WeGov Analysis Tools to connect Policy Makers with Citizens Online

Timo Wandhöfer, GESIS – Leibniz Institute for the Social Sciences, Knowledge Technologies for the Social Sciences, Unter Sachsenhausen 6-8, 55667 Köln, Germany
        timo.wandhoefer@gesis.org

Catherine Van Eeckhaute, Gov2u, Square Eugène Plasky 92-94, b1, 1030 Brussels, Belgium
        catherine@gov2u.org

Steve Taylor, IT Innovation Centre, Enterprise Road, Southampton SO16 7NS, UK
        sjt@it-innovation.soton.ac.uk

Miriam Fernandez, Open University, Milton Keynes MK7 6AA, Knowledge Media Institute, UK
        M.Fernandez@open.ac.uk

Abstract

WeGov[1] is an online-accessible toolbox for policy makers that provide tools and techniques to engage with citizens on well-established social networking sites (abr. SNS). It provides three broad categories of functions, enabling the policy maker to search for discussions, topics and opinions from different SNS; analyse and summarize these discussions to determine the themes and important posts; and finally to help them inject information into the SNS. The different analysis components allow users and posts to be identified with respect to a particular topic. The topics and opinions analysis identify groups of words that represent the topics within a discussion. The prediction of user activity shows posts that are going to generate more attention and the modelling of user behaviour analysis classifies users according to their behaviour and interactions within the SNS.

Keywords: Social networking sites, e-Government, policy maker, e-Society, citizen, dialogue.

1   WeGov Approach

The WeGov project addresses the connection of citizens with governmental policy makers through social networks like Twitter and Facebook. Before the project dedicated citizen participation platform were used for political engagement, but these were often unpopular, so WeGov has taken the approach of using tools and platforms the citizens already use as the channel to connect with policy makers. WeGov sees itself as a feasibility study to exploit the potential of social networks for policy making, by providing easy access to citizen opinions for policy makers. The approach chosen consists of developing a website including tools that support the political decision-makers in the analysis of social networks. In terms of methodology, WeGov has had close participation of potential users (e.g. policy makers, communities, organizations) in the development process of the software to determine requirements and provide feedback on prototypes. The challenge is to reconcile as much as possible the requirements of these user groups in terms of social media analysis with the technical feasibility of the WeGov analysis models. WeGov has developed diverse analytical approaches that are currently tested and improved, as a basis for integration in the policy maker's daily workflow.

---

[1] EU project WeGov – Where eGovernment meets the eSociety: http://wegov-project.eu (Retrieved April 2012)

Wandhöfer et al.
WeGov Analysis Tools to connect Policy Makers with Citizens Online

1

2   WeGov Toolbox

The WeGov toolbox is a web application to support policy makers' everyday interactions with citizens on SNS. Politicians login with a username and password to see the landing page showing updated search and analysis results in a set of widgets that can be customized around topics and groups of their choice (Cp. 1st screenshot of Figure 1). Widgets are made in a number of different types, for example analysis of topics that SNS users discuss on a Facebook page. Some widgets can be geographically restricted to the current location of the politician and additional locations can be added to the WeGov toolbox. In addition to widgets on the landing page, the toolbox supports a search page with more detailed search functionality and analysis results (Cp. 2nd screenshot of Figure 1).
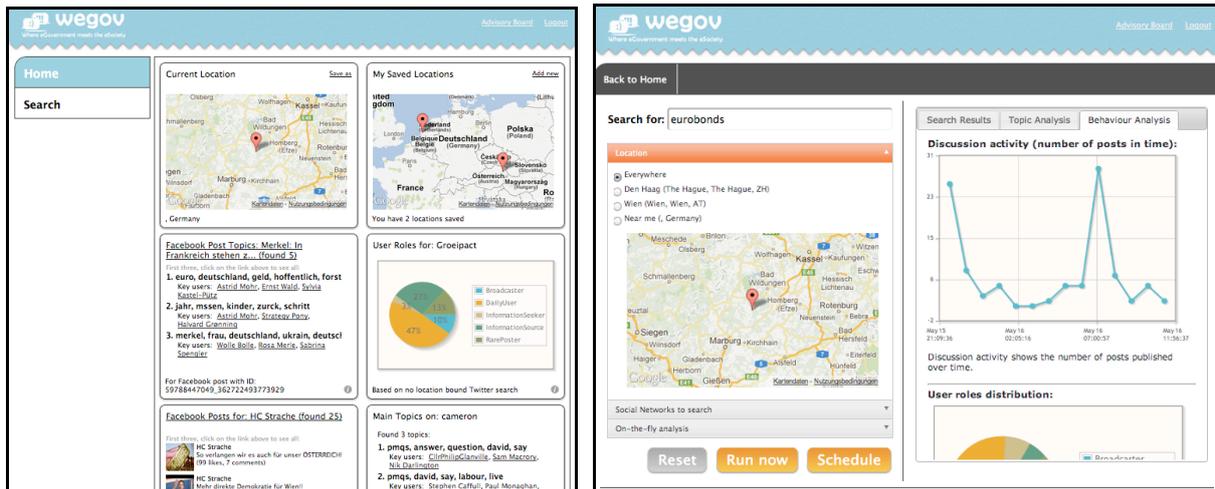
Figure 1   Toolbox concept.

