Automated Analysis of e-Participation Data by Utilizing Associative Networks, Spreading Activation and Unsupervised Learning

Peter Teufl[1], Udo Payer[1], and Peter Parycek[2]

[1] Institute for Applied Information Processing and Communications -IAIK,
Graz University of Technology
[2] Donau Universitaet Krems

Abstract. According to [1], the term e-participation is defined as "the use of information and communication technologies to broaden and deepen political participation by enabling citizens to connect with one another and with their elected representatives". This definition sounds quite simple and logical, but when considering the implementation of such a service in a real world scenario, it is obvious that it is not possible to evaluate messages, which are generated by thousands of citizens, by hand. Such documents need to be read and analyzed by experts with the required in-depth domain knowledge. In order to enable this analysis process and thereby to increase the number of possible e-particpation applications, we need to provide these experts with automated analysis tools that cluster, pre-screen and pre-evaluate public opinions and public contributions. In this paper we present a framework based on Machine Learning-(ML) and Artificial Intelligence-(AI) techniques that are capable of various analysis mechanisms such as unsupervised clustering of yet unread documents, searching for related concepts within documents and the description of relations between terms. To finish, we show how the proposed framework can be applied to real world data taken from the Austrian e-participation platform mitmachen.at.

1 Introduction

E-participation presents an important possibility for citizens to actively take part in democracy. As a fundamental principle of democracy, participation in the broader sense includes engagement in acts of representative democracy. Public participation in its different forms can be legally institutionalized in all governmental powers: for example, petitions for a referendum in legislation, lay judges, juries in jurisdiction, and in large-scale administration projects which require official approval (as can be seen with planning laws or building regulations). Beside these regulated forms, there are various types of informal participation to be found; particularly in public administration, where individuals and lobby groups are engaged in projects, regional planning and developments in the public sector.

The main problem when dealing with large-scale, nationwide projects which lead to an enormous number of texts, is that they are simply too abundant to be analyzed without the use of technology. This framework was designed in order to gain a better

A. Macintosh and E. Tambouris (Eds.): ePart 2009, LNCS 5694, pp. 139–150, 2009.

picture of texts dealing with similar topics yet using different semantic structures. One main research focus was real time data analysis for e-participation platforms.

In 2006, the Working Group for e-democracy and e-participation was founded by the Austrian Federal Chancellery with the aim of collecting and comparing e-participation projects in Austria in order to develop a general policy. An e-democracy policy and strategy were developed in 2008 displaying standardized e-participation methods. This policy can be used for a wide range of participation projects ranging from local, neighbourhood projects to nation-wide involvement. Standardization and reusability are important issues for a number of reasons: It is easier for public administration to build on the experience of successful methods. Every completed project leads to further improvement and the users become accustomed to certain formats and procedures. Reuse of existing platforms brings financial advantages and projects can be initiated quickly and with greater ease. Especially nationwide projects need automated analysis of e-participation data. More Austrian projects will be launched in 2009, and the coming years will show the extent to which citizens want to become politically involved and make use of the new technologies for this purpose. All in all, we will know soon whether technology can actually change democracy. One main research focus is the real time data analysis of e-participation information. As the first result of this research focus, we propose a framework for the automated analysis of e-participation data. The framework employs different algorithms from the areas of Machine Learning (ML) techniques and Artificial Intelligence (AI) and introduces so-called *activatio*n *pattern*s that are used to represent sentences and documents. These *activatio*n *pattern*s take the semantic relation between terms into account and thereby improve the quality of the analysis. We show the capabilities of the framework by analyzing data from the Austrian e-participation project mitmachen.at.

