SHAPING OPINIONS: VISUAL ANALYSIS OF PUBLIC DISCOURSE ON ENVIRONMENTAL ISSUES

WOLFGANG KIENREICH[1]
Know-Center, Competence Center for Knowledge-Based Applications and Systems, Inffeldgasse 21a, 8010 Graz, Austria

Abstract. The public discourse on environmental issues employs the news media and the emerging consumer generated media as its primary communication channels. Analyzing the usage of these channels by the various discourse participants yields valuable insight into the status of opinion formation on environmental problems. This paper outlines common methods for the monitoring and visualization of public discourse in the news media and proposes requirements for the application of such methods to environmental discourse. The integration of geospatial visualizations with semantic dimensions and numeric data is identified as the key challenge in visualizing public discourse on environmental issues. A showcase application which addresses this challenge is briefly presented.

Keywords: Public Discourse, Environmental Issues, Visualization.

1. Introduction

Public discourse refers to the process of collaborate opinion formation and decision making which enables consensual problem solving on a societal level. In practice, the media serve as the prime communication channel for this process. The role of the news media and of consumer generated media is particularly pronounced in public discourse on environmental issues, as outlined in section 2.

---

[1] Kienreich, Wolfgang, Know-Center, Competence Center for Knowledge-Based Applications and Systems, Inffeldgasse 21a, 8010 Graz, Austria; e-mail: wkien@know-center.at

The observation of public discourse in the news media relies primarily on automated media monitoring. Information retrieval techniques and content analysis methods are applied to process news content into quantitative results, as discussed in section 3. Several visualizations of media attention, issue coverage and other results of content analysis are presented in section 4. Public discourse on environmental issues shows several peculiarities which require modification and extension of these standard visualizations. Section 5 formulates some requirements for visualizations of environmental discourse and presents a showcase application.

2.   Public Discourse On Environmental Issues

Considering common definitions of political discourse (Johnson and Johnson 2000), public discourse on environmental issues can be understood as the communication process which enables collaborative opinion formation and decision making with regard to environmental problems. Pragmatic environmental discourse is conducted by domain experts, as for instance environmental scientists, by representatives of economic interest and by affected citizens and their representatives. (Dryzek 2005)

Public discourse on environmental issues is characterized by the exchange of views between these agents through the news media and, recently, through consumer generated media. The role of the media in this process is emphasized by the international nature of issues, which restrains agents from assessing matters based on personal experience, and by the complexity of the decision basis, which requires elaboration and explanation in order to be comprehensible to non-experts.

In the wake of the United Nation's Agenda 21 programme for sustainable development, many nations and international organizations have adopted regulations which ensure free access to environmental information. Such legislation is intended to provide all actors of the environmental discourse with a consistent decision basis. For example, individuals and organizations have the right to access environmental information generated by public authorities in the European Union.

Considering the outlined role of the media in the environmental discourse, such legislation is insufficient if the media fail to communicate relevant information to the general public in a comprehensible way (Bell 2004) or set agendas which do not reflect the real-world priority of environmental problems (Nas 2000).

The constant monitoring of media coverage by means accessible to the general public is therefore in the best interest of all participants of the public discourse on environmental issues.

3. Monitoring Public Discourse

First attempts to gauge informal public discourse using polls date from the early 19th century. A firm statistical basis for opinion research has been developed in the 20th century. Opinion polls yield explicit, quantitative information on the state of opinion formation for a specific set of questions. Results can immediately be analyzed and visualized. However, opinion polls are less suitable for monitoring the formal exchange of views in the media which precedes and affects opinion formation. It is interesting to note that recent trends towards consumer-generated media, as, for instance, blogs and discussion forums, have opened an alternative window into informal public discourse.

Public discourse in the news media has been monitored by dedicated agencies since the end of the 19th century, when "clipping services" started to collect and process news articles. In the second half of the 20th century, content analysis has been developed as a standard methodology for the analysis of news articles. (Holsti 1969) Content analysis relies on quantitative, statistical properties of large text corpora, as for instance the frequency of keywords, to identify patterns and structures in communication. This approach has been criticized for excessive emphasis upon quantification, and alternatives focusing on qualitative features have been proposed (Goffman 1974) Content analysis continues to provide the basis for the automated processing of large corpora because it is comparably easy to realize using information technology.

