6], google image search[7] and picalike[8]. Last.fm is a music discovery service that gives personalised recommendations based on the music the user is listening to or looking for and thus combine Direct and Navigational Search. Retrievr enables the user to define a vague search query by using sketches. Google image search and picalike are search by example approaches that allow a similarity search based on pictures and belong to direct search interfaces. Picalike also incorporates a product search that detects products similar to the reference object. The results can be filtered by various criteria such as brand and price. Hence, the user can switch to a search paradigm using Navigational Search.

Both the GtS and StG paradigms are well-suited for motive-based search scenarios to narrow down or broaden the scope. Combined, they could become a powerful tool to support the user during an exploratory search process with frequently changing information needs.

[4]https://www.flickr.com/photos/tags/search/clusters/, last accessed: 23.06.2014
[5]http://www.last.fm, last accessed: 23.06.2014
[6]http://labs.systemone.at/retrievr, last accessed: 23.06.2014
[7]http://images.google.de/, last accessed: 23.06.2014
[8]http://www.picalike.com/products/similarity-search.php, last accessed: 23.06.2014

3 Visual Exploration Systems for Product Searches

In this chapter, we present three product search interfaces different from conventional form-based interfaces with text boxes, radio buttons and drop down menus which are developed in our research project "VISEA - Visual Search Interfaces"[9]. Each of the suggested approaches support a particular search strategy to find the desired product in a very large database with structured or semistructured data. Further on, we consider different search domains such as travel search and financial products. Since both domains entail different user requirements and expectations, the resulting interface concepts support emotional or analytical decisions.

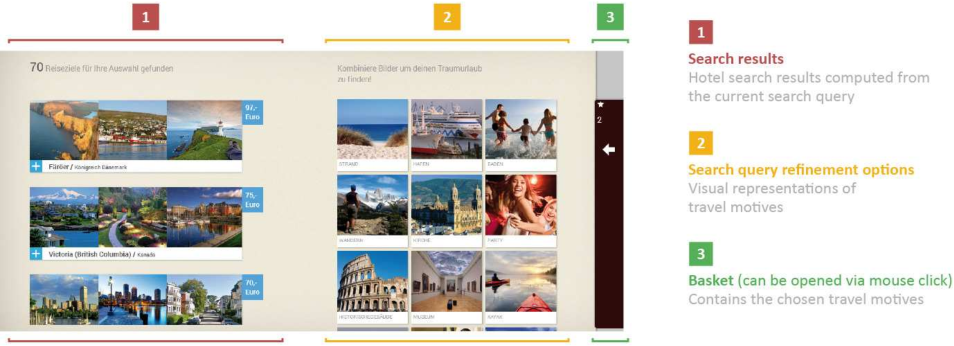

Figure 1: getInspired: interface with visual concepts for travel searches

3.1 Recommendation-based Search Approach

Figure 1 shows the interface 'getInspired' for travel search. We designed this interface to support and trigger emotionally driven decisions. Instead of a direct query on the attributes of the result set, the user communicates his preferences to the system through a selection of concepts represented by expressive images. Based on the previous decision of the user, the system decides which concepts are presented in the next refinement step. For each step, the interface provides the best concepts in terms of a fast reduction of the result set. In order to implement this interface, the results need to be tagged with concepts from a domain topology that serves as knowledge base. After each refinement step, the most applicable results are presented on the left side of the interface. These results can be explored or further specified by selecting more appealing concepts on the right side. The interface focuses on easy self-explanatory interactions and the use on mobile and touch interfaces. An initial user study with 29 participants and 12 different search-tasks was conducted to compare the recommendation-based approach to a strict navigational concept selection. The study indicates that the time and interactions needed to solve a given search task can be improved significantly. In order to validate these claims in real conditions, a more

[9] http://visea-projekt.de

extensive knowledge base as well as an automatic method for the classification of images are needed. These improvements are necessary to extend the users possibilities to express his needs and the systems ability to respond to users actions.