2 *mitmachen.at*

This section gives a short introduction to the e-participation project *mitmachen.at*. For a more detailed report on the project we refer to [2]. Enabling democratic political processes requires (1) a relationship and a dialogue between politics and citizens, and (2), citizens who are willing to participate in a democratic process [3]). The project mitmachen.at – move your future, is a youth e-participation project led by the Austrian Federal Computing Centre (BRZ, Bundesrechenzentrum) with the aim of motivating young people to participate in a political discussion about important topics for the future in Austria. In Austria, 90.8% of young people believe in the value and importance of their political participation, but only 25.4% actually know how to join in [4]. The name of the project reveals the objective: in German, the verb mitmachen means to join in, to participate. It was directed at providing young people aged 15-25 in Austria with the opportunity to participate in a 4-step process of presenting and voicing their concerns about the future using the Internet. Mitmachen.at was one of the biggest e-participation projects in Austria. The BRZ worked together with different organisations including youth institutions, software companies, various Think Tanks, and the relevant public authorities to develop a democratic participation process. Aside from revealing a number of interesting results and conclusions, the actual project itself proves that e-participation represents a cross sectional subject, which can be part of a procedure on its own, but could also be useful in many other areas of application [5]. The aim of this project was to investigate and test the general electronic participation processes, but it also examined the technical implementation and the (technical) framework, which make such participation processes possible. Portals are important for simplifying the vertical and horizontal integration of e-government [6]. The virtual portal used for this project included both the necessary instruments for participation as well as two different user levels (administrative and end-user).

3 Algorithms and Techniques

Semantic/Associative networks: The concept of associative or semantic networks was presented in 1968 by Quillain [7]. Such networks are directed or undirected graphs that store information in the network nodes and use edges to present the relation between these nodes. Good examples for semantic networks are WordNet [8] for the English language and Germanet [9] for the German language. Both networks store relations between synonym sets (synsets), which are cognitive synonyms that group nouns, verbs, adjectives and adverbs. Synsets are interlinked by means of conceptual-semantic and lexical relations. Examples for such links are hypernymy/hyponomy, antonyms and derivationally related forms. Associative networks represent a more general concept than semantic networks, since they use unlabeled links to represent relations between nodes. Typically, these links are weighted according to the strength of associations. In this paper, we use the frequency between two co-occurring nodes as weight. Spreading Activation (SA): Spreading activation (SA) algorithms are used for searching associative or semantic networks. In order to do so, one or more nodes are activated within the network. This activation is then spread over the network to neighboring nodes within one or more iterations. The optimal number of iterations depends on the data that is represented by the underlying network and the type of search operation. The activation values of neighboring nodes depend on the weights of the links to these nodes and the employed activation function. The results gained from SA algorithms are influenced by the employed strategies:

-Activation function, activation threshold: The activation value of a node is determined by the input it receives from neighboring nodes. The decision if the node fires and therefore spreads its activation value during the next iteration depends on the employed activation function and a pre-set activation threshold**.**
-Decay, Iterations: During each iteration, the activation spreads further over the network. In order to avoid the activation of a large part of the network, we need to constrain the spreading by employing a decay factor that reduces the activation energy during each iteration or limiting the number of iterations**.**
-Fanout: Nodes that are connected to a large number of other nodes (e.g. the term **be**) do not provide information and decrease the accuracy of the analysis process. Their influence needs to be constrained by utilizing a fanout factor.

Activation patterns: Given is an associative/semantic network with n nodes and an arbitrary number of edges that describe the relations between these nodes. In addition, let us assume to activate a sub-set of these nodes by using the SA algorithm to spread the activation. Then, we define an activation pattern as a vector that contains the activation values of the n nodes stored within the associative/semantic network. In the area of text classification, such an activation pattern represents a concept that is based on the relations between different terms. By defining an arbitrary distance measurement between these activation patterns, we are able to calculate the similarity between different concepts and can use this similarity to run unsupervised clustering and searching mechanisms to search for related concepts.

Unsupervised learning: For the automated analysis of e-participation data, unsupervised learning algorithms play an important role. Due to their unsupervised nature, they are able to find clusters of related documents without requiring any label/class information for these documents. In general, we can apply any unsupervised learning algorithm to the activation patterns extracted from the associative network. Examples for such algorithms are Neural Gas based algorithms [10]), Self Organizing Maps (SOM), Hierachical Agglomerative Clustering (HAC), or Expectation Maximation (EM). For the unsupervised analysis of the mitmachen.at activation patterns we employ the Robust Growing Neural Gas algorithm (RGNG) [10]. should be capitalized (i.e., nouns, verbs, and all other words except articles, prepositions, and conjunctions should be set with an initial capital) and should, with the exception of the title, be aligned to the left. Words joined by a hyphen are subject to a special rule. If the first word can stand alone, the second word should be capitalized. The font sizes are given in Table 1.

