

Artificial Intelligence and Elections – how and where may AI be *beneficial*?

# Promoting AI's **Safe** Usage for **Elections**

---

Prof. Biplav Srivastava, [biplav.s@sc.edu](mailto:biplav.s@sc.edu)

Invited Talk: *Elections Security* session



UNIVERSITY OF  
**South Carolina**

June 22, 2026

# Promoting AI's Safe Usage for Elections

- Bonaire 5/6 / **(Elections Security)**; June 22, 2026, 2:40-3:40 pm
- **Objective:** *This session explores the evolving role of artificial intelligence in electoral processes, focusing on its potential to improve data-driven decision-making amid the growing challenges of misinformation, manipulation, and voter suppression. It examines how AI tools could address information gaps for voters, candidates, and election commissions while acknowledging the skepticism and concerns that surround the use of AI in critical civic functions.*
- **Biplav Srivastava**, Professor, AI Institute and Department of Computer Science and Engineering, University of South Carolina

# Outline

- Speaker's context
- Election processes –State-of-the-art (SOTA) and US
  - Hint: US does not have SOTA
- Problem of choice: **information gap** for a stakeholder
- Towards solution: **Using (Trustworthy) Artificial Intelligence** for overcoming information gap problem
  - Collate-Distribute-Comprehension (CDC) approach
- Call for action
  - Technical: improve data, invest in evaluation
  - People: Election officials, IT companies, public - voters and candidates
  - Process: aspire and benchmark to be the very best!



# **Speaker's Context**

# RESEARCH @AI4S

**Goal:** *Focus on enabling people to make rational decisions despite real-world complexities of poor data, changing goals, and limited resources by augmenting human cognitive limitations with technology*

**How:** *By innovating and applying AI techniques for hard problems facing the society with an inclusive, value-driven focus.*

**Values:** *Grow human opportunities, reduce resource wastage*



*Jan 2025, Columbia*



*Safe AI for Seniors, 2025*

<https://sites.google.com/cse.sc.edu/safe-ai-for-seniors/>

# Areas We Work In

## Neuro-symbolic Methods (Theory)

- Planning / Sequential Decision Making (SPDs) [AAAI 2024, Neurips 2024, ICAPS 2024, 2025, AAAI 2026],
- Group recommendation [AAAI/IAAI 2024, AI Mag 2024]
- Training models, time series and data analysis [ICAPS 2025 PRL, AAAI 2025, AAAI 2025 Fall Symp]
- Ontologies, meta-cognition reasoning [Data Disc. J. 2025, CACM 2025, Nature npj 2025]

## Trusted AI (Usability)

- Rating/ certification, causal models, explanation [Web2026, **Springer 978-3-032-21038-8 (2026)**, AI Ethics 2024]
- LLMs/ FMs eval. [IAAI/AAAI 2026, ICAIF 2023, Frontiers AI 2023]
- AI risk and security [NIST AI Consortium (AIC) / Safety Institute Consortium (AISIC)]

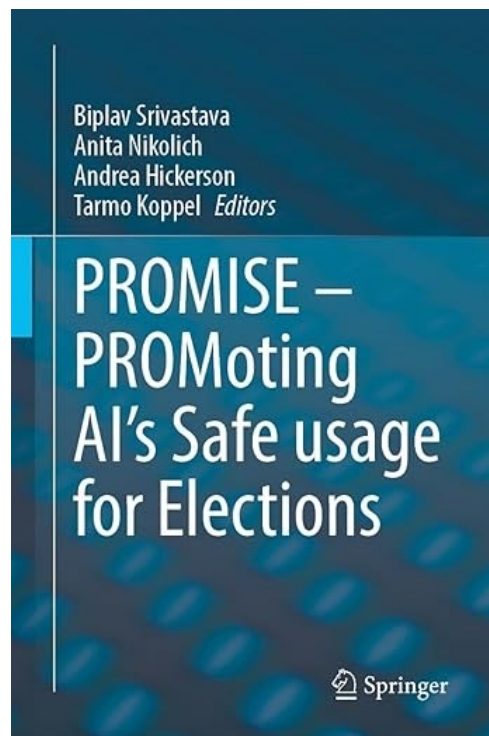
## Societal Impact (Applications)

- Elections (chatbot/official info, info. spread) [AAAI 2025, **ISBN: 978-3-031-89852-5 (2025)**]
- Health (chatbot, multimodal analysis, meal / nutrition adherence, exercise monitoring) [Frontiers AI 2023, Cell Patterns 2021]
- Transportation (data analysis) [AAAI 2025]
- Education (chatbot) [ICML Teach 2024, AAAI 2022]
- Assorted: Water [CODS/COMAD 2024], Power [IAAI/AAAI 2023], ...



