

Wheel of Life, an initial investigation: Topic-Related Polarity Visualization in Personal Stories

Henrique D. P. Santos, Greice P. D. Molin, Jackson Pinheiro, Renata Vieira

¹Faculdade de Informatica
Pontificia Universidade Catolica do Rio Grande do Sul
90619-900 – Porto Alegre – RS – Brazil

{henrique.santos.003, greice.molin, jackson.pinheiro}@acad.pucrs.br

***Abstract.** User-generated content is a rich source of information regarding human behavior in Internet social media. Sentiment analysis is a powerful tool to understand human psychological meanings in text. Visualizing these sentiments and knowledge about users is crucial to figure out the trends in data and to then use this information to make decisions. This work presents an initial investigation about a visualization chart considering topic-related polarities in personal stories by Brazilian bloggers. Visualizing these sentiments allows specialists to rapidly understand user-affected areas of life.*

1. Introduction

Since the Internet first appeared two decades ago, it has changed the way we interact and manage information. It is now possible to gather user-generated information from multiple sources and process this rich amount of data. One way to read this data is visualizing information aggregated in charts and interaction systems that summarize and organize them. [Ward et al. 2010] shows how important it is to choose the right visualization technique, so that users can easily understand the information data through 2D charts, maps or 3D systems.

Sentiment analysis is an increasingly important topic in Information Retrieval and Web data analysis. Visualizing this kind of information is also a trend, where identifying user opinions, sentiments and polarities about a topic is already possible through many algorithms and techniques. Now, the further step is to define how to properly show this information to specialists and decision makers.

In this initial work, we show a well-known figure, the Wheel of Life, applied to topic-related polarity in personal stories from Brazilian bloggers. In Section 2 we present some visualization techniques for sentiment data, while in Section 3 we explain the corpus used in our experiments. The origins of wheel visualization and our proposal are discussed in Section 4. Finally, we present our conclusions and further work in Section 5.

2. Related work

The interest in the area of visualizing sentiments in text has been growing lately. SocialHelix [Cao et al. 2015] and Twitter Collaborative Visualization [Brooks et al. 2014] have presented a similar approach, showing sentiment in a line chart, where each line represents the tweets' polarity and the variation of the sentiment about a topic over time

in a Twitter dataset. The SocialHelix system deals with differences of opinion in social networks. Therefore, it is a complex subject, since it involves people, communities, and events. The system works with data mining techniques for data analysis and knowledge discovery about differing feelings that change over time. In addition, there is the Emotion-Watch system, [Kempton et al. 2014] which deals with the granularity of feelings, since there is a concern over knowing about an event beyond the polarized feelings, considering the system deals with 20 categories of feelings and one in which no feelings exist. Both systems use Plutchik's wheel of emotions as a basis.

Another work approach was PEARL [Zhao et al. 2014], which uses a timeline chart together with a dense line for each sentiment (joy, anger, disgust, fear, anticipation, sadness, trust, surprise) to summarize and reveal emotional patterns over time. They use this visualization in relation to user feeds in Twitter and Facebook datasets.

On the other hand, researchers present Twitter sentiment analysis in a circle format. Senticompass [Wang et al. 2015] built a sentiment compass that shows the Valence/Arousal model in a circle and a radial timeline in time intervals separated by inner circles. Sentibank [Borth et al. 2013] uses Plutchik's wheel of emotions to segment all sentiment types in a Twitter corpus; users are able to change to a tree map view of the same sentiments.

Although the above works have proposed rich visualization tools, they fail to associate sentiments to topics of users' life. VIBES [Wensel and Sood 2008] tries to achieve this with visualizations about personal stories in LiveJournal blog platforms. They gather a timeline with users' posts and show the emotion graph in lines, the polarity related to the emotional topic in gauge charts, and finally an emotion cloud with polarity topics.

In this article, we show a different approach, using a well-known way to relate topics in people's lives and the polarity concerned with each one: the wheel of life.

3. Corpus of personal stories

In a previous work [dos Santos et al. 2017], we built a corpus with posts from the Blogspot platform ¹ written in Portuguese by Brazilian bloggers. The corpus contained 1,346,858 posts, 37,746 of which feature evidence of personal stories: subjectivity text with self references. We used [Moraes et al. 2016]'s approach to select subjectivity and [Benites et al. 2016]'s findings to filter self references. Finally, we selected 1,000 random posts to do a hand-annotation task in the Crowdfunder platform ².