4 Framework for the Analysis of e-Participation Data

For the evaluation of the *mitmachen.a*t data and the development of the framework we use the following strategy:

Finding suitable unsupervised learning algorithms: Typically, it is difficult to evaluate the quality of unsupervised learning algorithms since we do not have any label information for the analyzed data, that could be used for the evaluation of the performance. Furthermore, due to the high dimensionality of text data, we require an algorithm to have a model that is easy to understand and visualize. The visualization of the trained data is a vital component, since it helps to analyze unknown data. Therefore, in the first step we create simple bag-of-words models[1] and evaluated the suitability of unsupervised models. Neural Gas based algorithms are employed due to their simple model and the ability to visualize such models. The results of the first unsupervised analysis of the bag-of-words data with such an algorithm is visualized by employing an Actionscript[2] animation[3]. The clusters and the distances are extracted directly from the trained Neural Gas model. Integration of semantic information: The first results showed us that the simple bag-of-words approach is adequate for rough

[1] In such a model each document is presented by an *n*-dimensional vector that stores the frequency of each term within the document. n is the number of distinct terms within the whole data-set.

[2] http://www.adobe.com/devnet/actionscript/

[3] http://apps.egiz.gv.at/Mitmachen/bin/DocumentVisualizerNew.swf

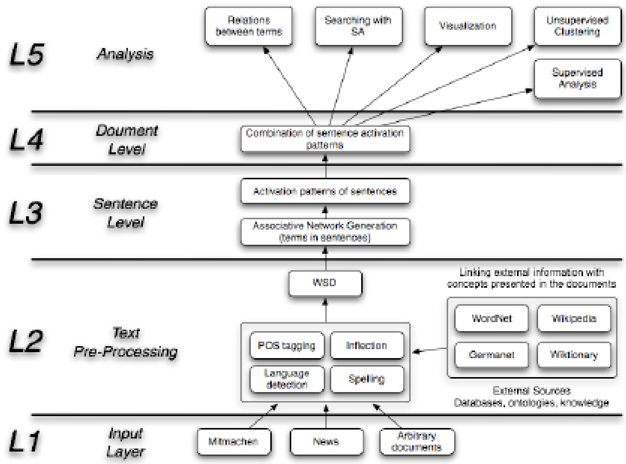


Fig. 1. Overview over the analysis framework

unsupervised clustering but cannot not be used for accurate searching within the data-set. The reason is that the bag-of-words approach does not cover any relations between the terms within the sentences or documents and thus, we are limited to simple keyword matching for finding related concepts within the data.

The results of the initial *mitmachen.at* analysis lead to the development of the presented framework that is depicted in Figure 1 and described in detail within the subsequent sections.

L1 -Input/L2 -Text pre-processing: The input layer consists of various input plug-ins that read documents from arbitrary sources. Before we can utilize sophisticated techniques for the analysis of the given text, we apply several preprocessing steps in L2:

–**POS tagging:** In order to decide which terms should be used for the subsequent analysis, we need to determine their part-of-speech (POS) tags. Such tags provide more information about the way a term is used within a sentence and how it is related to other terms. Currently, we only use a limited number of tags for the subsequent analysis: nouns, verbs, adjectives and adverbs. All other terms are dropped.

–**Word Sense Disambiguation (WSD):** WSD is the process of finding the sense of a term within a given sentence. Typically, WSD algorithms take the other terms within the sentence into consideration in order to determine the sense of the term.

–**Lexical analysis:** Currently, we only use a limited number of POS tags for the subsequent analysis. However, for future versions we plan to integrate the structural information about sentences gained by lexical parsers. A good example for such a parser is the Stanford parser [11].