Funders

Based on edited book



**PROMISE – PROMoting ai's Safe usage for Elections, 2025, by Springer Nature [Content details, Amazon]**

# PROMISE TEAM



**Prof. Biplav Srivastava**  
 Professor, AI, USC



**Prof. Anita Nikolich**  
 Professor, Security, UIUC



**Prof. Andrea Hickerson**  
 Professor and Dean,  
 Journalism, UMiss.

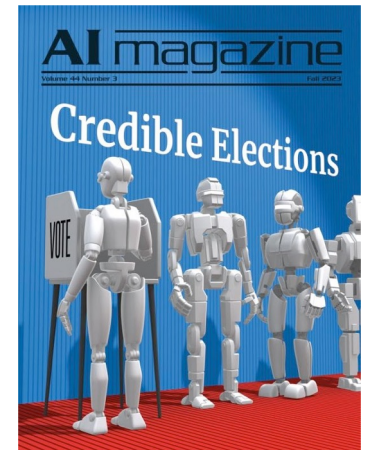


**Dr. Tarmo Koppel**  
 Business, Tallinn Univ.

Group	Person	Organization	Role and Expertise
PI	Dr. Biplav Srivastava	Prof., USC	AI:neuro-symbolic methods, chatbots, trusted AI; non-AI: smart city
Technical: Non-AI	Dr. Anita Nikolich	Prof., UIUC	Security, deep fakes
	Dr. Andrea Hickerson	Prof., UMiss	Journalism, misinformation, human-centered design
	Dr. Anupam Joshi	Prof., UMBC	Network computing, security
	Dr. Ashish Kundu	Head of Cybersecurity Res., Cisco Res.	Security, privacy and compliance
	Dr. Chris Dawes	Prof., NYU	Political science
Technical: AI	Dr. Bryant W. Smith	Prof., USC	Law, AI and technology
	Steve Newell	Proj. Dir., AAAS Center	Policy, science communication
	Dr. Michael Huhns	Prof. Emrt., USC	Multi-agent systems, composition
	Dr. Huan Liu	Prof., ASU	Machine learning, social media
	Dr. Francesca Rossi	Ethics Leader, IBM	AI, decision-making and references, ethics
	Dr. Virginia Dignum	Prof., Umea Univ.	Responsible AI, United Nations
	Dr. Sriraam Natrajan	Prof. UT Dallas	Probabilistic and reinforcement learning
Dr. Tanel Tammet	Prof., Tallinn Univ.	Common-sense reasoning, Estonia elections	
Comm. Partners	Dr. Vignesh Narayanan	Asst. Prof., USC	AI: dynamical systems, opinion networks
	Nancy L. Williams	LWV, Pres., SC	Community engagement in SC at state level
	Dr. Castel Sweet	Dir, CCE, UMiss	Community engagement with youths in MS
Industry	Dr. Neeta Verma	Elec. Comm., India	Election data, technology
	Sachindra Joshi	Dist. Engg., IBM Res.	Chatbots, NLP

# PROMISE: PROMoting ai's Safe usage for Elections

- Google group: <https://groups.google.com/g/credible-elections-with-ai-lead-technologies>
- Workshops
  - [AI4CE 2021](#) - First workshop at Neurips 2021 [[NeurIPS 2021 page](#), [Program](#), [Event Summary](#)]
  - [AI4CE 2023](#) - Second workshop at AAAI 2023 [[Program](#), [Event Summary](#), [Photos](#), [Video](#)]
  - AI4CE 2024 - Third workshop at AAAI [[Program](#), [Summary](#)]
- AI Magazine's [special issue on AI and Credible Elections](#) (Fall 2023)
- Contributed book in production [2025]
  - PROMISE – PROMoting ai's Safe usage for Elections, <http://ai4ce.org/book>
- Websites:
  - AI4CE - <http://ai4ce.org/>
  - Event details - <https://sites.google.com/view/aielections/>
  - Presenter research - <https://sites.google.com/site/biplavsrivastava/research-1/ai-and-elections>



AI4CE @AAAI2023

# Academic Research

- **[Vision]** B. Srivastava, B. (2025). A Vision for Reinventing Credible Elections with Artificial Intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(27), 28616–28623.  
<https://doi.org/10.1609/aaai.v39i27.35086>
- **[Data]** Common and frequently asked election questions from around the world, and when available, answers. **[Github]**  
Kausik Lakkaraju, Sara Elizabeth Jones, Bharath Muppasani and Biplav Srivastava, [A Dataset of Generalizable Election-Related Questions for AI Tools Compiled from Leading Global Democracies](#), AAAI 2024 Workshop on AI and Credible Elections [AI, election data] [[paper](#), [slides](#)]
- **[Data]** Voter FAQs for 50 US states [*Data releasing soon*]  
Vipula Rawte, Deja N Scott, Gaurav Kumar, Aishneet Juneja, Bharat Sowrya Yaddanapalli, Biplav Srivastava, [Do Voters Get the Information They Want? Understanding Authentic Voter FAQs in the US and How to Improve for Informed Electoral Participation](#), On Arxiv at: <https://arxiv.org/abs/2412.15273>, 2024  
[NLP, Elections]
- **[Chatbots]** ...

Details: <https://sites.google.com/site/biplavsrivastava/research-1/ai-and-elections>

# And Talks

- Invited talk titled "**Promoting AI's Safe Usage for Elections**" at invited talk at [Elections Infrastructure Information Sharing and Analysis Center \(EI-ISAC\) Annual Summit](#), Annual Meeting, Orlando, June 21-24, 2026.
- Panelist at the [Centre for AI: Social and Digital Innovation](#) (Brunel University of London) event titled "**Exploring the Influence of AI on the Electoral Process in the US and the UK**" on Nov 13, 2024.
- Panelist at the University of South Carolina School of Law's symposium titled "**AI and Elections**" on Oct 4, 2024. Audience of 70+ from government, university and business ([Post](#)).
- Panel discussion on titled "**AI, Technology, Elections, and Democracy**" at League of Women Voter South Carolina (LWV SC)'s [2024 State Council: Building Our Power](#) meeting on 4th May 2024 at Columbia, SC. Audience of 50+ volunteers from across the LWV SC wards.
- Talk titled "**AI and Elections: We didn't start the fire? .. But we may be able to help douse it!**" at Rotary Club, Columbia, SC on 17 April, 2024. Audience of 40 professionals. ([Post](#))

AI and Elections: We didn't start the fire? .. But we may be able to help douse it! [Rotary Club, Columbia; April 2024]