Modern automated media monitoring systems collect and merge content from dedicated news databases and online sources as, for instance, websites or blogs, into a single, large text corpus which is annotated with metadata like author, date and source. Information retrieval methods as, for instance, lexical analysis, stop word removal, stemming and indexing are applied to facilitate searching the corpus and to prepare the corpus for further processing (Baeza-Yates and Ribeiro-Neto 1999). Named entity recognition based on linguistic patterns or statistical models is performed to identify entities as, for instance, organizations, persons and geographical references. Once named entities have been identified for a corpus, classical content analysis can be applied and results can then be visualized as outlined in the next section.

It is interesting to note that the processing steps described so far can be executed in near-realtime for a typical search result set which contains several thousand news articles. Therefore, it is possible to interactively obtain media monitoring results for arbitrary issues as long as the formulation of an adequate search query is feasible.

4.   Visualizing Public Discourse

The term visualization commonly refers to the use of visual representations to support communication and comprehension of complex information. Applications of visualization techniques in the context of public discourse can be traced back to the 19th century. The development of real-time computer graphic systems in the late 20th century has enabled the creation of dynamic and interactive visualizations. Information visualization has emerged as the scientific field investigating the use of computer-supported visual representations to aid understanding of abstract data (Card et al. 1999). Findings in Information Visualization have soon been applied to the visualization of news articles (Renisson 1994) (Andrews et al. 2002). Today, a plethora of web-based visualization systems backed by automated media observation systems provides visual analysis of public discourse to a wide audience.

Attention analysis investigates the frequency and extent of media coverage allocated to specific concepts (as, for example, individuals or products). The visual representation of media attention towards public actors (as, for example, election candidates) constitutes a prime application area of public discourse visualization. For small, coherent sets of actors, attention values are often displayed using bar charts, pie charts or line graphs, which are supplemented with descriptive labels and graphics. The United States presidential election of 2008 provides a number of examples, for instance (US Election 2008 Web Monitor). For large, diverse sets of actors, attention values are often displayed using figurative graphical representations which place actors in a visual context designed to ease comprehension and interpretation (compare Figure 1).
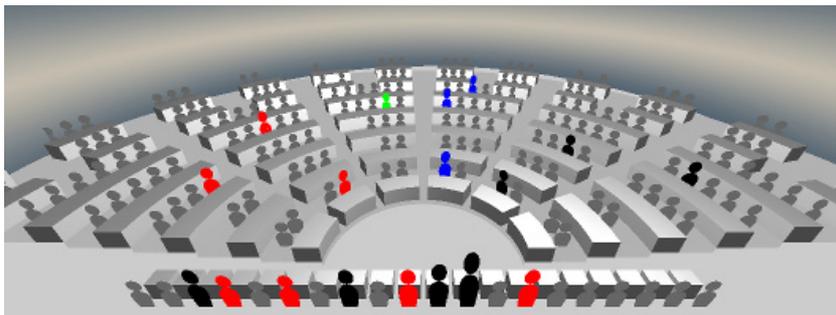

Figure 1. Visualizing media attention towards members of the Austrian parliament: Each Member of Parliament is represented by an icon which is correctly positioned within a stylized model of Austria's National Council. Icon size indicates relative attention value and icon color indicates party affiliation. (Kienreich et al. 2008)

Issue analysis extends the idea of attention analysis by investigating the context of the media coverage allocated to specific concepts. From a technical point of view, issue analysis amounts to a co-occurrence analysis of concepts. One prominent application in the domain of public discourse is the visualization of the issues most often associated with a public actor (compare Figure 2). Again, the United States presidential election of 2008 provides a number of examples, for instance (Washington Post 2008).

