Generative Social Choice: The Next Generation

ICML 2025

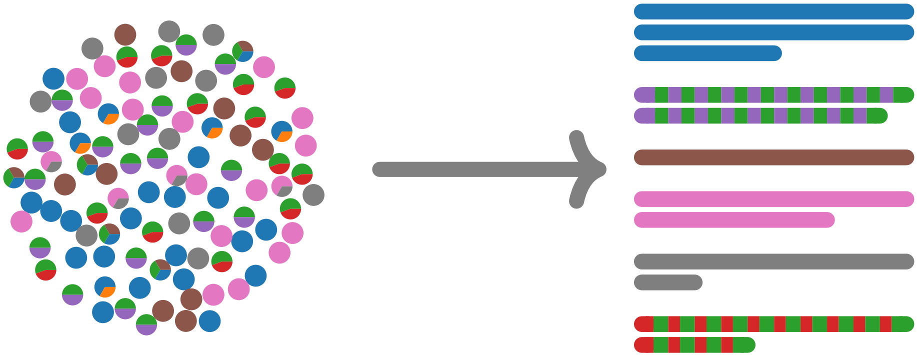
Niclas Boehmer, Sara Fish, Ariel Procaccia

July 17, 2025

Link to paper

Problem Statement

Given a large dataset of diverse opinions, how can we $\underbrace{\textbf{proportionally summarize}}_{x\% \text{ of users ``control''} \simeq x\% \text{ of words}, \forall x}$ them?



Motivation

Proportional summarization has a diverse array of potential applications.

Motivating application: **AI and democracy**; specifically, *collective response* systems like **Polis**.

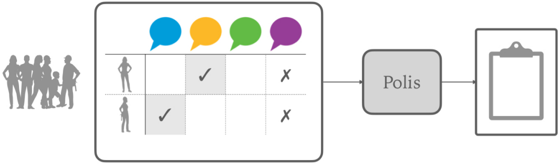


Figure source: Daniel Halpern

LLM-based methods allow for greater flexibility in such processes (for both inputs and outputs).

However: ad hoc LLM-based methods lack reliability, robustness, and interpretability.
$\hookrightarrow$ *E.g., how to ensure the LLM does not suppress or overweight fringe opinions?*

Approach: Generative Social Choice

Our goal: **proportionally summarize** user opinions using a $\underbrace{\text{trustworthy, LLM-based}}_{\text{How to unite these conflicting objectives?}}$ process.

Our approach (initiated by [FGPPRSW'23]): **generative social choice** query framework.
Related to broader literature on AI alignment with guarantees, see e.g. Wu and Hartline (2024).

Theory: Specify process using (black-box) queries

Process $P :=$ algorithm that uses queries $\square$, $\bigcirc$.

Thm. Under assumptions about $\square$, $\bigcirc$, process P satisfies proportionality/runtime/... guarantees.

Our work: social choice $\rightarrow$ proportionality guarantees.

Instantiation: Implement and empirically test queries

Implement $\square$, $\bigcirc$, typically leveraging LLMs

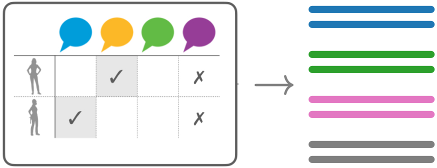
Evaluate $\square$, $\bigcirc$ using relevant data (e.g. Polis)

Our work: PROSE, a general-purpose implementation.

Key observation: establishing trust in the queries is sufficient for trust in the whole process.

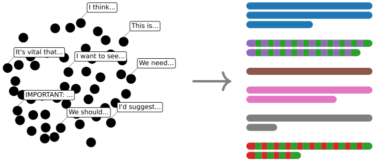
Generative Social Choice... The Next Generation

Generative Social Choice [FGPPRSW'23]



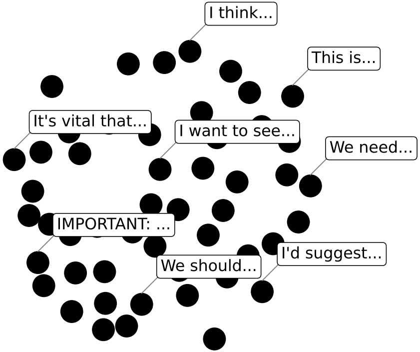
1. User sets number of statements k
2. Each statement represents $1/k$ of users
3. Guarantees only for perfect query results
4. Implementation for structured user data

The Next Generation



1. Algorithm adaptively chooses k
2. Variable statement lengths (support $\propto$ length)
3. Process and guarantees for noisy query results
4. Flexible implementation compatible with unstructured user data