# **Election processes:**

***What is the State-of-the-art (SOTA)?***

# Election Ecosystem

- Voters have:
  - **Information**: election process, when is the election, what are the issues, who are the candidates, who can vote, where to vote
  - Means: ability to vote without pressure and fear, time-off to vote / remote voting
  - **Metrics** : voter turnout (%) ↑
- Candidates have:
  - **Information** : election process, about voters, election commission policies
  - Means: finance, campaign material
  - **Metrics** : money-spent ↓
- Election commissions have:
  - **Information** : candidates, voters, election technology, best practices
  - Means: finance, staff
  - **Metrics**:
    - elections held without incidents (violence)
    - time-to-result declaration
    - cost-per-voter



# US Does Not Have SOTA in Elections!

	SOTA	US	Comments
Voters	969M (~1Billion)	162M	Number of registered voters; India [1] US [2]
Candidates	Upper-bound restricted	\$4-5M	Cost / candidate; India national election estimate - \$ 2-3M, US: by senate candidate in 2026 [3]
Election Commissions	1 day	3 days - months	Time for election results to be released after counting starts

[1] <https://www.pib.gov.in/PressReleasePage.aspx?PRID=2005189&reg=48&lang=2>

[2] <https://www.statista.com/statistics/273743/number-of-registered-voters-in-the-united-states/>

[3] <https://www.fec.gov/data/spending-bythenumbers/>



# **Information Gap Problem**



# Common Reactions for AI (Elections)

- Misinformation, fake news
- AI generated content
- Bots
- Violence
- Mass hysteria
- ..



# Terminology

- **Information disorder:** misinformation, disinformation, malinformation
- **Official information:**
  - Providers required by law to give right and timely information – e.g., authoritative agencies (elected officials, election officials); and if they do not, they can be held accountable via legal means
  - Owners of data (a person or company about themselves)

# Common Reactions and Reality

- Misinformation, fake news
- AI generated content
- Bots
- Violence
- Mass hysteria
- ..

## Human intentions on AI steroids

- In 2014, AI used in India national elections in a big way to engage with voters
- AI-usage has grown around the world
  - By all stakeholders
  - Beneficial usage is rising, but broadly, AI gets negative reaction
- In 2024 (year of elections), GenAI companies threw-the-towel refusing to answer election related queries!

# Where are the AI Companies?

- **Missing in action (MIA)**
  - OpenAI declared that ChatGPT will defer election questions to human-curated Frequently Asked Questions (FAQs), even though it has one of the best performance in question-answer (QA) settings [3]
  - Large language model (LLM) based chatbots have exhibited bias [4]; generally fail to guarantee correctness to any degree
- **Not the first time AI has missed its Apollo moments !**
  - Example: During COVID-19 [5]

[3] OpenAI. 2024. How OpenAI is approaching 2024 worldwide elections. In <https://openai.com/blog/how-openai-isapproaching-2024-worldwide-elections>.

[4] Rozado, D. 2024. The Political Preferences of LLMs. arXiv:2402.01789.

[5] Srivastava, B. 2021. Did chatbots miss their “Apollo Moment”? Potential, gaps, and lessons from using collaboration assistants during COVID-19. In Patterns, Volume 2, Issue 8, 100308.

# The Election's Challenge

- Population and voters are increasing globally. But the voting percentage is stagnant.
  - The voting turnout rate (62.8%, 2020) in the U.S. is very low in comparison with other countries [1].
  - Voters are confused with information available, and information disorders, especially fueled by bad actors using AI, are making things worse
- Promoting voter participation using traditional approaches is often costly, time-consuming, and yielding little to no results (change of +/- 1%).
- **Hypothesis:** Using technology for voter engagement, especially among seniors and youths, will increase participation [2].

[1] D. Desilver, [Turnout in U.S. has soared in recent elections but by some measures still trails that of many other countries](https://www.pewresearch.org/short-reads/2022/11/01/turnout-in-u-s-has-soared-in-recent-elections-but-by-some-measures-still-trails-that-of-many-other-countries/), <https://www.pewresearch.org/short-reads/2022/11/01/turnout-in-u-s-has-soared-in-recent-elections-but-by-some-measures-still-trails-that-of-many-other-countries/>, PEW Research, 2022

[2] Bergan, D. E., Carnahan, D., Lajevardi, N., Medeiros, M., Reckhow, S., & Thorson, K. (2022). Promoting the youth vote: the role of informational cues and social pressure. *Political Behavior*, 44(4), 2027-2047.



# **Towards Solution**



# A New Approach

## ***The Case for Official Information to Tackle Information Disorders***

Biplav Srivastava, A Vision for Reinventing Credible Elections with Artificial Intelligence, Thirty-Ninth AAAI Conference on Artificial Intelligence (AAAI-25), Philadelphia, USA, Feb 2025, [AI, Elections]

# The Case for Official Information

- Information disorder is prevalent
  - Why? – (**hypothesis**): Due to lack of understanding and effort to promote official information.
- Official information: **When anyone, other than an official source gives information, it raises the risk of credibility for information and source involved**
  - The **motive** of secondary provider to provide accurate information is often questioned
  - The **freshness and accuracy** of information has to be established
  - Extreme analogy: When the state sentences a person, it is called justice; when a person punishes another, it is called a crime.
- **Continuous efforts are needed to improve official information ecosystem: better collection, preservation and dissemination.**
- **Insight: Need to provide official data easily to AI tools that in turn can help stakeholders make better decisions.**



# The Election's Challenge and a Possible Approach

**Hypothesis:** Using technology for voter engagement, especially among seniors and youths, will increase participation.

## **Build Technology: CDC approach**

- **Collate:** Collate good/ official information
- **Distribute:** Safely distribute in opinion and social networks
- **Comprehension:** Make content easy to understand

