Time Varying Predominance Tag Maps

Martin Reckziegel*
Leipzig University

Stefan Jänicke†
Southern University of Denmark

ABSTRACT

Visually conveying time-dependent changes in tag maps is insufficiently addressed by current approaches. Typically, for each time range a tag map is determined, and the change between tag maps of subsequent time ranges is progressively visualized. Our method compares tag maps locally in order to enable a continuous display of geographical topic changes among subsequent time ranges. We further provide an alternate tag map variant focusing on frequency changes instead of relative frequency values to visualize the geospatial-temporal rise and fall of topics.

Keywords: visualization, geo-spatial, time-dependent, point-based data, data aggregation

1 INTRODUCTION

Tag maps are geographical maps having a thematic layer in the form of a tag cloud, in which tags are placed in close vicinity to the geographical location they are associated with [16]. By placing tags instead of colored glyphs on the map, the richness of categorial data sets can be preserved and communicated to the observer [24]. Nevertheless, tag map algorithms require to aggregate tags and to hide less relevant tags in order to avoid occlusions. Means of filtering on the basis of diverse metadata can serve to deliver a more precise tag map representation. For example, for data sets having temporal alongside with geographical information, current approaches offer to select a time range that filters the tags on the map to be displayed [18]. In addition, time-dependent selections can be animated using a sliding time window [29]. But, for each time range the tag map is recomputed and changes between subsequent time ranges are neither determined nor visually communicated.

We fill this gap in order to easen the visual analysis of time-dependent tag maps. Based on Predominance Tag Maps [24], we identify locally related tags among subsequent time ranges in order to enable a seamless visualization of temporally changing tag maps. In addition, we propose a variant tag map approach that focuses on tag frequency changes instead of relative frequency values in order to illustrate the geospatial-temporal rise and fall of topics.

2 RELATED WORK

Many visualizations are designed to analyze time-stamped, georeferenced and/or tagged data. Focusing on tagged data, we observe related works in three categories.

Space & Tags *Tag maps* visualize tags representing georeferenced data items on a map. Two basic approaches for generating tag maps exist [24]. *Tag-cloud-driven Tag Maps* aggregate the frequencies of tags for a specific geospatial area, and a tag cloud layout algorithm is used to position tags on the map. Maple [13] determines tag distributions for specific geographical locations and, originating from that location, a Wordle [30] is computed and used as a thematic map layer. Other methods make use of polygonal geographical boundaries. Taggram [21] aggregates all tags associated to a single country and places tags arbitrarily within the country's bounds. In contrast, Geo word clouds [3] take actual averaged locations of tags into account. *Location-driven Tag Maps* position tags at their respective geographical locations, and algorithms are designed to avoid occlusions. Jaffe et al. [16] cluster data items geographically, and a representative tag is placed for each cluster leading to a sparse tag map in which tags may occlude. Thom et al. [27] detect occluding tags that are merged if they are equal, or displaced using an Archimedian spiral, which leads to moving the tag away from its dedicated geographical context. Predominance Tag Maps [24] are designed to prevent from potential misinterpretations of that kind. Each tag on the map represents the relative majority of data items enclosed by the tag's bounding box. In this paper, we present a method that extends the design of Predominance Tag Maps to communicate and to analyze time-dependent, geographical topic changes.

Time & Tags The visualization of time-dependent, tagged data has been subject to several works in order to illustrate predominant tags for specific time ranges and to support the visual analysis of trends. Parallel Tag Clouds [6] list the most important tags per time range vertically, and changes of frequencies of words between time ranges are visually indicated by thickening or thinning horizontal connections. In contrast, SparkClouds [19] display a time chart alongside each tag in the tag cloud to easen the comparative temporal analysis among tags. Other approaches display tags in a time chart representation. While many approaches only show tags on demand when selecting certain time ranges, e.g., [7, 26], some methods position tags directly within large whitespace areas of streams of stacked graphs [10, 25]. WordStream [8] uses all available space by reconstructing the streams with tags associated to the corresponding category and time. Other visualizations generate spatializations to show the evolution of tagged data over time. Mashima et al. [20] generate an animated tag cloud like map together with a heat map to indicate trends, Gansner et al. [12] and Chen et al. [5] overlay map spatializations with keywords or integrate them into the map metaphor [4].