Figure 2. Visualizing issues associated with the top candidates for the Austrian Parliamentary Election of 2008: Each candidate is represented by an icon and a place card. Icon size indicates relative attention value and icon color indicates party affiliation. A tag cloud displays issues associated with the selected candidate. (Kienreich et al. 2008)

To enable the use of domain-specific visual representations, the results obtained from issue analysis can be restricted by narrowing the investigated type of context. Limiting the context to geographical references facilitates a geospatial visualization revealing geographical patterns in public discourse (compare Figure 3). Recursive analysis within the context of public actors yields a directed graph of individuals which can be displayed using social network visualization techniques (Freeman 2000).
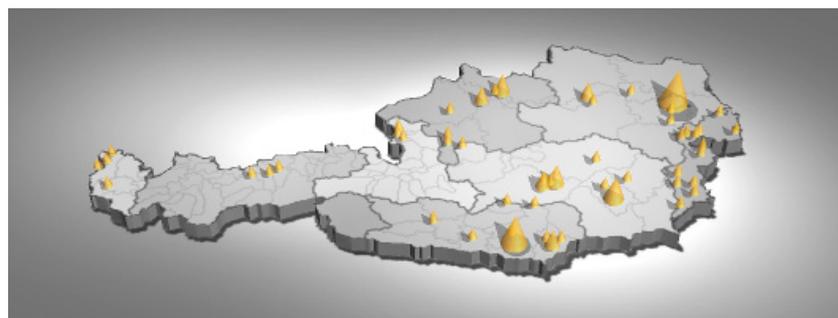

Figure 3. Visualizing geographical references associated with one of the top candidates for the Austrian Parliamentary Election of 2008: Each geographical reference is represented by yellow cone located on a stylized map of Austria. The size of each cone indicates relative association frequency. (Kienreich et al. 2008)

Sentiment analysis investigates the attitude of media coverage allocated to specific concepts by ranking references to a concept on a scale ranging from positive to neutral to negative. Reliable detection of sentiment is problematic because intricate language structures (as, for instance, irony or double negatives) are often used in relevant text segments. Successful applications of sentiment analysis to the domain of public discourse have been based on heuristics and term dictionaries (Scharl and Pollach 2003). Suitable visual representations are analogous to those described for attention and issue analysis. In the absence of a focus set of concepts, global surveys of public discourse yield massive amounts of relevant information. Corresponding visualizations usually present an initial overview and provide interactive facilities for zooming and filtering, in accordance with the "mantra of visual information seeking" (Shneiderman and Plaisant 2004). Adequate visualization strategies include, for instance, tree maps (Bederson et al. 2002), night sky metaphors (Andrews et al. 2002) and tag clouds (Seifert et al. 2008) (compare Figure 4).

Figure 4. Visualizing an overview of public discourse: Named entities have been extracted from news articles and arranged into areas representing topics. Each area contains a tag cloud of entities. Each entity displays an icon denoting entity type and a descriptive label. Label size denotes entity frequency in the corpus. (Seifert et al. 2008)

5.   Visualizing Environmental Discourse

The visual representations discussed in the previous section required significant modification and extension in order to yield comprehensible and helpful visualizations of public discourse on environmental issues. In particular, environmental concepts subjected to attention and issue analysis must be presented in a (visual) context which enables correct interpretation and verification.

For instance, an issue analysis of the concept "ozone" yields a large number of geographical references, which have to be presented in a geospatial visualization to enable recognition of patterns. Said analysis also yields a large number of technical terms, which have to be explained and put into context for non-experts, for instance by using a semantic graph. Attempts to retrace arguments about "ozone" involve numeric data which has either been present in the media articles or has to be acquired from scientific literature. This data has to be visualized, too.

Some requirements for the design of comprehensible visualizations of environmental discourse can be devised from peculiarities of the domain as outlined in the presented example:

- A geospatial visual representation should be provided if at all possible. Geospatial services are readily available. It is estimated that at least 20% of all web pages contain accessible geographic references (Delboni 2005). Most news articles contain at least one geographic reference. Thus, this requirement should be easy to implement.