The annotation process has an agreement percentage for each annotated text, so we only used the posts with higher agreement to train a machine learning model using the multinomial naive bayes algorithm. In our experiments, TF-IDF was the best index to find personal story texts, with 78% accuracy. Running this personal story model against the 37,746 posts, we found another 30,542 texts with high chances of having personal story content. The raw and annotated dataset can be retrieved in the author's Github page ³.

¹<http://www.blogspot.com>

²<http://www.crowdfunder.com>

³<https://github.com/heukirne/brazilian-blog-dataset>

4. Wheel of Life: overview

In this section we discuss the origins of the wheel of life and its application using a psycholinguist dictionary on the annotated corpus.

4.1. The origins of the wheel of life

The idea of using a wheel to represent areas of life comes from Bhavacakra, the central theory of Buddhism, the Wheel (*cakra*) of Life (*bhava*). Bhavacakra originally has three levels, and each level has three subdivisions: the first consists of the three basic unhealthy mental factors; the second represents the six worlds; and the third features the twelve *Nidānas*, which mean causes or motivations, the outbound sections in the wheel [Du Pre 1993].

Nowadays, a common exercise used in a personal coaching context is the Wheel of Life. It consists of filling the wheel areas based on life satisfaction first; then, the subject chooses some areas of low activation to improve, with objectives, goals and deadlines [Yamashita and Kato 2012].

4.2. Wheel of life as a visualization tool

The first step is to extract topics from the text and use them in a visualization tool. This requires text processing and sentiment analysis. For this work we use LIWC, a psycholinguistic dictionary that classifies English words according to 64 categories [Pennebaker et al. 2007]. The 2007 version of the LIWC public dictionary was also translated into Portuguese and used in a sentiment analysis experiment [Balage Filho et al. 2013]. Using the Portuguese version, we match LIWC categories to the most common areas of life used in the coaching wheel. Only the bottom hierarchy of LIWC categories were selected, the ones with no subcategories.

For this experiment, we split each text in sentences and match their topic and sentiment. For instance, if a sentence has more than one topic, each topic is associated with the polarity emotions expressed in the text. This naive approach was used to build the vector for each topic section in the wheel.

In Figure 1 we show main components of the Wheel of Life: (a) each circle section was described by a LIWC psychological category, called topic; (c) the color of each topic has its saturation variance; (b) besides the circle border, some topic percentages have a radial guide; (e) the medium color saturation refers to the references related to the positive topics in the sentence; (f) the dark color saturation refers to the references related to the negative topics; and (d) when a topic only has negative references to a topic, we keep a shell layer with medium saturation.

The wheel chart is still in development. It will have additional features. For example, each area could be clickable, showing how many sentences have the specific polarity in the topic, and the clicked area could display users' sample sentences of the topic-related polarity.

The processed corpus with the LIWC categories is available at the author's github page⁴. The chart draft using D3 library and the demo can also be reached in this github page.

⁴<https://heukirne.github.io/wheel-of-life/>

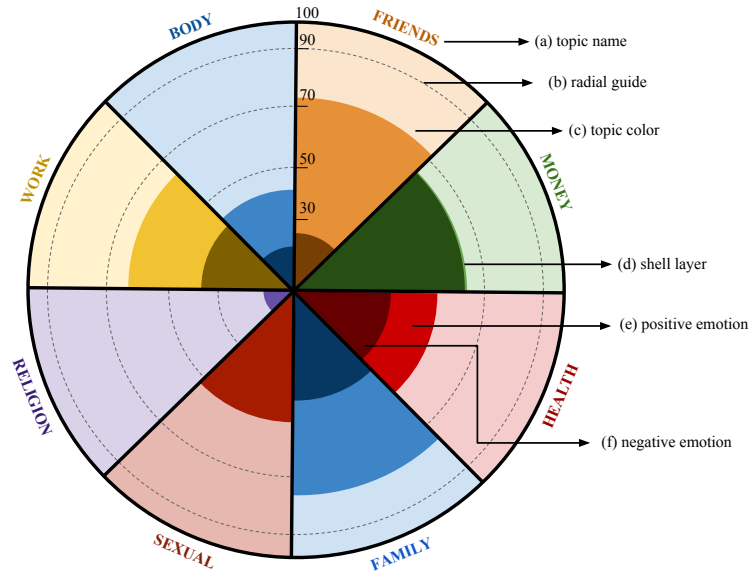


Figure 1. Wheel of Life: LIWC psychological categories related to its emotion valency in a radial bar chart