3.2 Analytical Approach using Parallel Coordinates

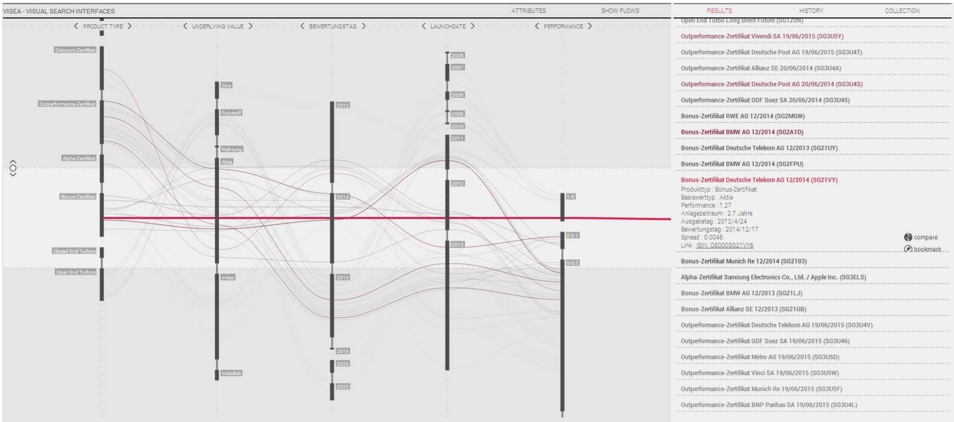

Figure 2: Interface concept based on Parallel Coordinates: the Similarity Search centers the selected product; a fuzzy filter (light grey) reduces the amount of polylines to focus on similar products

The interface presented in Figure 2 [KHB+14] is designed for an analytical aproach and elaborates ideas of the product search interfaces from the Product Explorer [ROF12]. It combines principles of Faceted Browsing with the visualization method of parallel coordinates [ID90] to support additional analytical tasks (cp. Elastic Lists in section 2). Each parallel axis represents one facet of the multi-dimensional data set and is spatially separated in container elements representing the facet values. The height of each container, in reference to the overall height of the axis, shows the distribution of data items within the facets and allows the analysis of predominant values. A product is visualized by a polyline that intersects each axis at the appropriate value. In addition to the analytical benefits of the parallel coordinate visualization presented in [ROF12], this interface supports facet filters (GtS) as well as a Similarity Search (StG). Multiple filters can be activated on one or different axes to create complex filters as known from Faceted Browsing to narrow the scope. With the Similarity Search approach, the user starts his search by selecting a product of interest matching his expectations in one or more characteristics. The function automatically centers all intersections of the selected product to the median line of the visualization. Hence, products with similar properties are arranged close to the center and can be easily identified. The interface is developed to explore financial data. In this use case, the user needs to get an overview of variety, range, and distribution of the provided dataset and their attributes. Consequently, the interface has a strong analytical focus and combines the StG and GtS paradigms to explore the information space.