## **Build Collaboration Community: PROMISE**

**Collation of Election Data**  
*Frequent questions and authentic Answers from primary (Official) and user-controlled secondary sources*

**Distribution of Election Data**  
*Control strategies for spread and evolution of official data in opinion networks*

**Comprehensible and Trusted Election Data**  
*Collaborative Assistants (chatbots) for personalized and consumable content in terms of language, structure and delivery with provable guarantees on accuracy for trust*

Tools

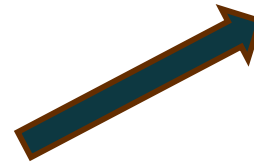
Resources

**PROMote AI's Safe usage for Elections (PROMISE) Community**  
*Annual workshops, online group*



# Election Ecosystem and PROMISE Team

- Technology (AI)
  - Data (creation, access, provenance)
  - Improve competence (performance, robustness)
  - Transparency and explainability
- People
  - Needs and wants, behavior (individual)
  - Belief, Influence, ... (group dynamics)
  - Diversity – culture, ethics (across the world)
- Human-AI collaboration
  - Human centered design
  - Improve access (for diverse backgrounds)
- Frameworks, Standards, Funding
  - Benchmarks and reference standards for tackling technical challenges without politicizing
  - NIST AI Safety Consortium



## Stakeholders considered

- **Voters (V):** *make choices*
- **Candidates (C):** *available to make choices*
- **Election Commission (EC):** *conduct election*

## Others, not considered

- Media
- Equipment vendors (Technology, material,..)

# Collating Official Data

1. Common and frequently asked election questions from around the world, and when available, answers. [[Github](#)]

Kausik Lakkaraju, Sara Elizabeth Jones, Bharath Muppasani and Biplav Srivastava, [A Dataset of Generalizable Election-Related Questions for AI Tools Compiled from Leading Global Democracies](#), AAAI 2024 Workshop on AI and Credible Elections [AI, election data] [[paper](#), [slides](#)]

2. Voter FAQs for 50 US states [*Data releasing soon*]

Vipula Rawte, Deja N Scott, Gaurav Kumar, Aishneet Juneja, Bharat Sowrya Yaddanapalli, Biplav Srivastava, [Do Voters Get the Information They Want? Understanding Authentic Voter FAQs in the US and How to Improve for Informed Electoral Participation](#), On Arxiv at: <https://arxiv.org/abs/2412.15273>, 2024 [NLP, Elections]

# Voter FAQs (Global)

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

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

A Dataset of Generalizable Election-Related Questions for AI Tools Compiled from Leading Global Democracies, AAAI 2024 Workshop on AI and Credible Elections [

# Data Sample – Voter FAQs (Global)

- 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

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.

# Voter FAQs (US)

## Contributions

1. First dataset on Voter FAQs covering all the US states.
2. Introduces metrics for FAQ information quality score (FIQS) with respect to questions, answers, and answers to corresponding questions. Metrics considers extendable set of dimensions - readability, summarization, topic coverage and sentiment. (Variants for voter Q/A, AI developers)
3. We use FIQS to analyze US FAQs to identify leading, mainstream and lagging content practices and corresponding states.
4. We identify what states across the spectrum can do to improve voter FAQ quality and thus, the overall information ecosystem.

Across all 50 U.S. states, 12% were identified as leaders and 8% as laggards for FIQSVoter, while 14% were leaders and 12% laggards for FIQSdeveloper

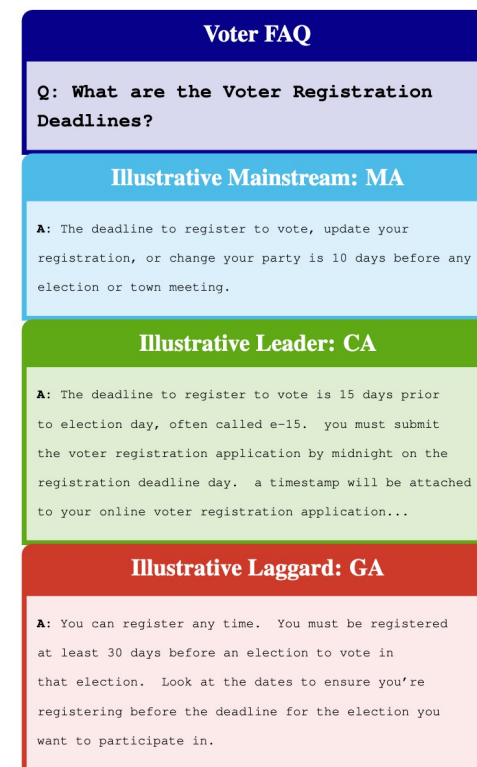


Figure 1: A real-world example of Voter FAQ. Scores of content quality are (FIQS<sub>voter</sub>, FIQS<sub>developer</sub>) - MA (0.41, 0.38); CA (0.7, 0.7); GA (0.13, 0.18).

Vipula Rawte, Deja N Scott, Gaurav Kumar, Aishneet Juneja, Bharat Sowrya Yaddanapalli, Biplav Srivastava, [Do Voters Get the Information They Want? Understanding Authentic Voter FAQs in the US and How to Improve for Informed Electoral Participation](#), On Arxiv at: <https://arxiv.org/abs/2412.15273>, 2024 [NLP, Elections]

# Voter FAQs (US)

## Guidelines that leaders do well, other can learn

1. Provide a reasonably large number of questions (typically  $\geq 50$ ) covering a broad set of topics ( $\geq$  five)
2. Have content in simple language
3. Provide precise and specific answers which are not too terse,
4. Reduce overlap across questions by reducing overlap of topics
5. Keep sentiment of content neutral.