Space, Time & Tags Many visual interfaces visualize temporal, geographical and thematic information simultaneously in different views [9, 15], but only few approaches cater for communicating temporal changes on the basis of tag maps. Nguyen et al. [22] use text style or visual glyphs in the background of tags to to illustrate temporal changes of tags in (static) Taggrams [21]. In order to analyze time-dependent changes in tag maps, Scatterblogs [27–29] provide a timeline where data items can be selected by interactively defining a time range. By sliding this time range, the tag map changes accordingly. Similarly, Bird's Eye [18] actualizes the tag map once the temporal selection changes. Likewise, Hao et al. [14] use a calendar view for temporal filtering prior to updating the tag map. In all cases, tag maps are just recomputed according to the new temporal selection, and relations between tags of subsequent time ranges are not taken into account. We fill this gap using the Predominance Tag Map algorithm [24] to identify time-dependent, geographical changes.

3 TIME VARYING PREDOMINANCE TAG MAPS

We base our method on the Predominance Tag Map (PTM) algorithm. This algorithm takes as input a set of two dimensional points

*e-mail: reckziegel@informatik.uni-leipzig.de
†e-mail: stjaenicke@vizcovery.de

each associated to one labeled category as well as a range of font sizes $[f_{min}, f_{max}]$ the final set of non-overlapping tags should lie inside. The algorithm consists of three major steps: (S1) A set of seed positions is derived from the given point set. (S2) Each seed position serves as center point for one tag candidate. For every candidate, a font size, label category, and score is computed. As the seed positions are generated relatively close to each other with respect to the user given font size bounds, the tag candidates overlap heavily. (S3) In the final step, a set of non-overlapping tags is selected by greedily placing the tags in descending order of font size and rejecting those that would produce an overlap. For more details on the original PTM algorithm we refer the reader to [24].

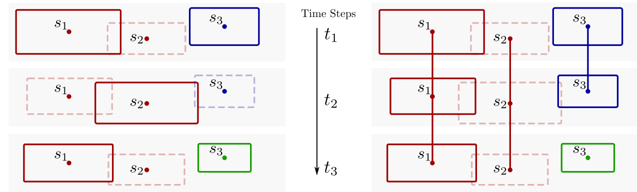
We modify the algorithm in order to generate an animation showing the development of the predominant categories over time. Instead of two dimensional points, the input consists of a set of labeled three dimensional positions (latitude, longitude and a time stamp). We divide this three dimensional input data into user specified time slices of equal duration. The straight forward approach is to apply the original PTM algorithm for each time slice, generating one frame of the animation. However, there are a few issues with this naive approach for which we present solutions below.

3.1 Font Size Comparability

In general, font size as indicator for the importance of tags is one of the most important visual features in tag clouds (and tag maps in particular). In the original PTM algorithm, the font sizes of the tag candidates are calculated in S2. This is done using a bisection optimization that correlates the font size of a tag candidate with its score. The score is computed by a configurable function S given the histogram of categories present in the candidates aggregation area—which itself is font size dependent—as input. To keep the font sizes comparable across different time steps, we slightly modify this step of the original algorithm such that tags with the same font size in different time steps represent the same underlying score. The original bisection method assigns the tag candidates the font size f_{min} which have the score value $score_{min}$, the minimum of all occurring scores computed at f_{min}. Analogously f_{max} is assigned to the candidates having the score value $score_{max}$ computed at f_{max}. All other candidates are assigned font sizes lying as closely as possible on the line spanned by the two points $(score_{min}, f_{min})$ and $(score_{max}, f_{max})$ in the respective score vs. font size Euclidean space. Instead of determining $score_{min}$ and $score_{max}$ independently for each time step, we calculate the global minimum and global maximum of all time steps in a pre-processing step. In each time step, we use these global extrema instead, resulting in the same line the bisection method optimizes the font sizes to, thus, resulting in comparable font sizes across the time steps.

3.2 Visual Stability

Another issue we encountered were lots of small movements of tags between time steps resulting in visual clutter and an unsteady animation. We gain a steady animation by modifying S1 and S3 of the original PTM algorithm. Instead of calculating the seed positions independently for each time step in S1 we use the same, unified set of seeds for all time steps. Fig. 1a shows the situation of the naive approach for a minimal example consisting of three time steps, each completed S2. As such a tag candidate, depicted by its bounding box, is calculated for each of the three seed positions. In the first time step t_1, S3 of the unmodified greedy method will choose the tag candidate at s_1 to be displayed, since it has the largest font size, as well as the candidate at position s_3. The candidate at s_2 will be omitted as it overlaps with the tag at s_1. In t_2 however, only s_2 will be selected, while in t_3 again s_1 and s_3 will be chosen. Since all red tags share the same label, it appears as if this label slightly moves from s_1 to s_2 and back to s_1. This demonstrates an example of such unnecessary visual clutter.