L3 -Sentence Processing: This layer analyzes each sentence, generates an associative network that stores the relation between terms within sentences and generates the *activation patterns* for all sentences.

Generation of the associative network/semantic network: For each different term (sense) within the analyzed text corpus we create a node within the associative network. The edges between nodes and their weights are determined in the following way: All senses within a sentence are linked within the associative network. Newly generated edges get an initial weight of 1. Every time senses co-occur together, we increase the weight of their edges by 1. In addition, we store the type of connection for each edge. Examples for these types are noun-to-noun links, noun-to-verb links or adjective-to-adverb links. By using this information when applying SA algorithms, we are able to constrain the spreading of activation values to certain types of terms.

Network processing: The output of the last step is an associative network with nodes representing the disambiguated terms and edges representing the relation between terms. The weight of the edges is determined by the number of times terms co-occur within sentences. In order to apply SA algorithms to this network, these weights need to be normalized, so that the maximum weight is equal to 1.0.

Determination of *activatio*n *pattern*s for sentences: We can now utilize SA algorithms and the information stored in the associative network to extract the *activatio*n *pattern*s from the sentences. For each sentence we have the POS tagged senses of nouns, verbs, adjectives and verbs[4]. The terms (senses) correspond to nodes within the network and get an initial activation value of 1.0. We apply the SA algorithm to the associative network for at maximum two iterations and extract the *activatio*n *pattern*s from the associative network.

L4 -Document processing: We could now apply unsupervised clustering algorithms or search algorithms based on SA to the extracted *activatio*n *patterns*. However, these patterns only represent information on the sentence level and we still need a method that allows us to represent whole documents. In order to do this, we simply sum up the *activatio*n *pattern*s of each sentence within a document and use the resulting pattern as a representation for the given document. While we are able to achieve good results with this technique, this may be due the relatively short length of the analyzed *mitmachen.a*t documents. We need to revise this process for larger documents that contain various different concepts.

L5 -Analysis: The *activatio*n *pattern*s generated in the previous layers are the basis for applying supervised and unsupervised Machine Learning algorithms. Furthermore, we can implement search algorithms that are based on SA algorithms.

Unsupervised analysis: Unsupervised analysis plays an important role for the analysis of text, since it allows us to cluster unread documents according to their content.

Search with Spreading Activation (SA): In order to search for related concepts within the analyzed text sources, we apply the following procedures:

1. The user enters the search query, which could be a combination of terms, a complete sentence or even a document containing multiple sentences.
2. We determine the POS tags for every term within the search query.
3. Optionally, we now make use of an external knowledge source (e.g. WordNet or Germanet) to find related terms and concepts for the terms in the query. Since, we currently do not employ WSD techniques we need to be careful when selecting the appropriate synsets from the external sources.

[4] Depending on the analysis we could filter out various types and constrain spreading to certain links.

4. We activate the nodes corresponding to the terms of the search query and the related terms extracted from the external knowledge source.
5. We use the SA algorithm to spread the activation over the associative network.
6. We extract the *activation pattern* of the associative network and compare it to the document or sentence patterns that were extracted during the training process. The patterns are sorted according to their similarity with the search pattern.

Relations between terms: The associative network trained in L3 contains information about relations between terms that co-occur within sentences. By activating one or more nodes within this network and applying the SA algorithm, we can find related terms.

5 Evaluation of the mitmachen.at Dataset

For the evaluation of the *mitmachen.at* data, we have applied the presented framework[5]:

–**L1 -Input:** Each *mitmachen.at* entry is stored in an XML file, that is read by an input plugin. Several transformations are applied to the raw text which include the replacement of characters specific to the German language and the removal of punction marks, other than period.

–**L2 -Text preprocessing:** For each sentence we apply:

- **POS tagging:** For POS tagging we use the LingPipe[6] API trained with text from Tiger Corpus [12]. Only nouns, verbs and adjectives are kept.
- **Lemmatization:** For each term we determine the lemma by using information extracted from the morphological analyzer MORPHY [13].
- **WSD:** Currently, we do not apply a WSD algorithm. However, we plan to integrate the WSD technique presented in [14]. It uses Wordnet to disambiguate English terms, but can easily be adapted to the German language by replacing Wordnet with Germanet.

