TagSpheres: Visualizing Hierarchical Relations with Tag Clouds

Keywords: Tag Clouds, Text Visualization, Hierarchical Data

Abstract: Tag clouds are widely applied, popular visualization techniques as they illustrate summaries of textual data in an intuitive, lucid manner. Many layout algorithms for tag clouds have been developed in the recent years, but none of these approaches is designed to reflect the notion of hierarchical distance. For that purpose, we introduce a novel tag cloud layout called *TagSpheres*. By arranging tags on various hierarchy levels and applying appropriate colors, the importance of individual tags to the observed topic gets assessable. To explore relationships among various hierarchy levels, we aim to place related tags closely. Three usage scenarios from the digital humanities, sports and aviation, and an evaluation with humanities scholars exemplify the applicability and point out the benefit of TagSpheres.

1 INTRODUCTION

The usage of tag clouds to visualize textual data is a relatively novel technique, which was rarely applied in the past century. In 1976, Stanley Milgram was one of the first scholars who generated a tag cloud to illustrate the mental map of Paris, for which he conducted a psychological study with inhabitants of Paris, aiming to analyze their mental representation of the city (Milgram and Jodelet, 1976). In 1992, a German edition of “Mille Plateaux”, written by the French philosopher Gilles Deleuze, appeared with a tag cloud printed on the cover to summarize the book’s content (Deleuze and Guattari, 1992). This idea to present a visual summary of textual data can be seen as the primary purpose of tag clouds (Sinclair and Cardew-Hall, 2008). But the popularity of tag clouds nowadays is attributable to its frequent usage in the social web community in the 2000s as overviews of website contents. Although there are known theoretical problems concerning the design of tag clouds (Viégas and Wattenberg, 2008), they are generally seen as a popular social component perceived as being fun (Hearst and Rosner, 2008). With the simple idea to encode the frequency of terms to a given topic, tag clouds are intuitive, comprehensible visualizations – and now, widely used metaphors to display summaries of textual data, to support analytical tasks such as the examination of text collections, or even as interfaces for navigation purposes on databases.

In the recent years, the visualization community developed various algorithms to compute effective tag cloud layouts in an informative and readable manner. One of the most popular tools is Wordle (Viegas et al., 2009). It computes compact, intuitive tag clouds and can be generated on the fly using a web-based interface.[1] Although the produced results are very aesthetic, the different used colors do not transfer an information and the final arragement of tags depends only on the scale, and not on the content of tags or potential relationships among them. Some approaches attend to the matter of visualizing more information than the frequency of terms with tag clouds, most often to compare textual summaries of different data facets with each other.

In this paper, we present the tag cloud design called TagSpheres, which endeavors to effectively visualize hierarchies of textual summaries. The motivation arised from research on philology – humanities scholars wanted to analyze the clause function of desired Latin terms. Querying the large database, the scholars often face a numerous results in the form of text passages. When only plain lists are provided to interact with the results, the analysis of the contexts in which the chosen term was used becomes laborious. To support this task, we provide summaries of text passages in the form of interactive tag clouds that group terms in accordance to their distance to the search term. So, the humanities scholar gets an overview and is able to retrieve text passages of interest on demand.

We designed TagSpheres in a way that various types of text hierarchies can be visualized in an intuitive, comprehensible manner. To emphasize the wide applicability of TagSpheres, we list several examples from the digital humanities, sports and aviation.

[1] http://www.wordle.net/

Figure 1: Wordle of Edgar Allan Poe's *The Raven*.

2 RELATED WORK

Although tag clouds rather became popular in the social media, research in visualization attended to the matter of developing various layout techniques in the last years. A basic tag cloud layout is a simple list of words placed on multiple lines (Viegas et al., 2007). In such a list, tags are typically ordered by their importance to the observed issue – encoded by variable font size (Murugesan, 2007). An alphabetical order is also often used, but a study on the utility of tag clouds revealed that the alphabetic order is not obvious for the observer (Hearst and Rosner, 2008). Later, more sophisticated tag cloud layout approaches that rather emphasize aesthetics than meaningful orderings were developed. A representative technique is Wordle (Viegas et al., 2009), which produces compact aesthetic layouts with tags in different colors and orientations, but both features do not transfer any additional information. A Worlde showing the most important terms in Edgar Allan Poe's *The Raven* is given in Figure 1.

Various approaches highlight relationships among tags by forming visual groups. In thematically clustered or semantic tag clouds, the detection of tags belonging to the same topic is supported by placing these tags closely (Lohmann et al., 2009). Traditional, semantic word lists place clustered tags subsequently (Schrammel and Tscheligi, 2014), more sophisticated layout methods often use force directed approaches with semantically close terms attracting each other (Cui et al., 2010; Wang et al., 2014; Liu et al., 2014). After force directed tag placement, tag cloud layouts can be compressed by removing occurring whitespaces (Wu et al., 2011).

Some methods generate individual tag clouds for each group of related tags, and combine the resultant multiple tag clouds to a single visual unity afterwards. An example is the Star Forest method (Barth et al., 2014), which applies a force directed method to pack

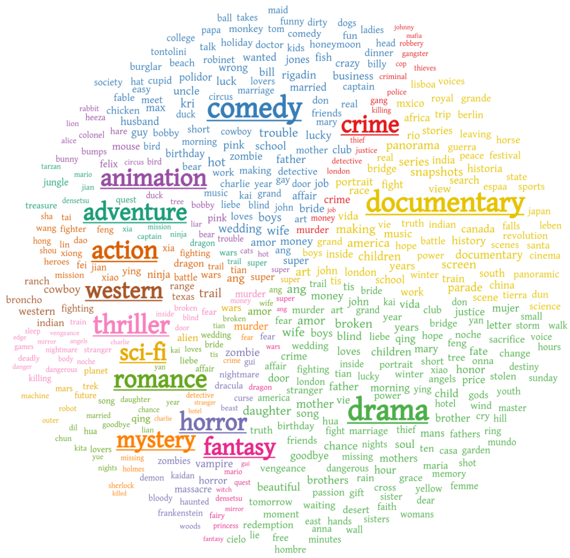


Figure 2: Typical tags for 14 different movie genres.