5. Conclusion and further work

By pre-processing the data, sorting it, grouping it through their techniques, data mining visualization tools enrich users' experiences when they are dealing with huge amounts of data. This early work shows an interesting way to access topic-related polarity in posts with personal stories. The Wheel of Life allows users to easily understand the author's valence in particular areas of their life. In order for lay users to be able to view and easily interpret their emotions during certain events, we find it interesting to use the EmotionWatch and SocialHelix tools to confront results, since the idea is to use the same dataset. With this, we can improve the initial work with the wheel of life. This interaction would be independent of specialized users, because it would be easier to understand the results. We will probably be able to develop an interactive tool that will contain the strengths of both models of visualization studied, resulting in a new, improved way of visualizing and understanding feelings.

References

- Balage Filho, P. P., Pardo, T. A., and Aluisio, S. M. (2013). An evaluation of the brazilian portuguese liwc dictionary for sentiment analysis. In *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology (STIL)*, pages 215–219.
- Benites, D., Gauer, G., and Gomes, W. B. (2016). Personal journal blogs as manifest internal conversation toward self-innovation: A semiotic phenomenological analysis. *Estudos de Psicologia (Campinas)*, 33(3):431–442.

- Borth, D., Chen, T., Ji, R., and Chang, S.-F. (2013). Sentibank: large-scale ontology and classifiers for detecting sentiment and emotions in visual content. In *Proceedings of the 21st ACM international conference on Multimedia*, pages 459–460. ACM.
- Brooks, M., Robinson, J. J., Torkildson, M. K., Aragon, C. R., et al. (2014). Collaborative visual analysis of sentiment in twitter events. In *International Conference on Cooperative Design, Visualization and Engineering*, pages 1–8. Springer.
- Cao, N., Lu, L., Lin, Y.-R., Wang, F., and Wen, Z. (2015). Socialhelix: visual analysis of sentiment divergence in social media. *Journal of Visualization*, 18(2):221–235.
- dos Santos, H. D. P., Woloszyn, V., and Vieira, R. (2017). Portuguese personal story analysis and detection in blogs. In *Web Intelligence (WI), 2017 IEEE/WIC/ACM International Conference on*, Leipzig, Germany.
- Du Pre, G. (1993). *SCIENCE AND THE WHEEL OF LIFE*. Motilal Banarsidass Publishes.
- Kempter, R., Sintsova, V., Musat, C. C., and Pu, P. (2014). Emotionwatch: Visualizing fine-grained emotions in event-related tweets. In *ICWSM*.
- Moraes, S. M., Santos, A. L., Redecker, M., Machado, R. M., and Meneguzzi, F. R. (2016). Comparing approaches to subjectivity classification: A study on portuguese tweets. In *International Conference on Computational Processing of the Portuguese Language*, pages 86–94. Springer.
- Pennebaker, J. W., Booth, R. J., and Francis, M. E. (2007). Linguistic inquiry and word count: Liwc [computer software]. *Austin, TX: liwc. net*.
- Wang, F. Y., Sallaberry, A., Klein, K., Takatsuka, M., and Roche, M. (2015). Senticompass: Interactive visualization for exploring and comparing the sentiments of time-varying twitter data. In *Visualization Symposium (PacificVis), 2015 IEEE Pacific*, pages 129–133. IEEE.
- Ward, M. O., Grinstein, G., and Keim, D. (2010). *Interactive data visualization: foundations, techniques, and applications*. CRC Press.
- Wensel, A. M. and Sood, S. O. (2008). Vibes: Visualizing changing emotional states in personal stories. In *Proceedings of the 2Nd ACM International Workshop on Story Representation, Mechanism and Context*, SRMC '08, pages 49–56, New York, NY, USA. ACM.
- Yamashita, H. and Kato, S. (2012). The wheel of language learning: A tool to facilitate learner awareness, reflection and action. *Advising in language learning: Dialogue, tools and context*, pages 164–169.
- Zhao, J., Gou, L., Wang, F., and Zhou, M. (2014). Pearl: An interactive visual analytic tool for understanding personal emotion style derived from social media. In *Visual Analytics Science and Technology (VAST), 2014 IEEE Conference on*, pages 203–212. IEEE.