2.1   Search

The WeGov toolbox is connected with SNS throughout their application-programming interface (abr. API) to retrieve data from, and post data to, SNS. Currently Facebook and Twitter are interlinked with the toolbox because they are widely used accepted by many citizens. The implementation of Google maps functionality within the toolbox allows the policy maker to see the current location that is identified by Google on the map. Figure 2 shows on the left hand side Uxbridge, in United Kingdom, as current location. The map on the right hand side of Figure 2 shows two locations: London and Berlin. These locations were set by the policy maker as their favourite locations and can be used in different searches to make them local. Within other widgets the current location will be used to show analysis results on local search queries (for example a politician's constituency).
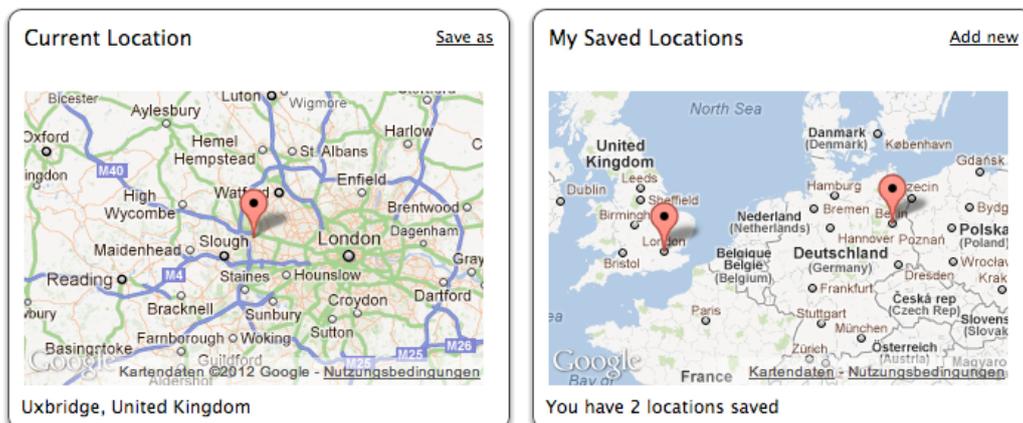
Figure 2   Google location.