multiple tag clouds. Other methods use predefined tag cloud containers, e.g., user-defined polygonal spaces in the plane (Paulovich et al., 2012), polygonal shapes of countries (Nguyen et al., 2011), or Voronoi tesselations (Seifert et al., 2011). For the comparison of the tags of various text documents, a ConcentriCloud divides an elliptical plane into sectors that list the shared main tags of several subsets of the underlying texts (Lohmann et al., 2015). Due to the rather independent computation of the individual tag clouds – which often leads to large whitespaces areas in the final composition step – the above mentioned methods can be seen as sophisticated small multiples. A rather traditional small multiples approach is Words Storms (Castellà and Sutton, 2014) that supports the visual comparison of textual summaries of documents.

Tag clouds also have been used to visualize trends. Parallel Tag Clouds generate alphabetically ordered tag lists as columns for a number of time slices and highlights the temporal evolution of a tag placed in various columns on mouse interaction (Collins et al., 2009). In contrast, SparkClouds attach a graph showing the tag's evolution over time (Lee et al., 2010). Other approaches overlay time graphs with tags characteristic for certain time ranges (Shi et al., 2010).

Only few approaches generate multifaceted tag cloud layouts in a single, continuous flow that includes the positioning of all tags belonging to various groups. RadCloud visualizes tags belonging to various groups within a shared elliptical area (Burch et al., 2014). In Compare Clouds, tags of two media frames (MSM, Blogs) are comparatively visualized in a single cloud (Diakopoulos et al., 2015). To support the comparative analysis of multiple tag groups, TagPies are arranged in a pie chart manner (Jänicke et al., 2015a). An example visualizing typical tags for different movie genres is shown in Figure 2.

Although techniques like TagPies or Parallel Tag Clouds are capable of visualizing sequences of tag groups, none of the mentioned approaches endeavors to visually encode generic hierarchical information intuitively in a single, compact, aesthetic tag cloud. TagSpheres – presented in this paper – are designed to fill this gap.

3 DESIGNING TAGSPHERES

The central idea of TagSpheres is the visualization of textual summaries that comprise hierarchical information. This paper provides three usage scenarios that exemplify hierarchies in textual data (see Section 4). An overview of the characteristics of these examples is given in Table 1.

Given n hierarchy levels $H_1, \ldots, H_n$, the top hierarchy level H_1 contains tags representing the focus of interest of a usage scenario. All other tags are divided into $n-1$ groups in dependency on their hierarchical distance according to the observed topic, or tag(s) on H_1. Each tag t in TagSpheres has a weight $w(t)$ reflecting its importance, and an optional predecessor tag $p(t)$ representing a relationship to a higher hierarchy levels.

3.1 Design Decisions

When designing TagSpheres, we use the following, well-established design features for tag clouds:

- **Font size**: Evaluated as the most powerful property (Bateman et al., 2008), font size encodes the weight $w(t)$ of a tag.
- **Orientation**: As rotated tags are perceived as “unstructured, unattractive, and hardly readable” (Waldner et al., 2013), we do not rotate tags to keep the layout easily explorable.
- **Color**: Being the best choice to distinguish categories (Waldner et al., 2013), various colors are assigned to tags belonging to different hierarchy levels. As TagSpheres encode the distance to a given topic, the usage of a categorial color map is inappropriate. Unfortunately, suitable sequential color maps as provided by the ColorBrewer (Harrower and Brewer, 2003) produce less distincitve colors even for a small number of hierarchy levels, so that adjacent tags belonging to different hierarchy levels are hard to classify. Following the suggestions illustrated in (Ware, 2013), we defined a divergent cold-hot color map using red for the first hierarchy level and blue for tags belonging to the last hierarchy level n. To avoid uneven visual attraction of tags, we only chose saturated colors that are in contrast to the white background. Example color maps for up to eight hierarchy levels are shown in Figure 3(a).

3.2 Layout Algorithm

In preparation, the tags are sorted by increasing hierarchy level, so that all tags within the same hierarchical distance to H_1 are placed subsequently. The tags of each hierarchy level are ordered by increasing value to ensure that important tags are circularly well distributed.

To avoid large whitespaces, a problem addressed by Seifert (Seifert et al., 2008), our method follows the idea of the Wordle algorithm (Viegas et al., 2009) – permitting overlapping tag bounding boxes if the tag's letters do not occlude – to determine the positions of tags. So, we obtain compact, uniformly looking, aesthetic tag clouds for the underlying hierarchical, textual data. That TagSpheres remain easily readable, we use a minimal padding between letters of different tags.

As shown in Figure 3(b), we aim to visually compose tags of the same hierarchy level in the form of spheres around the tag cloud origin at $(0,0)$. Initially, we iteratively determine positions for the tags of H_1 in the central sphere using an archimedean spiral originating from $(0,0)$. An example is given in Figure 4(a). For each tag t of the remaining hierarchy levels $H_2, \ldots, H_n$, we also use $(0,0)$ as spiral origin, if $p(t)$ is not provided (see Figure 4(b)). If $p(t)$ is defined, we use the predecessor's position as spiral origin. As a consequence, hierarchically related tags are placed closely and visually compose in the form of rays originating from $(0,0)$ as shown in Figure 7. In contrast to other spiral based tag cloud algorithms, we avoid to cover whitespaces with tags of H_i within spheres of already processed hierarchy levels $H_1, \ldots, H_{i-1}$. Dependent on the quadrant in the plane, in which a tag shall be placed, we search for already placed tags intersecting two vectors originating from the dedicated position as illustrated in Figure 3(c). If no intersections are found, we place the tag. This approach coheres all tags of a hierarchy level as a visual unity outside the inner bounds of the previously layouted hierarchy levels' spheres.

