

A Dataset of Generalizable Election-Related Questions for AI Tools Compiled from Leading Global Democracies

Kausik Lakkaraju, Sara Elizabeth Jones, Bharath Muppasani, Biplav Srivastava

AI Institute, University of South Carolina, Columbia, USA



UNIVERSITY OF
South Carolina

The AAAI 2024 Third
Workshop on AI for
Credible Elections: A
Call To Action with
Trusted AI

CONTENTS

- **Chapter-1:** Introduction
 - Problem
 - Contributions
 - Related Work
- **Chapter-2:** Data Collection and Applications
 - Data Collection
 - Applications
- **Chapter-3:** Discussion and Conclusion
 - Gaps and Potential Solutions
 - Conclusion

Chapter – 1

INTRODUCTION

Problem

- Elections are crucial for a society and they serve as a collective voice of citizens in shaping the government thereby, shaping their future.
- Hence, it is important for the election administration and the candidates contesting the elections to know the concerns of the voters to improve the current polling methods to engage with the voters effectively.
- To organize the elections successfully, it is important to know the various concerns different stakeholders might have.

Contributions

Our contributions are that:

- We provide a common set of parameterized election questions.
- Furthermore, we hypothesize that if they are answered they will help drive the voter participation in any given region.
- We categorize these questions into various categories (based on the stakeholders involved) and various subcategories with identified parameters for each so that they can be generalized to other regions.
- We identify further gaps and pave the way for the future work in this area.

Related Work

- In political sciences, lot of work was done on election poll performance and internet or e-voting [1].
- Recent literature assessed the impact of social media posts on voting results [2].
- Despite the vast research in these areas, there is still a gap in realizing the concerns of various stakeholders such as voters, candidates, journalists, and election administration involved in electoral process.

References:

1. Carney, K. 2022. The effect of social media on voters: experimental evidence from an Indian election. Job Market Paper, 1–44.
2. Chen, S.; and Kortner, J. 2022. Did “Trumpists” perform worse in US Senate election polls? Significance, 19(3): 28–31.

Chapter – 2

Data Collection and Applications

Data Collection

- We analyzed the 'most googled' election questions per country, gathering 181 specific and 79 adaptable question – 59 of which are customizable for particular regions. Each question includes adaptable parameters.
- **Example:**
 - **Original query:** *“What happens if I cannot attend the polling station?”*
 - **Modified query:** *“What happens if I cannot come to the polling station (in <region>)?”*.
 - *<region>* serves as an optional parameter, allowing for customization based on the voter’s region.

Number of Queries and Sources

Country	No. of Queries	Sources
India (IND)	11	https://www.vice.com/en/article/gy4x3m/10-most-googled-questions-about-indian-election
United States of America (USA)	39	https://www.usa.gov/how-to-vote ; https://www.cnn.com/interactive/2020/politics/voting-questions-answers/
Indonesia (IDN)	6	https://www.vice.com/en/article/neawv8/most-googled-questions-about-indonesian-election-2019
Ireland (IRE)	11	https://www.citizensinformation.ie/en/government-in-ireland/elections-and-referenda/types-of-elections-and-referendums/
United Kingdom (UK)	7	https://fullfact.org/election-2019/election-questions-answered/
Estonia (EST)	50	https://www.valimised.ee/en/internet-voting/frequently-asked-questions/
South Africa (SA)	47	https://www.elections.org.za/content/for-voters/FAQ-For-Voters/
Canada (CA)	11	https://www.tvo.org/article/your-federal-election-questions-answered

Table 1: The number of queries collected from each country with the corresponding country name and source.

Data Sample

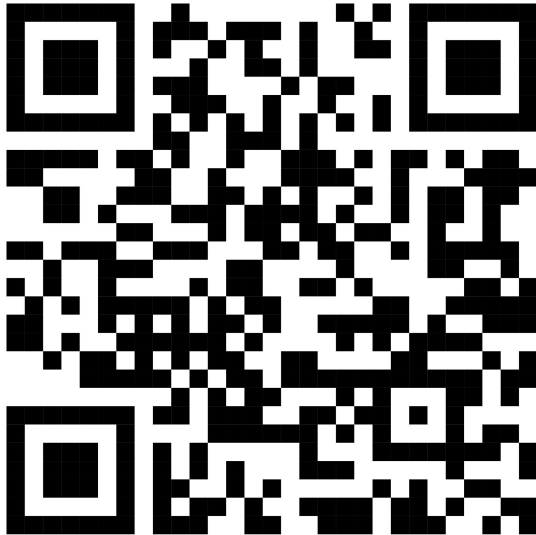
- Query categories based on the stakeholders involved: voters, candidates, and Election administration / Journalists.
 - Sub-categories: 'What', 'How', 'When', and 'Who' questions.