2.2   Analysis: Topics and Opinions

The "topics and opinions" component identifies groups of words that represent several areas of discussions that arise within a wider debate. This analysis is used for classifying individual comments and users by words or phrases they have in common – similar to a summary of subtopics that are discussed within the debate. Each group of words composes one topic. For each topic, key users and key posts (the strongest users and posts) are displayed. The analysis also considers content related factors such as word frequency and the use of hash tags (#), but non content-related factors such as the number of re-tweets are ignored. The first frame within Figure 3 shows the topic analysis running on the user comments of a post on David Cameron's Facebook page. The second frame within Figure 3 shows the analysis results applied on a Twitter query on the search term "nuclear power". This analysis is not limited to Facebook or Twitter and can be applied on other text documents as well. (Naveed et al. 2011, Sizov 2010)
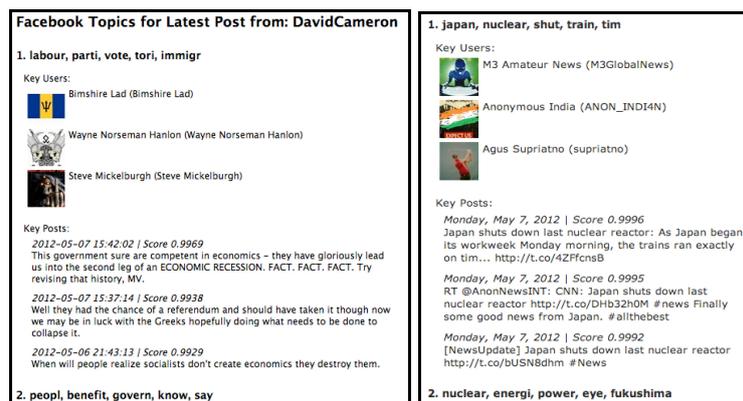
Figure 3   Topics with Key Users and Key Posts.

2.3   Analysis: Frequency of Posts

The first screenshot of Figure 4 shows the outcome of a Twitter query with the search term "European citizens initiative" entered on April 16, which generated 62 posts. The outcome is a list of posts that is ordered by the date and time the posts were written (most recent first). Another representation for the same set of posts shows the second screenshot in Figure 4. Here the analysis component visualizes how the tweets are concentrated in the concerned time period. The graph shows that the 62 posts were published over one week, with activity peaking on the first day. This analysis component is included within the search page of the WeGov toolbox. A policy maker from the German Bundestag mentioned that this visualization could be used for a long-term analysis – to display how the discussion activity changes on one particular topic over a longer time frame, and would be especially useful in combination with sentiment analysis.
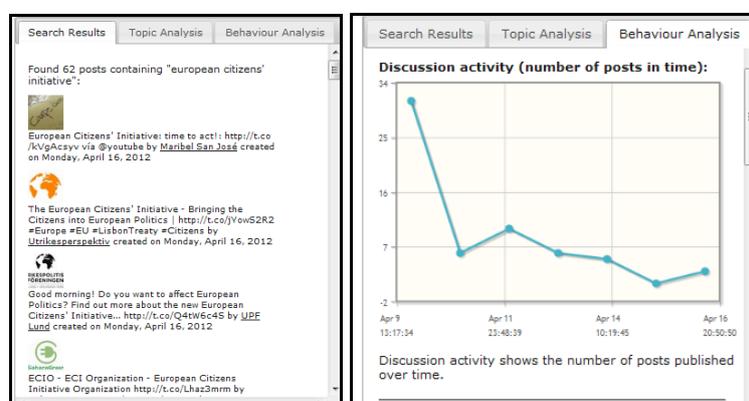
Figure 4   Frequency of posts.

2.4 Analysis: Prediction of Discussion Activity

The purpose of this analysis is to predict which posts are expected to generate more attention. In particular, the posts that generate high levels of attention generally fit the following characteristics:

- not written in the afternoon;
- written in a familiar language (the readability is high);
- written by people who follow many users and listen what they say (high out-degree[2]);
- the statement tends to be negative (stronger negative polarity).

In the WeGov toolkit the output of this analysis is translated into "Top Posts to Watch" (Cp. 1st screenshot of Figure 5). The "Top Users To Watch" (Cp. 2nd screenshot of Figure 5) are these users whose posts are often in the lists of "Top Post to Watch". Figure 5 shows the analysis results for a local search in Southampton, UK, for a current hot topic about a new technology for monitoring Football goal-lines. This component is available on the search page of the WeGov toolbox. (Rowe et al. 2011a, Rowe et al. 2011b)

Figure 5   Top 5 Users to Watch and Top 5 Posts to Watch.

2.5 Analysis: Modelling User Behaviour

The purpose of this analysis is to classify users according to their behaviour and interactions within the SNS. For this analysis we use the following features: in-degree[3] and out-degree, which represent the social network properties of the users; and post count, post rate and age, which represent the participation information of the user. Within WeGov the following behaviour roles are considered:

- Broadcaster is someone who posts with high daily rate and has a very high following (in-degree). However he follows very few people (out-degree).
- Information Source is someone who posts a lot, is followed a lot but follows more people than the Broadcaster. His engagement in the social network is higher than the Broadcaster.
- Daily User is an average user in relation to the number of posts, followers and the people he follows himself.
- Information Seeker is someone who follows many users but do not post frequently themselves. An information seeker is generally interested in gathering information from others.
- Rare Poster is someone who hardly ever posts.