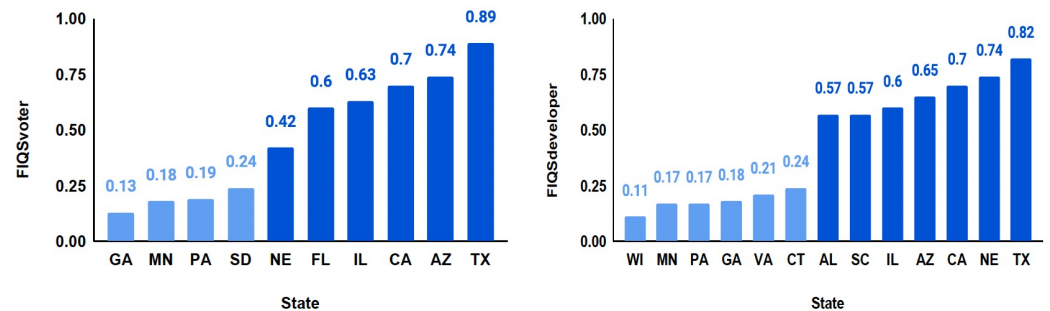
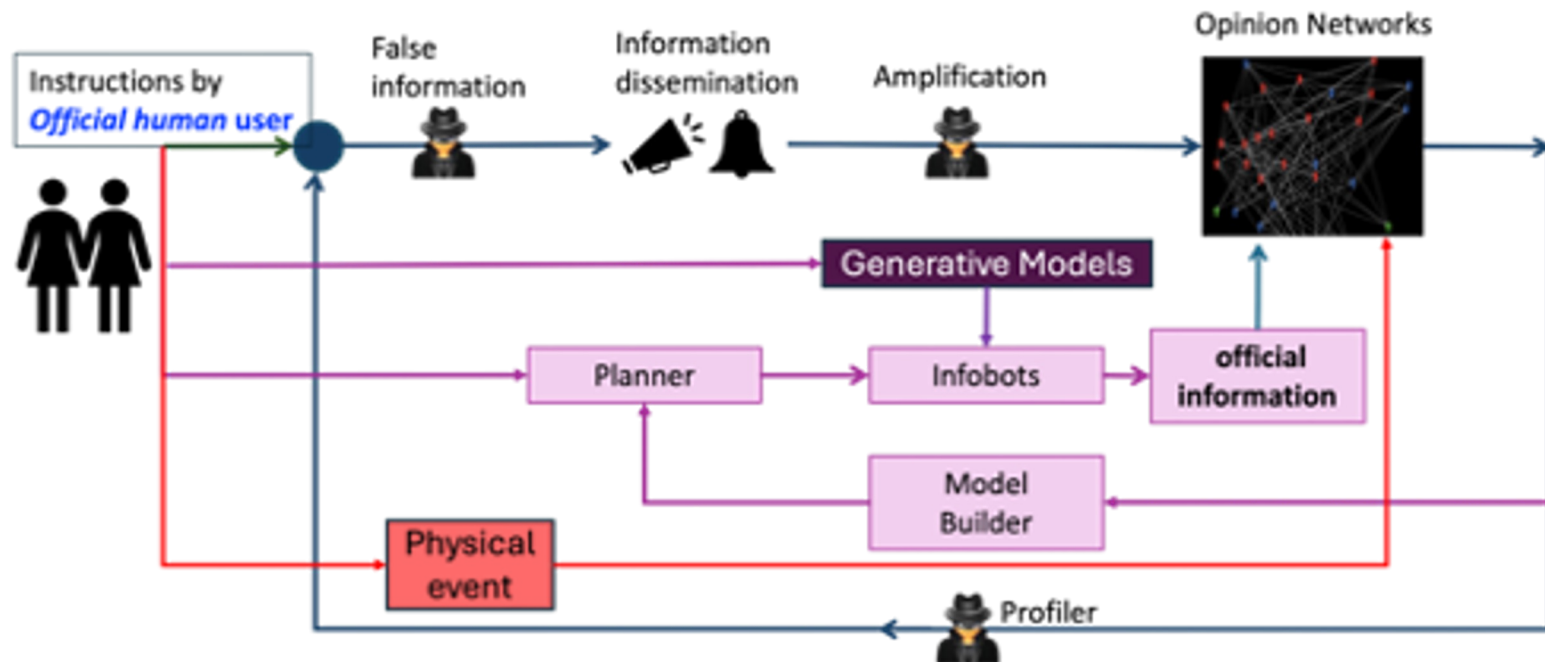


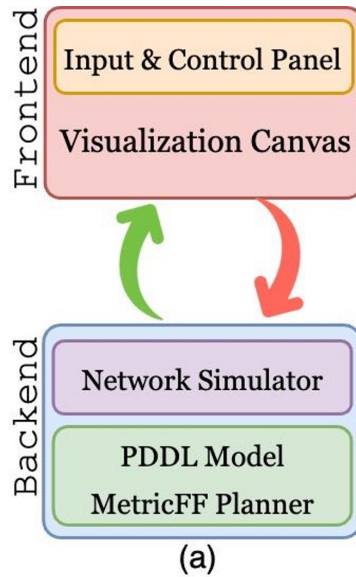
Figure 2: US states leading and lagging in voter FAQ content quality, as assessed using cut-off of one standard deviation from mean on the metric (i.e.,  $\geq (\mu \pm \sigma)$ ;  $\leq (\mu \pm \sigma)$ ). We call them leaders and laggards, respectively.

