

The Governance Tracking System: Proposed Online Forum for Interaction between Citizens and Representatives

Kshitij Desai¹ Aniket Salunkhe² Torana Kamble³

^{1,2}Student ³Assistant Professor

^{1,2,3}Department of Computer Engineering

^{1,2,3}Bharati Vidyapeeth College of Engineering, Navi Mumbai

Abstract— Online discussion forums have become a popular web technology to build and support online communities for several engineering interest areas and practice. However, this technology is not been used for interaction between citizen and their representatives. In India we still rely on RTI Application to review what work has been done by MLA or MP and what kind of questions they have raised in assembly. Thus our portal will bridge this gap and serve as an online interaction forum for Citizens and Representatives. Instead of analyzing the data through manual means, we will utilize the Natural Language Toolkit (NLTK) to capture textual patterns and pull a Topic Modeling approach using Latent Dirichlet Allocation (LDA), to identify similar group of words.

Key words: Governance Tracking, LDA, NLTK, Forum, Interaction, Citizens, Representatives

I. INTRODUCTION

A. Need of Citizen and Representatives Interaction:

Periodic interaction between representatives and citizens is need of the democracy. The interest in direct democracy has increased in recent years for several reasons. For example see direct democracy as a means of overcoming the shortfalls of representative democracy such as declining turnout and increasing dissatisfaction with parties, politicians and governments. Recent referendums on European integration in a number of countries have also stimulated the debate about the pros and cons of direct democracy.

B. Concept of Online Forum:

Online forums are the first generation social networks. Online forums exist even before the term social networking was invented. Social media are online platforms that facilitate global collaboration and sharing amongst users [1]. An online forum is a general term for any online "discussion group" where users can post queries and expect to see responses to queries other users have posted. Or users can just read the messages. The first discussion groups were available on bulletin board systems. A lot of websites offer discussion boards so that users can share and discuss information, knowledge and opinions. Particular software is available that provides discussion board capability for a website.

C. Online Forum for Citizen and Representatives (MLAs) Interaction:

Currently there is no existence of any online forum which will bring citizens and representatives together. Our system will combine the concept of online forums and need of citizens and representatives interaction. Thus our objective of this project is to utilize the technology for making democracy healthy.

II. REVIEW OF EXISTING SYSTEM

Current system for Citizens and Representatives interaction rely on personal meetings. If any citizen has to review what works has been done by their representative, they may get the information through RTI (Right to Information.)

A. Personal Meetings:

1) Risk of Unavailability:

For personal meetings, citizens need to visit the residence or the public relation office of the representative. In this method citizens may have to get the appointment first. There is risk whether representative will be available for discussion or not.

2) Face-To-Face Communication:

This is also an important factor for avoiding personal meetings. Many citizens have no courage to face their representatives due to misconceptions. Instead of raising problems citizens tend to avoid it.

B. RTI Applications:

1) Where to Apply?

Citizens are not sure where to apply for RTI to review the work done by representatives. For filing the RTI one needs to have the knowledge of the category of the problem. Problems may fall into one of the following three lists: Union List, State List and Concurrent List.

2) Delay in Process:

Even if there are rules and regulations for in-time response of the RTI, there is delay in procedure due to above mentioned problem. Also RTI document is very complicated to understand by common citizens.

3) Rigorous Online R.T.I. Procedure:

Online 'Right To Information' procedure is also a rigorous one. Citizens first have to register on the website. After submitting the query, users have to choose payment methods [2]. All of the Indian citizens (mostly in rural areas) are still not comfortable with online payment methods or net banking procedures. Thus online RTI service does not turn up to be adapted by everyone.