Stakeholders	Sub-category	Question - general	Parameter(s)	Region(s) Applicable	Reference to country-specific tab
Voter	What	What counts as a "fair" election?	None	Any	IDN-Q6
	How	How would the voters of advanced age vote as actively as the young?	None	Any	EST-Q36
	When	When can the <region> begin to count absentee & mail-in ballots?	<region>	Any	USA-Q35
	Who	Who can run for <Head-of-state> of <region>?	<Head-of-state>, <region>	<Head-of-state>, <region> = King Queen, Canada, President, USA, India, South Africa, Indonesia	IRE-Q2
Candidate	What	Do <poll officers> have any guidelines about conditions when a candidate can request a recount in a close result?	<poll officers>	Any	UK-Q3
	How	How is it ensured that each voter votes themselves (independently) (in <region>)?	<region>	Any	EST-Q20, EST-Q43
Election Administration / Journalist questions	What	What goes into the cost of an election in <region>?	<region>	Any	CA-Q7
	How	How does <region> handle voter abstention?	<region>	Any	IDN-Q3
	Who	Who are Millennial voters?	None	Any	IDN-Q1

Table 2: Sample generalized queries extracted from different sources with corresponding stakeholders, sub-categories, parameters, regions applicable, and reference to the country-specific tab. The country codes for each country are shown in Table 1.

Link to Our Repository

Scan this QR code to go to our repository with the complete list of queries and the code to analyze the datasets.



Applications

Our dataset can be used for the following applications:

1. **Chatbot development:** To train chatbots that can answer diverse election-related queries.
2. **Policy formulation:** Can be used by policy-makers can better align their agendas with public interests.
3. **Understanding intercultural contrasts:** Such a diverse dataset helps us understand the differences between queries that stakeholders from different countries may have.
4. **Improving polling procedure:** Understanding stakeholders' concerns can help the election administration to improve the polling procedures.
5. **Candidate-voter engagement:** Can be used by candidates to address the issues the voters are facing.

Chapter – 3

Discussion and Conclusion

Gaps and potential solutions (1/2)

The following are some critical gaps in this area that need to be addressed:

- **Gap-1:** A discrepancy still exists between the election FAQs provided on official websites and the queries resulted from different independent surveys that were conducted.
 - **Solution:** Gap can be filled if the government takes the initiative to conduct such surveys to align the official resources with real-time concerns.
- **Gap-2:** Existing studies explore voter concerns, neglecting the concerns of election administration or potential candidates who would like to contest elections.
 - **Solution:** Can be addressed by conducting inclusive surveys that take into considerations various stakeholders' voices.

Gaps and potential solutions (2/2)

- **Gap-3:** Poor and non-inclusive data
 - **Solution:** In [1], the authors trained a chatbot on the FAQs data provided on official websites of South Carolina and Mississippi. Such efforts when combined with more inclusive and vast data help researchers build better chatbots or AI tools to promote voter participation.

References:

1. Muppasani, B.; Pallagani, V.; Lakkaraju, K.; Lei, S.; Srivastava, B.; Robertson, B.; Hickerson, A.; and Narayanan, V. 2023. On safe and usable chatbots for promoting voter participation. *AI Magazine*, 44(3): 240–247.

Conclusion

- Elections are the cornerstone for citizens in a country to express their opinions.
- We compiled relevant election queries, collecting 181 country-specific and 79 generalizable questions, 59 of which can be adapted for specific demographic or region.
- We identified a notable gap in FAQs, particularly in addressing common queries different stakeholders may have.
- We envision applications such as creating election chatbots, understanding stakeholder concerns, and improving polling procedures to make voting more accessible.

THANK YOU ALL

Contact Information

Biplav Srivastava – biplav.s@sc.edu

Kausik Lakkaraju – kausik@email.sc.edu

Learn more about our Team!