- Semantic (visual) representations should be employed to present the problem domain in a consistent way (as opposed to presenting separate, unrelated issues). Domain ontologies are available for many areas of environmental information. However, implementing this requirement may demand advanced analysis methods (for instance, ontology extraction and ontology alignment) in the monitoring process.

- Supplemental information (as, for instance, numeric data) should be integrated into the primary visual representation or presented in separate views, which should be coordinated with the primary visualization.

Partial solutions have been presented for some of the formulated requirements in related research areas. Design principles and applications of multiple coordinated views have been documented in information visualization and visual analytics. The integrated visualization of multidimensional data and geospatial information has recently been addressed in architectural heritage visualization (Blaise et al. 2006).

The Media Watch on Climate Change (Scharl et al. 2007) provides a showcase for some of the formulated requirements. This automated media monitoring system collects news articles relevant to the domain of climate change from 150 online sources. It extracts geographic references from article content and gradually generates a geospatial knowledge base. The results are presented in a rich internet application which employs multiple coordinated views. Available visualizations include a geographic map, a semantic map, a graph-based ontology view and a tag cloud.
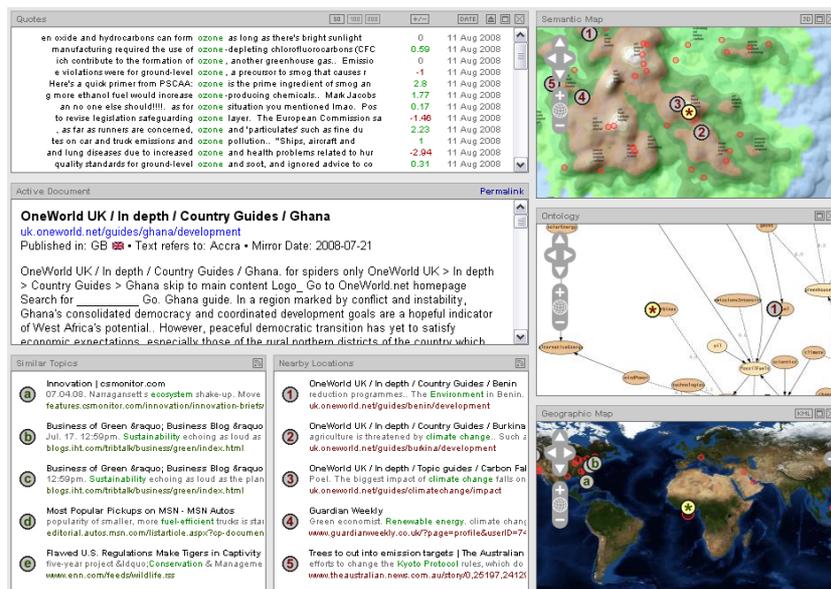
Figure 5. Visualizing public discourse on climate change: Similar topics and locations (bottom left) are presented for a news article (center left) selected from a list of search results (top left). Three synchronized visualizations present the search results as a semantic map (top right), an ontology graph (center right) and a geographical map (bottom right). The visualizations have been centered on the location of the selected article. (Scharl et al. 2007)

6.   Conclusions

We have described properties of public discourse on environmental issues and outlined common techniques for monitoring and visualizing public discourse in the news media. Based on the particular properties of environmental discourse, we have proposed some requirements for automated media monitoring and visualization systems targeting the environmental domain and presented a showcase system.

Acknowledgements

7. References

Andrews, K., Kienreich, W., Sabol, V., Becker, J., Droschl, G., Kappe, F., Granitzer, M., Auer, P., and Tochtermann, K., 2002, The InfoSky visual explorer: exploiting hierarchical structure and document similarities, *Information Visualization*, **1**(3):166-181

Baeza-Yates, R., and Ribeiro-Neto, B., 1999, *Modern Information Retrieval*, ACM Press, Addison Wesley, New York, USA.

Bederson, B.B., Shneiderman, B., and Wattenberg, M., 2002, Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies, *ACM Transactions on Graphics*, **21**(4): 833-854.

Bell, A., 1994, Climate of Opinion: Public and Media Discourse on the Global Environment, *Discourse & Society*, **5**(1): 33-64.