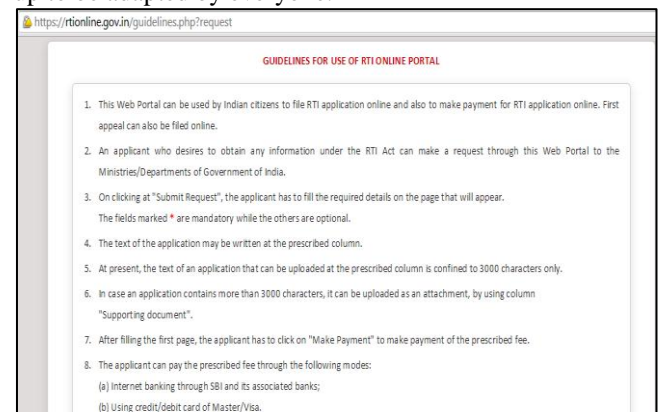


Fig. 2: Screenshot of the Guidelines displayed on the Government's Online R.T.I. filing website.

III. PROPOSED SYSTEM

The portal will be made available for citizens. Any query, issue or grievance that they want to raise shall be first entered in the given textbox by citizens. They need to enter some identification to prove the genuineness of query. The unique Issue Trailing Number will be generated for each query posted.

Citizens can view the status of their query by using the unique Issue Trailing Number given to them.

A. Step I:

Citizens will enter the problems and grievances. Input boxes will be available on the portal. Citizen need to enter their Aadhar Unique ID (which will be kept secret). Once it is posted each and every grievance will be issued a unique Issue Trailing Number.

B. Step II:

Techniques will be applied on the input data to sort out it according to respective departments. Following 2 methods will be applied on the input data:

- (1) Keyword Analysis
- (2) Topic Modelling

C. Step III:

Secretary of department will acknowledge the input problem and categorize the similar issues

D. Step IV:

After the assembly session, secretary will post a reply to the issue and attachment (assembly session video) if required.

IV. SYSTEM DESIGN

A. Citizen Input:

The issues or grievances raised by citizens which they will post on the forum. They just have to post the query irrespective of having knowledge of concerning department.

B. Unorganized Query:

The queries posted by citizens will not be in organized manner at beginning. Citizen can post grievances without categorizing it by themselves.

C. Sorting:

Techniques and methodologies [Further explained in Section V] will be applied on the unorganized queries to sort it.

D. Department Wise Query:

After sorting, every query will be directed to its corresponding department. Some queries may contain ambiguity because of which it can be directed to more than one department. Secretary of the department can then delete the irrelevant queries.

E. Forum:

The secretary of each department will acknowledge and reply to the queries which are directed at the concerning department. The queries not belonging to the department can be deleted by secretary.

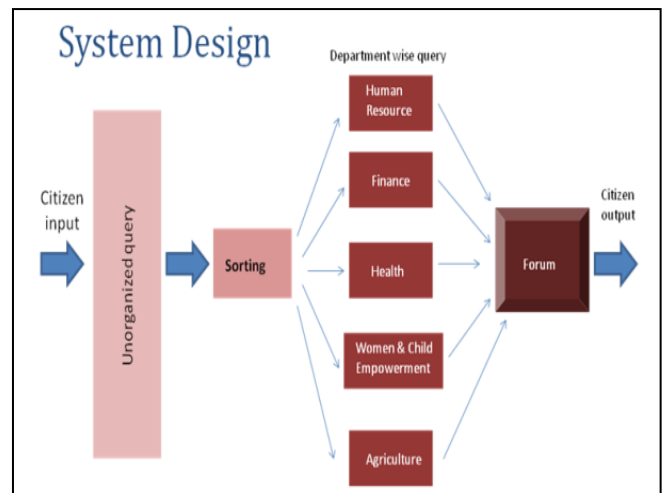


Fig. 2: System Design

F. Citizen Output:

The status of its query will be available to citizen. They can view the reply through searching the unique Issue Trailing Number.