(a) The naive approach treating each time step independently.

(b) Our proposed method, which groups consecutive tag candidates with identical category into atomic visual units.

Figure 1: An illustration of the visual stability issue when applying the PTM algorithm to each time slice independently. The tag candidates' bounding boxes for three seed positions s_1, s_2, s_3 and three consecutive time steps are shown. Solid boxes represent the tags which will be present in the final tag map, dotted boxes represent the tags which will be omitted. The color encodes the predominant category of each candidate.

Our proposed solution for this problem works by simply selecting a different, but still representative subset of non-overlapping tag candidates to render the final tag map in S3 of the PTM algorithm. After S2 is calculated for all time steps, we loop over all time steps for each seed position independently and chain together adjacent time steps of a seed position whose tag candidates share the same label. Fig. 1b depicts this for the example by solid colored, vertical lines. For s_1 and s_2 all three time steps belong to one chain, while for s_3 only the first two time steps connect, the last is a chain consisting of only one time step. Instead of deciding about the placement of the tag candidates for each time step independently in S3, we greedily place those chains as a whole, or rather their tag candidates, if space in respective time steps permit. The only thing left is to decide about the order the greedy algorithm processes the chains. We obtained good results by ordering descending by accumulated tag candidate score. As such, a chain with larger total score, which corresponds to larger accumulated font sizes, is placed before a chain having smaller accumulated font sizes. For equal accumulated scores, we order by number of time steps a chain has, then by label, then by position. In the example in Fig. 1b all tag candidates of the chain at s_1 will be placed first as it has the highest accumulated score. Then, the chain at s_2 will be omitted, as its tag candidates overlap with those of the chain at s_1. Finally, both chains at s_3 will be placed since space permits. As a result, the red labeled tags no longer move across time steps, but appear to change font size slightly, reflecting the change of score driven by the underlying area.

We incorporate two more measures to further increase the visual stability. First, it is desirable to suppress areas with a low density of data points. The probability that predominant categories change rapidly in those regions is high, compared to the low reliability of that information. One way to achieve this is determining low density areas and remove the data points of those regions. However, we propose another approach that employs our aggregation method. Before grouping the tag candidates in S3, we filter out those candidates with insufficient data points inside their bounding box, respectively with a low sum of histogram values derived from the label category distribution of the data points inside that box. Using the histogram for filtering allows for an abstraction useful for our trend visualization described in Sect. 4. Second, we implemented a sliding window approach to be able to smoothen the visualization over time. For each time step t, instead of using only the data points of t, we unify the data points of the time step range $[t, t+w]$ where w is the parameter of the window length and use this unified set.

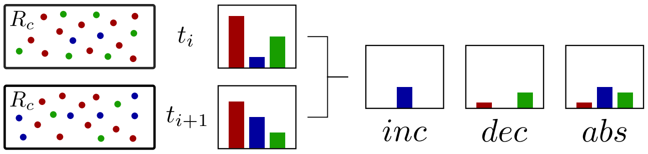


Figure 2: An illustration of the histogram calculation of a tag candidate for the three trend visualization modes showing increasing (inc), decreasing (dec) and absolute difference (abs) between two time steps t_i and t_{i+1}.

4 Trend Visualization using Differential Tag Maps

We model the data obtained after the division of our three dimensional input data into user-defined time slices of equal duration as a function

$$f : A \times T \to \mathbb{R}^n$$

that maps an arbitrary spatial region $a \in A$ for a given time step $t \in T$ into a vector, containing the number of occurrences[1] of each of the possible n categories. The predominant category is the component in that vector which has the largest value. A predominance tag map shows a subset $A' \subset A$—the bounding boxes of the tags—for a fixed time step t. As such our proposed animation shows the development of the predominant categories over T.