domain	digital humanities (see Section 4.1)	sports (see Section 4.2)	aviation (see Section 4.3)
task	analyzing the clause functions of a search term T	comparing the performances of nations participating the FIFA World Cup	observing all direct flights from an airport or a city
H_1	search term T	FIFA World Cup winners	departure airport/city
$H_2, \ldots, H_n$	co-occurrences in dependency on the word distance to T	teams knocked out in different tournament stages	direct federal (H_2), continental (H_3) and worldwide flights (H_4)
n	4	6	2..4
$w(t)$	number of (co-)occurrences of t	number of knock-outs in the corresponding tournament stage	inverse distance weighting between departure and arrival airports/cities
$p(t)$	equally labeled tag of a higher hierarchy level	equally labeled tag of a higher hierarchy level	previously placed tag of the same country/continent
strong tag relations	equally labeled tags	equally labeled tags	departure/arrival airports/cities
weak tag relations	spelling variants	N/A	airports/cities of the same country/continent
action on mouse click	text passages containing the selected tag and T are shown	N/A	redirection to Google Flights showing available flights for the selected connection

Table 1: Characteristics of usage scenarios for TagSpheres.

3.3 Limitations

The main objective of the presented layout algorithm is to combine a hierarchical information of textual data with the aesthetics of tag clouds. In contrast to the usual approach to always initialize an archimedian spiral at the tag cloud origin $(0,0)$ when determining the position of a tag, the usage of predecessor tags as spiral origins slightly affects the uniform appearance of the result in some cases. Occasionally, little holes occur, and – dependent on the hierarchical structure of the underlying data – the tag cloud boundaries get distorted.

The proposed hot-cold color map used to visually convey hierarchical distance generates well distinguishable colors when the number of hierarchy levels is small. For a larger number of hierarchies, closely positioned tags of different levels may become visually indistinct, especially when only few tags belong to a certain level.

The current TagSpheres design does not take the distribution of tags throughout different hierarchies into account. In use cases with a steadily increasing or decreasing number of tags per hierarchy level it gets possible, that a considerable proportion of the color maps' bandwidth gets used for a comparatively small portion of tags.

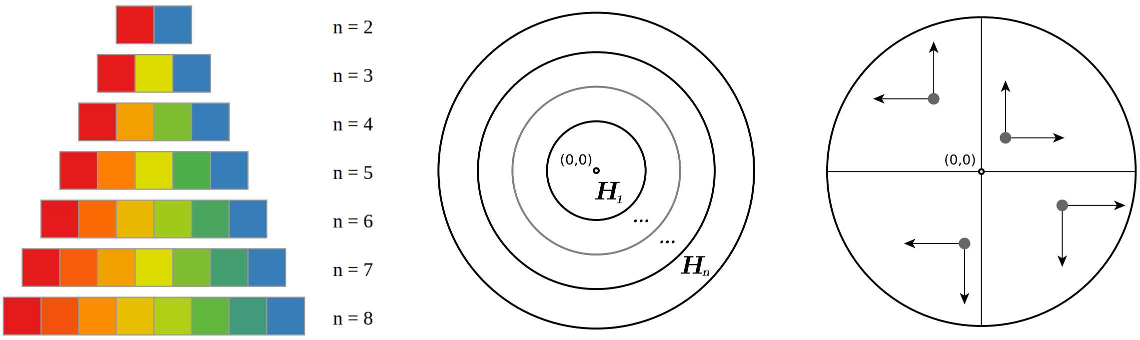
3.4 Interactive Design

Implemented as an OpenSource JavaScript library, TagSpheres can be dynamically embedded into web-based applications. With mouse interaction, we enable the user to detect hierarchically related tags quickly. Hovering a tag highlights strongly and weakly related tags. Strongly related tags are shown using a black font on transparent backgrounds having the hierarchy levels' assigned color as shown in Figure 7. In contrast, weakly related tags retain their saturated font color, but gray, transparent backgrounds indicate correlations.

TagSpheres provide a configurable tooltip displayed when hovering or clicking a tag to be used, e.g., to list all related tags and their weights. The mouse click function can be further used to link to an external source. Examples are given in Table 1.

3.5 Evaluation

TagSpheres are used by humanities scholars of a digital humanities project to analyze the clause functions of search terms (see Section 4.1 for details). In a small evaluation with seven project members, we asked for subjective ratings regarding intuitivity, aesthetics and the utility of TagSpheres for their work. The participants needed to choose a value on a Likert scale from

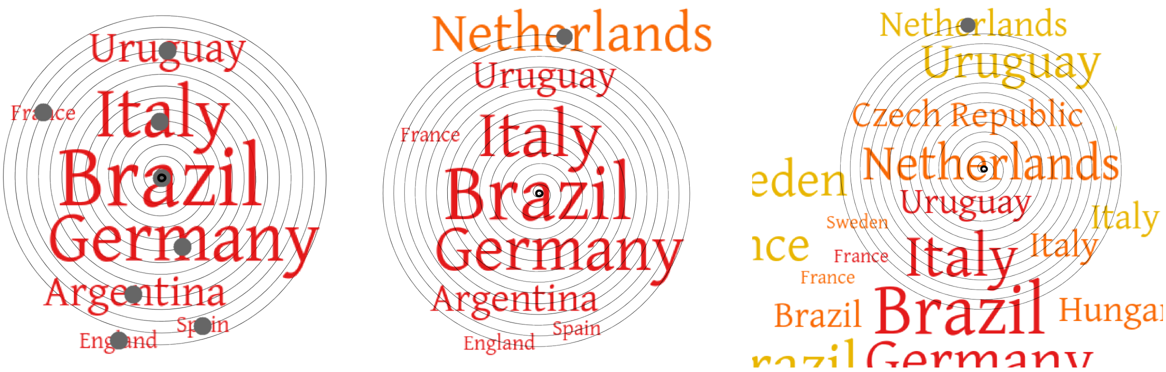


(a) Resultant color maps for $n = 2, \ldots, 8$ hierarchy levels.