Blaise, J.Y., Dudek, I. and De Domenico, F., Spatial distribution and visual analysis of architectural semantic features, in: *Proceedings of the 6th International Conference on Knowledge Management*, JUCS - Journal of Universal Computer Science, Graz, Austria

Card, S.K., Mackinlay, J.D. and Shneiderman, B., 1999, *Information Visualization: Using Vision to Think*, Morgan-Kaufmann, San Francisco, USA.

Delboni, T.M., Borges, K.A.V. and Laender, A.H.F., 2005, Geographic Web Search based on Positioning Expressions, in: *Proceedings of the 2005 Workshop On Geographic Information Retrieval,* Jones, C. and Purves, R., eds., ACM, Bremen, Germany, pp. 61

Dryzek, J.S., 2005, *The Politics of the Earth: Environmental Discourses,* Oxford University Press, New York, USA.

Freeman, L., 2000, Visualizing social networks, *Journal of Social Structure*, **1**(1)

Goffman, E., 1974, *Frame Analysis: An essay on the organization of experience*, Harvard University Press, Cambridge, England, pp. 21.

Harris, R.L., 1999, *Information Graphics: A Comprehensive Illustrated Reference*, Oxford University Press, New York, USA.

Holsti, O.R., 1969, *Content Analysis for the Social Sciences and Humanities*, Addison-Wesley, Reading, Massachusetts, pp. 235.

Johnson, W., and Johnson, D., 2000, Civil Political Discourse in a Democracy: The Contribution of Psychology, University of Minesota (August 19, 2008) http://www.co-operation.org/pages/contro-pol.html

Kienreich, W., Lex, E. and Seifert, C., 2008, APA Labs: An Experimental Web-Based Platform for the Retrieval and Analysis of News Articles, in: *Proceedings of the 1st International Conference on the Application of Digital Information and Web Technologies*, IEEE, Ostrava, Czech Republic

Nas, M., 2000, Sustainable Environment, Unsustained Attention: A study of attitudes, the media and the environment, in: *SCP reports on public opinion and the environment,* Social and Cultural Planning office of the Netherlands, The Hague, the Netherlands.

Rennison, E., 1994, Galaxy of News: An Approach to Visualizing and Understanding Expansive News Landscapes, in: *Proceedings of the 7th ACM Symposium on User Interface Software and Technology*, ACM Press, Marina del Rey, USA.

Scharl, A. and Weichselbraun, A., 2006, Web Coverage of the 2004 US Presidential Election, in: *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, Kilgarriff, A., and Baroni, M., eds., Association for Computational Linguistics, Trento, Italy, pp. 35-42.

Scharl, A., Weichselbraun, A., Hubmann-Haidvogel, A., Stern, H., Wohlgenannt, G. and Zibold, D., 2007, MediaWatch on Climate Change: Building and visualizing contextualized information spaces, in: *Proceedings of the 6th International Semantic Web Conference*, Noy, N., ed., Springer LNCS, Busan, Korea

Scharl, A. and Pollach, I., 2003, Determining the Semantic Orientation of Web-based Corpora, in: *Proceedings of the 4th International Conference Intelligent Data Engineering and Automated Learning*, Liu, J., Cheung, Y., and Yin, H., (eds), Springer LNCS, Berlin, Germany, pp. 840-849.

Seifert, C., Kump, B., Kienreich, W., Granitzer, G. and Granitzer, M., 2008, On the Beauty and Usability of Tag Clouds, in: *Proceedings of the 12th International Conference on Information Visualization*, IEEE Computer Society, London, England.

Shneiderman, B. and Plaisant, C., 2004, *Designing the User Interface: Strategies for Effective Human-Computer Interactions*, Addison-Wesley, Reading, Massachusetts

US Election 2008 Web Monitor, EcoResearch Network, 2008, (August 19, 2008), http://www.ecoresearch.net/election2008/technology

Washington Post Issue Coverage Tracker, Washington Post, 2008 (August 19th, 2008), http://www.washingtonpost.com/wp-srv/politics/interactives/campaign08/issues