The analysis of trends is an important aspect when dealing with time-dependent data [1]. The rise or fall of a variable over time may give hints to interesting events or properties. Examples for our data domain could be the rise or fall of the number of supermarkets of a certain brand in a certain area to indicate the change of its market share or the development of the number of birds of a species over time to indicate if the species consists of migrant birds and at which times the migration happens. With respect to our data model, even if we only consider one specific area—one tag, we find the development of n variables over time. Our tag map approach always selects the predominant variable per time step, which results in visualizing the development of the predominant categories. However, there might be categories that change more dramatically between time steps than the predominant does. We designed a variant of our approach that visualizes the category which changes most instead. The original PTM algorithm assigns a label to a tag candidate by evaluating the mode of the histogram reflecting the category distribution inside the tag candidates aggregation rectangle R_c (which equals its bounding box in the ePTM variant). For a given time step t_i and a given area of a tag candidate R_c, we adapt this label assignment by first calculating the difference between each component of the histograms of R_c at t_{i+1} and t_i. For this resulting difference histogram the mode is calculated, so the predominant category of the difference histogram, which corresponds to the category which changes most, is assigned as the label of a candidate. This histogram can be obtained in three different variants shown in Fig. 2.

Of course, in addition to generate an animation showing the development of most changing categories, we can generate a static image showing the difference between any (non adjacent) time steps.

5 Computational Complexity

Our proposed time varying animation changes the overall complexity roughly by a factor of t compared to the original PTM algorithm, where t is the number of time steps. Calculating the unified set of seed positions is still minor with respect to the other steps, as well as greedily placing the chains of tag candidates. This step, as well as accounting for comparable font sizes by using global extrema of scores is just a rearrangement of operations. Calculating the chains of tag candidates can be done in $O(t \cdot s)$, where s is the number of seed positions. The adaptations to the histogram calculation for the trend visualization do not change its complexity. So the overall complexity of the time dependent version of e.g. the ePTM variant is $O(t \cdot s \cdot l^2 \cdot log^2 n)$, where l is the total number of unique categories and n is the maximal number of data points per time step.

[1] Mapping to $\mathbb{R}$ is a generalization which could be useful e.g. to model proportional scores

Table 1: Visual stability analysis showing summary statistics of the visual similarity of successive time steps of an animation. For all three data sets two versions of animations showing the evolution of predominant categories over time were analyzed. The means for all three data sets increased – at most for the *synthetical* data set – using the optimized version. It can further be seen that the supermarket data set shows the less visual change of predominance, while our *synthetical* data set contains the most, respectively contains the most noise which we consciously provoked.

	naive		optimized	
	μ	σ	μ	σ
synthetical	0.60	0.007	0.79	0.010
supermarkets	0.91	0.052	0.97	0.026
musicians	0.72	0.076	0.78	0.079

6 Evaluation

To evaluate our approach, we use one synthetical and two real world data sets.

Noisy Synthetical Data (*synthetical*): We generated a data set containing three categories "A", "B", and "C". It spans a rectangular area and consists of 30 time steps. Each time step is generated randomly using Poisson disc sampling [2]. We keep the density of the categories "A" and "B" uniform, where "A" is twice as dense. The density of "C" follows a two dimensional Gaussian distribution with its center moving diagonal across the time steps. The density at this center point for category "C" is higher than that of "A".

Open Street Map Supermarket Names (*supermarkets*): From the historic Open Street Map data set [23], we filtered those nodes that are positioned inside a bounding box spanning roughly across Europe and which are marked as supermarkets. We used their respective name attribute as their category. As the remaining set contained more than 10,000 different categories – which our algorithm is currently not able to handle computationally – we further filtered those nodes associated to the top 100 occurring names. We divided the remaining set into time slices of one month duration each.

Musical Professions (*musicians*): The musiXplora [11] is a database that offers biographical information of more than 30,000 musicians from the past 2,000 years. As an ongoing research project (started in 2004), the musiXplora deals with historical and present sources of data of different facets of musicology. Associated to the persons in the database are their musical professions, relevant places as well as years of their activity. We generated time steps after ten years each, starting with the year 1500. For each time step, we generated a data point for a location, if that location is associated with a person active in the respective year and associated the musical profession of that person to the data point.

Visual Stability We quantitatively evaluated the improvements to the visual stability described in Sect. 3.2 by using the matching visual overlap introduced in [24]. This measure calculates the percentage of identically labeled tag bounding box area between two tag maps. For a given animated tag map, we calculate this measure between any two successive time steps. Table 1 shows the mean and standard deviation over all obtained percentages. A video showing the animations used as basis for the calculation is included in the supplemental material.