According to these behavioural characteristics, the most influential and engaged role is the Information Source. Policy makers may be therefore interested in the users who fall into this category. In the WeGov toolkit, the output of this analysis is translated into a cake diagram that shows the

---

[2] Out-degree refers to the Twitter users that a person follows and potentially reads their posts and reacts by responding or retweeting.
[3] In-degree refers to the direction of the followers, i.e. the circle of persons who potentially read a post of that particular user.

percentage of the five user roles and the user profiles for one particular role. In the toolbox the components are available as widget and on the search page as well. The left hand side of Figure 6 shows cake diagram on the search term "nuclear energy" and on the right hand side the users that were identified as "Information Sources". (Angeletou et al. 2011)
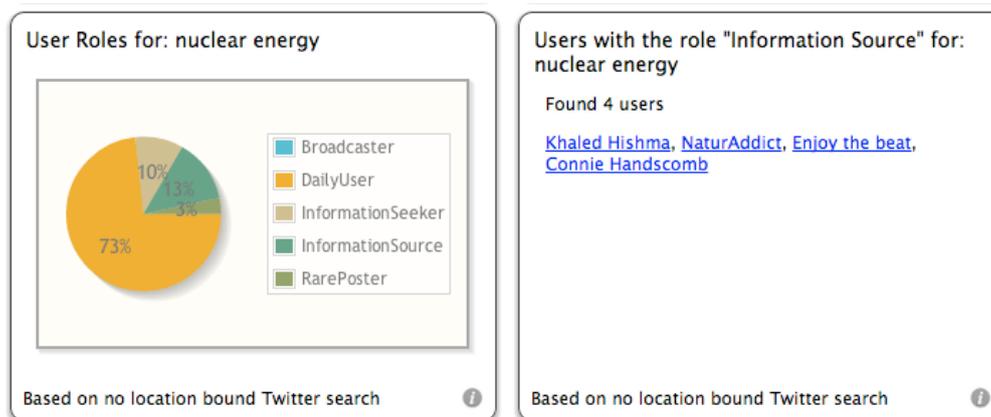
Figure 6   Visualization of User Roles.

3   POLICY MAKERS' FEEDBACK

Our set of end users was in effect self-selecting: of those approached, the ones that wished to engage with us were those with interest in social networking. However our end user group includes members from different kinds of parliaments as well as their office members responsible for public relation and presswork issues. These stakeholders come from the EU Parliament, the German Bundestag and from the State Parliament of Nordrhein-Westfalen in Germany. In addition the group includes stakeholders from parliamentary parties and other public organizations. Most of the stakeholders were recruited in public meetings and conferences where the WeGov approach was presented and discussed with the number of participants.

Initial use cases (Addis et al. 2010) and basic functionalities (Wandhöfer et al. 2011a) were identified and discussed with policy makers. During initial meetings with external end users, a particular need of the policy makers was described as a requirement. This is the gathering citizens' opinions as feedback to a particular statement by a politician. The first WeGov prototype covered this scenario as a basic use case. Here, the policy maker posts a statement into a social network, collects the citizens' feedback (where it is publicly available) and runs the analysis components on the feedback. The result is a summary of the key themes and opinions over the sum total of the citizens' comments.

This use case was implemented in the initial toolbox and was presented to 29 office employees working for a parliamentarian of the German Bundestag with the aim of gathering feedback for the further development process. During discussions with them, the consensus was that parliamentarians' posts are unlikely to solicit a large amount of feedback, unless the politician is high-profile: "ordinary" parliamentarians' posts typically generate below 100 comments. They confirmed that the requirement to test citizens' reactions to politicians' statements is important, but they need more comments to provide a statistically significant sample of opinions. A modification of the original use case was proposed by the Bundestag employees, where politicians' statements are covered on the internet through news articles, which are in turn disseminated and discussed by citizens. This resulted in the "Newspaper Story" (Geana et al. 2012), which capitalises on the effect of "indirect injections" (Joshi et al, 2010) – this means the politician's statement is disseminated by citizens rather than the politician. A key example of this is that a news article is written around the statement, and this is discussed over many different locations by citizens.