–**L3 -Sentence Level:** We generate the associative network and the *activation patterns* for the sentences according to the procedures described before.

–**L4 -Document Level:** For each *mitmachen.at* entry (considered as document here), we generate an *activation pattern* by summing up the *activation patterns* of its sentences.

–**L5 -Analysis:** The *activation patterns* for each *mitmachen.at* entry and the associative network generated in L3 can now be analyzed by applying different techniques described in the subsequent sections.

5.1 Unsupervised Analysis-Clustering

In this section, we analyze the *activation patterns* of the *mitmachen.at* entries (generated in L4) by applying the RGNG algorithm. Users of the *mitmachen.at* platform

---

[5] The complete results can be downloaded from http://apps.egiz.gv.at/eparticipation.tar.gz

[6] LingPipe -http://alias-i.com/lingpipe/index.html

were able to create or join discussions on 8 topics: **environment, health, education, security, infrastructure, social system, political system, employment**. The results of the unsupervised clustering are shown in Table 1. As excepted the found clusters broadly cover these different topics. Since the trained model contains more than 400 documents, we only provide a short description for the clusters. The column *German terms* contains keywords that were automatically extracted from the clusters. The English translation of these keywords is available in column *English terms.* The column *topic*s contains the *mitmachen.at* topics that describe the documents contained in the cluster. These topics were manually assigned in order to see the type of clusters the algorithm is able to find. The last column – *summar*y – contains a short summary of the documents contained in the clusters. This summary was manually added in order to give a short description of the cluster. For the complete document cluster data we refer to *cluster-results.txt*. By taking a closer look at the found clusters we make the following observations:

–We are able to find clusters that are correlated with the topics available in *mitmachen.at*. Several topics, such as health, environment partly have similar contents and can be found together in single clusters.

–**Cluster 1:** This cluster contains discussions related to security, the social system and employment. The reason for the combination of these topics is the term security since it plays a role for security in terms of social security and in terms of security provided by the police and the military. This is a good example why WSD algorithms make sense, since the term security has two different senses here. Furthermore, the problem might have been avoided by employing a fanout factor (see the discussion on cluster 10 for more details).

–**Cluster 6:** This cluster contains documents that are related to the discussion of the *mitmachen.a*t project itself.

–**Cluster 10:** This cluster contains documents from various topics. By inspecting the documents we can see that the term **Meinung** (opinion) is quite frequent within these documents. Since one can have an opinion on every topic, it is connected to a large number of nodes within the associative network which is depicted in Figure 2(b). Thus, the documents containing this term cause the activation of a large number of other nodes of arbitrary topics and therefore end up in the same cluster. This is a perfect example for the requirement of a fanout factor. Such a factor would simply penalize this node and suppress the activation energy spread by the node.

5.2 Search for Related Patterns

The *activation patterns* of the *mitmachen.at* data are represented by 5755-dimensional vectors. Given the *activation pattern* of an *mitmachen.at* entry, we can find related patterns by calculating the distance to the other *activation patterns* and sort the results according to similarity. For determining the distance we employ the cosine similarity.

Relations between terms: In order to find the relations of one or more input terms, we activate the corresponding nodes in the associative network and spread the activation by applying the SA algorithm. The strength of the activation of the other terms indicate the strength of the association with the input term. As example we use the term **Fahrzeug (vehicle)** and show the related terms in Figure 2(a). Some examples for related terms are: **Verschmutzung (pollution), Klimawandel (climate change) , Automobil (car), Fussgaenger (pedestrian)** and **Fussgaengeruebergang (pedestrian crossing)**.

Searching for related entries: In this case we select an existing document and calculate the distance to the other documents within the data-set. The other documents are then listed according to similarity. As example we use an input document that deals with the necessity to reduce the ticket price for public transport and the need to introduce a time schedule with shorter intervals. The complete list of related documents is available in *search-related.txt* and shows that the framework is able to retrieve similar documents that deal with issues related to traffic and public transport.