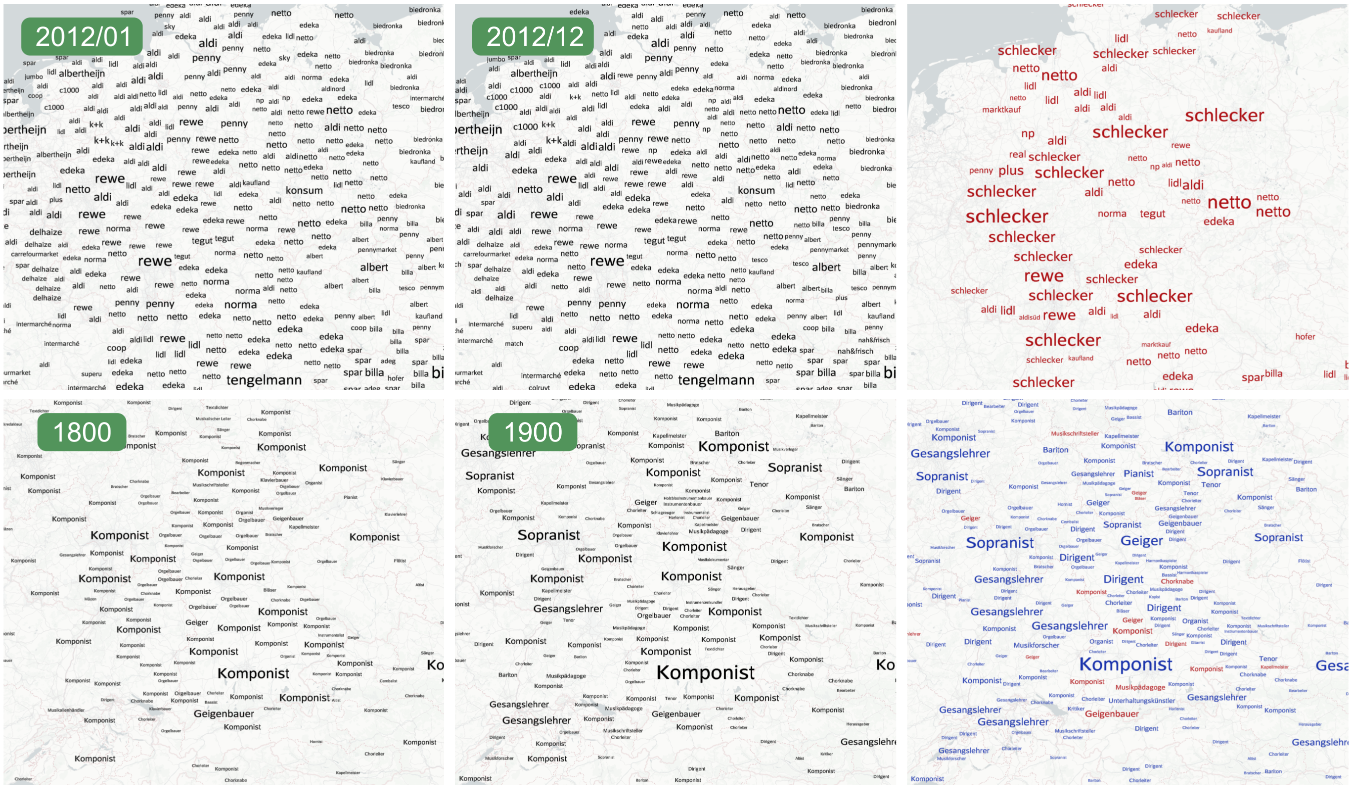


Figure 3: **Top row:** Results of the *supermarkets* data set. The left image is the frame of the animation showing the development of predominant categories at January 2012, the middle image is the frame at December 2012. The right image is the result of the trend visualization showing the most decreasing supermarkets (*dec*) between those time steps. We filtered out tag candidates with a low amount of change as described in Sect. 3.2. **Bottom row:** Results of the *musicians* data set. The left image shows the predominant categories at 1800, the middle image at 1900. The right image shows the trend visualization for both directions (*abs*). Increasing occurrences of professions between 1800 and 1900 are colored in blue, decreasing are colored in red.