A key recent requirement from policy makers is the need to address local issues. Many political topics are discussed on SNS, from local to international, and from the point of the WeGov stakeholders

interviewed, their need for knowing local topics and opinions is more important than on the global level - especially the policy maker's constituency is one important geographical restriction as definition for localization. This is due to the fact that many WeGov stakeholders act for their electorate as a member of parliament. This is valid for members of State Parliaments and as well for lots of members of the German Bundestag as WeGov interviewees mentioned. (Wandhöfer et al. 2011b)

4   DISCUSSION

The current version of the WeGov toolbox focuses on monitoring of SNS debates and the presentation of analysis results that makes them easy to understand for policy makers. A major item of further work is to be able to inject into SNS (e.g. respond to a citizen's comment). The technical work of making posts into SNS is not challenging, but the challenge is in determining where to make posts – which post or who to reply to, for example. Therefore we plan to link the injection side of the toolbox to the analyses that already have been implemented, so the user has guidance on where they might like to make posts into SNS. Another item of further work is to automate collection of SNS posts. Currently a single search on SNS does not produce many results, and searches are rate-limited. We plan to incorporate a scheduler that executes searches multiple times, whilst respecting the rate limits from the social networks. The result will be that the users and analyses have larger quantities of data to work from. Throughout the project, best practice of WeGov functionality has been considered, in particular concerning protection of the privacy of citizens whose posts are collected from social networks. This was initially discussed in Addis et al, 2010, and the final development of this will be delivered as a report at the end of the project, to guide future users and deployers of WeGov after the project. In addition, the validation of components' results, and final user evaluation is scheduled due the final version. The reason for this is that policy makers have mentioned important improvements within the current evaluation phase, and these are important to enable them to use the tool in a user-friendly and successful way.

Acknowledgement

References

Addis, Taylor, Fletcher, Wilson, Fallon, Alani, Mutschke, Wandhöfer 2010. 'New ways for policy makers to interact with citizens through open social network sites - a report on initial results'. *Internet, Politics, Policy 2010*.

Angeletou, Rowe, Alani 2011. 'Modelling and Analysis of User Behaviour in Online Communities'. *Proceedings International Semantic Web Conference*, Bonn, Germany.

Geana, Taylor, Wandhöfer 2012. 'Bringing Citizens' Opinions to Members of Parliament – The Newspaper Story'. *Proceedings of the international conference on e-democracy and open government,* Austria, Krems.

Joshi, Wandhöfer, Thamm, Mathiak, Van Eeckhaute 2011. 'Rethinking Governance via Social Networking: The case of direct vs. indirect stakeholder injection'. *Proceedings of 5th International Conference on Theory and Practice of electronic Governance*, New York 10121-0701: ACM PRESS, S. 429.

Naveed, Sizov, Staab, 2011. 'ATT: Analysing Temporal Dynamics of Topics and Authors in Social Media'. *Proceedings Web Science Conference*.

Rowe, Angeletou, Alani 2011a. 'Anticipating Discussion Activity on Community Forums'. *The Third IEEE International Conference on Social Computing*. Boston, USA.

Rowe, Angeletou, Alani 2011b. 'Predicting discussions on the social semantic web'. *Extended Semantic Web Conference*. Heraklion, Crete.

Sizov 2010. 'GeoFolk: Latent Spatial Semantics in Web2.0 Social Media'. *Proceedings Web Search and Data Mining*.

Wandhöfer, Thamm, Joshi 2011b. 'Politician 2.0 on Facebook: Information Behaviour and Dissemination on Social Networking Sites – Gaps and Best-Practice. Evaluation Results of a novel e-Participation toolbox to let politicians engage with citizens online'. *JeDEM - eJournal of eDemocracy and Open Government*, Vol. 3, No 2, S. 207-215.

Wandhöfer, Thamm, Mutschke 2011a. 'Extracting a basic use case to let policy makers interact with citizens on Social Networking Sites: a report on initial results'. *Proceedings of the international conference on e-democracy and open government*, Austria, Krems, S. 355-358.