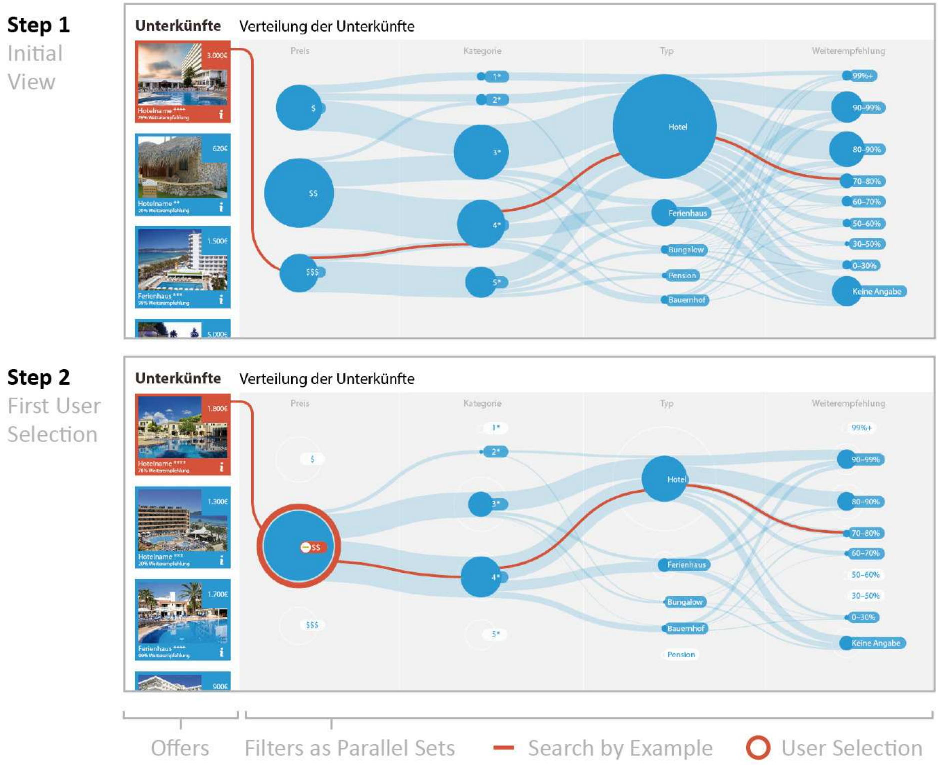

Figure 3: Faceted Browsing with Parallel Sets

As a variation of the parallel coordinates, parallel sets [KBH06] present sets of products instead of single polylines. Therefore, the containers are connected by polygonal streams representing the logical conjunction of the adjacent containers. The size of these streams indicate the frequency of items included in the conjunction. Figure 3 shows an interface designed for travel search based on the principle of parallel sets. Facets of each attribute are depicted as circles. The distribution of characteristic within an attribute is mapped to the size giving the user an impression on the most common values. In comparison to the parallel coordinates, parallel sets are independent to the number of visualized products and are a good possibility to gain insights in the overall structure of the dataset.

The user tests conducted in [KHB+14] have shown that parallel sets are faster and easier to understand with regard to faceted searches and analysis tasks, whereas parallel coordinates have advantages in comparison tasks and similarity-based searches.

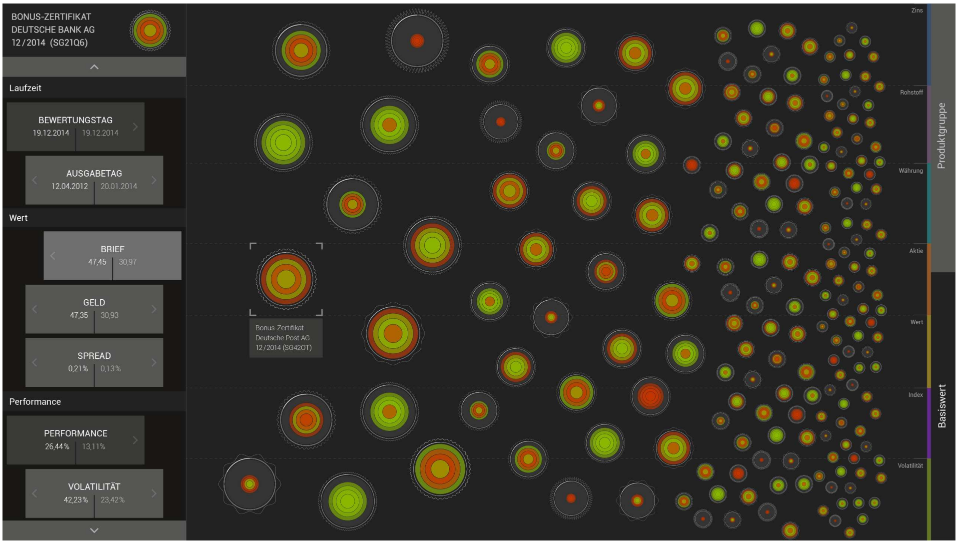

Figure 4: Visual Similarity Search: glyphs on the right side represent similar products, the attributes of the reference product (left side) can be selected or deselected to focus on properties of interest