(b) Using spheres for the tags of different hierarchy levels.

(c) Vectors for occlusion check to guarantee hierarchical coherence.

Figure 3: TagSpheres layout algorithm details.



(a) Placing all tags of H_1.

(b) Placing a tag without predecessor.

(c) Placing a tag with predecessor.

Figure 4: Determining tag positions using an archimedean spiral.

1 (very bad) to 7 (very good), and we also asked them to justify their decisions.

TagSpheres aesthetics was rated 5.57 on average, and the intuitivity to transmit a notion of hierarchical distance was also assessed as good (6). Especially, the chosen colors “clearly transmit the notion of distance between co-occurrences and the search term.” The readability of tags was justified as 5.14. Participants stated that TagSpheres are “easily understandable” and that “all important co-occurrences of the search term are visible at first glance.” Although related tags are positioned closely, it was not always easy for the humanities scholars to detect similar terms in different hierarchy levels (3). But all participants stated that the provided means of interaction facilitate this task (6) and overall foster the understanding of the visualization and the explorative analysis of results. Finally, the utility of TagSpheres to support the humanities scholars in examining research questions regarding the clause functions of search terms was also rated as good (5.36).

4 USE CASES

TagSpheres are applicable whenever statistics of unstructured text shall be visualized in the form of a tag cloud and a decent hierarchy among the tags exists. This sections illustrates usage scenarios of TagSpheres for text-based data from three different domains: digital humanities, sports and aviation.

4.1 Digital Humanities Scenario

Within the digital humanities project *eXChange*,[2] historians and classical philologists work with a database containing a large amount of digitized historical texts in Latin and ancient Greek. Usually, humanities scholars pose keyword based search queries and often receive numerous results, which are hard to revise individually. As a consequence, the generation of valueable hypotheses is a laborious, time-consuming process. To facilitate the humanities scholars workflows,

[2] http://exchange-projekt.de/

we develop visual interfaces that attempt to steer the analysis of search results into promising directions.

TagPies – developed within the *eXChange* project – are tag clouds arranged in a pie chart manner that support the comparison of multiple search query results (Jänicke et al., 2015a). Using a TagPie, humanities scholars analyze contextual similarities and differences of the observed terms. Unlike TagPies, TagSpheres provide an additional, hierarchical dimension, which supports approaching a further research interest of the humanities scholars: the analysis and classification of a term's co-occurrences according to their clause function. For this purpose, the scholars required four-level TagSpheres:

H_1 The first level contains the search term T.

H_2 The second level contains all co-occurrences with distance 1 to T.

H_3 The third level contains all co-occurrences with distance 2 to T.

H_4 The fourth level contains all co-occurrences with distance 3 up to distance n to T.

The font size of T on level 1 encodes how frequent the search term occurs in the underlying text corpus; the font sizes of all other terms reflect the number of co-occurrences with T in dependency on the corresponding distance. On level 4, font sizes are normalized in relation to the distance range $n-2$. A tag on hierarchy level i receives a predecessor tag if the corresponding term occurs on one of the previous layers $i-1,\ldots,1$.

A use case provided by one of the humanities scholars involved in the *eXChange* project shall illustrate the utility of TagSpheres to support the classification of a term's co-occurrences by their clause function. Analyzing the co-occurrences of ***morbo*** (disease), terms in similar relationships to the given topic were discovered and classified (see Figure 6). In large distances, the humanities scholar found objects in form of affected parts of the body, e.g., head (*caput*), soul (*animo*) and limbs (*membrorum*), affected persons, e.g., son (*filius*), woman (*mulier*) and king (*rex*), and related places, e.g., Rome (*romam*), church (*ecclesia*) and *villa*. Closer to *morbo* (most often with distance 1 or 2), typical attributes and predicates can be found. Whereas attributes describe the type or intensity of the disease, e.g., pestilential (*pestifero*), heavy (*gravi*), deadly (*exitiali*) and acute (*acuto*), the occurring predicates illustrate the disease's progress, e.g., seize (*correptus*), dissappear (*periit*) and worsening (*ingravescente*). Adjacent to *morbo*, specific terms for "moral" diseases, e.g., greediness (*avaritiae*), arrogance (*superbiae*) and lust (*concupiscentiae*), and actual diseases like jaundice ([morbo] *regio*), leprosy (*leprae*) and two common names for

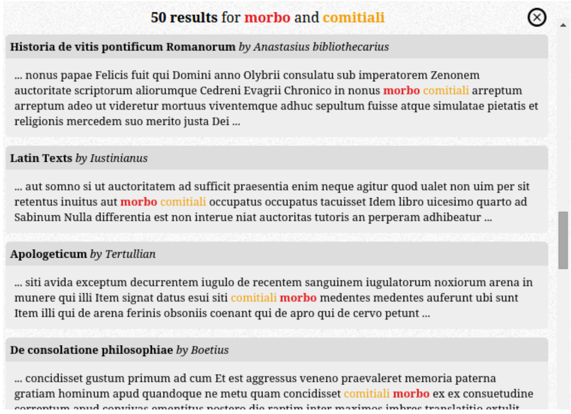


Figure 5: Close reading of text passages containing *morbo* and *comitiali* with distance 1.

epilepsy (*[morbo] comitiali, [morbo] sacro*) occur.