V. TECHNIQUES AND METHODOLOGIES

A. Keyword Analysis:

Keyword analysis is an organized technique that looks for repeat of words contained by a text compilation and computes the frequencies of occurrence for each word. This analysis technique allows one to observe the words that constantly come into view in a text collection. As many English words with the same origin will have similar meanings, there is a need to diminish words in the data to their origin or curtail form (for instance, both words "students" and "studying" will be reduced to study). In this research, the well-established Snowball word stemmer was used to perform the word stemming process [3].

Amongst a range of computational approaches used in keyword analysis, we select the Natural Language Toolkit (NLTK) which is an established open source Python toolkit with very proficient libraries for text processing [4]. NLTK has been well-utilized in educational research to further the understanding of learning and interaction amongst students. Haythornthwaite and Gruzd [5] utilized NLTK to explore noun phrases in an online bulletin board for a graduate class and based on the prevalence of words such as "thanks" and "agree", they suggest that the bulletin board was a supportive avenue for learning. On the other hand, Worsley and Blikstein [6] used NLTK to explore students' speech about electronics and mechanical devices and found that speech markers (such as adverbial modifiers) that are indicative of domain-specific expertise.

B. Topic Modeling:

Topic modeling refers to a statistical approach aimed at identifying semantic topics within text documents. [7]. Amongst quite a lot of topic modeling techniques, Latent Dirichlet Allocation (LDA) is a widely used model as it is an unsupervised model which does not require the production of additional instruction data sets. LDA can be described as a generative probabilistic model aimed at discovering hidden topics within textual content together with the words that are associated with those latent topics.

An illustrative example of the topic modelling is given as follows:

1) *Input Sentences:*

- (1) I like to eat mango and apples.
- (2) I ate an apple and spinach smoothie for breakfast.
- (3) Chinchillas and cats are cute.
- (4) My sister adopted a cat yesterday.
- (5) Look at this cute monkey munching on a piece of mango.

Now after applying the topic modelling technique which is LDA (Latent Dirichlet Allocation) we can categorize the input sentences into following Topics:

Sentences	Topic Distribution
Sentences 1 and 2	100% Topic A
Sentences 3 and 4	100% Topic B
Sentence 5	60% Topic A, 40% Topic B

Table 1: Output sentences of topic modelling example

Where, Topic A: 30% mango, 15% apples, 10% breakfast, 10% munching ... (at which point, we could interpret topic A to be about food)

And, Topic B: 20% chinchillas, 20% cats, 20% cute, 15% monkey ... (at which point, we could interpret topic B to be about animals)

VI. CONCLUSION

Our research raises an important point – Online interaction forums are supportive avenues for citizens who are seeking to bridge the gap between them and representatives.

Worldwide presence and open access of online forum will help reaching towards the healthy democracy.

REFERENCES

- [1] Cheung, Christy M.K., Lee, Matthew K.O. and Wagner Christian. "Introduction to Social Media and E-Business Transformation Minitrack", 2014 47th Hawaii International Conference on System Sciences (HICSS), p 550, January 2014
- [2] "Guidelines For Use Of RTI Online Portal" <https://rtionline.gov.in/guidelines.php?request>, December 2014
- [3] M. Porter. "Snowball: A language for stemming algorithms." <http://snowball.tartarus.org/texts/introduction.html>. 2001.
- [4] S. Bird, E. Loper and E. Klein. Natural Language Processing with Python. Sebastopol, CA: O'Reilly Media Inc. 2009.
- [5] C. Haythornthwaite and A. Gruzd. "A Noun Phrase Analysis Tool for Mining Online Community." Proceedings of the Third Communities and Technologies Conference, pp. 67-86, 2007.
- [6] M. Worsley, and P. Blikstein. "Towards the Development of Learning Analytics: Student Speech as an Automatic and Natural Form of Assessment." In Annual Meeting of the American Education Research Association, Denver. 2010.
- [7] D. Blei, A. Ng, and M. Jordan. "Latent Dirichlet allocation." Journal of Machine Learning Research, vol. 3, pp. 993–1022, January 2003.