Finding relations/Searching without external knowledge: For this example we use the search query vehicle but do not employ Germanet to find related terms.

Table 1. Summarizing the 15 clusters found by the RGNG algorithm

	Docs	German terms	English terms	Topics	Summary
1	12	Sicherheit Bundesheer Oesterreich	security army Austria	security social system	social security issues security related to public places crime, army
2	22	Sicherheit Polizei Land Bundesheer	security police country army	security	army related issues discussion about police crime
3	30	Umwelt Wasser	environment water	environment infrastructure	pollution of the environment, alternative energy, fossil fuels, global warming, transport
4	3	Abgas Luft	exhaust air	environment health	health related issuest protection of the environment
5	17	Automobil Transportmittel Strasse Bus Eisenbahn	exhaust vehicle road bus train	environment infrastructure	$CO2$ emissions solar energy fossil fuel quality of public trans-port traffic (trucks, cars)
6	4	Projekt Textanalyse	project text analysis	-	*mitmachen.at discussion*
7	26	Politik Mensch Politiker Umwelt	politics human politician	political system social system	discussion about politics, politi-cians

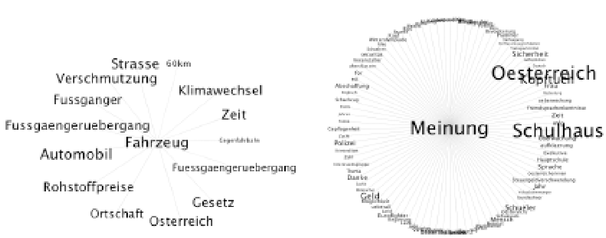
Table 1. (*continued*)

			environment		(un)employment political deciscions regarding health system, environment and
8	26	Bildung Lehrer Ausbildung	education teachers specific education	education social system	university, schools, teachers problems at schools costs of the education system
9	34	Schule	school	education	discussion about schools and teachers
10	23	Meinung Schueler Lehrer	opinion pupils teacher		various topics
11	10	Arzt Medizin Patient Geld	doctor medicine patient money	health	discussions about health and the health system
12	6	Gesundheit Medizin Gesundheitssystem Geld	health medicine health system money	health	discussions about health and the health system
13	13	Kosten Medizin Problem	costs medicine problem	health	discussions about health and the health system
14	12	Kosten Eltern Mensch	health parents human	health social system	discussions about health and the health system
15	9	Kinder Eltern Geld	children parents money	social system education	discussions about health the social system, children

We activate the node representing **vehicle** within the associative network and spread the activation. This results in the activation of other terms such as **pedestrian, car, road, price of resources**. The activation value depends on the strength of the relation between the terms. By taking a closer look at the activated terms we can draw several conclusions: Due to the co-occurence of **car** and **vehicle** within one or more sentences the term **car** is activated even if Germanet is not employed. Other terms such as **pedestrian or price of resources** are also found, since they co-occur with **vehicle** within sentences. Such relations could not be found with Germanet since they represent relations that are specific for the analyzed data.

The *activatio*n *patter*n generated out of the search query includes semantic information that is stored in the associative network. Now, we compare this pattern to the patterns of the other *mitmachen.a*t documents and sort the results according to similarity. The analyis of the results (see *fahrzeug-without-germanet.txt*) shows that we are also able to find documents, that do not contain **vehicle** but related terms.

Finding relations/Searching by including Germanet: In this case we specify a search query consisting of one or more terms and apply an SA algorithm to Germanet in order to find related terms. Germanet employs various semantic pointers[7] that define the relations between different synsets. For our search query expansion we only use the hyponym relation which links more general concepts to instances of these concepts. Again, we use the term **vehicle** as search query. By activating the node for **vehicle** within Germanet and applying the SA algorithm, we are able to find additional terms related to **vehicle: airplane, helicopter** and **car**. All of these terms are within the *mitmachen.a*t data. Other instances of **vehicle** are not taken since they are not within the analyzed documents. Now, we activate the nodes corresponding to the three terms within the associative network and spread the activation. Due to the activation of the terms found by Germanet additional terms are activated: **air, air pollution, greenhouse effect.** Due to **helicopter** we also get activations for **Galtuer** and **Eurofighter**. **Galtuer** is a village in Austria that was hit by avanlanches in 1999. After this event there were discussions about military helicopters that could be used for rescue and relieve efforts. In the *mitmachen.a*t data, there is a reference to this discussion, therefore the name of the village is linked to **helicopter**. Such relations cannot be found with Germanet. The complete search results are available in *fahrzeug-with-germanet.txt*.