In this usage scenario, the interaction capabilities of TagSpheres are tailored according to the needs of the humanities scholars. Hovering a tag opens a popup showing the term's number of occurrences on all hierarchy levels. Additionally, variant spellings or cases of the term are listed with their corresponding frequencies to support the analysis process. An important requirement for the humanities scholars was the discovery of potentially interesting text passages, but they desired a straightforward access to the underlying texts in general. This so-called *close reading* was often reported as an important component when designing visualizations for humanities scholars (Jänicke et al., 2015b). TagSpheres support close reading by clicking a tag, which displays the corresponding text passages containing the search term and the clicked term with the chosen distance. An example for text passages containing the adjacent terms *morbo* and *comitiali* is shown in Figure 5.

4.2 Championship Performances

This scenario illustrates how TagSpheres can be used to comparatively visualize performances in championships. Exemplariliy, we processed a dataset containing the results of all national teams ever qualified for the FIFA World Cup. We receive the following six-level hierarchy:

H_1 FIFA World Champions.

H_2 Second placed national teams.

H_3 National teams knocked out in the semifinal.

H_4 National teams knocked out in the quarterfinal.

H_5 National teams knocked out in the second round (second group stage or last 16).

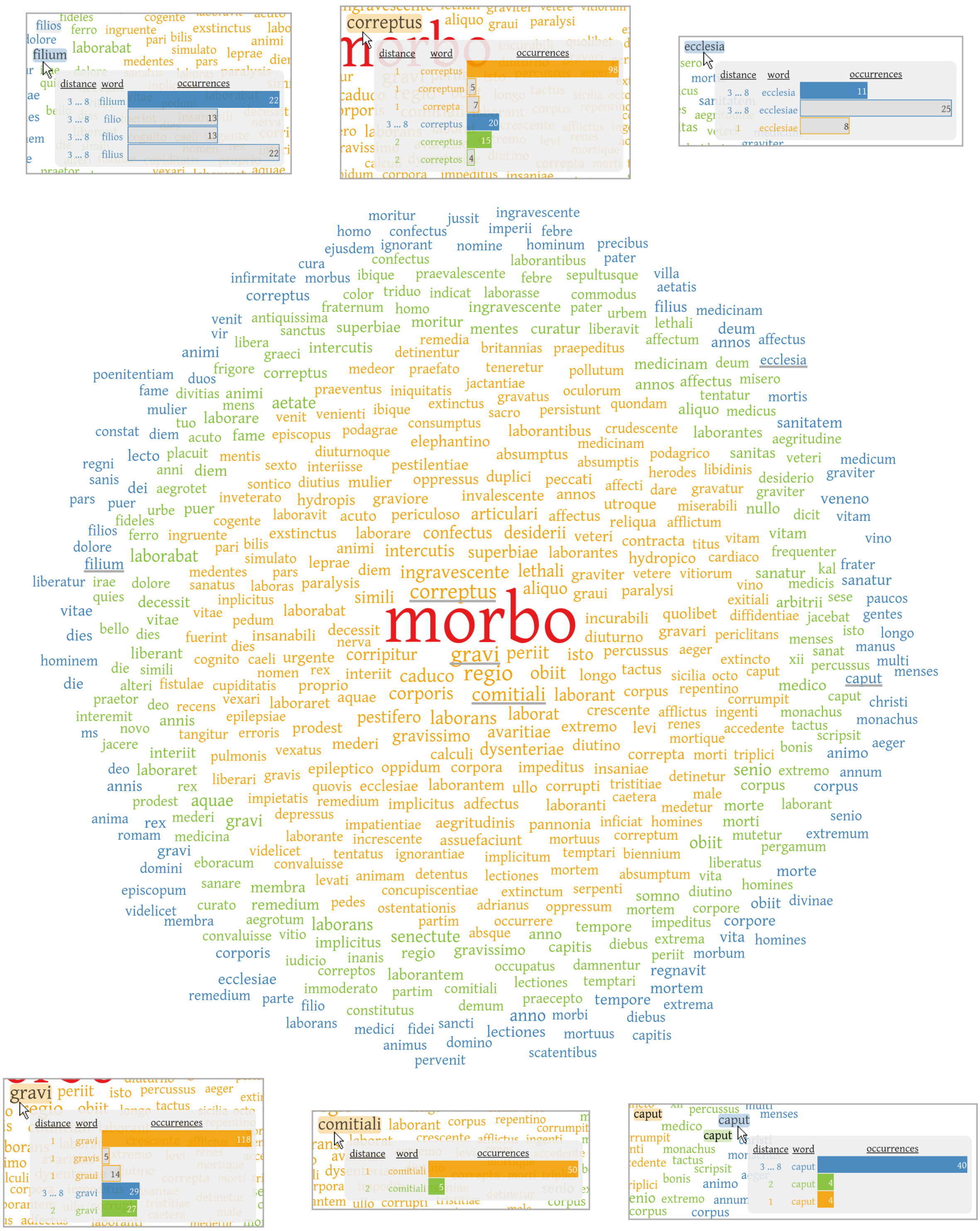


Figure 6: Morbus example.

$$G = 6378 \cdot \arccos\left(\sin(lat_d) \cdot \sin(lat_a) + \cos(lat_d) \cdot \cos(lat_a) \cdot \cos(lon_d - lon_a)\right) \quad (1)$$

H_6 National teams knocked out in the (first) group stage.

The nations are used as tags and font size encodes how often a national team partook a championship round without reaching the next level. If a tag for a nation was already placed at a higher hierarchy level, we use the corresponding tag as predecessor.

Figure 7 shows the resultant TagSphere. Especially this scenario illustrates the benefit of using the positions of predecessor tags as spiral origins for successor tags. In most cases, the various tags of a nation are closely positioned. Hovering a tag displays the all-time performance of the corresponding national team for all championship hierarchy levels in a popup. Expectedly, *Brazil* and *Germany* achieved very good results, especially in the last championship rounds. In contrast, *Italy* was often knocked out in the first round, but in case of reaching the semifinal (8x), *Italy* became FIFA World Champion four times. *England* and *Spain* show nearly equal performances. With the same number of appearances (38x) both nations reached the semifinal only twice. Few nations have a 100% success rate in the group stage. Qualified three times for the FIFA World Cup, *Senegal* always reached the quarterfinals. Most nations, e.g., *Sweden* and *Cameroon*, show the expected pattern “the higher the championship round, the smaller the number of appearances”.