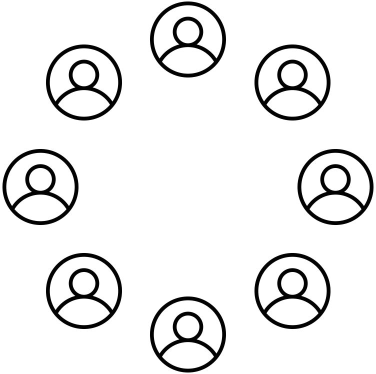
Input



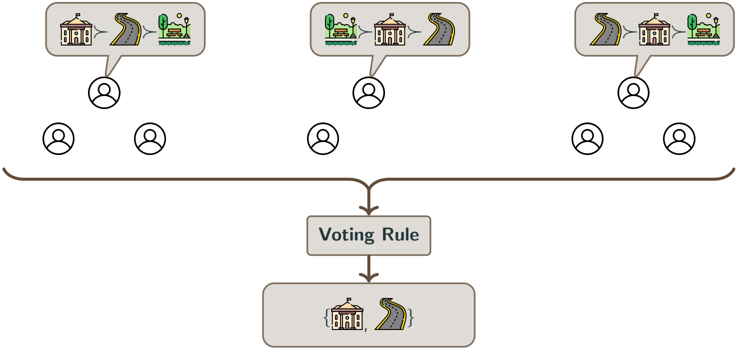
$+$ desired length B of output slate (in words)

Social Choice: Theory of Collective Decision Making

 *or*  *or*  *?*



Social Choice: Theory of Collective Decision Making



Proportional summarization

- Voters: participating users.
- Candidates: all possible statements.

Task: Select statements with total length $\leq B$.
$\rightsquigarrow$ Participatory budgeting
(theoretically very well understood)

Novelty Virtually infinite candidate space.

Slate Generation via Social Choice

Generative Query

Data of users, approval level r, length c

Most-liked statement of at most c words at level r

Discriminative Query

User data + statement

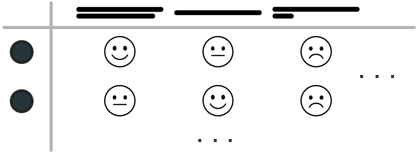
How much user agrees with statement

Slate Generation Algorithm

Initialize user set , slate .

For $\in\{$, ..., ,..., $\}$ and $c \in \{B, B-1, \dots, 1\}$

- Generate statements (, , c) for $\geq$
- Using discriminative query , compute:



- Pick with most $\geq$
- If (# with $\geq$)$\cdot\frac{B}{n}$ $\geq$wordcount( ):

Delete covered users + add statement to slate

PROSE and Experimental Setup

PROSE Query Implementations (GPT-4o-2024-11-20):

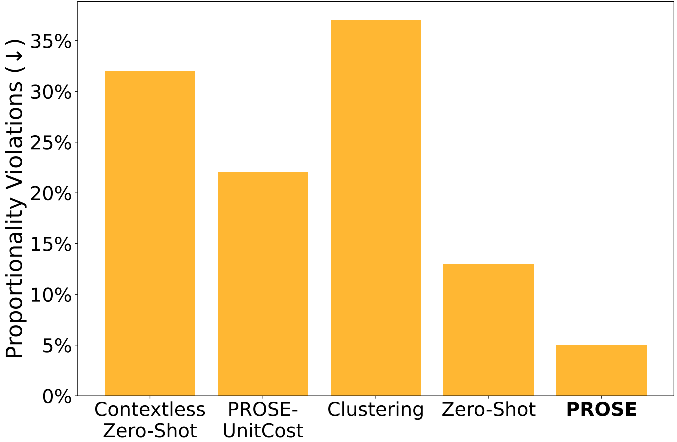
- user, statement $\longrightarrow$ Disc $\longrightarrow$ approval score computed based on two (fast) LLM calls
- users $\longrightarrow$ $\boxed{\text{Gen}}$ $\longrightarrow$ statement identify cohesive group (clustering) + generate consensus statement (LLM)

Datasets $\overbrace{\textsc{BirthControl}, \textsc{BirthControlSkew}, \textsc{Obesity}}^{\text{UCI Drugs Dataset}}$, $\overbrace{\textsc{Bowling Green}}^{\text{Polis}}$

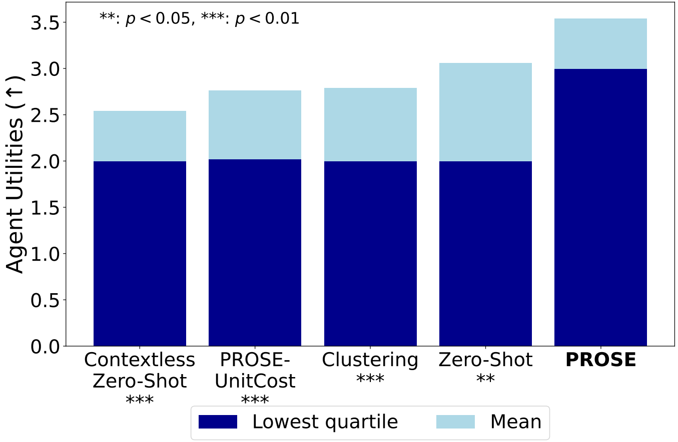
Baselines

- *Contextless Zero-Shot*: given topic and budget, generate slate in single response
- *Zero-Shot*: given topic, budget, and user data, generate slate in single response
- *Clustering*: clustering of embedded user data + LLM-generated cluster summaries
- *PROSE-UnitCost*: PROSE with unit-length statements (resembling [FGPPRSW'23]'s approach)

Experiments: Results on Bowling Green



Experiments: Results on Bowling Green



(Utilities computed using "independent", expensive CoT-based →Disc→)

High-Level Takeaways

(1) Increased trustworthiness of LLM-based algorithms via query framework

(2) LLMs can enable new forms of civic participation