(a) Fahrzeug (vehicle) (b) Meinung (opinion)

Fig. 2. Two examples for terms and their related terms

6 Conclusions and Outlook

We have proposed an analysis framework for e-participation data, that combines various algorithms from the areas of Machine Learning and Artificial Intelligence and we have introduced *activatio*n *pattern*s representing documents and sentences. We the applied the

[7] See http://www.sfs.uni-tuebingen.de/GermaNet/Pointers.html for a complete list and detailed description.

framework to real data from the Austrian e-participation project *mitmachen.a*t and provided examples for unsupervised clustering, searching for related concepts and the analysis of relations between terms. The analysis of the results shows that the inclusion of semantic information and external knowledge sources is of great importance for the quality of the results. In future, we want to include further techniques such as fanout factors for the associative networks, WSD algorithms and improved versions of the RGNG algorithm in the framework. Furthermore, we plan to apply the framework to various e-participation related data-sets and integrate it into a website that offers a support forum for the Austrian Citizen Card. By applying the framework to different data-sets we will be able to gain more knowledge that allows us to improve the current techniques.

References

1. Macintosh, A.: Eparticipation in policy-making: the research and the challenges. In: Exploiting the Knowledge Economy: Issues, Applications and Case Studies, pp. 364–369. IOS press, Amsterdam (2006)
2. Edelmann, N., Krimmer, R., Parycek, P.: Engaging youth through deliberative e-participation: a case study. International Journal of Electronic Governance 1/4, 385–399 (2008)
3. Lerberghe, V.: Are we moving towards a more participatory representative democracy? In: Proceedings of the Eastern European e-Gov Days 2007: Best Practice and Innovation (2007)
4. Serloth, A., Maerki, O.: Partizipation und information -ergebnisse einer jugendbefragung. Technical report, Austrian Ministry of Social Security Generations and Consumer Protection (2004)
5. Piswanger, C.M.: The participatory e-government strategy of the austrian federal computing centre. In: Proceedings of the Eastern European e-Gov Days 2007: Best Practice and Innovation (2007)
6. Moon, M.J.: The evolution of e-government among municipalities: Rhetoric or reality? Public Administration Review 62, 424–433 (2002)
7. Quillian, M.R.: Semantic memory (1968)
8. Fellbaum, C.: Wordnet: An electronic lexical database (language, speech, and communication). Hardcover (May 1998)
9. Hamp, B., Feldweg, H.: Germanet -a lexical-semantic net for german. In: Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications, pp. 9–15 (1997)
10. Qin, A.K., Suganthan, P.N.: Robust growing neural gas algorithm with application in cluster analysis. Neural Netw 17(8-9), 1135–1148 (2004)
11. Klein, D., Manning, C.D.: Fast exact inference with a factored model for natural language parsing. In: Advances in Neural Information Processing Systems, NIPS, vol. 15, pp. 3–10. MIT Press, Cambridge (2002)
12. Brants, S., Dipper, S., Hansen, S., Lezius, W., Smith, G.: The TIGER treebank. In: Proceedings of the Workshop on Treebanks and Linguistic Theories, Sozopol (2002)
13. Lezius, W., Rapp, R., Wettler, M.: A freely available morphological analyzer, disambiguator and context sensitive lemmatizer for german. In: Proceedings of the 17th international conference on Computational linguistics, Morristown, NJ, USA, pp. 743–748. Association for Computational Linguistics (1998)
14. Tsatsaronis, G., Vazirgiannis, M., Androutsopoulos, I.: Word sense disambiguation with spreading activation networks generated from thesauri (January 2007)