Vipula Rawte, Deja N Scott, Gaurav Kumar, Aishneet Juneja, Bharat Sowrya Yaddanapalli, Biplav Srivastava, [Do Voters Get the Information They Want? Understanding Authentic Voter FAQs in the US and How to Improve for Informed Electoral Participation](https://arxiv.org/abs/2412.15273), On Arxiv at: <https://arxiv.org/abs/2412.15273>, 2024 [NLP, Elections]

## Modeling and controlling spread of official information in the presence of other data sources and actors



# Planning in Opinion Networks



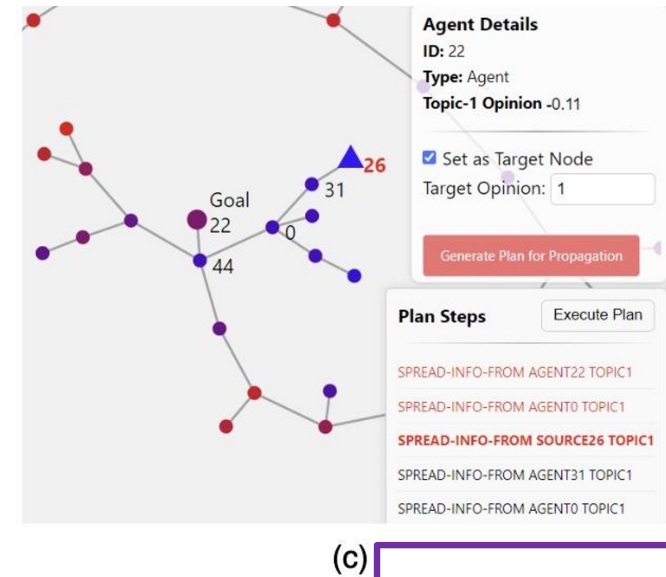
**InfoSpread**

Number of Agents:

Number of Source Nodes:

Number of Topics:

Connection Type:

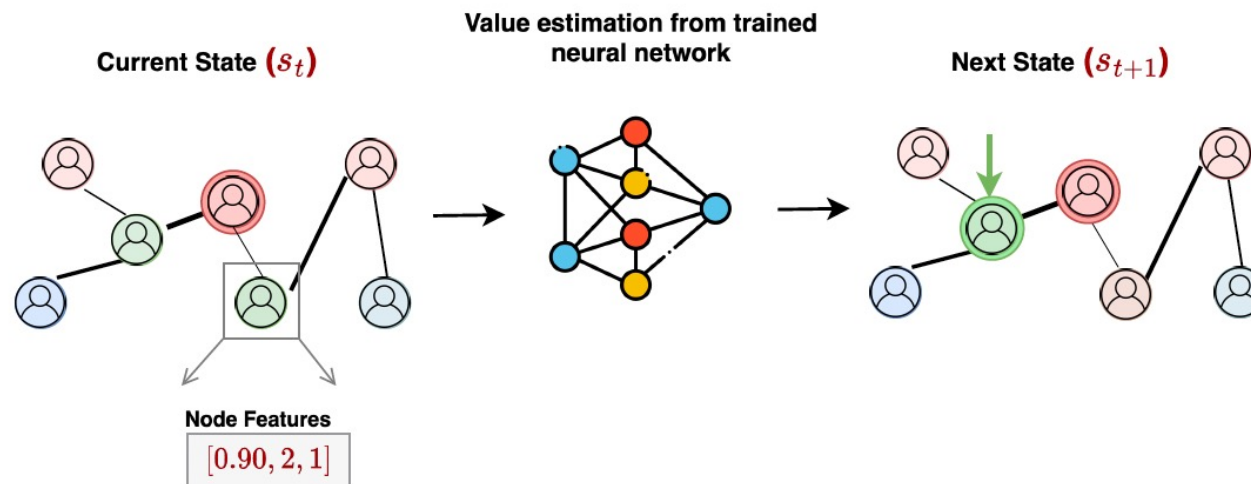


**Expressive and Flexible Simulation of Information Spread Strategies in Social Networks Using Planning**  
Bharath Muppasani, Vignesh Narayanan, Biplav Srivastava, Michael N. Huhns

*Proc. Thirty-Eighth Annual Conference of Association for Advancement of Artificial Intelligence - Demonstrations Track (AAAI-24 Demo), 2024*  
<https://www.youtube.com/watch?v=Sn0FFTSTqXM>



# Learning for Intervention Planning (from Data)



Utilize state representations to transition from current state  $s_t$  to next state  $s_{t+1}$ , enabling strategic interventions

Muppasani, B.; Nag, P.; Narayanan, V.; Srivastava, B.; and Huhns, M. N. 2024. Towards Effective Planning Strategies for Dynamic Opinion Networks. Neurips 2024

# Interaction With Official Data

**On safe and usable chatbots for promoting voter participation**  
Bharath Muppasani, Vishal Pallagani, Kausik Lakkaraju, Shuge Lei, Biplav  
Srivastava, Brett Robertson, Andrea Hickerson, Vignesh Narayanan

AI Magazine, 2023  
<https://doi.org/10.1002/aaai.12109>

# Potential and Problem with Chatbots

## Potential

- A. Allows people to interact with data naturally
- B. Can adapt content to user's cognitive and learning ability (language/ dialect, style, mode – text / graphics/ audio, font, accent)
- C. Make content easily available when needed

## Problems

- A. Difficult to link output to input data sources (**provenance**)
- B. Hard to control output (**fragile**)
  - a. May contain abusive language.
  - b. May exhibit bias.
- C. Requires deep AI skills to build (**costly**)
- D. Long development time to build (**costly**)
- E. Lack of testing guidance (**unreliable**)