Supermarkets By qualitatively analyzing the *supermarkets* data set using our method, we were able to detect and identify the change in market share of a number of supermarket brands. The top row of Fig. 3 depicts an example thereof. Our proposed trend visualization clearly shows the decreasing number of 'Schlecker' stores from the beginning to the end of 2012, which were not visible just by looking at the development of predominant categories. This decrease correlates with the time of the insolvency proceedings of the company in beginning of 2012[2] and the closing of the majority of their stores. The video in the supplemental material shows the development of predominance as well as a decreasing trend animation with which we were able to identify more of such events, e.g. the sale of the 'Tengelmann' (Southern Germany) brand in 2017 as well as the sale of 'C1000' (Netherlands) spanning from 2012 to 2015 or the insolvency of 'Zielpunkt' (Austria) in 2016.

Musical Professions Using our proposed animation of the temporal development of predominant tags and the proposed trend visualization we were able to strengthen and confirm findings related to the musiXplora database. Khulusi et al. [17] describe an evolution of musical institutions in the 19th century. Secularization processes lead to the foundation of institutions of higher education. At the same time the number of court orchestras decreased. It might be reasonable that such evolution coincides with the change of professions. In the bottom row of Fig. 3, we show the predominant professions at 1800, the predominant professions at 1900 as well as occurring trends by visualizing the difference between those time steps. A latent change in predominant professions can be seen in 1900 compared to 1800, apart from the overall increase in volume, such as the occurrence of 'Sopranist's (soprano) and 'Gesangslehrer' (singing teacher). In the right image a clearly upward trend of those and other professions can be seen, which gives an impression of this evolution. There are further interesting developments visible in the video included in the supplemental material, e.g. the rise and fall of the 'Meistersinger' (master singers) profession, which was predominant in a specific area in Germany around 1600.

7 Conclusion and Future Work

In this work, we focused on visualizing temporal changes in tag maps. We developed a geo-spatial, time-dependent visualization technique based on the Predominance Tag Map algorithm [24] that generates an animation optimized to reflect the development of predominant categories over time, while trying to avoid visual clutter by suppressing unnecessary visual changes. We further developed a new technique that visualizes trends among the occurrence of categories. The rise or fall of those categories are shown by generating differential tag maps between time steps. We evaluated both approaches using two data sets showing the development of supermarket stores and musical professions and we detected visual patterns that we could relate to real world events. While the algorithm is able to handle a large number of data points, it could be improved with respect to the number of different categories it can handle, as the *supermarket* data set indicates. Future work includes improving the scalability and the computational complexity of the approach.

[2]https://en.wikipedia.org/wiki/Schlecker

REFERENCES

[1] N. Andrienko and G. Andrienko. *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer-Verlag, Berlin, Heidelberg, 2005.

[2] R. Bridson. Fast poisson disk sampling in arbitrary dimensions. In *ACM SIGGRAPH 2007 Sketches*, SIGGRAPH '07. ACM, New York, NY, USA, 2007. doi: 10.1145/1278780.1278807

[3] K. Buchin, D. Creemers, A. Lazzarotto, B. Speckmann, and J. Wulms. Geo word clouds. In *2016 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 144–151, April 2016. doi: 10.1109/PACIFICVIS.2016.7465262

[4] S. Chen, S. Chen, L. Lin, X. Yuan, J. Liang, and X. Zhang. E-map: A visual analytics approach for exploring significant event evolutions in social media. In *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 36–47. IEEE, 2017.

[5] S. Chen, S. Chen, Z. Wang, J. Liang, Y. Wu, and X. Yuan. D-map+: Interactive visual analysis and exploration of ego-centric and event-centric information diffusion patterns in social media. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(1):11, 2018.

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

[7] W. Cui, S. Liu, Z. Wu, and H. Wei. How hierarchical topics evolve in large text corpora. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2281–2290, Dec 2014. doi: 10.1109/TVCG.2014.2346433

[8] T. Dang, H. N. Nguyen, and V. Pham. WordStream: Interactive Visualization for Topic Evolution. In J. Johansson, F. Sadlo, and G. E. Marai, eds., *EuroVis 2019 - Short Papers*. The Eurographics Association, 2019. doi: 10.2312/evs.20191178

[9] M. Dörk, S. Carpendale, C. Collins, and C. Williamson. Visgets: Coordinated visualizations for web-based information exploration and discovery. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1205–1212, Nov 2008. doi: 10.1109/TVCG.2008.175

[10] D. Duan, W. Qian, S. Pan, L. Shi, and C. Lin. VISA: A Visual Sentiment Analysis System. In *Proceedings of the 5th International Symposium on Visual Information Communication and Interaction*, VINCI '12, pp. 22–28. ACM, New York, NY, USA, 2012. doi: 10.1145/2397696.2397700

[11] J. Focht. MusiXplora, 2019. `www.home.uni-leipzig.de/mim` (Retrieved 2019-01-17).