3.3 Search by Example Approach

A visual approach for a similarity search in a financial product scenario is shown in Figure 4. Each circular glyph represents a single product and depicts its most important features, like performance over time and investment term. The horizontal positioning is used to express the similarity to a chosen reference object. Similar objects are displayed larger to show details (semantic zooming). The less significant results are smaller to give the user an overview of the data distribution. The global similarity calculation is determined by the ordinal attributes selected on the left side of the interface. The vertical positioning can be used to cluster products in categories aligned on the right side of the interface. The interface is designed for touch devices and a limited amount of items and attributes. An early prototypical implementation suggests that this kind of interface is well suited for the paradigm of search by example (StG) and the browsing of a large structured or semi structured dataset. Whilst the overall design of the interface may remain the same for any kind of structured dataset, the iconic representation should be adapted to the use case. With a more sophisticated version of our current prototype, we hope to create a flexible and playful way to find a suitable product and gain insights in distribution and structure in the complex area of financial products.

4 Conclusion

The presented interface concepts incorporate several search paradigms and strategies. The GtS paradigm is well-suited to narrow down large result sets of a known domain with the help of filters (Parallel Coordinates or Parallel Sets) or visual concepts (getInspired Interface). However, without a certain knowledge of the search domain, the user is often not able to specify abstract criteria to exploit the information space. Especially in the case of financial search, novices do not know which attributes such as performance or investment term would lead to a suitable financial product. For this use case, the StG paradigm (Search by example) is a promising approach. Here the user can start with one product to discover the meaning of the included product features and then generalize the interesting features to find similar results.

All three concepts can be classified into two groups: recommendation-based and analytical approaches. The recommendation-based approach (getInspired) hides the complex data structure of the search system and leads to a small result set in a few steps. The analytical approach (Parallel Coordinates or Parallel Sets) provides a deep insight into the data structure and enables an evaluation of the chosen product in comparison to the remaining information space. The presented search by example approach combines both strategies by offering products similar to the reference item, but also represents smaller glyphs to give an overview of the data distribution. The suggested usage of these strategies depends on the use case as well as the preferences and previous experience of the user. Whereas the recommendation-based approach allows to find a suitable product in very short time, the analytical approach requires more time and effort to explore and analyze the data set, but can lead to better and more trustworthy results.

The proposed examples indicate that the search experience and speed in product search scenarios can be significantly improved by visually appealing interfaces. Therefore, we suggest further research in this area. Future extensions need to focus on the backend infrastructure and the linking with sophisticated knowledge sources. An improved semantic evaluation of the users query and a better understanding of his current needs is the pathway towards an intelligent and user-adapted interface. Another challenge is to ensure the reliability of results and establish visual search interfaces as credible source of information and decision making in a big data environment with a huge variety of different information sources and formats. To this end, search tools have to find a good balance between deep insights in the underlying data sources and a user-friendly, easily comprehensible interface.

5 Acknowledgments

This work has been supported by the European Union and the Free State Saxony through the European Regional Development Fund (ERDF). The research presented in this article has been conducted in cooperation of the Chair of Media Design - Technische Universität in Dresden, Unister GmbH from Leipzig and queo GmbH from Dresden, Germany.

References

[BBBE11] Christian Bizer, Peter A. Boncz, Michael L. Brodie, and Orri Erling. The meaningful use of big data: four perspectives - four challenges. *SIGMOD Record*, 40(4):56–60, 2011.

[CD11] Aaron Clemmer and Stephen Davies. Smeagol: A âœSpecific-to-Generalâ Semantic Web Query Interface Paradigm for Novices. In Abdelkader Hameurlain, StephenW. Liddle, Klaus-Dieter Schewe, and Xiaofang Zhou, editors, *Database and Expert Systems Applications*, volume 6860 of *Lecture Notes in Computer Science*, pages 288–302. Springer Berlin Heidelberg, 2011.