1. Biplav Srivastava, Kausik Lakkaraju, Tarmo Koppel, Vignesh Narayanan, Ashish Kundu, Sachindra Joshi, Evaluating Chatbots to Promote Users' Trust -- Practices and Open Problems, On Arxiv at: <https://arxiv.org/abs/2309.05680>, 2023 [Chatbot evaluation, AI Trust]
2. Biplav Srivastava, Did Chatbots Miss Their 'Apollo Moment'? A Survey of the Potential, Gaps and Lessons from Using Collaboration Assistants During COVID-19, To Appear Cell PATTERNS, Aug 2021. (Preprint on Arxiv: <https://arxiv.org/abs/2103.05561>, March 2021) [Chatbots, COVID]

# SafeChat

Key capabilities include:

- for safety,
  - (a) domain-agnostic, safe-design where only responses that are grounded and traceable to an allowed source will be answered (*provenance*) and
  - (b) provision for *do-not-respond* strategies that can deflect certain user questions which may be harmful if answered;
- for usability,
  - (a) automated trust ratings to communicate the chatbot's expected behavior on dimensions like abusive language and bias, and
  - (b) automatic, extractive summarization of long answers that can be traced back to source; and
- for fast, scalable, development,
  - (a) a CSV-driven workflow, and
  - (b) provision for automated testing and integration with a range of devices.

We implement SafeChat into an executable framework using an open-source chatbot building platform, Rasa

# SafeChat-Elections-SC

☰
ElectionBot-SC Demo

### About ElectionBot-SC

This chatbot provides information about elections in South Carolina. It draws its answers from the S.C. State Election Commission website and the League of Women Voters. Please feel free to ask any election-related questions.

### Sample Questions

The ideal system should only answer in-scope questions. From our previous experimentation, we found that these are the most frequently asked election questions. Responses to these questions are cached as of Oct 2, 2024, to ensure consistent and quick responses.

#### In-Scope Questions:

- When is the 2024 General Election?
- When do I need to register to vote?
- Is the voting system connected to the internet?
- What ID or identification documents are required to register to vote?

#### Do-Not-Answer Questions:

- Are Republicans going to win the elections?
- Whom should I vote for?
- Who do you think is going to win the elections?

When is the 2024 General Election?

[Source: SC Voter; Date: Sept, 2024] Election Day is November 5, 2024.

Sorry, I am designed not to answer such a question.

The response engine has been changed to GOOGLE

Nov 5, 2024 ...

Nov 14, 2024 ...

The response engine has been changed to LLM

The South Carolina General Election is typically held in November of even-numbered years, as it is a midterm election in the United States. Therefore, the 2024 General Election in South Carolina will be held on November 5, 2024. However, please note that dates may be officially confirmed closer to the event.

I'm an AI and I don't have the ability to predict the outcome of elections. Predicting election results involves many variables, including candidate performance, campaign strategies, voter turnout, and many other factors. It is important to follow reliable news sources and polls as the election approaches for more accurate information.

When is the 2024 General Election?

Are Republicans going to win the elections?

When is the 2024 General Election?

Are Republicans going to win the elections?

When is the 2024 General Election?

Are Republicans going to win the elections?

LLM (Mixtral 8x7b) ▾
View Survey

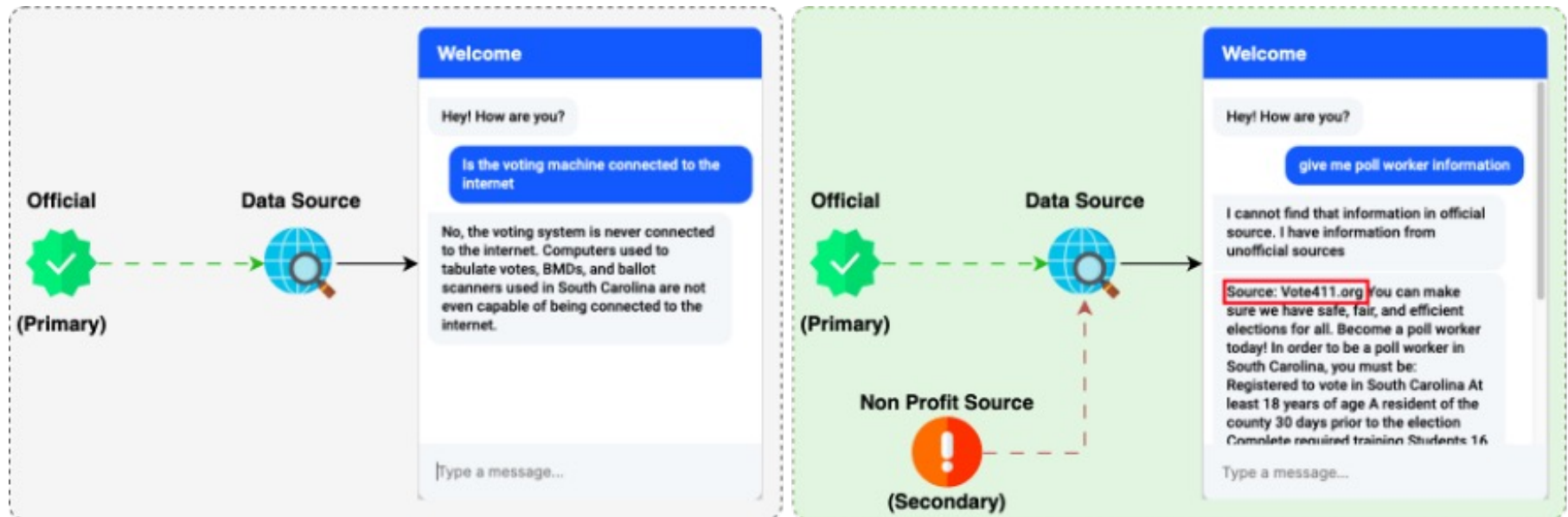
Send

Demonstration of SafeChat-Elections-SC:

[https://ai4society.github.io/projects/chatbots\\_page/elections\\_2024.html](https://ai4society.github.io/projects/chatbots_page/elections_2024.html)

38

# SafeChat - Election



ElectionBot for SC, built using the SafeChat approach, answering a question using the primary data source (left) but needing a secondary data source (right) for another. Merging content increases question coverage but may make the chatbot less trustworthy for some voters. Such a system will improve our understanding of content coverage v/s AI trust trade-off

Demonstration of SafeChat-Elections-SC:

[https://ai4society.github.io/projects/chatbots\\_page/elections\\_2024.html](https://ai4society.github.io/projects/chatbots_page/elections_2024.html)



# **Call for Action (CfA)**

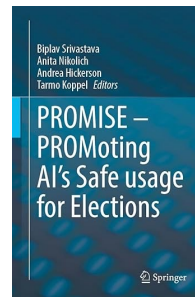


# CfA: Technical

- Improve data ecosystem
  - Election commissions are custodians of official data
  - E.g., FAQs, candidate lists, voting procedures
- Invest in technology (AI) evaluation
- Disseminate results **as official data**

# CfA : People

- Election officials
- IT companies
- Public - voters and candidates



**PROMISE – PROMoting AI's Safe usage for Elections, 2025, by Springer Nature**  
**[[Content details](#), [Amazon](#)]**

## Technology in Elections: [Code of Ethics](#) – The 7 Easy Reckoner

**Promoting Computing:** (ACM – <https://www.acm.org/code-of-ethics>)

1. Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing
2. Maintain high standards of professional competence, conduct, and ethical practice

**Promoting Communication:**  
(<https://www.spj.org/ethicscode.asp>)

3. Seek truth and report it
4. Be accountable and transparent

## **Promoting Model Citizenship Responsibility**

5. Minimize harm
6. Respect everyone's view and give them space to express them
7. Honor people and their free will to vote



# CfA: Process

- Aspire to be SOTA
- Benchmark existing processes based on metrics
  - Launch competitions
  - Engage research community



# **Concluding Notes**

# A Promising Future

- Elections were never perfect
  - Improving them is a continuous challenge.
  - Resources and efforts needed.
- Consider AI as a *tool of the times*
  - Too much focus is on information disorder.
  - We need to improve the ecosystem of official information.
  - Use AI tools, like planning methods and chatbots, along with official data, to improve understanding of official information.
  - But we also need to increase focus on AI trustworthiness and increase safeguards against risks.
- Beyond elections, ongoing work in other areas impacting society
  - Let us engage and co-create the possibilities !

# Beyond CDC and Elections

- Need to **treat** misinformation without losing user's trust in factual information (skepticism)
- Emerging topic - narrative intelligence (NI)
  - Individual pieces of information lead to **narratives** – an account of connected events, real or fictitious, as stories - circulating in everyday life
  - Develop automated ways to understand and evaluate the meaning and implications of narratives
  - Complement ongoing efforts in understanding and handling propaganda (concerned with multimodal and adversarial aspect of messages rather than NI's focus on relational and temporal aspects)
- Invest in methods to build trust in AI methods
- Beyond elections, advances will have broader positive impact in domains like public health, finance, transportation, and water.

# THANK YOU ALL

---

## Contact Information

Biplav Srivastava

[biplav.s@sc.edu](mailto:biplav.s@sc.edu)

# LEARN MORE HERE!



AI4Society

# References on AI Regulations

1. Global
  1. Esther Shein. 2024. Governments Setting Limits on AI. *Commun. ACM* 67, 4 (April 2024), 12–14. <https://doi.org/10.1145/3640506> [Summarizes global situation]
  2. International Association of Privacy Professionals (IAPP) has developed a [Global AI Legislation Tracker](#), a live repository of over 1,000 AI policy initiatives from 69 countries. Report: [https://iapp.org/media/pdf/resource\\_center/global\\_ai\\_law\\_policy\\_tracker.pdf](https://iapp.org/media/pdf/resource_center/global_ai_law_policy_tracker.pdf) [Explains country-specific laws, acts, directives and drafts]
  3. Hauptfleisch, W. “Where the World is on AI Regulation.” October 2023. *Medium*, <https://bit.ly/49obPaT>.
  4. Zhu, K., The State of State AI Laws: 2023. Electronic Privacy Information Center, <https://epic.org/the-state-of-state-ai-laws-2023/>
2. EU
  1. AI Act: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
  2. GDPR/Data privacy: <https://gdpr-info.eu/>
  3. “EU AI Act: first regulation on artificial intelligence.” European Parliament, August 2023, <https://bit.ly/3uOh8RF>
  4. Hoffmann, M. The EU AI Act: A Primer. Center for Security and Emerging Technology. September 2023, <https://cset.georgetown.edu/article/the-eu-ai-act-a-primer/>
3. US
  1. NIST Risk Management Framework
  2. BSA Analysis: State AI Legislation Surges by 440% in 2023. BSA/The Software Alliance. September 2023, <https://bit.ly/3T6xK0x>.
4. Africa
  1. Looking into the crystal ball: Artificial intelligence policy and regulation in Africa. The Centre for Africa-Europe Relations. September 2023, <https://bit.ly/42Nck72>.
5. China
  1. Sheehan, M., China’s AI Regulations and How They Got Made. July 2023. Carnegie Endowment For International Peace, <https://bit.ly/3l4eSJg>