Another example is given in Figure 8 that illustrates the success of football clubs ever played in Englands first league. The average rank at the end of the seasons is used to cluster 68 clubs into 17 hierarchy levels, and font size encodes the number of appearances.

4.3 Airport Connectivity

To analyze the federal, continental and worldwide connectivity of airports, we derived a dataset from the OpenFlights database,[3] which provides a list of direct flight connections between around 3,200 airports worldwide. With the selected departure airport d (or city) on H_1, all other airports (or cities) reachable with a non-stop flight cluster into three further hierarchy levels:

H_2 airports/cities in the same country as d,

H_3 airports/cities on the same continent as d, and

H_4 all other reachable worldwide airports/cities.

[3] http://openflights.org/data.html

Figure 8: Performances of English first league clubs from 1888/89 – 2014/15.

As tags we chose either airport names, the provided IATA codes,[4] or the corresponding city names. In this scenario, font size encodes the inverse geographical distance between departure airport $d = \{lat_d, lon_d\}$ and arrival airport $a = \{lat_a, lon_a\}$. To keep the deviation to the actual distance as small as possible, we apply the great circle distance G (Head, 2003), defined by Equation 1. Predecessor tags are used to place airports or cities of the same country or continent closely. For a tag t to be placed on H_3, we choose the first placed tag with the same associated country as predecessor, if existent; for H_4, we choose the first placed tag with the same associated continent.

Figure 9 shows TagSpheres for non-stop flights from various airports or cities. All examples show that airports/cities of the same countries/continents are placed closely in clusters. For Sydney, no tags are placed on H_3, and for Cagliari, no connections to non-European airports exist. When the user hovers a tag, the corresponding connection and the travel distance are shown in a tooltip. Clicking a tag redirects to Google Flights[5] listing possible flight connections.

5 CONCLUSION

We introduced TagSpheres that arrange tags on several hierarchy levels to transmit the notion of hierarchical distance in tag clouds. We accentuate

[4] http://www.iata.org/services/pages/codes.aspx

[5] https://www.google.com/flights/

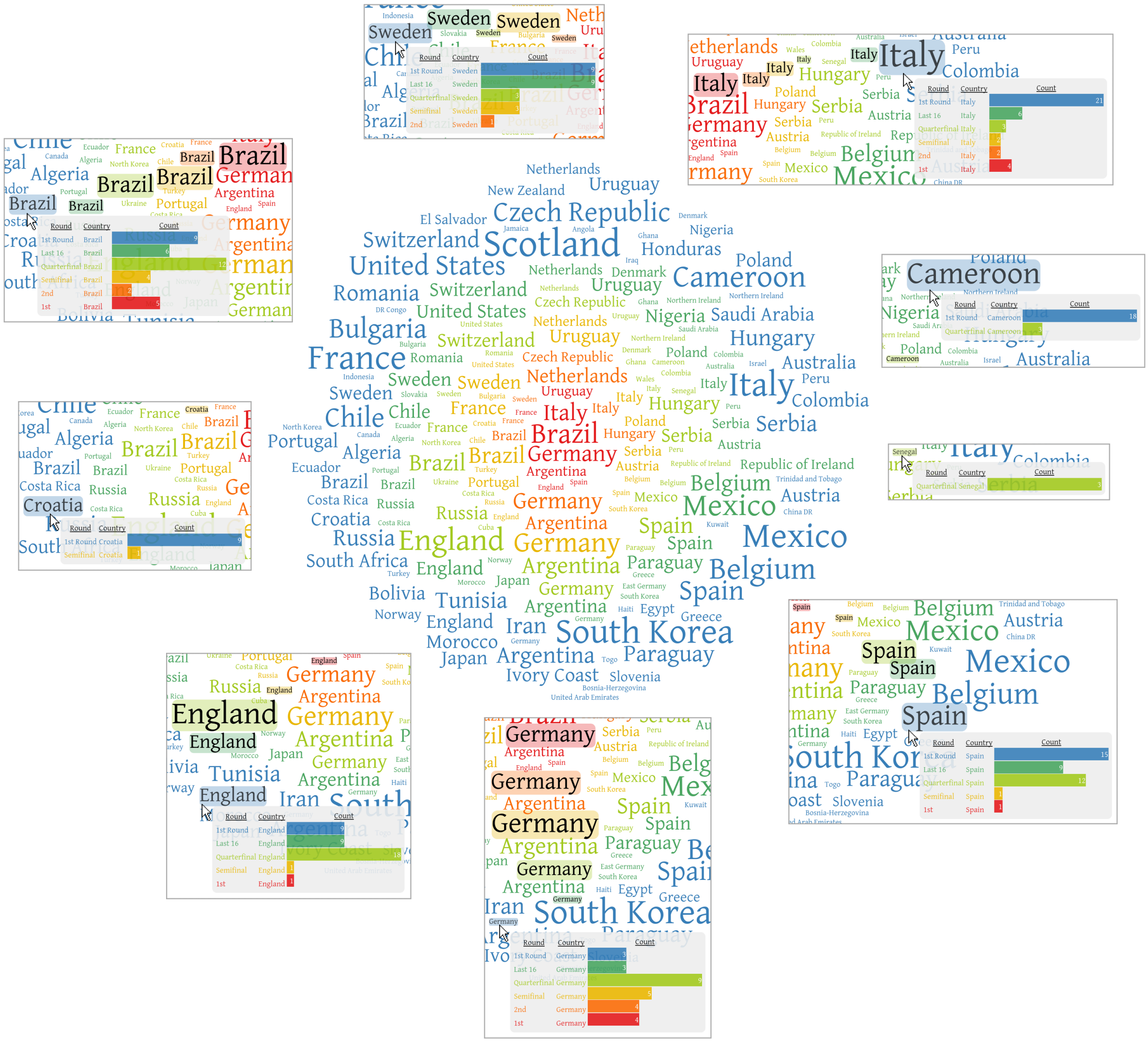


Figure 7: Performances of all nations qualified for the FIFA World Cup.

relationships between different hierarchy levels by placing hierarchically related tags closely. Applied within a digital humanities project, the design of TagSpheres was evaluated as aesthetic and intuitive, and the humanities scholars emphasized the utility of TagSpheres for their work. Further usage scenarios in sports and aviation outline the inherence of hierarchical textual information in various domains.