[CHJR+13] Diego Calvanese, Ian Horrocks, Ernesto Jimenez-Ruiz, Evgeny Kharlamov, Michael Meier, Mariano Rodriguez-Muro, and Dmitriy Zheleznyakov. On Rewriting and Answering Queries in OBDA Systems for Big Data (Short Paper). In *OWL Experiences and Directions Workshop (OWLED)*, 2013.

[ID90] Alfred Inselberg and Bernard Dimsdale. Parallel Coordinates: A Tool for Visualizing Multi-dimensional Geometry. In *Proceedings of the 1st Conference on Visualization '90*, VIS '90, pages 361–378, Los Alamitos, CA, USA, 1990. IEEE Computer Society Press.

[KBH06] R. Kosara, F. Bendix, and H. Hauser. Parallel Sets: interactive exploration and visual analysis of categorical data. *Visualization and Computer Graphics, IEEE Transactions on*, 12(4):558–568, July 2006.

[KHB+13] Mandy Keck, Martin Herrmann, Andreas Both, Ricardo Gaertner, and Rainer Groh. Improving Motive-Based Search: Utilization of Vague Feelings and Ideas in the Process of Information Seeking. In Norbert Streitz and Constantine Stephanidis, editors, *Distributed, Ambient, and Pervasive Interactions*, volume 8028 of *Lecture Notes in Computer Science*, pages 439–448. Springer Berlin Heidelberg, 2013.

[KHB+14] Mandy Keck, Martin Herrmann, Andreas Both, Dana Henkens, and Rainer Groh. Exploring Similarity: Improving Product Search with Parallel Coordinates. In *Human Interface and the Management of Information. Information and Knowledge in Applications and Services*, volume 8522 of *Lecture Notes in Computer Science*, pages 160–171. Springer, 2014.

[Mar06] Gary Marchionini. Exploratory Search: From Finding to Understanding. *Communications of the ACM*, 49(4):41–46, April 2006.

[ROF12] Patrick Riehmann, Jens Opolka, and Bernd Froehlich. The product explorer: decision making with ease. In Genny Tortora, Stefano Levialdi, and Maurizio Tucci, editors, *AVI*, pages 423–432. ACM, 2012.

[RR11] Tony Russell-Rose. Designing the Search Experience. In Pedro Campos, T. C. Nicholas Graham, Joaquim A. Jorge, Nuno Jardim Nunes, Philippe A. Palanque, and Marco Winckler, editors, *INTERACT (4)*, volume 6949 of *Lecture Notes in Computer Science*, pages 702–703. Springer, 2011.

[SGJR+13] Ahmet Soylu, Martin Giese, Ernesto Jimenez-Ruiz, Evgeny Kharlamov, Dmitry Zheleznyakov, and Ian Horrocks. OptiqueVQS – Towards an Ontology-based Visual Query System for Big Data. In *International Conference on Management of Emergent Digital EcoSystems (MEDES 2013)*. ACM, 2013.

[SM07] Moritz Stefaner and Boris Muller. Elastic Lists for Facet Browsers. In *Proceedings of the 18th International Conference on Database and Expert Systems Applications*, DEXA '07, pages 217–221, Washington, DC, USA, 2007. IEEE Computer Society.

[ST09] Giovanni Maria Sacco and Yannis Tzitzikas. *Dynamic Taxonomies and Faceted Search Theory, Practice, and Experience.* Springer, New York, 2009.

[Tun06] D. Tunkelang. Dynamic category sets: An approach for faceted search. *ACM SIGIR*, 6, 2006.

[YSLH03] Ka-Ping Yee, Kirsten Swearingen, Kevin Li, and Marti Hearst. Faceted Metadata for Image Search and Browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 401–408, New York, NY, USA, 2003. ACM.