[12] E. R. Gansner, Y. Hu, and S. C. North. Interactive visualization of streaming text data with dynamic maps. *J. Graph Algorithms Appl.*, 17(4):515–540, 2013.

[13] S. Hahmann and D. Burghardt. *Maple – a Web Map Service for Verbal Visualisation using Tag Clouds Generated from Map Feature Frequencies*, pp. 3–12. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. doi: 10.1007/978-3-642-19143-5_1

[14] M. C. Hao, C. Rohrdantz, H. Janetzko, D. A. Keim, U. Dayal, L. e. Haug, M. Hsu, and F. Stoffel. Visual sentiment analysis of customer feedback streams using geo-temporal term associations. *Information Visualization*, 12(3-4):273–290, 2013. doi: 10.1177/1473871613481691

[15] U. Hinrichs, B. Alex, J. Clifford, A. Watson, A. Quigley, E. Klein, and C. M. Coates. Trading Consequences: A Case Study of Combining Text Mining and Visualization to Facilitate Document Exploration. *Digital Scholarship in the Humanities*, 2015.

[16] A. Jaffe, M. Naaman, T. Tassa, and M. Davis. Generating Summaries for Large Collections of Geo-referenced Photographs. In *Proceedings of the 15th International Conference on World Wide Web*, WWW '06, pp. 853–854. ACM, New York, NY, USA, 2006. doi: 10.1145/1135777.1135911

[17] R. Khulusi, S. Jänicke, J. Kusnick, and J. Focht. An Interactive Chart of Biography. *Pacific Visualization Symposium (PacificVis), 2019 IEEE*, 2019.

[18] R. Krueger, Q. Han, N. Ivanov, S. Mahtal, D. Thom, H. Pfister, and T. Ertl. Bird's-Eye - Large-Scale Visual Analytics of City Dynamics using Social Location Data. *Computer Graphics Forum*, 2019. doi: 10.1111/cgf.13713

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

[20] D. Mashima, S. Kobourov, and Y. Hu. Visualizing dynamic data with maps. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1424–1437, Sep. 2012. doi: 10.1109/TVCG.2011.288

[21] D. Nguyen and H. Schumann. Taggram: Exploring geo-data on maps through a tag cloud-based visualization. In *2010 14th International Conference Information Visualisation*, pp. 322–328, July 2010. doi: 10.1109/IV.2010.52

[22] D. Q. Nguyen, C. Tominski, H. Schumann, and T. A. Ta. Visualizing tags with spatiotemporal references. In *Information Visualisation (IV), 2011 15th International Conference on*, pp. 32–39, July 2011. doi: 10.1109/IV.2011.43

[23] OpenStreetMap contributors. History Planet dump retrieved from https://planet.openstreetmap.org/planet/full-history/ . `https://www.openstreetmap.org`, 2019.

[24] M. Reckziegel, M. F. Cheema, G. Scheuermann, and S. Jänicke. Predominance tag maps. *IEEE transactions on visualization and computer graphics*, 24(6):1893–1904, 2018.

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

[26] G. Sun, Y. Wu, S. Liu, T. Peng, J. Zhu, and R. Liang. EvoRiver: Visual Analysis of Topic Coopetition on Social Media. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):1753–1762, Dec 2014. doi: 10.1109/TVCG.2014.2346919

[27] D. Thom, H. Bosch, S. Koch, M. Wrner, and T. Ertl. Spatiotemporal anomaly detection through visual analysis of geolocated twitter messages. In *2012 IEEE Pacific Visualization Symposium*, pp. 41–48, Feb 2012. doi: 10.1109/PacificVis.2012.6183572

[28] D. Thom, R. Krüger, and T. Ertl. Can twitter save lives? a broad-scale study on visual social media analytics for public safety. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1816–1829, July 2016. doi: 10.1109/TVCG.2015.2511733

[29] D. Thom, R. Krüger, T. Ertl, U. Bechstedt, A. Platz, J. Zisgen, and B. Volland. Can twitter really save your life? a case study of visual social media analytics for situation awareness. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 183–190. IEEE, 2015.

[30] F. B. Viegas, M. Wattenberg, and J. Feinberg. Participatory Visualization with Wordle. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1137–1144, Nov. 2009. doi: 10.1109/TVCG.2009.171