Despite few listed limitations, TagSpheres might be applicable to a multitude of further research questions from other areas. Also imaginable is the combination of TagSpheres and TagPies to support the comparative analysis of different hierarchical, textual summaries.

REFERENCES

Barth, L., Kobourov, S., and Pupyrev, S. (2014). Experimental Comparison of Semantic Word Clouds. In *Experimental Algorithms*, volume 8504 of *Lecture Notes in Computer Science*, pages 247–258. Springer International Publishing.

Bateman, S., Gutwin, C., and Nacenta, M. (2008). Seeing Things in the Clouds: The Effect of Visual Features on Tag Cloud Selections. In *Proceedings of the Nineteenth ACM Conference on Hypertext and Hypermedia*, HT '08, pages 193–202. ACM.

Burch, M., Lohmann, S., Beck, F., Rodriguez, N., Di Silvestro, L., and Weiskopf, D. (2014). RadCloud: Visualizing Multiple Texts with Merged Word Clouds. In *Information Visualisation (IV), 2014 18th International Conference on*, pages 108–113.

Castellà, Q. and Sutton, C. (2014). Word Storms: Multi-

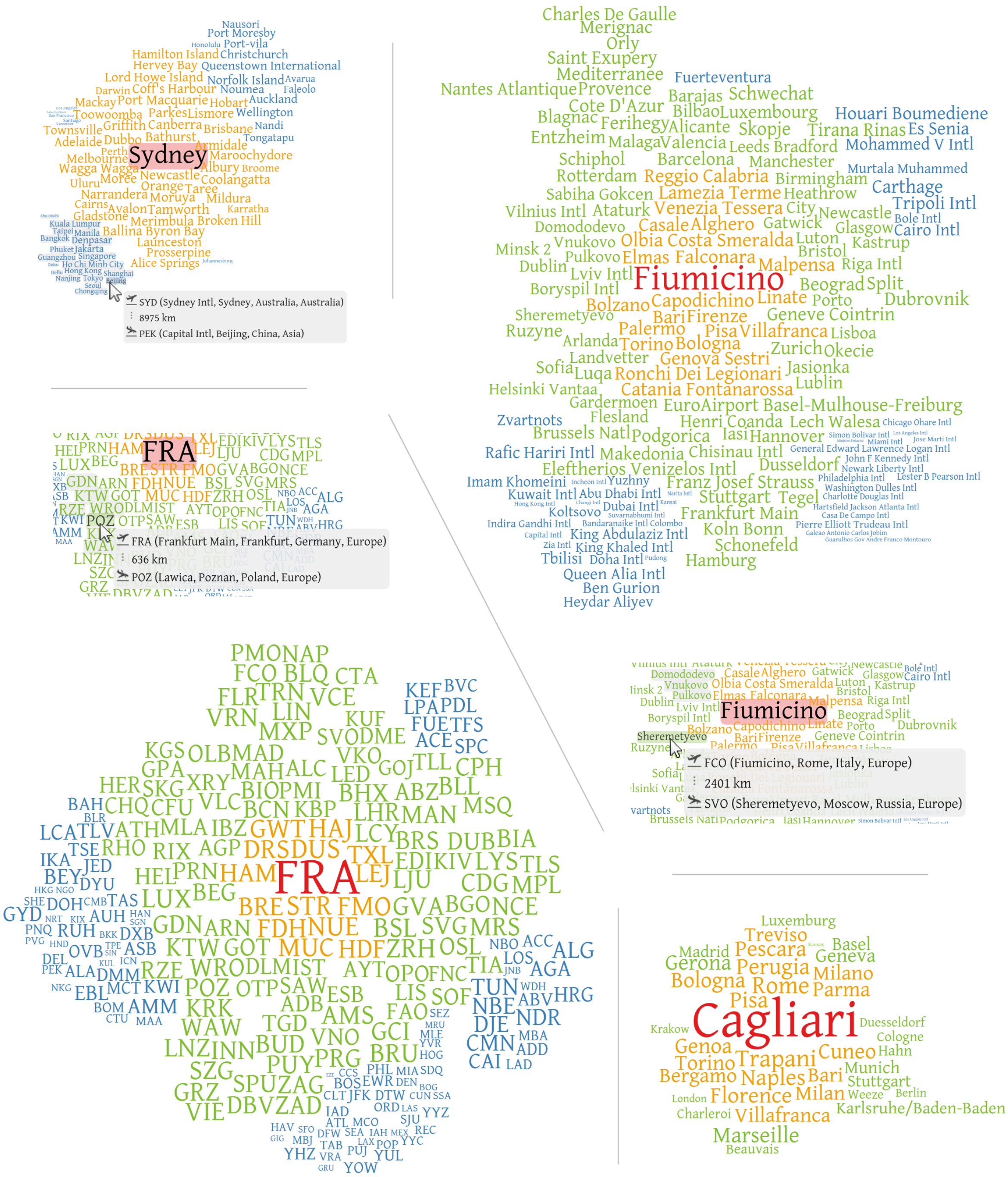


Figure 9: Direct flight connections from airports in Sydney, Rome, Frankfurt and Cagliari.

ples of Word Clouds for Visual Comparison of Documents. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, pages 665–676. ACM.

Collins, C., Viegas, F., and Wattenberg, M. (2009). Parallel Tag Clouds to explore and analyze faceted text corpora. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on*, pages 91–98.

Cui, W., Wu, Y., Liu, S., Wei, F., Zhou, M., and Qu, H. (2010). Context preserving dynamic word cloud visualization. In *Pacific Visualization Symposium (PacificVis), 2010 IEEE*, pages 121–128.

Deleuze, G. and Guattari, F. (1992). *Tausend Plateaus. Kapitalismus und Schizophrenie II*. Merve Verlag.

Diakopoulos, N., Elgesem, D., Salway, A., Zhang, A., and Hofland, K. (2015). Compare Clouds: Visualizing Text Corpora to Compare Media Frames. In *Proc. of IUI Workshop on Visual Text Analytics*.

Harrower, M. and Brewer, C. A. (2003). ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. *The Cartographic Journal*, 40(1):27–37.

Head, K. (2003). Gravity for Beginners. *University of British Columbia*, 2053.

Hearst, M. and Rosner, D. (2008). Tag Clouds: Data Analysis Tool or Social Signaller? In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, pages 160–160.

Jänicke, S., Blumenstein, J., Rücker, M., Zeckzer, D., and Scheuermann, G. (2015a). Visualizing the Results of Search Queries on Ancient Text Corpora with Tag Pies. *Digital Humanities Quarterly*. To appear.

Jänicke, S., Franzini, G., Cheema, M. F., and Scheuermann, G. (2015b). On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In Borgo, R., Ganovelli, F., and Viola, I., editors, *Eurographics Conference on Visualization (EuroVis) - STARs*. The Eurographics Association.

Lee, B., Riche, N., Karlson, A., and Carpendale, S. (2010). SparkClouds: Visualizing Trends in Tag Clouds. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1182–1189.

Liu, X., Shen, H.-W., and Hu, Y. (2014). Supporting multifaceted viewing of word clouds with focus+context display. *Information Visualization*.

Lohmann, S., Heimerl, F., Bopp, F., Burch, M., and Ertl, T. (2015). ConcentriCloud: Word Cloud Visualization for Multiple Text Documents. In *19th International Conference on Information Visualisation*.

Lohmann, S., Ziegler, J., and Tetzlaff, L. (2009). Comparison of Tag Cloud Layouts: Task-Related Performance and Visual Exploration. In *Human-Computer Interaction - INTERACT 2009*, volume 5726 of *Lecture Notes in Computer Science*, pages 392–404. Springer Berlin Heidelberg.

Milgram, S. and Jodelet, D. (1976). Psychological Maps of Paris. *Environmental Psychology*, pages 104–124.

Murugesan, S. (2007). Understanding Web 2.0. *IT Professional*, 9(4):34–41.

Nguyen, D. Q., Tominski, C., Schumann, H., and Ta, T. A. (2011). Visualizing Tags with Spatiotemporal References. In *Information Visualisation (IV), 2011 15th International Conference on*, pages 32–39.

Paulovich, F. V., Toledo, F., Telles, G. P., Minghim, R., and Nonato, L. G. (2012). Semantic Wordification of Document Collections. In *Computer Graphics Forum*, volume 31, pages 1145–1153. Wiley Online Library.

Schrammel, J. and Tscheligi, M. (2014). Patterns in the Clouds - The Effects of Clustered Presentation on Tag Cloud Interaction. In *Building Bridges: HCI, Visualization, and Non-formal Modeling*, Lecture Notes in Computer Science, pages 124–132. Springer Berlin Heidelberg.

Seifert, C., Kienreich, W., and Granitzer, M. (2011). Visualizing Text Classification Models with Voronoi Word Clouds. In *Proceedings of the International Conference Information Visualisation (IV), London*.

Seifert, C., Kump, B., Kienreich, W., Granitzer, G., and Granitzer, M. (2008). On the Beauty and Usability of Tag Clouds. In *Information Visualisation, 2008. IV '08. 12th International Conference*, pages 17–25.

Shi, L., Wei, F., Liu, S., Tan, L., Lian, X., and Zhou, M. (2010). Understanding text corpora with multiple facets. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*, pages 99–106.

Sinclair, J. and Cardew-Hall, M. (2008). The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34(1):15–29.

Viegas, F., Wattenberg, M., and Feinberg, J. (2009). Participatory Visualization with Wordle. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1137–1144.

Viegas, F., Wattenberg, M., van Ham, F., Kriss, J., and McKeon, M. (2007). ManyEyes: a Site for Visualization at Internet Scale. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6):1121–1128.

Viégas, F. B. and Wattenberg, M. (2008). TIMELINES: Tag Clouds and the Case for Vernacular Visualization. *interactions*, 15(4):49–52.

Waldner, M., Schrammel, J., Klein, M., Kristjánsdóttir, K., Unger, D., and Tscheligi, M. (2013). FacetClouds: Exploring Tag Clouds for Multi-dimensional Data. In *Proceedings of Graphics Interface 2013*, GI '13, pages 17–24. Canadian Information Processing Society.

Wang, J., Zhao, J., Guo, S., North, C., and Ramakrishnan, N. (2014). ReCloud: Semantics-based Word Cloud Visualization of User Reviews. In *Proceedings of the 2014 Graphics Interface Conference*, GI '14, pages 151–158. Canadian Information Processing Society.

Ware, C. (2013). *Information Visualization: Perception for Design*. Elsevier.

Wu, Y., Provan, T., Wei, F., Liu, S., and Ma, K.-L. (2011). Semantic-Preserving Word Clouds by Seam Carving. In *Computer Graphics Forum*, volume 30, pages 741–750. Wiley Online Library.